Anomalies in Quantum Field Theory: Lecture Course EPFL, Lausanne, May 28, June 4, 11, and 18

M. Shifman[a]

[a] *William I. Fine Theoretical Physics Institute, University of Minnesota, Minneapolis, MN 55455, USA*

Lecture 1

{11}

Introduction and Summary

Since the conception of quantum field theory (QFT) in the late 1920s and till 1960 it was believed that any respectable field theory (and the only one known at that time was quantum electrodynamics, QED) must have a single vacuum (ground state), and the *raison d'être* of QFT is quantization of small oscillations near the vacuum in the same manner as we quantize harmonic oscillator. The quantized fields in the Lagrangian represent particles which can scatter due to cubic and high order terms in the Lagrangian, and the ultimate goal of QFT is to predict the S matrix, in practice order by order in perturbation theory. The problem of renormalization and elimination of "infinities" was believed to be the most fundamental (and mysterious) problem of QFT.

Since then all that changed, with accumulation of evidence that QFT is in fact (a) an effective theory requiring ultraviolet (UV) completion, and (b) it has a much richer structure than was thought previously, in particular, physical states may have nothing to do with the fields that appear in the Lagrangian, the vacua need not be unique (or, even physically equivalent), small field oscillations by far do not exhaust relativistic quantum dynamics, etc. Especially fruitful where breakthrough developments in non-Abelian gauge theories (e.g. QCD), and supersymmetric theories in the 1970s and 80s. Approximately at the same time people learned how to quantize topologically stable solitons and, somewhat later, discovered weak-strong coupling dualities in QFT, i.e. distinct descriptions of the same range of phenomena in two different frameworks.

As we will see, the discovery of quantum anomalies in the late 1960s was a precursor and a hint to all later developments. Quantum anomalies appear if there are two symmetries in the classical action which come into conflict upon quantization, so that only one of two symmetries can be maintained at the quantum level.

Quantum anomalies can be classified as follows:

(a) local versus global;

(b) terminal (i.e. those which kill the theory making it internally inconsistent; also known as internal) versus external (i.e those which destroy a global "external" symmetry without destroying the theory under consideration);

(c) chiral, scale, and gravitational anomalies;

(d) subtle anomalies in supersymmetric theories, for instance, a fermion parity anomaly in two dimensions.

I plan to cover all these topics (in more or less detail) in my lecture course.

Implications of quantum anomalies are numerous. First and foremost, their knowledge is needed in order to avoid theories which look fully "legitimate" at the classical level, but become terminally sick upon quantization. For instance, suppose one would like to build an extension of the Standard Model, with additional fermions beyond the standard three generations. If the fermion content is chosen inappropriately, such an extension may well be internally inconsistent.

Second, the chiral quantum anomalies play an important role in the soft pion theory in QCD, and in the 't Hooft matching condition, which, in turn, presents the foundation for the Seiberg duality in supersymmetric QCD. The scale quantum anomalies which are typical of asymptotically free field theories (such as QCD) can be used for establishing a number of low-energy theorems. I plan to discuss all these implications too, emphasizing more recent developments.

Pedagogical Example: Chiral Anomaly in the Schwinger Model

33 Chiral Anomaly in the Schwinger Model

Our first encounter with the chiral anomalies in gauge theories occurred in Chapter 5. We invoked them, in a pragmatic way, in our deliberations more than once. The current chapter is designed to explain conceptual issues behind the anomalies. The questions to be asked are "why they appear?" and "what they imply?". Here I will address these questions on a more systematic basis.

This topic is important, since the phenomenon of anomalies plays a role in a number of subtle aspects of gauge dynamics. Our first task will be to understand the physical meaning of the phenomenon. This is best done in a simple example [1] of a two-dimensional model which can be treated at weak coupling — the Schwinger model on a spatial circle. It clearly demonstrates that (i) anomalies appear when two contradictory requirements clash, and we have to choose one of them as "sacred" (usually gauge invariance); (ii) anomalies have two faces: infrared and ultraviolet, and (iii) the infinite number of degrees of freedom in field theory is crucial. The chiral anomaly involves fermions. There is another anomaly in gauge theories, the scale anomaly. It takes place even in pure Yang–Mills theory, with no quarks. I will present a number of methods allowing us to derive both. Then we will pass to implications. I will discuss the 't Hooft matching condition, one of a few tools applicable to non-Abelian theories at strong coupling. We will prove that the chiral symmetry of QCD must be spontaneously broken, at least at large N. As an illustrative example of the usefulness of proper understanding of the anomalies we will calculate the $\pi^0 \to \gamma\gamma$ decay rate. Many more applications are known. They would be in order in a good textbook on particle theory. With regret, I have to leave them aside in this general field theory textbook.

