SPhT/01-???

ANOMALIES IN GAUGE THEORIES*

J. ZINN-JUSTIN

*CEA-Saclay, Service de Physique Théorique**, F-91191 Gif-sur-Yvette Cedex, FRANCE*

and

Université de Cergy–Pontoise

email: zinn@spht.saclay.cea.fr

ABSTRACT

To define a quantum field theory, one often starts from a classical lagrangian or action. When the classical theory has a symmetry, global or local like in gauge theories, formal manipulations lead to believe that the symmetry can also be implemented in the quantum theory, provided one uses the proper quantization rules. While this is often true, it is not a general property and therefore requires a proof because simple formal manipulations ignore the unavoidable divergences of perturbation theory.

The existence of invariant regularizations allows to solve the problem in most cases but the combination of gauge symmetry and chiral fermions leads to subtle issues. Depending on the specific group and field content, *anomalies* are found: obstructions to the quantization of gauge symmetries. Because anomalies take the form of local polynomials in the fields, are linked to local group transformations, but vanish for global (rigid) transformations one discovers that they have a *topological* character.

Divergences in quantum field theory have a double origin: they are a consequence of the infinite number of degrees of freedom in quantum fields, but they may also be generated in a local theory by the problem of ordering quantum operators. The latter problem arises in boson theories with derivative couplings and in many fermion theories.

---

*Contribution to the School "Topology and Geometry in Physics", Root 24–28 september 2001.

**Laboratoire de la Direction des Sciences de la Matière du Commissariat à l'Energie Atomique

Regularization is a temporary modification at short distance or large momentum of a quantum field theory which renders the perturbative expansion finite. When a group invariant regularization can be found, a symmetric renormalized quantum field theory can be defined. This is generally easy for global linear symmetries, because the group operations involve no commutations of quantum operators. It is less obvious for non-linear global symmetries or gauge symmetries due to the presence of derivative couplings.

For many years this problem had been discussed only within the framework of perturbation theory. New non-perturbative solutions based on lattice regularization have recently been proposed.

In these notes we review various perturbative and non-perturbative regularization techniques, and show that they leave room for possible anomalies when both gauge fields and chiral fermions are present. We determine the form of anomalies in simple examples, and exhibit gauge instantons which contribute to the anomaly. We briefly mention a few physical situations where anomalies are relevant.

For more details (including some aspects of euclidean field theory) and references see

J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Clarendon Press (Oxford 1989, third ed. 1996).

Contents

1 Symmetries and regularization 3
2 Momentum cut-off regularization 5
2.1 Matter fields 5
2.2 Abelian gauge theory 9
2.3 Non-abelian gauge theories 11
3 Other regularizations 12
3.1 Dimensional regularization 12
3.2 Lattice regularization 14
3.3 Boson field theories 14
3.4 Fermion and the doubling problem 16
4 The abelian anomaly 17
4.1 Abelian axial current and abelian vector gauge field 18
4.2 Explicit calculation 21
4.3 Non-abelian vector gauge fields and abelian axial current 27
4.4 Anomaly and eigenvalues of the Dirac operator 28
4.5 Anomaly and instantons 30
5 Non-abelian anomaly 32
5.1 General axial current 32
5.2 Obstruction to gauge invariance 33
5.3 Wess–Zumino consistency conditions 34
6 Lattice fermions: Ginsparg–Wilson relation 35
Bibliographical Notes 40

1 Symmetries and regularization

Divergences. Symmetries of the classical lagrangian do not always translate into symmetries of the corresponding quantum theory. Indeed quantum field theories are affected by UV divergences which invalidate simple algebraic proofs.

The origin of UV divergences in field theory is double. First a field contains an infinite number of degrees of freedom. The corresponding divergences are directly related to renormalization group and reflect the property that even in renormalizable quantum field theories degrees of freedom remain coupled on all scales.

However another of type of divergences can appear, which is related to the order between quantum operators, and the transition between classical and quantum hamiltonians. Such divergences are already present in perturbation theory in ordinary quantum mechanics, for instance in the quantization of the geodesic motion on a manifold (like a sphere). In the case of forces linear in the velocities (like a coupling to a magnetic field) one finds instead only finite ambiguities. In local quantum field theory the problem is even more severe because for a scalar field for example the commutation between field operator $\hat{\phi}$ and conjugate momentum $\hat{\pi}$ takes the form

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\hbar\, \delta^{d-1}(x-y).$$

In a local theory all operators are taken at the same point and thus the commutator is divergent except in quantum mechanics ($d = 1$ with our conventions).

Divergences of this nature thus are present as soon as derivative couplings are involved, or when fermions are coupled to gauge fields (because fermions are never classical). They reflect the property that the knowledge of the classical action is not sufficient, in general, to completely determine the quantized theory.

Regularization. Regularization is a useful intermediate step in the renormalization program, which consists in modifying the initial theory at short distance, large momentum or otherwise to render perturbation theory finite. Note that from the point of view of Particle Physics all these modifications alter in some essential way the physical properties of the theory, and thus can only be considered as intermediate steps in the removal of divergences.

When a regularization can be found which preserves the symmetry of the initial classical action, a symmetric quantum field theory can be constructed.

Various regularization schemes, like Pauli–Villars's regularization, are specifically designed to cut the infinite number of degrees of freedom by cutting large momenta. With some care these methods will preserve formal symmetries of the un-renormalized theory which correspond to global (space-independent) linear group transformations. Problems may, however, arise when the symmetries correspond to non-linear or local transformations, like in the examples of non-linear σ models or gauge theories, due to the unavoidable presence of derivative couplings. It is easy to verify that in this case regularizations which only cut momenta are in general not applicable or sufficient to provide a complete regularization. Other methods have to be explored. In many examples dimensional regularization solves the problem because then the commutator between field and conjugated momentum taken at the same point vanishes. However in the case of chiral fermions dimensional regularization fails because chiral properties are specific to even dimensions.

Of particular interest is the method of lattice regularization, because it can be used, beyond perturbation theory, either to discuss the existence of a quantum field theory, or to determine physical properties of field theories by non-perturbative numerical techniques. One verifies that such a regularization indeed specifies an order between quantum operators, and therefore, when applicable solves the ordering problem of non-linear σ-models or gauge theories. However, again it fails in presence of chiral fermions: the manifestation of this difficulty takes the form of a doubling of the fermion degrees of freedom. Until recently this has prevented a straightforward numerical study of chiral theories.

Anomalies. That no conventional regularization scheme can be found in the case of gauge theories with chiral fermions is not surprising since we know examples of theories with anomalies, i. e. theories in which a local symmetry present at the classical level cannot be implemented in the quantum theory. Note that this creates obstructions to the construction of chiral gauge theories because exact gauge symmetry, and thus the absence of anomaly, is essential for the physical consistency of a gauge theory.

Anomalies are local quantities because they are consequences of short distance singularities. They are responses to local (space-dependent) group transformations but vanish for a class of space-independent transformations. This gives them a topological character that is further confirmed by their relations with the index of the Dirac operator in a gauge background.

The recently discovered solutions of the Ginsparg–Wilson relation and the method of overlap fermions seem to provide an unconventional solution to the regularization problem in gauge theories with chiral fermions on the lattice. They evade the doubling problem of fermion because chiral transformations are no longer strictly local on the lattice, and relate the problem of anomalies with the invariance of the fermion measure. The absence of anomalies can then be verified directly on the lattice, and this seems to confirm that the theories which had been discovered anomaly-free in perturbation theory are also anomaly-free in the non-perturbative lattice construction. Therefore the specific problem of lattice fermions was in essence technical rather than reflecting an inconsistency of chiral gauge theories beyond perturbation theory, as one may have feared.

Finally since these new regularization schemes have a natural implementation in five dimensions, it again opens the door to speculations about higher space dimensions.

We first discuss the advantages and shortcomings, from the point of view of symmetries, of three regularization schemes, momentum cut-off, dimensional and lattice regularization. We show that they leave room for possible anomalies when both gauge fields and chiral fermions are present. We then recall the origin and the form of anomalies, beginning with the simplest example of the so-called abelian anomaly, i.e. the anomaly in the conservation of the abelian axial current in gauge theories. We relate anomalies to the index of a covariant Dirac operator in the background of a gauge field. We exhibit gauge instantons which contribute to the anomaly. We examine the form of the anomaly for a general axial current, and the obstruction to the construction of chiral gauge theories. A few classical physical applications are also briefly mentioned. Finally the formalism of overlap fermions and Ginsparg–Wilson relation are explained.

Conventions. Throughout these notes we work in *euclidean space* (with imaginary or euclidean time), and this also implies a formalism of *euclidean fermions.*

2 Momentum cut-off regularization

We first discuss methods which work in the continuum (compared to lattice methods) and at fixed dimension (unlike dimensional regularization). The idea then is to modify field propagators beyond a large momentum cut-off to render all Feynman diagrams convergent. However the regularization has to satisfy one important condition: the regularized propagator should remain a *smooth* function of the momentum $\mathbf{p}$. Indeed singularities in momentum variables generate, after Fourier transformation, contributions to the large distance behaviour of the propagator, and regularization should modify the theory only at short distance.

2.1 Matter fields

A simple modification of the propagator improves the convergence of Feynman diagrams at large momentum. For example the propagator $1/(m^2+p^2)$ of the scalar field corresponding to the action

$$\mathcal{S}(\phi)=\int \mathrm{d}^d x\left[\tfrac{1}{2}\phi(-\nabla^2+m^2)\phi+V(\phi)\right], \tag{2.1}$$

can be replaced by

$$\Delta_{\mathrm{B}}(p)=\left(\frac{1}{p^2+m^2}\right)_{\mathrm{reg.}},$$

with

$$\Delta_{\mathrm{B}}^{-1}(p)=(p^2+m^2)\prod_{i=1}^{n}(1+p^2/M_i^2). \tag{2.2}$$

The masses M_i are proportional to the momentum cut-off Λ,

$$M_i=\alpha_i\Lambda\,.$$

The degree n is chosen large enough to render all diagrams convergent. In the large cut-off limit, at parameters α fixed, the initial propagator is recovered. This is the spirit of Pauli–Villars's regularization.

Note that such a propagator cannot be derived from a hermitian hamiltonian. Hermiticity of the hamiltonian implies that if the propagator is, as above, a rational fraction, it must be a sum of poles with positive residues and thus cannot decrease faster than $1/p^2$.

While this modification can be implemented also in Minkowski space because the regularized propagators decreases in all complex p^2 directions, in euclidean time more general modifications are possible. Schwinger's proper time representation suggests:

$$\Delta_{\mathrm{B}}(p) = \int_0^\infty \mathrm{d}t\, \rho(t\Lambda^2)\, \mathrm{e}^{-t(p^2+m^2)}, \tag{2.3}$$

in which the function $\rho(t)$ is positive and satisfies the condition

$$|1-\rho(t)| < C\,\mathrm{e}^{-\sigma t}\ (\sigma > 0) \text{ for } t \to +\infty\,.$$

By choosing a function $\rho(t)$ that decreases fast enough for $t \to 0$ the behaviour of the propagator can be arbitrarily improved. If $\rho(t) = O(t^n)$ the behaviour (2.2) is recovered. Another example is:

$$\rho(t) = \theta(t-1), \tag{2.4}$$

$\theta(t)$ being the step function, which leads to exponential decrease:

$$\Delta_{\mathrm{B}}(p) = \frac{\mathrm{e}^{-(p^2+m^2)/\Lambda^2}}{p^2+m^2}. \tag{2.5}$$

As the example shows, it is thus possible to find in this more general class propagators without unphysical singularities, but they do not follow from a hamiltonian formalism because continuation to real time becomes impossible.

For spin 1/2 fermions similar methods are applicable. The inverse propagator $m + i\not{p}$ can be replaced by

$$\Delta_{\mathrm{F}}^{-1}(p) = (m + i\not{p}) \prod_{i=1}^{n} (1 + p^2/M_i^2), \tag{2.6}$$

with euclidean fermion conventions, and hermitian matrices γ_μ.

Momentum cut-off regularizations have several advantages: one can work at fixed dimension and in the continuum. However, two potential weakness have to stressed:

(i) The generating functional functional of correlation function $\mathcal{Z}(J)$ obtained by adding to the action a source term for fields can be written, for instance in the case of the self-coupled scalar field,

$$\mathcal{Z}(J) = \det{}^{1/2}(\Delta_{\rm B}) \exp\left[-\mathcal{V}_{\rm I}\left(\frac{\delta}{\delta J}\right)\right] \exp\left(\tfrac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}^d y\, J(x)\Delta_{\rm B}(x-y)J(y)\right), \tag{2.7}$$

where the determinant is generated by the gaussian integration. None of these methods can regularize the determinant. As long as the determinant is a divergent constant that cancels in normalized correlation functions this is not a problem, but in the case of a determinant in an external field (which generates a set of one-loop diagrams) this may become a serious issue.

(ii) This problem is related to another one: Some models have even in simple quantum mechanics divergences or ambiguities due to problem of order between quantum operators. A class of Feynman diagrams then cannot be regularized by this method. Quantum field theories where this problem occurs include models with non-linear or gauge symmetries.