33.1 Schwinger model on a circle

Two-dimensional QED with the massless Dirac fermion seems to be the simplest gauge model. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4e_0^2}\, F_{\mu\nu}F^{\mu\nu} + \bar\psi\, i\not{\mathcal{D}}\psi\,, \tag{33.1}$$

where $F_{\mu\nu}$ is the photon field strength tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,, \tag{33.2}$$

e_0 is the gauge coupling constant having the dimension of mass for $D = 2$. Moreover, $\mathcal{D}_\mu$ is the covariant derivative

Defining the covariant derivative in the Schwinger model

$$i\mathcal{D}_\mu = i\,\partial_\mu + A_\mu\,, \tag{33.3}$$

Consult Sections 12.3 and 45.2

and ψ is the two-component spinor field. Gamma matrices in Minkowski space can be chosen in the following way:

$$\gamma^0 = \sigma_2 \,, \qquad \gamma^1 = -i\,\sigma_1 \,, \qquad \gamma^5 = -\sigma_3 \,. \tag{33.4}$$

The spinor $\psi_{\rm L} = \binom{\psi_1}{0}$ will be called left-handed ($\gamma^5\psi_{\rm L} = -\psi_{\rm L}$), while the spinor $\psi_{\rm R} = \binom{0}{\psi_2}$ will be called right-handed ($\gamma^5\psi_{\rm R} = \psi_{\rm R}$), correspondingly. Note also that $\bar\psi = \psi^\dagger\gamma^0$.

In spite of the considerable simplification compared to the four-dimensional QED, the dynamics of the model (33.1) is still too complicated for our purposes. Indeed, the set of the asymptotic states in this model drastically differs from the fields in the Lagrangian. In the two-dimensional theory the photon, as is well-known, has no transverse degrees of freedom and essentially reduces to the Coulomb interaction.† The latter, however, grows linearly with the distance. The linear growth of the Coulomb potential results in confinement of the charged fermions in the Schwinger model irrespectively of the value of the coupling constant e_0. The model (33.1) was even used as a prototype for describing color confinement in QCD (see e.g. [2] and Section 41).

In order to simplify the situation further let us do the following. Consider the system described by the Lagrangian (33.1) on a finite spatial domain of length L. If L is small, $e_0 L \ll 1$, the Coulomb interaction never becomes strong and one can actually treat it as a small perturbation. In particular, in the first approximation its effect can be neglected altogether. We impose periodic boundary conditions on the field A_μ and antiperiodic ones on ψ. Thus, the problem to be considered below is the Schwinger model on the circle. Notice that the antiperiodic boundary conditions are imposed on the fermion field for convenience only. As will be seen, any other boundary condition (periodic, for instance) would do as well; nothing would change except minor technical details. Thus,

Boundary conditions

$$\begin{aligned} A_\mu(t\,,\, x{=}-\tfrac{1}{2}L) &= A_\mu(t\,,\, x{=}\tfrac{1}{2}L)\,, \\ \psi\,(t\,,\, x{=}-\tfrac{1}{2}L) &= -\,\psi\,(t\,,\, x{=}\tfrac{1}{2}L)\,. \end{aligned} \tag{33.5}$$

Equations (33.5) imply that the fields A_μ and ψ can be expanded in the Fourier modes, $\exp\left[ikx\frac{2\pi}{L}\right]$ for the bosons and $\exp\left[\,i(k+\frac{1}{2})\,x\,\frac{2\pi}{L}\right]$ for the fermions ($k = 0, \pm1, \pm2, ...$).

Now, let us recall the fact that the Lagrangian (33.1) is invariant under

† It is instructive to compare this assertion with those in Section 41.

the local gauge transformations

$$\psi \to e^{i\alpha(t,\,x)}\psi\,, \qquad A_\mu \to A_\mu + \partial_\mu\,\alpha(t,\,x)\,. \tag{33.6}$$

It is quite evident that all modes for the field A_1 except the zero mode (i.e. $k=0$) can be gauged away. Indeed, the term of the type $a(t)\sin\left[kx\frac{2\pi}{L}\right]$ in A_1 is gauged away by virtue of the gauge function

$$\alpha(t,\,x) = L\,(2\pi k)^{-1}\,a(t)\,\cos\left[kx\frac{2\pi}{L}\right]\,.$$

The latter is periodic on the circle and does not violate the conditions (33.5), as it should be. Thus, in the most general case we can treat A_1 as an x-independent constant.

This is not the end of the story, however, since the possibilities provided by gauge invariance are not yet exhausted. There exists another class of admissible gauge transformations — sometimes, they are referred to as "large" gauge transformations — with the gauge function which is not periodic in x,

Large gauge transformations

$$\alpha = \frac{2\pi}{L}\,nx\,, \quad n = \pm 1\,, \pm 2\,, \ldots\,, \tag{33.7}$$

where n is an integer. In spite of nonperiodicity, such a choice of the gauge function is also compatible with the conditions (33.5). This fact is readily verifiable: since $\frac{\partial\alpha}{\partial x}=\text{const}$ and $\frac{\partial\alpha}{\partial t}=0$ the periodicity for A_μ is not violated; the analogous assertion is also valid for the phase factor $e^{i\alpha}$ – the difference of phases at the endpoints of the interval $x\in[-\frac{L}{2},\,\frac{L}{2}]$ is equal to $2\pi n$.

As a result, we arrive at the conclusion that the variable A_1 (remember that in the sense of x-dependence A_1 is constant, it depends only on time) should not be considered on the whole interval $(-\infty,\,\infty)$. The points

$$A_1,\; A_1 = \pm\frac{2\pi}{L},\; A_1 = \pm\frac{4\pi}{L}\,, ...$$

are gauge equivalent and must be identified. The variable A_1 is an independent variable only on the interval $[0,\,\frac{2\pi}{L}]$. Going beyond these limits we find ourselves in the gauge image of the original interval. Following the commonly accepted terminology we say that A_1 lives on the circle of circumference $\frac{2\pi}{L}$.