Global linear symmetries. To implement symmetries of the classical action in the quantum theory we need a regularization scheme which preserves the symmetry. This requires some care but can always be achieved for linear global symmetries, i. e. symmetries which correspond to transformations of the fields of the form

$$\phi_R(x) = R\,\phi(x)\,,$$

where R is a constant matrix. The main reason is that in the quantum hamiltonian field operators and conjugate momenta are not mixed by the transformation, and therefore the order of operators is to some extent irrelevant. To take an example directly relevant here, a theory with massless fermions may, in four dimensions, have a chiral symmetry

$$\psi_\theta(x) = e^{i\theta\gamma_5}\psi(x), \qquad \bar\psi_\theta(x) = \bar\psi(x)\,\mathrm{e}^{i\theta\gamma_5}\ .$$

The substitution (2.6) (for $m=0$) preserves chiral symmetry. Note the importance here of being able to work in fixed dimension four because chiral symmetry is defined only in even dimensions. In particular the invariance of the integration measure $[\mathrm{d}\bar\psi(x)\mathrm{d}\psi(x)]$ relies on the property that $\operatorname{tr}\gamma_5 = 0$.

Regulator fields. Regularization in the form (2.2) or (2.6) has another equivalent formulation based on the introduction of regulator fields (Pauli–Villars's regularization).

The scalar case. In the case of scalar fields to regularize the action $\mathcal{S}(\phi)$ for the scalar field ϕ one introduces additional dynamical fields ϕ_k, $k = 1,\ldots,n$, and considers the modified action $\mathcal{S}_{\text{reg.}}(\phi, \phi_k)$:

$$\mathcal{S}_{\text{reg.}}(\phi, \phi_k) = \int \mathrm{d}^d x \left[\frac{1}{2}\phi\left(-\nabla^2 + m^2\right)\phi + \sum \frac{1}{2z_k}\phi_k\left(-\nabla^2 + M_k^2\right)\phi_k \right.$$
$$\left. + V(\phi + \textstyle\sum_k \phi_k)\right]. \tag{2.8}$$

With the action (2.8) any internal ϕ propagator is replaced by the sum of the ϕ propagator and all the ϕ_k propagators $z_k/(p^2 + M_k^2 t)$. For an appropriate choice of the constants z_k, after integration over the regulator fields, the form (2.2) is recovered.

Fermions. Note that the fermion inverse propagator (2.6) can be written

$$\Delta_{\mathrm{F}}^{-1}(p) = (m + i\not{p}) \prod_{i=1}^{n} (1 + i\not{p}/M_i)(1 - i\not{p}/M_i).$$

This indicates that again the same form can be obtained by a set of regulator fields $\bar{\psi}_{i\pm}\psi_{i\pm}$. One replaces the kinetic part of the action by

$$\int \mathrm{d}^d x\, \bar{\psi}(x)(\not{\partial} + m)\psi(x) \mapsto \int \mathrm{d}^d x\, \bar{\psi}(x)(\not{\partial} + m)\psi(x)$$
$$+ \sum_{k,\epsilon=\pm} \frac{1}{z_{k\epsilon}} \int \mathrm{d}^d x\, \bar{\psi}_{k\epsilon}(x)(\not{\partial} + \epsilon M_k)\psi_{k\epsilon}(x). \tag{2.9}$$

In the same way in the interaction the fields ψ and $\bar{\psi}$ are replaced by the sums

$$\psi \mapsto \psi + \sum_{k,\epsilon} \psi_{k\epsilon}\,, \qquad \bar{\psi} \mapsto \bar{\psi} + \sum_{k,\epsilon} \bar{\psi}_{k\epsilon}\,.$$

Again for a proper choice of the constants z_k, after integration over the regulator fields the form (2.6) is recovered. Note in particular that for $m = 0$ one finds $z_{k+} = z_{k-}$. This indicates how chiral symmetry is preserved by the regularization, although the regulators are massive: by fermion doubling. The fermions ψ_+ and ψ_- are chiral partners. For a pair $\psi \equiv (\psi_+, \psi_-)$, $\bar{\psi} \equiv (\bar{\psi}_+, \bar{\psi}_-)$ the action can be written

$$\int \mathrm{d}^d x\, \bar{\psi}(x)\left(\not{\partial} \otimes \mathbf{1} + M\mathbf{1} \otimes \sigma_3\right)\psi(x),$$

where the first matrix $\mathbf{1}$ and the Pauli matrix σ_3 act in $\pm$ space. The spinors then transform like

$$\psi_\theta(x) = e^{i\theta\gamma_5 \otimes \sigma_1}\psi(x), \qquad \bar{\psi}_\theta(x) = \bar{\psi}(x)\, \mathrm{e}^{i\theta\gamma_5 \otimes \sigma_1}\,,$$

because σ_1 anticommutes with σ_3.

2.2 Abelian gauge theory

The problem of matter in presence of a gauge field can be decomposed into two steps, first matter in an external gauge field, and then the integration over the gauge field. For gauge fields we choose a covariant gauge, in such a way that power counting is the same as for scalar fields.

Charged fermions in a gauge background. The new problem that arises in presence of a gauge field is that only covariant derivatives are allowed, because gauge invariance is essential for the physical consistency of the theory. As a consequence the higher order derivatives of the regularization generate new, more singular, gauge interactions and it is no longer clear whether the theory can be rendered finite. The regularized action in a gauge background now reads

$$\mathcal{S}\left(\bar{\psi}, \psi, A\right) = \int \mathrm{d}^d x\, \bar{\psi}(x)\left(m + \not{\mathrm{D}}\right) \prod_k \left(1 - \not{\mathrm{D}}^2/M_k^2\right) \psi(x), \tag{2.10}$$

where D_μ is the covariant derivative

$$\mathrm{D}_\mu = \partial_\mu + ieA_\mu\,.$$

Note that up to this point the regularization, unlike dimensional or lattice regularizations, preserves a possible chiral symmetry for $m = 0$.

Correlation functions in the gauge background are then generated by:

$$\mathcal{Z}(\bar{\eta}, \eta; A) = \int \left[\mathrm{d}\psi(x)\mathrm{d}\bar{\psi}(x)\right] \exp\left[-\mathcal{S}(\bar{\psi}, \psi, A) + \int \mathrm{d}^d x\, \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right], \tag{2.11}$$

where $\bar{\eta}, \eta$ are Grassmann sources. Integrating over fermions explicitly we obtain

$$\mathcal{Z}(\bar{\eta}, \eta; A) = \mathcal{Z}_0(A) \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}^d y\, \bar{\eta}(y)\Delta_{\mathrm{F}}(A; y, x)\eta(x)\right], \tag{2.12}$$

$$\mathcal{Z}_0(A) = \mathcal{N} \det\left[(m + \not{\mathrm{D}}) \prod_k \left(1 - \not{\mathrm{D}}^2/M_k^2\right)\right],$$

where $\mathcal{N}$ is a gauge field-independent normalization and $\Delta_{\mathrm{F}}(A; y, x)$ the fermion propagator in an external gauge field.

Diagrams constructed from $\Delta_{\mathrm{F}}(A; y, x)$ belong to loops with gauge field propagators, and therefore can be rendered finite if the gauge field propagator, which we check below. The other problem involves the determinant that generates closed fermion loops in gauge background. Using $\ln\det = \mathrm{tr}\ln$ we find

$$\ln \mathcal{Z}_0(A) = \mathrm{tr}\ln\left(m + \not{\mathrm{D}}\right) + \sum_k \mathrm{tr}\ln\left(1 - \not{\mathrm{D}}^2/M_k^2\right) - (A = 0),$$

or using the anticommutation of γ_5 with $\not{D}$

$$\det(\not{D}+m)=\det\gamma_5(\not{D}+m)\gamma_5=\det(m-\not{D}),$$

$$\ln\mathcal{Z}_0(A)=\tfrac{1}{2}\operatorname{tr}\ln\left(m^2-\not{D}^2\right)+\sum_k\operatorname{tr}\ln\left(1-\not{D}^2/M_k^2\right)-(A=0),$$

We see that the regularization has no effect from the point of view of power counting on the determinant, and therefore on one-loop diagrams of the form of fermion closed loops with external gauge fields, a problem that requires an additional regularization.

The fermion determinant. The fermion determinant can finally be regularized by adding to the action a bosonic regulator field with fermion spin, and therefore a propagator similar to Δ_{F} but with different masses

$$\mathcal{S}_{\mathrm{B}}(\bar{\phi},\phi;A)=\int\mathrm{d}^dx\,\bar{\phi}(x)\left(M_0^{\mathrm{B}}+\not{D}\right)\prod_{k=1}\left(1-\not{D}^2/(M_k^{\mathrm{B}})^2\right)\phi(x). \qquad (2.13)$$

The integration over the boson ghost fields $\bar{\phi},\phi$ adds to $\ln\mathcal{Z}_0$ the quantity

$$\delta\ln\mathcal{Z}_0(A)=-\tfrac{1}{2}\operatorname{tr}\ln\left((M_0^{\mathrm{B}})^2-\not{D}^2\right)-\sum_{k=1}\operatorname{tr}\ln\left(1-\not{D}^2/(M_k^{\mathrm{B}})^2\right)-(A=0).$$

Expanding in inverse powers of $\not{D}$ one adjusts the masses to cancel as many powers as possible. However, the unpaired initial fermion mass m is the source of a problem. The corresponding determinant can only be regularized with an unpaired boson M_0^{B}. In the chiral limit $m=0$ we have two options: either we give a chiral charge to the boson field and the mass M_0^{B} breaks chiral symmetry, or we leave it invariant in a chiral transformation. Then we have a transformed determinant of an operator

$$\mathrm{e}^{i\theta\gamma_5(x)}\not{D}\,\mathrm{e}^{i\theta\gamma_5(x)}(\not{D}+M_0^{\mathrm{B}})^{-1}.$$

For $\theta(x)$ constant $\mathrm{e}^{i\theta\gamma_5}$ anticommutes with $\not{D}$ and cancels. Otherwise a non-trivial contribution remains. The method thus is not applicable with chiral fermions.

Since actually the problem reduces to the study of a determinant in an external background one can study it directly, as we will starting with Section 4. One verifies whether it is possible to define some regularized form in a way consistent with chiral symmetry. One then inserts the one-loop renormalized diagrams in the general diagrams regularized by the preceding cut-off methods.

The gauge propagator. For the free gauge action in a covariant gauge usual derivatives can be used because in an abelian theory the gauge field is neutral. The tensor $F_{\mu\nu}$ is gauge invariant and the action for the scalar $\partial_\mu A_\mu$ is arbitrary. Therefore the large momentum behaviour of the gauge field propagator can be arbitrarily improved.

2.3 Non-abelian gauge theories

Compared with the abelian case, the new features of the non-abelian gauge action are the presence of gauge field self-interactions and ghost terms. For future purpose we define our notation. We introduce the covariant derivative, as acting on matter field,

$$\mathbf{D}_\mu = \partial_\mu + \mathbf{A}_\mu(x)\,, \tag{2.14}$$

where $\mathbf{A}_\mu$ is an anti-hermitian matrix, and the curvature tensor $\mathbf{F}_{\mu\nu}$

$$\mathbf{F}_{\mu\nu} = [\mathbf{D}_\mu, \mathbf{D}_\nu] = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu]. \tag{2.15}$$

The pure gauge action then is

$$\mathcal{S}(\mathbf{A}_\mu) = -\frac{1}{4g^2}\,\mathrm{tr}\int \mathrm{d}^d x\ \mathrm{tr}\,\mathbf{F}_{\mu\nu}(x)\mathbf{F}_{\mu\nu}(x), \tag{2.16}$$

In the covariant gauge

$$\mathcal{S}_{\mathrm{gauge}}(\mathbf{A}_\mu) = -\frac{1}{2\xi}\int \mathrm{d}^d x\ \mathrm{tr}(\partial_\mu \mathbf{A}_\mu)^2,$$

the ghost field action takes the form:

$$\mathcal{S}_{\mathrm{ghost}}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}) = -\int \mathrm{d}^d x\ \mathrm{tr}\,\bar{\mathbf{C}}\,\partial_\mu\left(\partial_\mu \mathbf{C} + [\mathbf{A}_\mu, \mathbf{C}]\right).$$

The ghost fields thus have a simple δ_{ab}/p^2 propagator and canonical dimension one in four dimensions.

The problem of regularization in non-abelian gauge theories has many features in common with the abelian case, as well as with the non-linear σ-model. The regularized gauge action takes the form:

$$\mathcal{S}\left(\mathbf{A}_\mu\right) = -\int \mathrm{d}^d x\ \mathrm{tr}\,\mathbf{F}_{\mu\nu} P\left(\mathbf{D}^2/\,\Lambda^2\right)\mathbf{F}_{\mu\nu}\,, \tag{2.17}$$

in which P is a polynomial of arbitrary degree. In the same way the gauge function $\partial_\mu \mathbf{A}_\mu$ is changed into:

$$\partial_\mu \mathbf{A}_\mu \longmapsto Q\left(\partial^2/\,\Lambda^2\right)\partial_\mu \mathbf{A}_\mu\,, \tag{2.18}$$

in which Q is a polynomial of same degree as P. As a consequence both the gauge field propagator and the ghost propagator can be arbitrarily improved. However, as in the abelian case, the covariant derivatives generate new interactions which are more singular. It is easy to verify that the power counting of one-loop diagrams is unchanged while higher order diagrams can be made convergent by taking the degrees of P and Q large enough. Regularization by higher derivatives takes care of all diagrams except, as in all geometric models, some one-loop diagrams (and thus subdiagrams).

As with charged matter the one-loop diagrams have to be examined separately. For fermion matter it is however still possible as, in the abelian case, to add a set of regulator fields, massive fermions and bosons with spin. This procedure breaks a possible chiral symmetry. In the chiral situation the problem of the compatibility between the gauge symmetry and the quantization is reduced to an explicit verification of the WT identities for the one-loop diagrams. Note that the preservation of gauge symmetry is necessary for the cancellation of unphysical states in physical amplitudes, and thus essential to the physical relevance of the quantum field theory.

3 Other regularizations

The other regularization schemes we now discuss have the common property that they modify in some essential way the structure of space–time, dimension regularization because it relies on defining Feynman diagrams for non-integer dimensions, lattice regularization because continuum space is replaced by a discrete lattice.

3.1 Dimensional regularization

Dimensional regularization involves continuation of Feynman diagrams in the parameter d (d is the space dimension) to arbitrary complex values, and therefore seems to have no meaning outside perturbation theory. However this regularization very often leads to the simplest perturbative calculations.

In addition it solves with the problem of commutation of quantum operators in local field theories. Indeed commutators for example in the case of a scalar field take the form

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\hbar\, \delta^{d-1}(x-y) = i\hbar(2\pi)^{1-d} \int \mathrm{d}^{d-1}p\ \mathrm{e}^{ip(x-y)}\,,$$

where $\hat{\pi}(x)$ is the momentum conjugated to the field $\hat{\phi}(x)$. As we have already stressed, in a local theory all fields are taken at same point, and therefore a commutation in the product $\hat{\phi}(x)\hat{\pi}(x)$ generates a divergent contribution (for $d > 1$) proportional to

$$\delta^{d-1}(0) = (2\pi)^{1-d} \int \mathrm{d}^{d-1}p\,.$$

The rule of dimensional regularization imply the consistency of the change of variables $p \mapsto \lambda p$ and thus $\int \mathrm{d}^d p = 0$, in contrast to momentum regularization where it is proportional to a power of the cut-off. Therefore the order between operators becomes irrelevant because the commutator vanishes. Dimensional regularization thus is applicable to geometric models where these problems of quantization occur, like non-linear σ models or gauge theories.

Its use, however, requires some care in massless theories. For instance in a massless theory in two dimensions integrals of the form $\int \mathrm{d}^d k/k^2$ are met. They again vanish in dimensional regularization for the same reason. However here they correspond to an unwanted cancellation between UV and IR logarithmic divergences.

More important here, it is not applicable when some essential property of the field theory is specific to the initial dimension. An example is provided by theories containing fermions in which parity symmetry is violated.

Fermions. For fermions belonging to the fundamental representation of the spin group Spin(d) the strategy is the same. The spin problem can be reduced to the calculation of traces of γ matrices. Therefore only one additional prescription for the trace of the unit matrix is needed. There is no natural continuation since odd and even dimensions behave differently. Since no algebraic manipulation depends on the explicit value of the trace, any smooth continuation in the neighbourhood of the relevant dimension is satisfactory. A convenient choice is to take the trace constant. In even dimension as long as only γ_μ matrices are involved no other problem arises. However no dimensional continuation which preserves all properties of γ_5 can be found. This leads to serious difficulties if some diagrams involve γ_5 and if it becomes necessary to use the identity relating γ_5 to the other γ matrices

$$4!\,\gamma_5 = -\epsilon_{\mu_1\ldots\mu_4}\gamma_{\mu_1}\ldots\gamma_{\mu_4}\,, \tag{3.1}$$

where $\epsilon_{\mu_1\cdots\mu_4}$ is the complete antisymmetric tensor. Problems therefore arise in the case of gauge theories with chiral fermions, because the special properties of γ_5 are involved as we recall below. This difficulty is the source of chiral anomalies.

Since perturbation theory involves the calculation of traces, one possibility is to define γ_5 near four dimensions by

$$\gamma_5 = E_{\mu_1\ldots\mu_4}\gamma_{\mu_1}\ldots\gamma_{\mu_4}\,, \tag{3.2}$$

where $E_{\mu\nu\rho\sigma}$ is a completely antisymmetric tensor, which reduces to $-\epsilon\mu\nu\rho\sigma/4!$ in four dimensions. It is then easy to verify that, with this definition, γ_5 anticommutes with the other γ_μ matrices only in four dimensions. If for example we evaluate the product $\gamma_\nu\gamma_5\gamma_\nu$ in d dimensions, we find:

$$\gamma_\nu\gamma_5\gamma_\nu = (d-8)\gamma_5\,.$$

Anticommuting properties of the γ_5 would have led to a factor $-d$ instead. Multiplied by pole in $1/(d-4)$ due to UV divergences the additional contribution proportional to $d-4$ may lead a finite difference.

The other option would be to keep the anticommuting property of γ_5, but the previous example shows that this is contradictory with a form (3.2). Actually one verifies that the only consistent prescription for generic dimensions then is that the traces of γ_5 with any product of γ_μ matrices vanishes.

3.2 Lattice regularization

We have explained that Pauli–Villars's regularization does not work for field theories in which the action has a definite geometric character like models on homogeneous spaces (for example the non-linear σ-model) or gauge theories. In these theories some divergences are related to the problem of quantization and order of operators, which already appears in simple quantum mechanics. Other regularization methods then are needed. In many cases lattice regularization may be used.

Lattice field theory. To each site x of the lattice are attached field variables corresponding to fields in the continuum. To the action $\mathcal{S}$ in the continuum corresponds a lattice action, the energy of lattice field configurations in the language of classical statistical physics. The functional integral becomes a sum over configurations and the regularized partition function is the partition function of a lattice model.

Note that all expressions will be written for a hypercubic lattice and we denote by a the lattice spacing,

The advantages of lattice regularization are:

(i) Lattice regularization indeed correspond to a specific choice of quantization.

(ii) It is the only established regularization which has a meaning outside perturbation theory. For instance the regularized functional integral can be calculated by numerical methods, like stochastic methods (Monte-Carlo type calculations) or strong coupling expansions.

(iii) It preserves most global or local symmetries with the exception of the space $O(d)$ symmetry which is replaced by a hypercubic symmetry (but this can be shown to be not too serious) and fermion chirality, which turns out to be a more serious problem, as we will show.

The main disadvantage is that it leads to very complicated perturbative calculations.

3.3 Boson field theories

Scalar fields. To the action (2.1) for a scalar field ϕ in the continuum corresponds a lattice action, which is obtained in the following way: The euclidean lagrangian density becomes a function of lattice variables $\phi(x)$, where x now is a lattice site. Locality can be implemented by the property that the lattice lagrangian should depend on a site and its neighbours. Derivatives $\partial_\mu \phi$ of the continuum are replaced by finite differences, for example:

$$\partial_\mu \phi \mapsto \nabla_\mu \phi = \left[\phi(x + a n_\mu) - \phi(x)\right]/a\,, \tag{3.3}$$

where a is the lattice spacing and n_μ the unit vector in the μ direction. The lattice action then is the sum over lattice sites.

With the choice (3.3) the propagator $\Delta_a(p)$ for the Fourier components of a massive scalar field is given by

$$\Delta_a^{-1}(p) = m^2 + \frac{2}{a^2}\sum_{\mu=1}^{d}\bigl(1-\cos(ap_\mu)\bigr). \tag{3.4}$$

It is a periodic function of the components p_μ of the momentum vector with period $2\pi/a$. In the small lattice spacing limit the continuum propagator is recovered:

$$\Delta_a^{-1}(p) = m^2 + p^2 - \tfrac{1}{12}\sum_{\mu} a^2 p_\mu^4 + O\left(p_\mu^6\right). \tag{3.5}$$

In particular hypercubic symmetry implies $O(d)$ symmetry at order p^2.

Gauge theories. Lattice regularization defines unambiguously a quantum theory. Therefore, once one has realized that gauge fields should be replaced by link variables corresponding to parallel transport along links of the lattice, one can regularize a gauge theory.

The link variables U_{xy} are group elements associated with the links joining the sites x and y on the lattice. The regularized form of $\int \mathrm{d}x\, F_{\mu\nu}^2$ is the product of link variables on a closed curve on the lattice, the simplest being a square on a hypercubic lattice, leading to the well-known plaquette action, each square forming a plaquette. The typical gauge invariant lattice action corresponding to the continuum action of a gauge field coupled to scalar bosons then has the form:

$$\mathcal{S}(U,\phi^*,\phi) = \beta \sum_{\text{plaquettes}} \operatorname{tr} U_{xy}U_{yz}U_{zt}U_{tx} + \kappa \sum_{\text{links}} \phi_x^* U_{xy}\phi_y + \sum_{\text{sites}} V(\phi_x^*\phi_x), \tag{3.6}$$

where x, y,... denotes lattice sites, and β and κ coupling constants. The action (3.6) is invariant under independent group transformations on each lattice site, lattice equivalents of the gauge transformations of the continuum theory. The measure of integration over the gauge variables is the group invariant measure on each site. Note that on the lattice and in a finite volume the gauge invariant action leads to a well-defined partition function because the gauge group is compact. However in the continuum limit the compact character of the group is lost. It is therefore necessary to fix the gauge on the lattice in order to be able to construct a regularized perturbation theory.

We finally note that on the lattice the difficulties with the regularization do not come from the gauge field directly, but involve the gauge field only through the integration over chiral fermions.

3.4 Fermion and the doubling problem

We now review a few problems specific to relativistic fermions on the lattice. We consider the free action for a Dirac fermion:

$$\mathcal{S}\left(\bar{\psi},\psi\right)=\int \mathrm{d}^d x\,\bar{\psi}(x)\left(\not{\partial}+m\right)\psi(x).$$

To regularize this action by a lattice and preserve chiral properties in the massless limit one can replace $\partial_\mu\psi(x)$ by

$$\nabla_\mu\psi(x)=\left[\psi\left(x+an_\mu\right)-\psi\left(x-an_\mu\right)\right]/2a.$$

Then the inverse of the fermion propagator Δ for the Fourier components $\tilde{\psi}(p)$ of the field is:

$$\Delta^{-1}(p)=m+i\sum_\mu \gamma_\mu\frac{\sin ap_\mu}{a} \tag{3.7}$$

a periodic function of the components p_μ of the momentum vector. A problem then arises: the equations relevant to the small lattice spacing limit,

$$\sin(a\,p_\mu)=0$$

have each two solutions $p_\mu=0$ and $p_\mu=\pi/a$ within one period, i.e. within the Brillouin zone $2\pi/a$. Therefore the propagator (3.7) propagates 2^d fermions. To remove this degeneracy it is possible to add to the regularized action an additional scalar term $\delta\mathcal{S}$ involving second derivatives:

$$\delta\mathcal{S}\left(\bar{\psi},\psi\right)=\frac{M}{2}\sum_{x,\mu}\left[2\bar{\psi}(x)\psi(x)-\bar{\psi}\left(x+an_\mu\right)\psi(x)-\bar{\psi}(x)\psi\left(x+an_\mu\right)\right]. \tag{3.8}$$

After Fourier transformation the modified Dirac operator D_W reads

$$D_W(p)=m+M\sum_\mu\left(1-\cos ap_\mu\right)+\frac{i}{a}\sum_\mu\gamma_\mu\sin ap_\mu\,. \tag{3.9}$$

The fermion propagator becomes:

$$\Delta(p)=D_W^\dagger(p)\left(D_W(p)D_W^\dagger(p)\right)^{-1},$$

with:

$$D_W(p)D_W^\dagger(p)=\left[m+M\sum_\mu\left(1-\cos ap_\mu\right)\right]^2+\frac{1}{a^2}\sum_\mu\sin^2 ap_\mu\,.$$

Therefore the degeneracy between the different states has been lifted. For each component p_μ which takes the value π/a the mass is increased by M. If M is of order $1/a$ the spurious states are eliminated in the continuum limit. This is the recipe of Wilson's fermion.

However a problem arises if one wants to construct a theory with massless fermions and chiral symmetry. Chiral symmetry implies for the Dirac operator $D(p)$ anticommutation with γ_5

$$\{D(p), \gamma_5\} = 0\,, \tag{3.10}$$

and therefore both the mass term and the term (3.8) are excluded. It remains possible to add various counter-terms and try to adjust them to recover chiral symmetry in the continuum limit. But then there is no *a priori* guarantee that this is indeed possible and moreover calculations are plagued by fine tuning problems and cancellations of unnecessary UV divergences.

One could also think about modifying the fermion propagator by adding terms connecting fermions separated by more than one lattice spacing. But it has been proven that this does not solve the doubling problem. In fact this doubling of the number of fermion degrees of freedom is directly related to the problem of anomalies.

Since the most naive form of the propagator yields 2^d fermion states, one tries in practical calculations to reduce this number to a smaller multiple of two. The idea of staggered fermions introduced by Kogut and Susskind is often used: first by modifying the action one is able to decrease the multiplication factor from 2^d to $2^{d/2}$ with respect to form (3.7). Then the remaining degeneracy is interpreted as the reflection of an internal symmetry $SU(2^{d/2})$ of the action. Recently the global picture has been modified by the discovery of the properties of overlap fermions and solutions of the Ginsparg–Wilson relation, a topic we postpone, and on which we will come back at the end.