A_1 is an angle-type variable

Everybody knows that the gauge invariance of electrodynamics is closely interrelated with conservation of the electric charge. Indeed, the Lagrangian (33.1) (with finite as well as infinite L) admits multiplication of the fermion field by a constant phase,

$$\psi \to e^{i\alpha}\psi\,, \qquad \psi^\dagger \to \psi^\dagger e^{-i\alpha}.$$

Using the standard line of reasoning one easily derives from this phase invariance the conservation of the electric current

$$j^\mu = \overline{\psi}\,\gamma^\mu\,\psi\,, \quad \dot{Q}(t) = 0\,, \quad Q = \int dx\, j^0(x,\,t)\,.$$

The vanishing of the the divergence $\partial\mu\, j^\mu$ follows from the equations of motion.

The Lagrangian (33.1) exhibits the second conservation law. Observe that the classical Lagrangian (33.1) is invariant under another phase rotation, the global axial transformation

$$\psi \to e^{-i\,\alpha\gamma^5}\psi\,, \qquad \psi^\dagger \to \psi^\dagger e^{i\,\alpha\gamma^5}\,,$$

which multiplies the left- and right-handed fermions by the opposite phases (remember, $\gamma^5 = -\sigma_3$) . At the classical level the axial current

$$j^{\mu 5} = \overline{\psi}\,\gamma^\mu\gamma^5\,\psi$$

Conservation laws for chiral fermions (at the classical level)

is conserved just in the same way as the electromagnetic one. One can readily check using the equations of motion that $\partial_\mu j^{\mu 5} = 0$. If the axial charge of the left-handed fermions is $Q_5 = +1$, for the right-handed fermions $Q_5 = -1$. The conservation of Q and Q_5 is equivalent to the conservation of the number of the left-handed and right-handed fermions separately. This fact is quite obvious for any Born (tree) graph. Indeed, in all such graphs the fermion lines are continuous, the photon emission does not change their chirality, and the number of ingoing fermion legs is equal to that of the outgoing legs. In the exact answer including all quantum effects, however, only the sum of the chiral charges is conserved, only one of two classical symmetries survives quantization of the theory.

As will be seen below, the characteristic excitation frequencies for A_1 are of order of e_0 while those associated with the fermionic degrees of freedom are of order L^{-1}. Since $e_0 L \ll 1$ the variable A_1 is adiabatic with respect to the fermionic degrees of freedom. Consequently, the Born–Oppenheimer approximation is justified in our case. In the next subsection we will analyze in more detail the fermion sector assuming temporarily that A_1 is a fixed (time-independent) quantity. From Eqs. (33.10) and (33.11) it is evident that fermionic frequencies are indeed of order L^{-1}. Calculation of the A_1 frequencies is carried out later, see (33.31).

For our pedagogical purposes we can confine ourselves to the study of the limit $e_0 L \ll 1$. Those readers who would like to know about the solution of the Schwinger model on a circle with arbitrary L should turn to the original publications (e.g. [3]).

33.2 Dirac sea: the vacuum wave function

Following the standard prescription of the adiabatic approximation we freeze the time dependence of the photon field A_μ and consider it as "external." As for the $\mu = 0$ component of the photon field, it is responsible for the Coulomb interaction between the charges; the corresponding effect is of the order $e_0 L \ll 1$ and does not show up in the leading approximation to which we will limit ourselves in the present section. Thus, we can put $A_0 \approx 0$. The difference between these two components lies in the fact that the fluctuations of A_0 are small, while this is not the case for A_1. The wave function is not localized in A_1 in the vicinity of $A_1 = 0$. It is just this phenomenon – delocalization of the A_1 wave function and the possibility of penetration to large values of A_1 – that will lead to observable manifestations of the chiral anomaly.

In two-dimensional electrodynamics the Dirac equation determining the fermion energy levels has the form

$$\left[i \frac{\partial}{\partial t} - \sigma_3 \left(i \frac{\partial}{\partial x} + A_1 \right) \right] \psi = 0 . \tag{33.8}$$

For the k^{th} stationary state $\psi \sim \exp(-iE_k t)\, \psi_k(x)$, and the energy of this state is

$$E_k \, \psi_k(x) = \sigma_3 \left(i \frac{\partial}{\partial x} + A_1 \right) \psi_k(x) . \tag{33.9}$$

Furthermore, the eigenfunctions are proportional to

$$\psi_k \sim \exp\left[i \, (k + \tfrac{1}{2}) \tfrac{2\pi}{L} \, x \right] , \quad k = 0 , \pm 1 , \pm 2 , \ldots . \tag{33.10}$$

The extra term $\frac{1}{2}\frac{2\pi}{L}x$ in the exponent ensures the antiperiodic boundary conditions, see Eqs. (33.5). As a result, we conclude that the energy of the k^{th} level for the left-handed fermions is

$$E_{k(\mathrm{L})} = - \left(k + \frac{1}{2} \right) \frac{2\pi}{L} + A_1 , \tag{33.11a}$$

while for the right-handed fermions

$$E_{k(\mathrm{R})} = \left(k + \frac{1}{2} \right) \frac{2\pi}{L} - A_1 . \tag{33.11b}$$

The energy level dependence on A_1 is displayed in Fig. 33.1. The dashed lines show the behavior of $E_{k(\mathrm{L})}$ and the solid lines are $E_{k(\mathrm{R})}$. At $A_1 = 0$ the energy levels for the left-handed and right-handed fermions are degenerate. If A_1 increases, the degeneracy is lifted and the levels are split. At the point $A_1 = \frac{2\pi}{L}$ the overall structure of the energy levels is precisely the same

Level flow. Rearrangement of levels in gauge equivalent points.