4 The abelian anomaly

We have pointed out that none of the standard regularization methods can deal in a straightforward way with one-loop diagrams in the case of gauge fields coupled to chiral fermions. We now show that indeed gauge theories with massless fermions and chiral symmetry can be found where the axial current is not conserved. The divergence of the axial current, when it does not vanish, is called an *anomaly*. This leads in particular to obstructions to the construction of gauge theories where the gauge field couples differently to the two fermion chiral components. Several examples are physically important like the theory of weak electromagnetic interactions, the electromagnetic decay of the π_0 meson, or the $U(1)$ problem. We first discuss the abelian axial current, in four dimensions (the generalization to all even dimensions is then straightforward), and then the general non-abelian situation.

4.1 Abelian axial current and abelian vector gauge field

The only possible source of anomalies are one-loop fermion diagrams in gauge theories when chiral properties are involved. This reduces the problem to the discussion of fermions in background gauge fields, or equivalently to the properties of the determinant of the gauge covariant Dirac operator.

We thus consider the QED-like fermion action $\mathcal{S}(\bar{\psi},\psi)$ for massless Dirac fermions $\psi,\bar{\psi}$ in the background of an abelian gauge field A_μ

$$\mathcal{S}\left(\bar{\psi},\psi\right)=-\int \mathrm{d}^4x\,\bar{\psi}(x)\not{D}\psi(x),\quad \not{D}\equiv\not{\partial}+ie\not{A}\,, \tag{4.1}$$

and the corresponding functional integral

$$\mathcal{Z}(A_\mu)=\int\left[\mathrm{d}\psi\mathrm{d}\bar{\psi}\right]\exp\left[-\mathcal{S}(\psi,\bar{\psi})\right]=\det\not{D}\,. \tag{4.2}$$

In what follows we denote by $\langle\bullet\rangle$ expectation values with respect to the measure $\mathrm{e}^{-\mathcal{S}}$.

We can find regularizations which preserve gauge invariance,

$$\psi(x)=\mathrm{e}^{i\Lambda(x)}\,\psi'(x),\quad \bar{\psi}(x)=\mathrm{e}^{-i\Lambda(x)}\,\bar{\psi}'(x),\quad A_\mu(x)=-\frac{1}{e}\partial_\nu\Lambda(x)+A'_\mu(x), \tag{4.3}$$

and since the fermions are massless, chiral symmetry. We would therefore naively expect the corresponding axial current to be conserved. However the proof of current conservation involves space-dependent chiral transformations, and therefore steps which cannot be regularized, for the same reasons as for chiral gauge theories.

The coefficient of $\partial_\mu\theta(x)$ in the variation of the action under a space-dependent chiral transformation

$$\psi_\theta(x)=\mathrm{e}^{i\theta(x)\gamma_5}\,\psi(x),\qquad \bar{\psi}_\theta(x)=\bar{\psi}(x)\,\mathrm{e}^{i\theta(x)\gamma_5}, \tag{4.4}$$

yields the axial current $J^5_\mu(x)$. For the action (4.1) one finds

$$\delta\mathcal{S}=\int \mathrm{d}^4x\,\partial_\mu\theta(x)J^5_\mu(x)\ \text{with}\ J^5_\mu(x)=i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x). \tag{4.5}$$

After the transformation (4.4) $\mathcal{Z}(A_\mu)$ becomes:

$$\mathcal{Z}(A_\mu,\theta)=\det\left[\mathrm{e}^{i\gamma_5\theta(x)}\,\not{D}\,\mathrm{e}^{i\gamma_5\theta(x)}\right]. \tag{4.6}$$

Note that $\ln[\mathcal{Z}(A_\mu,\theta)]$ is the generating functional of connected $\partial_\mu J^5_\mu$ correlation functions in an external A field.

Since $e^{i\gamma_5\theta}$ has a determinant which is unity, one would naively conclude that $\mathcal{Z}(A_\mu,\theta)=\mathcal{Z}(A_\mu)$ and therefore that the current $J^5_\mu(x)$ is conserved. This is a conclusion we now check by an explicit calculation of the expectation value of $\partial_\mu J^5_\mu(x)$ in the case of the action (4.1).

Remarks.

(i) For any regularization which is consistent with the hermiticity of γ_5 we find that

$$|\mathcal{Z}(A_\mu,\theta)|^2=\det(\not{D}\not{D}^\dagger),$$

and thus is independent of θ. Therefore an anomaly can appear only in the imaginary part of $\ln\mathcal{Z}$.

(ii) We have shown that one can find a regularization with regulator fields such that gauge invariance is maintained, and the determinant independent θ for $\theta(x)$ constant.

(iii) If the regularization is gauge invariant $\mathcal{Z}(A_\mu,\theta)$ is also gauge invariant. Therefore a possible anomaly will also be gauge invariant.

(iv) $\ln\mathcal{Z}(A_\mu,\theta)$ is connected and 1PI. Short distance singularities thus take the form of local polynomials in the fields and sources. Since a possible anomaly is a short distance effect it must also take the form of local polynomial of A_μ and $\partial_\mu\theta$ constrained by power counting, both having dimension one and no mass being available, and parity,

$$\ln\mathcal{Z}(A_\mu,\theta)-\ln\mathcal{Z}(A_\mu,0)=i\int \mathrm{d}^4x\,\mathcal{L}(A,\partial\theta;x),$$

where $\mathcal{L}$ is the sum of monomials of dimension four. At order θ only one is available:

$$e^2\epsilon_{\mu\nu\rho\sigma}\partial_\mu\theta(x)A_\nu(x)\partial_\rho A_\sigma(x),$$

which a simple integration by parts and anti-symmetrization shows to be gauge invariant.

The coefficient of θ yields the average of the divergence of the axial current, $\partial_\mu J^5_\mu(x)$, in an external gauge field It is thus determined up to a multiplicative constant,

$$\left\langle\partial_\lambda J^5_\lambda(x)\right\rangle\propto e^2\epsilon_{\mu\nu\rho\sigma}\partial_\mu A_\nu(x)\partial_\rho A_\sigma(x)\propto e^2\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\,,$$

$F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$ being the electromagnetic tensor. Since the general form of the possible anomaly is independent of the regularization, it must indeed be gauge invariant.

To find the multiplicative factor, which is the only regularization dependent feature, it is sufficient to calculate the coefficient of term quadratic in A in the expansion of $\left\langle\partial_\lambda J^5_\lambda(x)\right\rangle$ in powers of A. We define the three-point function:

$$\begin{aligned}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)&=\frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)}\left\langle J^5_\lambda(k)\right\rangle\Big|_{A=0}\,,\\&=\frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)}i\,\mathrm{tr}\left[\gamma_5\gamma_\lambda\not{D}^{-1}(k)\right]\Big|_{A=0}\,.\end{aligned}\tag{4.7}$$

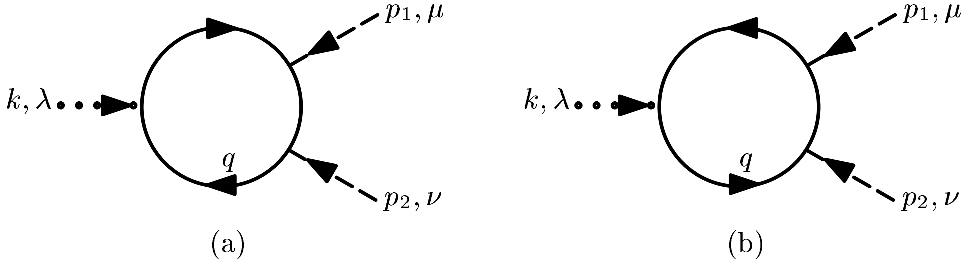

Fig. 1 Anomalous diagrams.

$\Gamma^{(3)}$ is the sum of the two Feynman diagrams of figure 1.

The contribution of diagram (a) is:

$$(\mathrm{a}) \mapsto \frac{e^2}{(2\pi)^4} \operatorname{tr}\left[\int \mathrm{d}^4 q\, \gamma_5 \gamma_\lambda \left(\not{q} + \not{k}\right)^{-1} \gamma_\mu \left(\not{q} - \not{p}_2\right)^{-1} \gamma_\nu \not{q}^{-1}\right], \tag{4.8}$$

and the contribution of diagram (b) is obtained by exchanging $p_1, \gamma_\nu \leftrightarrow p_2, \gamma_\nu$.

Power counting tells us that the function $\Gamma^{(3)}$ may have a linear divergence which, due to the presence of the γ_5 factor, must be proportional to $\epsilon_{\lambda\mu\nu\rho}$, symmetric in the exchange $p_1, \gamma_\nu \leftrightarrow p_2, \gamma_\nu$, and thus proportional to

$$\epsilon_{\lambda\mu\nu\rho} \left(p_1 - p_2\right)_\rho . \tag{4.9}$$

On the other hand by commuting γ_5 we notice that $\Gamma^{(3)}$ is formally a symmetric function of the three sets of external arguments. A divergence breaks the symmetry between external arguments. Therefore a symmetric regularization of the kind we shall adopt leads to a finite result. The result is not ambiguous because a possible ambiguity again is proportional to (4.9).

In the same way if the regularization is consistent with vector gauge invariance the vector current is conserved

$$p_{1\mu} \Gamma^{(3)}_{\lambda\mu\nu}(k; p_1, p_2) = 0 . \tag{4.10}$$

Applied to the divergent part the equation yields

$$-p_{1\mu} p_{2\rho} \epsilon_{\lambda\mu\nu\rho} = 0 ,$$

which is impossible. Therefore the sum of the two diagrams is finite. Different regularizations may still differ by finite quantities of the form (4.9) but again all regularizations consistent with vector gauge invariance must give the same answer.

Therefore there are two possibilities:

(i) The divergence $k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k; p_1, p_2)$ in a regularization respecting the symmetry between the three arguments vanishes. Then $\Gamma^{(3)}$ is gauge invariant and the axial current is conserved.

(ii) The divergence of the symmetric regularization does not vanish. Then it is possible to add to $\Gamma^{(3)}$ a term proportional to (4.9) to restore gauge invariance but this term breaks the symmetry between external momenta: the axial current is not conserved, an anomaly is present.

4.2 Explicit calculation

Momentum regularization. The calculation can be done using one of the various gauge invariant regularizations, for example momentum cut-off regularization or dimensional regularization with γ_5 being defined as in dimension four and thus no longer anticommuting with other γ matrices. Instead we choose a regularization which preserves the symmetry between the three external arguments and global chiral symmetry, but breaks gauge invariance, replacing the fermion propagator:

$$(\not{q})^{-1} \longmapsto (\not{q})^{-1}\rho(\varepsilon q^2),$$

where ε is the regularization parameter ($\varepsilon \to 0_+$), $\rho(z)$ is a positive differentiable function such that $\rho(0) = 1$, and decreasing fast enough for $z \to +\infty$, at least like $1/z$.

Then the compatibility between current conservation and gauge invariance implies that $k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$ vanishes.

It is convenient to consider directly the contribution $C^{(2)}(k)$ of order A^2 to $\left\langle k_\lambda J^5_\lambda(k)\right\rangle$ which sums the two diagrams:

$$\begin{aligned} C^{(2)}(k) = e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\bigl(\varepsilon(q+k)^2\bigr)\rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr) \\ \times \operatorname{tr}\left[\gamma_5 \not{k}(\not{q}+\not{k})^{-1}\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right], \end{aligned} \tag{4.11}$$

because the calculation then suggests how the method generalizes to arbitrary even dimensions.

We first transform the expression, using the identity:

$$\not{k}(\not{q}+\not{k})^{-1} = 1 - \not{q}(\not{q}+\not{k})^{-1}. \tag{4.12}$$

Then

$$C^{(2)}(k) = C^{(2)}_1(k) + C^{(2)}_2(k),$$

with

$$\begin{aligned} C^{(2)}_1(k) = e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\bigl(\varepsilon(q+k)^2\bigr)\rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr) \\ \times \operatorname{tr}\left[\gamma_5 \gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right], \end{aligned} \tag{4.13}$$

and

$$\begin{aligned} C^{(2)}_2(k) = -e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\bigl(\varepsilon(q+k)^2\bigr)\rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr) \\ \times \operatorname{tr}\left[\gamma_5 \not{q}(\not{q}+\not{k})^{-1}\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right]. \end{aligned} \tag{4.14}$$

In $C_2^{(2)}(k)$ we use the cyclic property of the trace and the *commutation* of $\not q(\not q + \not k)^{-1}$ and γ_5 to cancel the propagator $\not q^{-1}$ and obtain

$$C_2^{(2)}(k) = -e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q+k)^2\big) \rho\big(\varepsilon(q-p_2)^2\big) \rho\big(\varepsilon q^2\big)$$
$$\times \operatorname{tr}\left[\gamma_5 \gamma_\mu (\not q - \not p_2)^{-1} \gamma_\nu (\not q + \not k)^{-1}\right]. \tag{4.15}$$

We then interchange (p_1, μ) and (p_2, ν), shift $q \mapsto q + p_1$ and commute γ_5,

$$C_2^{(2)}(k) = -e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q-p_2)^2\big) \rho\big(\varepsilon q^2\big) \rho\big(\varepsilon(q+p_1)^2\big)$$
$$\times \operatorname{tr}\left[\gamma_5 \gamma_\mu (\not q - \not p_2)^{-1} \gamma_\nu \not q^{-1}\right]. \tag{4.16}$$

We see that the two terms $C_1^{(2)}$ and $C_2^{(2)}$ would cancel in the absence of regulators. This corresponds to the formal proof of current conservation. However without regularization the integrals diverge and previous manipulations are not legitimate.

Instead here we find a non-vanishing sum due to the regulating factors which are different:

$$C^{(2)}(k) = e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q-p_2)^2\big) \rho\big(\varepsilon q^2\big)$$
$$\times \operatorname{tr}\left[\gamma_5 \gamma_\mu (\not q - \not p_2)^{-1} \gamma_\nu \not q^{-1}\right] \left[\rho\big(\varepsilon(q+k)^2\big) - \rho\big(\varepsilon(q+p_1)^2\big)\right]. \tag{4.17}$$

After evaluation of the trace $C^{(2)}$ becomes (using (3.1)):

$$C^{(2)}(k) = -4e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q-p_2)^2\big) \rho\big(\varepsilon q^2\big)$$
$$\times \epsilon_{\mu\nu\rho\sigma} \frac{p_{2\rho} q_\sigma}{q^2 (q-p_2)^2} \left[\rho\big(\varepsilon(q+k)^2\big) - \rho\big(\varepsilon(q+p_1)^2\big)\right]. \tag{4.18}$$

Contributions coming from finite values of q cancel in the $\varepsilon \to 0$ limit. Due the cut-off the relevant values of q are of order $\varepsilon^{-1/2}$. We therefore rescale q accordingly $q\varepsilon^{1/2} \mapsto q$ and find

$$C^{(2)}(k) = -4e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big((q - p_2\sqrt{\varepsilon})^2\big) \rho\big(q^2\big)$$
$$\times \epsilon_{\mu\nu\rho\sigma} \frac{p_{2\rho} q_\sigma}{q^2 (q - p_2\sqrt{\varepsilon})^2} \frac{\rho\big((q + k\sqrt{\varepsilon})^2\big) - \rho\big((q + p_1\sqrt{\varepsilon})^2\big)}{\sqrt{\varepsilon}}.$$

Taking the $\varepsilon \to 0$ limit we obtain a finite result:

$$C^{(2)}(k) = -4e^2 \epsilon_{\mu\nu\rho\sigma} \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) I_{\rho\sigma}(p_1, p_2),$$

with

$$I_{\rho\sigma}(p_1,p_2) \sim \int \frac{\mathrm{d}^4 q}{(2\pi)^4 q^4} p_{2\rho} q_\sigma \rho^2\left(\varepsilon q^2\right)\rho'\left(\varepsilon q^2\right)\left[2\varepsilon q_\lambda (k-p_1)_\lambda\right]. \tag{4.19}$$

The identity:

$$\int \mathrm{d}^4 q\, q_\alpha q_\beta f(q^2) = \tfrac{1}{4}\delta_{\alpha\beta}\int \mathrm{d}^4 q\, q^2 f(q^2),$$

transforms the integral into:

$$I_{\rho\sigma}(p_1,p_2) \sim -\tfrac{1}{2} p_{2\rho}(2p_1+p_2)_\sigma \int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2}\rho^2\left(\varepsilon q^2\right)\rho'\left(\varepsilon q^2\right). \tag{4.20}$$

The remaining integral can be calculated explicitly (we recall $\rho(0)=1$)

$$\int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2}\rho^2\left(\varepsilon q^2\right)\rho'\left(\varepsilon q^2\right) = \frac{1}{8\pi^2}\int_0^\infty \varepsilon q \mathrm{d}q\, \rho^2\left(\varepsilon q^2\right)\rho'\left(\varepsilon q^2\right) = -\frac{1}{48\pi^2},$$

and yields a result independent of the function ρ. We finally obtain:

$$\left\langle k_\lambda J_\lambda^5(k)\right\rangle = -\frac{e^2}{12\pi^2}\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, p_{1\mu} A_\nu(p_1) p_{2\rho} A_\sigma(p_2). \tag{4.21}$$

From the definition (4.7) it follows that:

$$k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2}\epsilon_{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma}\,. \tag{4.22}$$

This non-vanishing result implies that any definition of the determinant $\det \not{D}$ breaks at least either current conservation or gauge invariance. Since gauge invariance is essential to the consistency of a gauge theory we choose to break current conservation. Exchanging arguments, we obtain the value of $p_{1\mu}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$:

$$p_{1\mu}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2}\epsilon_{\lambda\nu\rho\sigma} k_\rho p_{2\sigma}\,. \tag{4.23}$$

If instead we had used a gauge invariant regularization, the result for $\Gamma^{(3)}$ would have differed by a term $\delta\Gamma^{(3)}$ proportional to (4.9):

$$\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = K\epsilon_{\lambda\mu\nu\rho}(p_1-p_2)_\rho\,. \tag{4.24}$$

The constant K then is determined by the condition of gauge invariance

$$p_{1\mu}\left[\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) + \delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)\right] = 0\,,$$

which yields

$$p_{1\mu}\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = -\frac{e^2}{6\pi^2}\epsilon_{\lambda\nu\rho\sigma}k_\rho p_{2\sigma} \;\Rightarrow\; K = e^2/(6\pi^2). \tag{4.25}$$

This gives an additional contribution to the divergence of the current

$$k_\lambda\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{3\pi^2}\epsilon_{\mu\lambda\rho\sigma}p_{1\rho}p_{2\sigma}\,. \tag{4.26}$$

Therefore in a QED-like gauge invariant field theory with massless fermions the axial current is not conserved: this is called the chiral anomaly. For any gauge invariant regularization one finds

$$k_\lambda\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \left(\frac{e^2}{2\pi^2} \equiv \frac{2\alpha}{\pi}\right)\epsilon_{\mu\nu\rho\sigma}p_{1\rho}p_{2\sigma}\,. \tag{4.27}$$

Equation (4.27) can be rewritten after Fourier transformation as an axial current non-conservation equation:

$$\partial_\lambda J^5_\lambda(x) = -i\frac{\alpha}{4\pi}\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\,. \tag{4.28}$$

Since global chiral symmetry is not broken, the integral over the whole space of the anomalous term must vanish. This condition is indeed verified since the anomaly can immediately be written as a total derivative:

$$\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = 4\partial_\mu(\epsilon_{\mu\nu\rho\sigma}A_\nu\partial_\rho A_\sigma). \tag{4.29}$$

The space integral of the anomalous term only depends on the behaviour of the gauge field at boundaries, and this property already indicates a relation between *topology and anomalies*.

Equation (4.28) also implies:

$$\ln\det\left[\mathrm{e}^{i\gamma_5\theta(x)}\,\not{D}\,\mathrm{e}^{i\gamma_5\theta(x)}\right] = \ln\det\not{D} - i\frac{\alpha}{4\pi}\int \mathrm{d}^4x\,\theta(x)\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}(x)F_{\rho\sigma}(x). \tag{4.30}$$

Remark. One might be surprised that in the calculation the divergence of the axial does not vanish, though the regularization of the fermion propagator seems to be consistent with chiral symmetry. The reason is simple: if we add for example higher derivative terms in the action, the form of the axial current is modified and the additional contributions cancel the term we have found.

In the form we have presented it the calculation generalizes without difficulty to general even dimensions $2n$. Note simply that the permutation $(\mathbf{p}_1,\mu) \leftrightarrow (\mathbf{p}_2,\nu)$

in equation (4.16) is replaced by a cyclic permutation. The anomaly in the divergence of the axial current $J^S_\lambda(x)$ in general is

$$\partial_\lambda J^S_\lambda(x) = -2i\frac{e^n}{(4\pi)^n n!}\epsilon_{\mu_1\nu_1\ldots\mu_n\nu_n}F_{\mu_1\nu_1}\ldots F_{\mu_n\nu_n}\,, \tag{4.31}$$

where $\epsilon_{\mu_1\nu_1\ldots\mu_n\nu_n}$ is the completely antisymmetric tensor.

Boson regulator. We have seen that we could also regularize by adding massive fermions and bosons with spin, the unpaired boson breaking chiral symmetry. Then the total contribution to the divergence of the axial current comes from this boson, whose mass we denote by M, leading to

$$\begin{aligned} C^{(2)}(k) = -2iMe^2 \int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\,A_\mu(p_1)A_\nu(p_2)\int\frac{\mathrm{d}^4q}{(2\pi)^4} \\ \times\,\mathrm{tr}\left[\gamma_5(\not{q}+\not{k}-iM)^{-1}\gamma_\mu(\not{q}-\not{p}_2-iM)^{-1}\gamma_\nu(\not{q}-iM)^{-1}\right]. \end{aligned} \tag{4.32}$$

The apparent divergences of this contribution are subtracted by formally vanishing diagrams that we do not write, but which justify the following formal manipulations.

Taking the trace one see that the formal divergences cancel and one obtains

$$\begin{aligned} C^{(2)}(k) \underset{M\to\infty}{\sim} 8M^2e^2\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\,p_{1\rho}p_{2\sigma}A_\mu(p_1)A_\nu(p_2) \\ \times\frac{1}{(2\pi)^4}\int\frac{\mathrm{d}^4q}{(q^2+M^2)^3}. \end{aligned}$$

The limit $M\to\infty$ corresponds to remove the regulator. The limit is finite because after rescaling of q the mass can be eliminated. One finds:

$$C^{(2)}(k) \underset{M\to\infty}{\sim} \frac{e^2}{4\pi^2}\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\,(p_2)p_{1\rho}p_{2\sigma}A_\mu(p_1)A_\nu\,,$$

in agreement with equation (4.27).

Point-splitting regularization. Another calculation, based on regularization by point splitting, gives further insight into the mechanism which generates the anomaly. We thus consider the non-local operator

$$J^5_\mu(x,a) = i\bar\psi(x-a/2)\gamma_5\gamma_\mu\psi(x+a/2)\exp\left[ie\int_{x-a/2}^{x+a/2}A_\lambda(s)\mathrm{d}s_\lambda\right], \tag{4.33}$$

in the limit $|a|\to 0$. To avoid a breaking of rotation symmetry by the regularization, before taking the limit $|a|\to 0$ we will average over all orientations of the vector a. The multiplicative gauge factor (parallel transporter) ensures gauge invariance of the regularized operator (transformations (4.3)). The divergence of the operator for $|a| \to 0$ then becomes

$$\partial_\mu^x J_\mu^5(x,a) \sim -e a_\lambda \bar{\psi}(x-a/2)\gamma_5\gamma_\mu F_{\mu\lambda}(x)\psi(x+a/2) \times \exp\left[ie\int_{x-a/2}^{x+a/2} A_\lambda(s)\mathrm{d}s_\lambda\right], \tag{4.34}$$

where the $\psi, \bar{\psi}$ field equations have been used. We now expand the functional average of the equation in powers of A. The first term vanishes. The second term is quadratic in A and yields

$$\left\langle \partial_\mu^x J_\mu^5(x,a)\right\rangle \sim ie^2 a_\lambda F_{\mu\lambda}(x)\int \mathrm{d}^4y\, A_\nu(y+x)\,\mathrm{tr}\,\gamma_5\Delta_\mathrm{F}(y-a/2)\gamma_\nu\Delta_\mathrm{F}(-y-a/2)\gamma_\mu\,,$$

where $\Delta_\mathrm{F}(y)$ is the fermion propagator:

$$\Delta_\mathrm{F}(y) = -\frac{i}{(2\pi)^4}\int \mathrm{d}^4k\,\frac{\not{k}}{k^2} = \frac{1}{2\pi^2}\frac{\not{y}}{y^4}\,. \tag{4.35}$$

We now take the trace. The propagator is singular for $|y| = O(|a|)$ and therefore we can expand $A_\nu(x+y)$ in powers of y. The first term vanishes for symmetry reasons ($y \mapsto -y$), and we obtain

$$\left\langle \partial_\mu^x J_\mu^5(x,a)\right\rangle \sim \frac{ie^2}{\pi^4}\epsilon_{\mu\nu\tau\sigma}a_\lambda F_{\mu\lambda}(x)\partial_\rho A_\nu(x)\int \mathrm{d}^4y\,\frac{y_\rho y_\sigma a_\tau}{|y+a/2|^4|y-a/2|^4}\,.$$

The integral over y gives a linear combination of $\delta_{\rho\sigma}$ and $a_\rho a_\sigma$ but the second term gives a vanishing contribution due to ϵ symbol. It follows

$$\left\langle \partial_\mu^x J_\mu^5(x,a)\right\rangle \sim \frac{ie^2}{3\pi^4}\epsilon_{\mu\nu\tau\rho}a_\lambda a_\tau F_{\mu\lambda}(x)\partial_\rho A_\nu(x)\int \mathrm{d}^4y\,\frac{y^2-(y\cdot a)^2/a^2}{|y+a/2|^4|y-a/2|^4}\,.$$

After integration we then find

$$\left\langle \partial_\mu^x J_\mu^5(x,a)\right\rangle \sim \frac{ie^2}{4\pi^2}\epsilon_{\mu\nu\tau\rho}\frac{a_\lambda a_\tau}{a^2}F_{\mu\lambda}(x)F_{\rho\nu}(x).$$

Averaging over the a directions we see that the divergence is finite for $|a| \to 0$, and thus finally

$$\lim_{|a|\to 0}\left\langle \partial_\mu^x \overline{J_\mu^5(x,a)}\right\rangle = \frac{ie^2}{16\pi^2}\epsilon_{\mu\nu\lambda\rho}F_{\mu\lambda}(x)F_{\rho\nu}(x),$$

and we recover the result (4.28).