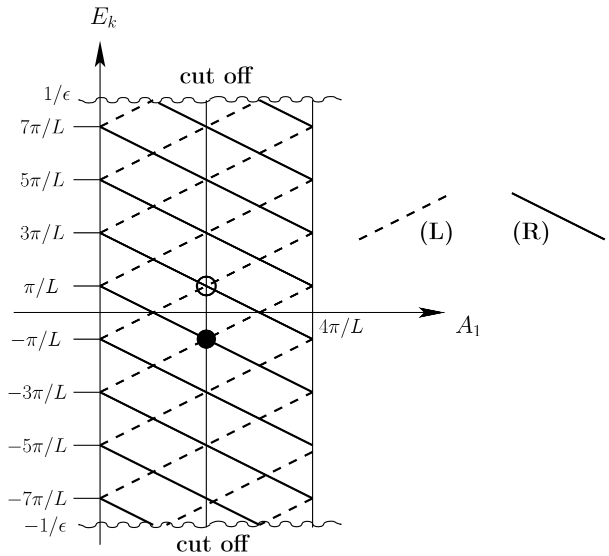

Fig. 33.1. Fermion energy levels as a function of A_1 .

as for $A_1\!=\!0$; the degeneracy takes place again. The identity of the points $A_1\!=\!0$ and $A_1\!=\!\frac{2\pi}{L}$ is the remnant of the gauge invariance of the original theory (see the discussion in Section 33.1).

We note that the identity is achieved in a non-trivial way; in passing from $A_1 = 0$ to $A_1 = \frac{2\pi}{L}$ a restructuring of the fermion levels takes place. All left-handed levels are shifted upwards by one interval while all right-handed levels are shifted downwards by the same one interval. This phenomenon, the restructuring of the fermion levels, is the essence of the chiral anomaly, as will become clear shortly.

Let us proceed from the one-particle Dirac equation to field theory. The first task is the construction of the ground state, the vacuum. To this end, following the well-known Dirac prescription we fill up all levels lying in the Dirac sea, leaving all positive-energy levels empty. The following notations will be used below for filled and empty levels with a given k: $|1_{\rm L,R}\,,k\rangle$ and $|\,0_{\rm L,R}\,,k\rangle$, respectively. The subscript L(R) indicates that we deal with the left-handed (right-handed) fermions.

Recall that A_1 is a slowly varying adiabatic variable; the corresponding quantum mechanics will be considered later. At first, the value of A_1 is fixed in the vicinity of zero, $A_1 \approx 0$. Then, the fermion wave function of the vacuum, as seen from Fig. 33.1 , reduces to

$$\Psi_{\text{ferm. vac.}} = \left(\prod_{k=0,1,2,\ldots} |1_{\rm L}\,,\,k\rangle\right)\left(\prod_{k=-1,-2,\ldots} |\,0_{\rm L}\,,\,k\rangle\right) \times \left(\prod_{k=-1,-2,\ldots} |1_{\rm R}\,,\,k\rangle\right)\left(\prod_{k=0,1,2,\ldots} |\,0_{\rm R}\,,\,k\rangle\right). \tag{33.12}$$

The Dirac sea, or all negative-energy levels, are completely filled. Now let A_1 increase adiabatically from 0 to $\frac{2\pi}{L}$. The same figure shows that at $A_1 = \frac{2\pi}{L}$ the wave function (33.12) describes a state which, from the standpoint of the normally filled Dirac sea, contains one left-handed particle and one right-handed hole (small circles on Fig. 33.1) .

Do the *quantum numbers* of the fermion sea change in the process of the transition from $A_1 = 0$ to $A_1 = \frac{2\pi}{L}$? Answering this question, we would say that the appearance of the particle and the hole does not change the electric charge since the electric charges of the particle and the hole are obviously opposite. In other words, the electromagnetic current is conserved. On the other hand, the axial charges of the left-handed particle and the right-handed hole are the same ($Q_5 = -1$) and, hence, in the transition at hand

$$\Delta Q_5 = -2\,. \tag{33.13}$$

A more formal analysis, to be carried out shortly, will confirm this assertion.

Equation (33.13) can be rewritten as $\Delta Q_5 = -\frac{L}{\pi}\Delta A_1$. Dividing by Δt , the transition time, we get

$$\dot{Q}_5 = -\frac{L}{\pi}\dot{A}_1\,, \tag{33.14}$$

which implies, in turn, that the conserved quantity has the form

$$\int dx\left(j^{\,05} + \frac{1}{\pi}\,A_1\right). \tag{33.15}$$

Anomaly in the axial current derived from the level flow

The current corresponding to the charge (33.15) is obviously

$$\widetilde{j}^{\mu 5} = j^{\mu 5} + \frac{1}{\pi}\,\varepsilon^{\mu\nu}A_\nu\,, \quad \partial_\mu\,\widetilde{j}^{\mu 5} = 0\,, \quad \partial_\mu\,j^{\mu 5} = -\frac{1}{2\pi}\,\varepsilon^{\mu\nu}F_{\mu\nu}\,, \tag{33.16}$$