Chiral gauge theory. A gauge theory is consistent only if the gauge field is coupled to a conserved current. An anomaly which affects the current destroys gauge invariance at the quantum level. Therefore the theory with axial gauge symmetry, where the action in the fermion sector reads

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\, \bar{\psi}(x)(\not{\partial} + ig\gamma_5 \not{B})\psi(x), \tag{4.36}$$

is inconsistent. Indeed current conservation applies to the AAA vertex at one-loop order. Because now the three point vertex is symmetric the divergence is given by the expression (4.21), and thus does not vanish.

More generally the anomaly prevents the construction of a theory which would have both an abelian gauge vector and axial symmetry, where the action in the fermion sector would read

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\, \bar{\psi}(x)(\not{\partial} + ie\not{A} + i\gamma_5 g\not{B})\psi(x). \tag{4.37}$$

A way to solve both problems is to cancel the anomaly by introducing another fermion of opposite chiral coupling. With more fermions other combination of couplings are possible. Note, however that in the purely axial gauge theory it is easy to verify that a theory with two fermions of opposite chiral charges can be rewritten as a vector theory by combining differently the chiral components of both fermions.

4.3 Non-abelian vector gauge fields and abelian axial current

We still consider an abelian axial current but now in the framework of a non-abelian gauge theory. The fermion fields transform non-trivially under a gauge group G and $\mathbf{A}_\mu$ is the corresponding gauge field. The action is:

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\, \bar{\psi}(x)\not{\mathbf{D}}\psi(x), \tag{4.38}$$

with:

$$\not{\mathbf{D}} = \not{\partial} + \not{\mathbf{A}}\,. \tag{4.39}$$

The axial current J^5_μ

$$J^5_\mu(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x),$$

is still gauge invariant. Therefore no new calculation is needed; the result is completely determined by dimensional analysis, gauge invariance and the previous calculation which yields the term of order $\mathbf{A}^2$, :

$$\partial_\lambda J^5_\lambda(x) = -\frac{i}{16\pi^2}\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma}\,, \tag{4.40}$$

in which $\mathbf{F}_{\mu\nu}$ now is the corresponding curvature (2.15). Again this expression must be a total derivative. One verifies:

$$\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma} = 4\,\epsilon_{\mu\nu\rho\sigma}\partial_\mu\,\mathrm{tr}(\mathbf{A}_\nu\partial_\rho\mathbf{A}_\sigma + \tfrac{2}{3}\mathbf{A}_\nu\mathbf{A}_\rho\mathbf{A}_\sigma). \tag{4.41}$$

4.4 Anomaly and eigenvalues of the Dirac operator

We assume that the spectrum of $\not{\mathbf{D}}$, the Dirac operator in a non-abelian gauge field (equation (4.39)), is discrete (putting temporarily the fermions in a box if necessary) and call d_n and $\varphi_n(x)$ the corresponding eigenvalues and eigenvectors:

$$\not{\mathbf{D}}\varphi_n = d_n\varphi_n \,. \tag{4.42}$$

For a unitary or orthogonal group the massless Dirac operator is antihermitian; therefore the eigenvalues are imaginary and the eigenvectors orthogonal. In addition we choose them with unit norm.

The eigenvalues are gauge invariant, because in a gauge transformation of unitary matrix $\mathbf{g}(x)$ the Dirac operator becomes

$$\not{\mathbf{D}} \mapsto \mathbf{g}^{-1}(x)\not{\mathbf{D}}\mathbf{g}(x) \;\Rightarrow\; \varphi_n(x) \mapsto \mathbf{g}(x)\varphi_n(x).$$

The anticommutation $\not{\mathbf{D}}\gamma_5 + \gamma_5\not{\mathbf{D}} = 0$ implies

$$\not{\mathbf{D}}\gamma_5\varphi_n = -d_n\gamma_5\varphi_n \,. \tag{4.43}$$

Therefore either d_n is different from zero, and $\gamma_5\varphi_n$ is an eigenvector of $\not{\mathbf{D}}$ with eigenvalue $-d_n$, or d_n vanishes. The eigenspace corresponding to the eigenvalue 0 then is invariant under γ_5 which can be diagonalized: the eigenvectors of $\not{\mathbf{D}}$ can be chosen eigenvectors of definite chirality, i.e. eigenvectors of γ_5 with eigenvalue ± 1,

$$\not{\mathbf{D}}\varphi_n = 0\,, \quad \gamma_5\varphi_n = \pm\varphi_n \,.$$

We call n_+ and n_- the dimensions of the eigenspace of positive and negative chirality respectively.

We now consider the determinant of the operator $\not{\mathbf{D}} + m$ regularized by mode truncation (mode regularization):

$$\det{}_N(\not{\mathbf{D}} + m) = \prod_{n\le N} (d_n + m), \tag{4.44}$$

keeping the N lowest eigenvalues of $\not{\mathbf{D}}$ (in modulus), with $N - n_+ - n_-$ even, in such a way that the corresponding subspace is γ_5 invariant.

The regularization is gauge invariant because the eigenvalues of $\not{\mathbf{D}}$ are gauge invariant.

Note that in the truncated space the trace γ_5 does not vanish if $n_+ \neq n_-$

$$\operatorname{tr}\gamma_5 = n_+ - n_- \,, \tag{4.45}$$

a situation which endangers axial current conservation.

In a chiral transformation (4.4) the determinant of $(\not{\mathbf{D}} + m$ becomes

$$\det{}_N(\not{\mathbf{D}} + m) \mapsto \det{}_N\bigl(\mathrm{e}^{i\theta\gamma_5}(\not{\mathbf{D}} + m)\,\mathrm{e}^{i\theta\gamma_5}\bigr).$$

We now consider the various eigenspaces.

If $d_n \neq 0$ the matrix γ_5 is represented by the Pauli matrix σ_1 in the sum of eigenspaces corresponding to the two eigenvalues $\pm d_n$ and $\not{D} + m$ by $d_n\sigma_3 + m$. The determinant in the subspace then is

$$\det\left(\mathrm{e}^{i\theta\sigma_1}(d_n\sigma_3 + m)\,\mathrm{e}^{i\theta\sigma_1}\right) = \det \mathrm{e}^{2i\theta\sigma_1} \det(d_n\sigma_3 + m) = m^2 - d_n^2,$$

because σ_1 is traceless.

In the eigenspace of vanishing eigenvalues d_n with eigenvectors with positive chirality, γ_5 is diagonal with eigenvalue 1, of dimension n_+, and thus

$$m^{n_+} \mapsto m^{n_+}\,\mathrm{e}^{2i\theta n_+},$$

and in the same way in the eigenspace of chirality -1

$$m^{n_-} \mapsto m^{n_-}\,\mathrm{e}^{-2i\theta n_-}\,.$$

We conclude

$$\det{}_N\left(\mathrm{e}^{i\theta\gamma_5}(\not{D} + m)\,\mathrm{e}^{i\theta\gamma_5}\right) = \mathrm{e}^{2i\theta(n_+ - n_-)} \det{}_N(\not{D} + m),$$

The ratio of both determinants is independent of N. Taking the limit $N \to \infty$ we find

$$\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D} + m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D} + m)^{-1}\right] = \mathrm{e}^{2i\theta(n_+ - n_-)}\,. \tag{4.46}$$

Note that the l.h.s. of equation (4.46) is obviously 1 when $\theta = n\pi$, which implies that the coefficient of 2θ in the r.h.s. must indeed be an integer.

The variation of $\ln\det(\not{D} + m)$

$$\ln\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D} + m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D} + m)^{-1}\right] = 2i\theta\,(n_+ - n_-)\,,$$

at first order in θ is related to the variation of the action (4.1) and thus to the functional average of the integral of the divergence of the axial current $\left\langle \int \mathrm{d}^4x \partial_\mu J^5_\mu(x)\right\rangle$. In the limit $m = 0$ it is thus related to the space integral of the chiral anomaly (4.40):

$$-\frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4x\ \mathrm{tr}\,\mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma} = n_+ - n_-\,. \tag{4.47}$$

Concerning this result several comments can be made:

(i) At first order in θ in the absence of regularization we have calculated ($\ln\det = \mathrm{tr}\ln$)

$$\ln\det\left[1 + i\theta\left(\gamma_5 + (\not{D} + m)\gamma_5(\not{D} + m)^{-1}\right)\right] \sim 2i\theta\,\mathrm{tr}\,\gamma_5\,,$$

where we have used the cyclic property of the trace. Since the trace of the matrix γ_5 vanishes we could expect naively a vanishing result. But trace here means trace in γ space and in coordinate space, and γ_5 really stands here for $\gamma_5\delta(x-y)$. The mode regularization yields a well-defined finite result for the ill defined product $0\times\delta^d(0)$.

(ii) The property that the integral (4.47) is quantized shows that the form of the anomaly is related to topological properties of the gauge field since the integral does not change when the gauge field is deformed continuously. The integral of the anomaly over the whole space thus depends only on the behaviour at large distances of the curvature tensor $\mathbf{F}_{\mu\nu}$ and the anomaly must a total derivative as equation (4.41) confirms.

(iii) One might be surprised that $\det \not{\mathbf{D}}$ is not invariant under global chiral transformations. However we have just established that when the integral of the anomaly does not vanish, $\det \not{\mathbf{D}}$ vanishes. This explains that to give a meaning to the r.h.s. of equation (4.46) we have been forced to introduce a mass to find a non-trivial result. The determinant of $\not{\mathbf{D}}$ in the subspace orthogonal to eigenvectors with vanishing eigenvalue, even in presence of a mass, is chiral invariant by parity doubling, but for $n_+\neq n_-$ not the determinant in the eigenspace of eigenvalue zero because the trace of γ_5 does not vanish in the eigenspace (equation (4.45)). In the limit $m\to 0$ the complete determinant vanishes but not the ratio of determinants for different values of θ because the powers of m cancel.

4.5 Anomaly and instantons

Gauge field configurations exist for which the r.h.s. of equation (4.47) does not vanish. A specially interesting example is provided by instantons, i.e. finite action solutions of euclidean field equations. If we write the gauge action

$$\mathcal{S}(A_\mu)=-\tfrac{1}{4}\int \mathrm{d}^4x\,\mathrm{tr}\,\mathbf{F}^2_{\mu\nu}(x),$$

the field equations are

$$\mathbf{D}_\nu\mathbf{F}_{\nu\mu}=0\,.$$

It is convenient to introduce here the notation

$$\tilde{\mathbf{F}}_{\mu\nu}=\tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}\mathbf{F}_{\rho\sigma}\,.$$

Then the inequality

$$-\int \mathrm{d}^4x\,\mathrm{tr}\left[\mathbf{F}_{\mu\nu}(x)\pm\tilde{\mathbf{F}}_{\mu\nu}(x)\right]^2\geq 0\,, \tag{4.48}$$

yields

$$-\int \mathrm{d}^4x\,\mathrm{tr}\,(\mathbf{F}_{\mu\nu})^2\geq\left|\int \mathrm{d}^4x\ \mathrm{tr}\,\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}_{\mu\nu}\right|, \tag{4.49}$$

an inequality that relates the action to the anomaly.

The finiteness of the action $\mathcal{S}(A_\mu)$ implies that the classical solution asymptotically must become a pure gauge

$$\mathbf{A}_\mu(x) \underset{|x|\to\infty}{\sim} \mathbf{g}(x)\partial_\mu \mathbf{g}^{-1}(x) = -\partial_\mu \mathbf{g}(x)\mathbf{g}^{-1}(x). \tag{4.50}$$

In the example of the $SU(2)$ gauge group, the group element $\mathbf{g}$ can be parametrized in terms of Pauli σ matrices

$$\mathbf{g} = u_0 + i\mathbf{u}\cdot\sigma\,,$$

where $(u_0, \mathbf{u})$ is a four-component real vector belonging the unit sphere S_3:

$$u_0^2 + \mathbf{u}^2 = 1\,.$$

Since $SU(2)$ is topologically equivalent to the sphere S_3, the pure gauge configurations on a sphere of large radius $|x| = R$ define a mapping from S_3 to S_3. Such mappings belong to different homotopy classes which are characterized by an integer called the *winding number*.

The simplest mapping with winding number one corresponds to an element $\mathbf{g}(x)$ of the form (we denote the four coordinates x by $(\mathbf{x}, x_4)$)

$$\mathbf{g}(x\ ) = \frac{x_4 + i\mathbf{x}\cdot\sigma}{r}, \quad r = (x_4^2 + \mathbf{x}^2)^{1/2}, \tag{4.51}$$

and thus for $r \to \infty$

$$\mathbf{A}_i \sim -\frac{i}{r^2}(x_4\sigma_i + \epsilon_{ijk}x_j\sigma_k), \quad \mathbf{A}_4 \sim \frac{i}{r^2}\mathbf{x}\cdot\sigma\,.$$

Integrating inside a sphere of large radius R, using equation (4.41) and Stokes theorem one finds

$$Q = -\int \mathrm{d}^4x\ \mathrm{tr}\,\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}_{\mu\nu} = R^3\int_{r=R}\mathrm{d}\Omega\ \frac{x_\mu}{r}V_\mu\,,$$

in which $\mathrm{d}\Omega$ is the measure on the sphere, x_μ/r the unit vector normal to the sphere and

$$V_\mu = -2\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}(\mathbf{A}_\nu\partial_\rho\mathbf{A}_\sigma + \tfrac{2}{3}\mathbf{A}_\nu\mathbf{A}_\rho\mathbf{A}_\sigma).$$

When $\mathbf{A}_\mu$ approaches a pure gauge as in (4.50), for $R \to \infty$

$$V_\mu \sim \tfrac{2}{3}\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\mathbf{A}_\nu\mathbf{A}_\rho\mathbf{A}_\sigma\,.$$

Parametrizing $\mathbf{A}_\mu$ in terms of real components A_μ^i

$$\mathbf{A}_\mu = iA_\mu^i\sigma_i\,,$$

we find

$$V_\mu \sim \tfrac{4}{3}\epsilon_{\mu\nu\rho\sigma}\epsilon_{ijk}A^i_\nu A^j_\rho A^k_\sigma = 8x_\mu r^2 = 8n_\mu R^3\,.$$

The powers of R cancel and since $\int \mathrm{d}\Omega = 2\pi^2$ we finally obtain

$$Q = 16\pi^2. \tag{4.52}$$

If we compare this result with equation (4.47) we see that we have indeed found the minimal non-vanishing value since in general we expect

$$\int \mathrm{d}^4x\,\mathbf{F}_{\mu\nu}\cdot\tilde{\mathbf{F}}_{\mu\nu} = 16\pi^2 n\,.$$

The inequality (4.49) then implies

$$\mathcal{S}(A_\mu) \ge 4\pi^2\,. \tag{4.53}$$

The equality, which corresponds to a local minimum of the action, is obtained for fields satisfying the self-duality equations

$$\mathbf{F}_{\mu\nu} = \pm\tilde{\mathbf{F}}_{\mu\nu}\,, \tag{4.54}$$

which are first order partial differential equations. The corresponding one-instanton solution, which depends on a scale parameter λ, is

$$\mathbf{A}_i = -\frac{i}{r^2+\lambda^2}(x_4\sigma_i + \epsilon_{ijk}x_j\sigma_k), \quad \mathbf{A}_4 = \frac{i}{r^2+\lambda^2}\mathbf{x}\cdot\sigma\,.$$

We have shown above that if massless fermions are coupled to such gauge field configurations the determinant resulting from the fermion integration necessarily vanishes.

5 Non-abelian anomaly

We first consider the problem of conservation of a general axial current in a non-abelian vector gauge theory, and then the issue of obstruction to gauge invariance in chiral gauge theories.

5.1 General axial current

We now discuss the problem of the conservation of a general axial current in the example of a fermion action which has a $G\times G$ chiral symmetry, in the background of non-abelian vector gauge fields. The generators of the gauge group may or may not be related to the diagonal subgroup G of $G\times G$ which correspond to vector currents.

We call t^α the generators of G. The current then has the form:

$$J_\mu^{5\alpha}(x) = -\bar{\psi}\gamma_5\gamma_\mu t^\alpha\psi\,. \tag{5.1}$$

When the gauge group is connected with the chiral group, the current conservation equation involves the gauge covariant derivative:

$$\mathbf{D}_\mu \mathbf{J}_\mu^5 = 0\,, \tag{5.2}$$

where the covariant derivative in the adjoint representation takes the form

$$\mathbf{D}_\mu = \partial_\mu + [A_\mu, \bullet].$$

In the calculation of the contribution to the anomaly coming from terms quadratic in the gauge fields the only modification in the previous results is the appearance of a different geometrical factor. Then the complete form of the anomaly is dictated by gauge covariance. One finds:

$$\mathbf{D}_\lambda J_\lambda^{5\alpha}(x) = -\frac{i}{16\pi^2}\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,t^\alpha \mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma}\,. \tag{5.3}$$

In particular if the gauge group is disconnected from the chiral group the anomaly is proportional to $\mathrm{tr}\,t^\alpha$ and therefore only different from zero for the abelian factors of G.

5.2 Obstruction to gauge invariance

We now want to consider a non-abelian gauge field coupled to left or right-handed fermions:

$$\mathcal{S}\left(\bar{\psi},\psi\right) = -\int \mathrm{d}^4x\,\bar{\psi}\left(x\right)\tfrac{1}{2}\left(1+\gamma_5\right)\not{\mathbf{D}}\psi\left(x\right), \tag{5.4}$$

($\frac{1}{2}\left(1-\gamma_5\right)$ is treated in the same way). We can construct a consistent gauge theory only if the partition function

$$\mathcal{Z}\left(\mathbf{A}_\mu\right) = \int \left[\mathrm{d}\psi \mathrm{d}\bar{\psi}\right]\exp\left[-\mathcal{S}\left(\psi,\bar{\psi}\right)\right] \tag{5.5}$$

is gauge invariant.

If we introduce the generators t^α of the gauge group in the fermion representation, we can write the corresponding current $\mathbf{J}_\mu$ as:

$$J_\mu^\alpha(x) = -\bar{\psi}\tfrac{1}{2}\left(1+\gamma_5\right)\gamma_\mu t^\alpha\psi\,. \tag{5.6}$$

The invariance of $\mathcal{Z}(\mathbf{A}_\mu)$ under an infinitesimal gauge transformation again leads to a covariant conservation equation for the current:

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu \rangle = 0\,.$$

The calculation of the term of degree two in the gauge field of the anomaly is straightforward: the regularization adopted for the calculation in section 4.2 is also suited to the present case since the current-gauge field three-point function is symmetric in the external arguments. The group structure is reflected by a simple geometrical factor. The global factor can be taken from the abelian calculation. It differs from result (4.21) by a factor 1/2 which comes from the projector $\frac{1}{2}(1+\gamma_5)$. The general form of the term of degree three in the gauge field can also easily be found, but the calculation of the global factor is somewhat tedious. We argue in the next section that it can be obtained from consistency conditions. The complete expression reads:

$$(\mathbf{D}_\mu \mathbf{J}_\mu(x))^\alpha = -\frac{i}{24\pi^2}\partial_\mu \epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\left[t^\alpha\left(\mathbf{A}_\nu\partial_\rho\mathbf{A}_\sigma + \tfrac{1}{2}\mathbf{A}_\nu\mathbf{A}_\rho\mathbf{A}_\sigma\right)\right]. \qquad (5.7)$$

If the projector $\frac{1}{2}(1+\gamma_5)$ is replaced by $\frac{1}{2}(1-\gamma_5)$ the sign of the anomaly changes.

Unless this term vanishes identically there is an obstruction to the construction of the gauge theory. It is easy to verify, taking into account the antisymmetry of the ϵ tensor, that the group factor is:

$$d_{\alpha\beta\gamma} = \tfrac{1}{2}\,\mathrm{tr}\left[t^\alpha\left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right]. \qquad (5.8)$$

For a unitary representation the generators t^α are, with our conventions, anti-hermitian. Therefore the coefficients $d_{\alpha\beta\gamma}$ are purely imaginary:

$$d^*_{\alpha\beta\gamma} = \tfrac{1}{2}\,\mathrm{tr}\left[t^\alpha\left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right]^* = -d_{\alpha\beta\gamma}\,. \qquad (5.9)$$

For all real (the t^α antisymmetric) or "pseudo-real" ($t^\alpha = -S\,{}^T t^\alpha S^{-1}$) representations these coefficients vanish. It follows that the only non-abelian groups which can lead to anomalies in four dimensions are: $SU(N)$ for $N \geq 3$, $SO(6)$ and E_6.

5.3 Wess–Zumino consistency conditions

In section 5.2 we have calculated the part of the anomaly which is quadratic in the gauge field and asserted that the remaining part could be obtained from geometric arguments. The method is based on BRS transformations. The anomaly is the variation of a functional under an infinitesimal gauge transformation. This implies compatibility conditions, which here are constraints on the form of the anomaly.

One convenient way to express these constraints is to express the nilpotency of BRS transformations. In a BRS transformation the variation of the gauge field $\mathbf{A}_\mu$ takes the form:

$$\delta_{\mathrm{BRS}} \mathbf{A}_\mu(x) = \mathbf{D}_\mu \mathbf{C}(x) \bar{\varepsilon} \,, \tag{5.10}$$

where $\mathbf{C}$ is a (fermion) ghost field and $\bar{\varepsilon}$ an anticommuting constant. The corresponding variation of $\ln \mathcal{Z}(\mathbf{A}_\mu)$ is:

$$\delta_{\mathrm{BRS}} \ln \mathcal{Z}(\mathbf{A}_\mu) = -\int \mathrm{d}^4 x \, \langle \mathbf{J}_\mu(x) \rangle \, \mathbf{D}_\mu \mathbf{C}(x) \bar{\varepsilon} \,. \tag{5.11}$$

The anomaly equation has the general form:

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu(x) \rangle = \mathcal{A}\left(\mathbf{A}_\mu ; x\right) . \tag{5.12}$$

In terms of $\mathcal{A}$ the equation (5.11), after an integration by parts, can be rewritten:

$$\delta_{\mathrm{BRS}} \ln \mathcal{Z}(\mathbf{A}_\mu) = \int \mathrm{d}^4 x \, \mathcal{A}\left(\mathbf{A}_\mu ; x\right) \mathbf{C}(x) \bar{\varepsilon} \,. \tag{5.13}$$

We recall that a BRS variation is invariant under BRS transformations, $\delta^2_{\mathrm{BRS}} = 0$. To calculate the BRS variation of $\mathcal{A}\mathbf{C}$ we also need the BRS transformation of the fermion ghost $\mathbf{C}(x)$:

$$\delta_{\mathrm{BRS}} \mathbf{C}(x) = \bar{\varepsilon} \, \mathbf{C}^2(x). \tag{5.14}$$

By expressing that $\mathcal{A}C$ is BRS invariant one obtains a constraint on the possible form of anomalies. One verifies that this condition determines the term cubic in $\mathbf{A}$ in the r.h.s. of equation (5.7).

6 Lattice fermions: Ginsparg–Wilson relation

Notation. We now return to the problem of lattice fermions discussed in section 3.4. For convenience we set the lattice spacing $a = 1$ and use for the fields the notation $\psi(x) \equiv \psi_x$.

Ginsparg–Wilson relation. Recently a decisive advance on the problem of chiral fermions has been achieved. It had been noted, many years ago, that a potential way to avoid the doubling problem while still retaining chiral properties in the continuum limit was to construct a lattice Dirac operator $\mathbf{D}$ which, instead of anticommuting with γ_5, would satisfy the relation

$$\{\mathbf{D}^{-1}, \gamma_5\} = \gamma_5 \tag{6.1}$$

where γ_5 in the r.h.s. means $\gamma_5 \delta_{xy}$, where δ_{xy} is the identity for lattice sites. More generally the r.h.s. can be replaced by any local operator on the lattice. This modification being local, it is expected that it does not affect correlation functions at large distance and that chiral properties are recovered in the continuum limit. Note that when $\mathbf{D}$ is the Dirac operator in a gauge field the condition (6.1) is gauge invariant.

However, only recently have lattice Dirac operators solutions to the Ginsparg–Wilson relation (6.1) been discovered, because the demands that both $\mathbf{D}$ and the anticommutator $\{\mathbf{D}^{-1}, \gamma_5\}$ should be local, seemed difficult to satisfy, specially in the most interesting case of gauge theories.

Let us briefly discuss the main properties of a Dirac operator satisfying relation (6.1). Using the relation, quite generally true for an euclidean Dirac operator (consequence of hermiticity and reflection symmetry),

$$\mathbf{D}^{\dagger} = \gamma_5 \mathbf{D} \gamma_5 \,, \tag{6.2}$$

one can rewrite the relation (6.1)

$$\mathbf{D}^{-1} + \left(\mathbf{D}^{-1}\right)^{\dagger} = \mathbf{1} \,,$$

and therefore

$$\mathbf{D} + \mathbf{D}^{\dagger} = \mathbf{D}\mathbf{D}^{\dagger} = \mathbf{D}^{\dagger}\mathbf{D} \,. \tag{6.3}$$

This implies that the lattice operator $\mathbf{D}$ has an index, and in addition

$$\mathbf{S} = \mathbf{1} - \mathbf{D} \,, \tag{6.4}$$

is unitary

$$\mathbf{S}\mathbf{S}^{\dagger} = \mathbf{1} \,. \tag{6.5}$$

The eigenvalues of $\mathbf{S}$ lie on the unit circle. The eigenvalue one corresponds to the pole of the Dirac propagator.

It is then possible to construct lattice actions that have a chiral symmetry which corresponds to local but non point-like transformations. In the abelian example,

$$\delta \psi_x = i\theta \sum_y \gamma_5 \left(\mathbf{1} - \mathbf{D}\right)_{xy} \psi_y \,, \tag{6.6}$$

$$\delta \bar{\psi}_x = i\theta \bar{\psi}_x \gamma_5 \,. \tag{6.7}$$

Indeed, from relation (6.1),

$$\delta(\bar{\psi}\mathbf{D}\psi) = i\theta \bar{\psi} \left[\gamma_5 \mathbf{D} + \mathbf{D}\gamma_5 (1 - \mathbf{D})\right] \psi = 0 \,.$$

Note that in this form the transformation generalizes immediately to non-abelian groups.