where $\varepsilon^{\mu\nu}$ is the asymmetric tensor, $\varepsilon^{01} = -\,\varepsilon^{10} = 1$. (Notice that $\varepsilon_{01} = -1$.) The last equality in (33.16) represents the famous axial anomaly in the Schwinger model. We succeeded in deriving it by the "hand-waving" arguments, by inspecting the picture of motion of the fermion levels in the external field $A_1(t)$. It turns out that in this language the chiral anomaly presents an extremely simple and widely known phenomenon: the crossing of the zero point in the energy scale by this or that level (or by a group of levels). The presence of the infinite number of levels and the Dirac "multiparticle" interpretation, according to which the emergence of the filled level from the sea means the appearance of the particle while the submergence of the empty level into the sea is equivalent to production of a hole — an antiparticle — are the most essential elements of the whole construction. With the finite number of levels, when there is no place for such an interpretation, there can be no quantum anomaly.

I would like to draw the reader attention to a somewhat different (although intimately related with the previous) aspect of the picture. The fermion levels move parallel to each other through the bulk of the Dirac sea. Therefore, the disappearance of the levels beyond the zero-energy mark occurs simultaneously with the disappearance of the "copies" beyond the ultraviolet cut-off, which is always implicitly present in the field theory (below we will introduce it explicitly). Because of this fact the heuristic derivation of the anomaly given in this section and a more standard treatment based on the ultraviolet regularization are actually one and the same. Often it turns out more convenient to trace just the crossing of the ultraviolet cut-off by the levels from the Dirac sea. Beyond toy models, in QCD-like theories, the latter approach becomes an absolute necessity, not a question of convenience, due to the notorious "infrared slavery." The connection between the ultraviolet and infrared interpretations of the anomaly is discussed in more detail in Sections 33.3 and 33.7. The interested reader is referred to the original work [4] where all subtle points are exhaustively analyzed.

33.3 Ultraviolet regularization

In spite of the transparent character of this heuristic derivation almost all of the "evident" points above can be questioned by the careful reader. Indeed, why is the wave function (33.12) the appropriate choice? In what sense is the energy of this state minimal, taking into account the fact that according to (33.11),

$$E \sim -\sum_{k=0}^{\infty}(k+\tfrac{1}{2})\,\frac{2\pi}{L}\ ,$$

and the sum is ill-defined (the series is divergent)? Moreover, it is usually asserted that the quantum anomalies are due to the necessity of the ultraviolet regularization of the theory. If so, why speak of the Dirac sea and the crossing of the zero-energy point by the fermion levels?

Surprisingly, all these questions are connected with each other. Probably, it will be most instructive to start with the last one. Now I will explain that although the ultraviolet regularization was not even mentioned thus far, actually, it is the key element. More than that, the derivation sketched above tacitly assumes quite a specific regularization.

The fermion levels stretch in the energy scale up to indefinitely large energies, positive or negative. The wave function (33.12) describing the fermion sector at $A_1 \approx 0$ contains, in particular, the direct product of an infinitely large number of the filled states $|\,1_{\rm R}\,,\,k\,\rangle\,,\ |\,1_{\rm L}\,,\,k\,\rangle$ with the negative energy. It is clear that such an object – the infinite product – is ill-defined, and one cannot do without some regularization in calculating physical quantities. The contribution corresponding to large energies (momenta) should be somehow cut off.

At first sight, it seems it would be sufficient simply to throw away the terms with $|k|>|k|_{\rm max}$ ($|k|_{\rm max}$ is a fixed number independent of A_1). This *is* a regularization, of course, but, clearly enough, the prescription will inevitably lead to a violation of the gauge invariance and the electric charge nonconservation. Indeed, in the gauge theories the momentum p always appears only in the combination $p+A$, not p (or, what is the same, k).

Making the cut-off in a gauge invariant manner

In order to preserve gauge invariance, it is possible and convenient to use the regularization called in the literature the Schwinger, or ϵ, splitting. This regularization will provide a more solid mathematical basis to the heuristic derivation presented above. Instead of the original currents

$$j^\mu = \bar\psi(t,x)\,\gamma^\mu\,\psi(t,x)\,, \qquad j^{\mu 5} = \bar\psi(t,x)\,\gamma^\mu\,\gamma^5\,\psi(t,x)\,, \tag{33.17}$$

we introduce the regularized objects

$$\begin{aligned} j^\mu_{\rm Reg} &= \bar\psi(t,x+\epsilon)\,\gamma^\mu\,\psi(t,x)\,\exp\left(i\int\limits_x^{x+\epsilon} A_1\,dx\right), \\ j^{\mu 5}_{\rm Reg} &= \bar\psi(t,x+\epsilon)\,\gamma^\mu\,\gamma^5\,\psi(t,x)\,\exp\left(i\int\limits_x^{x+\epsilon} A_1\,dx\right). \end{aligned} \tag{33.18}$$

It is implied that $\epsilon\to 0$ in the final answer for the physical quantities. At the intermediate stages, however, all computations are performed with fixed ϵ. The exponential factor in (33.18) ensures gauge invariance of the "split" currents. Without this factor, multiplying $\psi(t,x)$ by an x-dependent phase,

$\psi(t,x) \rightarrow \exp[\,i\alpha(x)]\,\psi(t,x)$, yields

$$\begin{aligned}\psi^\dagger_\alpha(t,x+\epsilon)\,\psi_\beta(t,x) \quad &\rightarrow \quad \exp[-i\alpha(x+\epsilon)+i\alpha(x)] \\ &\times \quad \psi^\dagger_\alpha(t,x+\epsilon)\,\psi_\beta(t,x)\,. \end{aligned} \tag{33.19}$$

The gauge transformation (33.6) of A_1 compensates for the phase factor in Eq. (33.19) .

Now, there is no difficulty in calculating the electric and axial charges of the state (33.12) "scientifically." If

$$Q = \int dx\, j^0_{\rm Reg}(t,x)\,, \qquad Q_5 = \int dx\, j^{05}_{\rm Reg}(t,x)\,, \tag{33.20}$$

then for the vacuum wave function we, evidently, get

$$Q = Q_{\rm L} + Q_{\rm R}\,, \qquad Q_5 = -Q_{\rm L} + Q_{\rm R}, \tag{33.21}$$

$$\begin{aligned} Q_{\rm L} &= \sum_k \exp\left\{-i\epsilon\left[\left(k+\tfrac{1}{2}\right)\tfrac{2\pi}{L} - A_1\right]\right\}, \\ Q_{\rm R} &= \sum_{k'} \exp\left\{-i\epsilon\left[\left(k'+\tfrac{1}{2}\right)\tfrac{2\pi}{L} - A_1\right]\right\}, \end{aligned} \tag{33.22}$$

where k and k' run over all filled levels. In the limit $\epsilon \rightarrow 0$ both charges, $Q_{\rm L}$ and $Q_{\rm R}$, turn into the sum of unities, each unity representing one energy level from the Dirac sea. Equations (33.22) once again demonstrate the gauge invariance of the Schwinger regularization. Indeed, the cut-off suppresses the states with $|\,p + A_1| \gtrsim \epsilon^{-1}$.

If it were not for the phase factor in Eqs. (33.18), the suppressing function would not contain the desired combination, $p + A$.

We hasten to add here that although superficially Eqs. (33.22) do not differ from each other, actually they do not coincide because the summation runs over different values of k. What the particular values are is easy to establish from Fig. 33.1.† Let $|A_1| < \frac{\pi}{L}$. Then in the "left-handed" sea the filled levels have $k = 0,\,1\,,\,2\,,\,\ldots$. In the "right-handed" sea the filled levels correspond to $k = -1\,,\,-2\,,\,\ldots$. Thus, if $|A_1| < \frac{\pi}{L}$ we have

$$\begin{aligned} Q_{\rm L} &= \sum_{k=0}^{\infty} \exp\left\{i\,\epsilon\,E_{k({\rm L})}\right\}, \\ Q_{\rm R} &= \sum_{k=-1}^{-\infty} \exp\left\{-i\,\epsilon\,E_{k({\rm R})}\right\}. \end{aligned} \tag{33.23}$$

† See also Eq. (33.12) .

Performing the summation and expanding in ϵ we arrive at

$$(Q_{\rm L})_{\rm vac} = -(Q_{\rm R})_{\rm vac} = \frac{e^{i\epsilon A_1}}{2i\,\sin(\epsilon\pi/L)}$$

$$= \frac{L}{2\pi i\epsilon} + \frac{L}{2\pi}A_1 + {\rm O}(\epsilon)\,, \qquad (33.24)$$

We pause here to summarize our results. Equations (33.24) show that under our choice of the vacuum wave function (33.12) the charge of the vacuum vanishes, $Q = Q_{\rm L} + Q_{\rm R} = 0$. Moreover, there is no time dependence, the charge is conserved. The axial charge consists of two terms: the first term represents an infinitely large *constant* and the second one gives a linear A_1 dependence. In the transition $(A_1 \approx 0) \to (A_1 \approx \frac{2\pi}{L})$ the axial charge changes by (minus) two units (cf. Eq. (33.21)).

These conclusions are not new for us. We have found just the same from the illustrative picture described in Section 33.2 in which the electric and axial charges of the Dirac sea are determined intuitively. Now we learned how to sum up the infinitive series $\sum_k 1$, the charges of the "left-handed" and "right-handed" seas, by virtue of the well-defined procedure which automatically cuts off the levels with $|\,p + A_1| \gtrsim \epsilon^{-1}$.

The procedure suggests an alternative language for describing the axial charge nonconservation in the transition $(A_1 \approx 0) \to (A_1 \approx \frac{2\pi}{L})$. Previously we thought that the nonconservation is due to the level crossing of the zero-energy point. It is equally correct to say – as we see now – that the nonconservation is explained by the following: one right-handed level from the sea leaves the "fiducial domain" via the lower boundary (the cut-off $-\epsilon^{-1}$) and one new left-handed level appears in the sea through the same boundary (Fig. 33.1). Both phenomena – the crossing of the zero-energy point and the departure (arrival) of the levels via the ultraviolet cut-off – occur simultaneously though, and represent, actually, two different facets of one and the same anomaly, which admits both, the infrared and ultraviolet interpretations.

Gauge invariance should be maintained by all means!