The problem is that these transformations no longer leave the integration measure over the fermion fields,

$$\prod_x \mathrm{d}\psi_x \mathrm{d}\bar{\psi}_x \,,$$

automatically invariant. Indeed an infinitesimal change of variables

$$\begin{aligned}\frac{\partial \psi'_x}{\partial \psi_y} &= \delta_{xy} + i\theta\gamma_5 \left(\mathbf{1}-\mathbf{D}\right)_{xy} \\ \frac{\partial \bar{\psi}'_x}{\partial \bar{\psi}_y} &= \delta_{xy}\left(1+i\theta\gamma_5\right),\end{aligned}$$

leads to the jacobian J

$$J = 1 + i\theta \operatorname{tr} \gamma_5 (2-\mathbf{D})\,,$$

where trace means trace in the space of γ matrices and in the lattice indices. This leaves the possibility of generating the expected anomalies, when the simple Wilson–Dirac operator of the free theory is replaced by the covariant operator in the background of a gauge field.

Eigenvalues of Dirac operator in a gauge background. We briefly show that the Dirac operator has an index. We assume that its spectrum is discrete (this is certainly true on a finite lattice where it is a matrix). We denote by $|n\rangle$ its n^{th} eigenvector,

$$\mathbf{D}\,|n\rangle = (1-\mathrm{e}^{i\theta_n})\,|n\rangle \ \Rightarrow \mathbf{D}^\dagger\,|n\rangle = (1-\mathrm{e}^{-i\theta_n})\,|n\rangle \ .$$

Then using equation (6.2), we infer

$$\gamma_5 \mathbf{D}\,|n\rangle = (1-\mathrm{e}^{-i\theta_n})\gamma_5\,|n\rangle \,.$$

The discussion which follows then is analogous to the discussion of Section 4.4 to which we refer for details. We note that when the eigenvalues are not real, $\theta_n \neq 0$ (mod π), $\gamma_5\,|n\rangle$ is an eigenvector different from $|n\rangle$ because the eigenvalues are different. Instead in the two subspaces corresponding to the eigenvalues 0 and 2, we can choose eigenvectors with definite chirality

$$\gamma_5\,|n\rangle = \pm\,|n\rangle \,.$$

We call below $n_\pm$ the number of eigenvalues 0, and $\nu_\pm$ the number of eigenvalues 2 with chirality ± 1.

Note that on a finite lattice δ_{xy} is a finite matrix, and thus

$$\operatorname{tr} \gamma_5 \delta_{xy} = 0\,.$$

Therefore

$$\operatorname{tr}\gamma_5(2-\mathbf{D}) = -\operatorname{tr}\gamma_5\mathbf{D}\,,$$

which implies

$$\sum_n \langle n|\,\gamma_5(2-\mathbf{D})\,|n\rangle = -\sum_n \langle n|\,\gamma_5\mathbf{D}\,|n\rangle\,. \tag{6.8}$$

In the equation all complex eigenvalues cancel because the vector $|n\rangle$ and $\gamma_5\,|n\rangle$ are orthogonal. The sum reduces to the subspace of real eigenvalues, where the eigenvectors have definite chirality. In the l.h.s. only the eigenvalue 0 contributes, and in the r.h.s. only the eigenvalue 2. We find

$$n_+ - n_- = -(\nu_+ - \nu_-).$$

The equation tells us that the difference between the number of states of different chirality in the zero eigenvalue sector is cancelled by the difference in the massive sector of eigenvalue two.

Jacobian and lattice anomaly. The variation of the jacobian can now be evaluated

$$\operatorname{tr}\gamma_5(2-\mathbf{D}) = \sum_n \langle n|\,\gamma_5(2-\mathbf{D})\,|n\rangle = 2(n_+ - n_-).$$

The quantity $\operatorname{tr}\gamma_5(2-\mathbf{D})$ is a sum of terms which are local, gauge invariant, pseudoscalar and topological as the continuum anomaly.

Remark. It is interesting to note the relation between the spectrum of $\mathbf{D}$ and $\gamma_5\mathbf{D}$ which from relation (6.2) is a hermitian matrix

$$\gamma_5\mathbf{D} = \mathbf{D}^\dagger\gamma_5 = (\gamma_5\mathbf{D})^\dagger,$$

and thus is diagonalizable with real eigenvalues. It is easy to verify the two equations, the second being obtained by changing θ in $\theta + 2\pi$,

$$\begin{aligned}
\gamma_5\mathbf{D}(1 - i\,\mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle &= 2\sin(\theta_n/2)(1 - i\,\mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle\,,\\
\gamma_5\mathbf{D}(1 + i\,\mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle &= -2\sin(\theta_n/2)(1 + i\,\mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle\,,
\end{aligned}$$

which implies that the eigenvalues are paired $\pm\sin(\theta_n/2)$ except for $\theta_n = 0$ (mod π) where $|n\rangle$ and $\gamma_5\,|n\rangle$ are proportional. For $\theta_n = 0$ $\gamma_5\mathbf{D}$ has also eigenvalue 0. For $\theta_n = \pi$ $\gamma_5\mathbf{D}$ has eigenvalue ± 2 depending on the chirality of $|n\rangle$.

In the same way

$$\begin{aligned}
\gamma_5(2-\mathbf{D})(1 + \mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle &= 2\cos(\theta_n/2)(1 + \mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle\,,\\
\gamma_5(2-\mathbf{D})(1 - \mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle &= -2\cos(\theta_n/2)(1 - \mathrm{e}^{i\theta_n/2}\gamma_5)\,|n\rangle\,.
\end{aligned}$$

Finally the matrix $\gamma_5(1-\mathbf{D})$ is both hermitian and unitary and has thus eigenvalues ± 1.

Non-abelian generalization. We now assume that we have several fermions on each site and a vector symmetry such that the fields transform like

$$\delta\psi_x = \theta\psi_x \,, \tag{6.9}$$
$$\delta\bar\psi_x = -\bar\psi_x\theta \,, \tag{6.10}$$

where θ are matrices belonging to the representation of a Lie algebra. The invariance of the fermion action implies the commutation relation

$$[\mathbf{D},\theta] = 0\,.$$

The general chiral transformations now read

$$\delta\psi_x = \theta\sum_y \gamma_5\left(\mathbf{1}-\mathbf{D}\right)_{xy}\psi_y \,, \tag{6.11}$$
$$\delta\bar\psi_x = \bar\psi_x\theta\gamma_5 \,. \tag{6.12}$$

Indeed

$$\begin{aligned}\delta(\bar\psi\mathbf{D}\psi) &= \bar\psi\theta\gamma_5\mathbf{D}\psi + \bar\psi\theta\gamma_5\left(\mathbf{1}-\mathbf{D}\right)\psi \\ &= \bar\psi\theta\gamma_5\left[\mathbf{D}+\mathbf{D}^\dagger(1-\mathbf{D})\right]\psi \,,\end{aligned}$$

where the relation (6.2) has been used. The latter expression vanishes as a consequence of the relation (6.3).

Explicit construction. Overlap fermions. In the absence of gauge fields, an explicit solution can be derived from operators which share the properties of the Wilson–Dirac operator $\mathbf{D}_\mathrm{W}$ of equation (3.9), i.e. which avoid doublers at the price of breaking chiral symmetry:

$$D_W(p) = \alpha(p) + i\gamma_\mu\beta_\mu(p), \tag{6.13}$$

where $\alpha(p)$ and $\beta_\mu(p)$ are real, periodic, smooth functions such that

$$\beta_\mu(p) \underset{|p|\to 0}{\sim} p_\mu \,, \quad \alpha(p) \ge 0 \underset{|p|\to 0}{\propto} p^2,$$

and $\alpha(p) \ge 2M > 0$ for all values of p_μ which correspond to doublers, i.e. such that $\beta_\mu(p) = 0$ for $|p| \ne 0$.

Setting

$$\mathbf{A} = \mathbf{1} - \mathbf{D}_\mathrm{W}\,,$$

one takes (the idea of overlap fermions)

$$\mathbf{S} = \mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2} \;\Rightarrow\; \mathbf{D} = \mathbf{1} - \mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2}. \tag{6.14}$$

With this ansatz $\mathbf{D}$ has a zero eigenmode when $\mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2}$ has the eigenvalue one. This can happen when $\mathbf{A}$ and $\mathbf{A}^\dagger$ have the same eigenvector with a *positive* eigenvalue. In the case of the operator (6.13) a short calculation shows that the determinant of $\mathbf{D}$ vanishes when

$$\left[\sqrt{\left(1-\alpha(p)\right)^2+\beta_\mu^2(p)}-1+\alpha(p)\right]^2+\beta_\mu^2(p)=0\,.$$

This implies $\beta_\mu(p)=0$, an equation which necessarily admits doubler solutions, and

$$|1-\alpha(p)|=1-\alpha(p).$$

The solutions to this equation depend on the value of $\alpha(p)$ with respect to one for the doubler modes. If $\alpha(p)\le 1$ the equation is automatically satisfied and the corresponding doubler survives. If instead $\alpha(p)>1$ the equation implies $\alpha(p)=1$ which is impossible. Therefore by rescaling, if necessary, $\alpha(p)$ we can choose $2M>1$ and thus keep the wanted $p_\mu=0$ mode, while eliminating all doublers which then correspond to the eigenvalue two for $\mathbf{D}$, and the doubling problem is in principle solved.

In presence of a gauge field the construction can be generalized and works provided the plaquette action is sufficiently close to one.

Remark. Finally let us stress that, if it seems that the doubling problem has been solved from the formal point of view, from the numerical point of view the calculation of the operator $\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2}$ in a gauge background represents a major challenge.

Acknowledgments. Useful discussions with T.W. Chiu, P. Hasenfratz and H. Neuberger are gratefully acknowledged.

Bibliographical Notes

For an early reference on momentum cut-off regularization see

W. Pauli and F. Villars, *Rev. Mod. Phys.* 21 (1949) 434.

Dimensional regularization has been introduced by:

J. Ashmore, *Lett. Nuovo Cimento* 4 (1972) 289; G. 't Hooft and M. Veltman, *Nucl. Phys.* B44 (1972) 189; C.G. Bollini and J.J. Giambiaggi, *Phys. Lett.* 40B (1972) 566, *Nuovo Cimento* 12B (1972).

Its use in problems with chiral anomalies has been proposed in

D.A. Akyeampong and R. Delbourgo, *Nuovo Cimento* 17A (1973) 578.

For an early review see

G. Leibbrandt, *Rev. Mod. Phys.* 47 (1975) 849.

The consistency of the lattice regularization is proven (except for theories with chiral fermions) in

T. Reisz, *Commun. Math. Phys.* 117 (1988) 79, 639.

The doubling phenomenon for lattice fermions has been proven quite generally by

H.B. Nielsen and M. Ninomiya, *Nucl. Phys.* B185 (1981) 20.

Wilson's solution to the fermion doubling problem is described in

K.G. Wilson in *New Phenomena in Subnuclear Physics*, Erice 1975, A. Zichichi ed. (Plenum, New York 1977).

Staggered fermions have been proposed in

T. Banks, L. Susskind and J. Kogut, *Phys. Rev.* D13 (1976) 1043.

For other schemes see also

E.R. Speer in *Renormalization Theory*, Erice 1975, G. Velo and A.S. Wightman eds. (D. Reidel, Dordrecht, Holland 1976).

The problem of anomalies is discussed in

J.S. Bell and R. Jackiw, *Nuovo Cimento* A60 (1969) 47; S.L. Adler, *Phys. Rev.* 177 (1969) 2426; D.J. Gross and R. Jackiw, *Phys. Rev.* D6 (1972) 477; H. Georgi and S.L. Glashow, *Phys. Rev.* D6 (1972) 429; C. Bouchiat, J. Iliopoulos and Ph. Meyer, *Phys. Lett.* 38B (1972) 519; E. Witten, *Phys. Lett.* 117B (1982) 324.

See also the lectures

M. Peskin, *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, R. Stora and J.-B. Zuber eds. (North-Holland, Amsterdam 1984),

and the volume

S.B. Treiman, R. Jackiw, B. Zumino and E. Witten, *Current Algebra and Anomalies* (World Scientific, Singapore 1985) and references therein.

Several authors have found Dirac operators for chiral fermions by various approaches, and studied the consequences

H. Neuberger, *Phys. Lett.* B417 (1998) 141, hep-lat/9707022, *ibidem* B427 (1998) 353, hep-lat/9801031; P. Hasenfratz, V. Laliena, F. Niedermayer, *Phys. Lett.* B427 (1998) 125, hep-lat/9801021; M. Lüscher, *Phys. Lett.* B428 (1998) 342, hep-lat/9802011, hep-lat/9811032.

They have been found to provide solutions to the Ginsparg–Wilson relation

P.H. Ginsparg and K.G. Wilson, *Phys. Rev.* D25 (1982) 2649.

For more references see the contributions to Proceedings of the workshop “Chiral 99”, *Chinese Journal of Physics*, 38 (2000) 521–743.