One last remark concerning the axial charge is in order. Instead of Eqs. (33.18) one could regularize the axial charge in a different way, so that $\partial_\mu j^{\mu 5} = 0$ and $\Delta Q_5 = 0$. (A nice exercise for the reader!) Under such a regularization, however, the expression for the axial current would not be gauge invariant. Specifically, the conserved axial current, apart from Eqs. (33.18), would include an extra term $\frac{1}{\pi}\varepsilon^{\mu\nu}A_\nu$, cf. Eqs. (33.16). As has already been mentioned, there is no regularization ensuring simultaneously gauge invariance and conservation of $j^{\mu 5}$.

33.4 The theta vacuum

Compare with Section 18.2

Now, we leave the issue of charges and proceed to the calculation of the fermion-sea energy, the problem which could not be solved at the naive level, without regularization. Fortunately, all necessary elements are already prepared.

The fermion part of the Hamiltonian, cf. Eqs. (33.9) ,

$$H = \int_{-L/2}^{L/2} dx\, \psi^\dagger(t,x)\, \sigma_3 \left(i\, \frac{\partial}{\partial x} + A_1 \right) \psi(t,x)\,, \tag{33.25}$$

reduces after the ϵ splitting to

$$H_{\rm Reg} = \int_{-L/2}^{L/2} dx\, \psi^\dagger(t,x+\epsilon)\, \sigma_3 \left(i\, \frac{\partial}{\partial x} + A_1 \right) \psi(t,x) \exp\left(i \int_x^{x+\epsilon} A_1\, dx \right). \tag{33.26}$$

This formula implies, in turn, the following regularized expression for the energies of the "left-handed" and "right-handed" seas:

$$E_{\rm L} = \sum_{k=0}^{\infty} E_{k({\rm L})} \exp\{ i\,\epsilon\, E_{k({\rm L})} \}\,, \quad E_{\rm R} = \sum_{k=-1}^{-\infty} E_{k({\rm R})} \exp\{ -i\,\epsilon\, E_{k({\rm R})} \}\,, \tag{33.27}$$

where the energies of the individual levels $E_{k({\rm L,\,R})}$ are given in (33.11) and summation runs over all levels with the negative energy. The concrete values of the summation indices in Eqs. (33.27) correspond to $|A_1| < \frac{\pi}{L}$. Expressions (33.27) have an absolutely obvious meaning: in the limit $\epsilon \to 0$ they simply reduce to the sum of the energies of all filled fermion levels from the Dirac sea. The additional exponential factors guarantee the convergence of the sums.

Dirac sea energy

Furthermore, we notice that $E_{\rm L}$ and $E_{\rm R}$ can be obtained by differentiating the expressions (33.23)and (33.24) for $Q_{\rm L,\,R}$ with respect to ϵ. (Equation (33.23) presents a geometrical progressions and is trivially summable.) Expanding in ϵ we get

$$\begin{aligned} E_{\rm sea} = E_{\rm L} + E_{\rm R} = \tfrac{L}{2\pi} \left({A_1}^2 - \tfrac{\pi^2}{L^2} \right) \\ + \text{(a constant independent of } A_1)\,. \end{aligned} \tag{33.28}$$

In the expression above we omit the infinite A_1 independent constant term on the second line and choose the constant term in the braces on the first line in such a way that the sea energy vanishes at the points $A_1 = \pm\frac{\pi}{L}$ (see Fig. 33.2).

Two remarks are in order here. First, it is instructive to check that the Born–Oppenheimer approximation, accepted from the very beginning, is indeed justified. In other words, let us verify that the dynamics of the variable A_1 is slow in the scale characteristic of the fermion sector. The effective Lagrangian determining the quantum mechanics of A_1 is

I promised to do this check in Section 33.1

$$\mathcal{L} = \frac{L}{2\,e_0^2}\,\dot{A}_1^2 - \frac{L}{2\pi}\,A_1^2\,. \tag{33.29}$$

This is the ordinary harmonic oscillator with the ground state wave function

$$\Psi_0(A_1) = \left(\frac{L}{e_0\pi^{3/2}}\right)^{1/4} \exp\left(-\frac{L\,A_1^2}{2e_0\sqrt{\pi}}\right)\,, \tag{33.30}$$

and the level splitting

$$\omega_A = \frac{e_0}{\sqrt{\pi}}\,. \tag{33.31}$$

The characteristic frequencies in the fermion sector are $\omega_{\,\rm ferm} \sim L^{-1}$. Hence,

$$\omega_A/\omega_{\,\rm ferm} \sim e_0 L \ll 1\,. \tag{33.32}$$

The second remark concerns the structure of the total vacuum wave function. We have convinced ourselves that

$$\Psi_{\rm vac} = \Psi_{\rm ferm.\,vac.}\,\Psi_0(A_1) \tag{33.33}$$

is the eigenstate of the Hamiltonian of the Schwinger model on the circle in the Born–Oppenheimer approximation. The wave function (33.33) is quite satisfactory from the point of view of the "small" gauge transformations, i.e. those continuously deformable to the trivial (unit) transformation. (More exactly, Eq. (33.33) refers to the specific gauge in which the gauge degrees of freedom associated with A_1 are eliminated and A_1 is independent of x.) This wave function, however, is not invariant under the "large" gauge transformations $A_1 \to A_1 + \frac{2\pi}{L}k$, where $k = \pm1\,, \pm2\,, \ldots$.

The essence of the situation becomes clear if we return to Fig. 33.1. When A_1 performs small and slow oscillations in the vicinity of zero, the Dirac sea is filled in such a way as shown in Eq. (33.12) . But A_1 can oscillate as well in the vicinity of the gauge equivalent point $A_1 = \frac{2\pi}{L}$. In this case if we do *not* restructure the fermion sector and leave it just as in Eq. (33.12) , then the configuration of Eq. (33.12) is obviously *not* the vacuum – it corresponds to one particle plus one hole. This assertion is confirmed, in particular, by the plot showing the Dirac sea energy as a function of A_1 (Fig. 33.2). In order to get the configuration with the lowest energy in the vicinity of $A_1 = \frac{2\pi}{L}$ it is necessary to fill the fermion levels as follows:

$$\prod_{k=1,2,3,\ldots} |1_{\rm L}\,,\,k\rangle \prod_{k=0,-1,-2,\ldots} |1_{\rm R}\,,\,k\rangle$$

(the empty levels are not shown explicitly, cf. Eq. (33.12)).

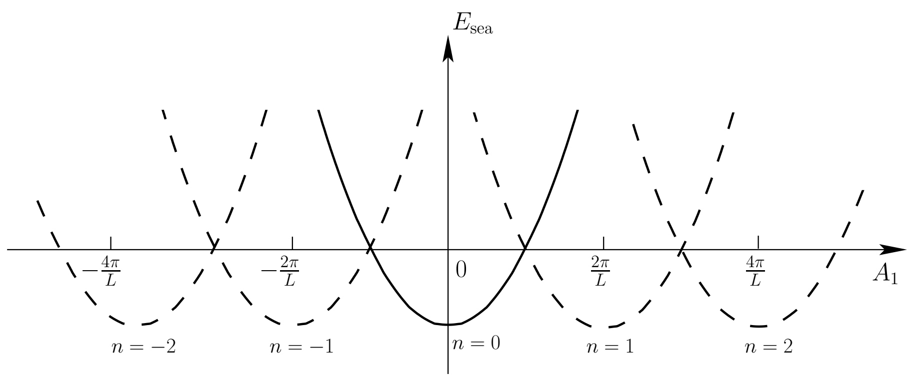

Fig. 33.2. Energy of the Dirac sea in the Schwinger model on a circle. The solid line corresponds to Eq. (33.12). Dashed lines reflect restructuring of the Dirac sea necessary if $|A_1| > \frac{\pi}{L}$.

n-th prevacuum

Thus, the Hilbert space is naturally split into distinct sectors corresponding to different structure of the fermion sea. The wave function of the ground state in the $n^{\rm th}$ sector has the form

$$\Psi_n = \left(\prod_{k=n}^{\infty} |1_{\rm L}\,,\,k\rangle\right) \left(\prod_{k=n-1}^{-\infty} |1_{\rm R}\,,\,k\rangle\right)\, \Psi_0\left(A_1 - \frac{2\pi}{L}\,n\right)\,,$$

$$n = 0\,,\ \pm 1\,,\ \pm 2\,,\ \ldots\,. \tag{33.34}$$

The organization of the fermion sea is correlated with the position of the "center of oscillation" of A_1. It is quite evident that if $n \neq n'$ then Ψ_n and $\Psi_{n'}$ are strictly orthogonal to each other due to the fermion factors.

Is it possible to construct the vacuum wave function invariant under the "large" gauge transformations $A_1 \to A_1 + \frac{2\pi}{L}k$ (with the simultaneous renumbering of the fermion levels)? The answer is yes. Moreover, such a wave function is not unique. It depends on a new hidden parameter θ which is often called the vacuum angle in the literature. Consider the linear combination

$$\Psi_{\theta\,\rm vac} = \sum_n e^{in\,\theta}\Psi_n. \tag{33.35}$$

This linear combination is also the eigenfunction of the Hamiltonian with the lowest energy, just in the same way as Ψ_n. But unlike Ψ_n the "large" gauge transformations leaves $\Psi_{\theta\,\mathrm{vac}}$ essentially intact. More exactly, under $A_1 \to A_1 + \frac{2\pi}{L}$ the wave function (33.35) is multiplied by $e^{i\theta}$. This overall phase of the wave function is unobservable; all physical quantities resulting from averaging over the θ-vacuum are invariant under the gauge transformations.

Summarizing, we have now become acquainted with another model in which the notion of the vacuum angle θ, as well as the θ-vacuum, is absolutely transparent: the Schwinger model on the spatial circle. The presence of the vacuum angle θ in the wave function is imitated in the Lagrangian language by adding the so-called topological density to the Lagrangian. In the Schwinger model the topological density is

Previously we discussed the θ-vacuum in Chapter 5

$$\Delta\mathcal{L}_\theta = \frac{\theta}{4\pi}\varepsilon^{\mu\nu}F_{\mu\nu}\,. \tag{33.36}$$

This extra term in the action is the integral over the full derivative: it does not affect the equations of motion and gives a vanishing contribution for any topologically trivial configuration $A_\mu(t,x)$. The topological density $\Delta\mathcal{L}_\theta$ shows up only if

$$\int_{-L/2}^{L/2} dx\,\Big[A_1(t=+\infty\,,\,x) - A_1(t{=}-\infty\,,\,x)\Big] = 2\pi k\,,\quad |k| = 1\,,\,2\,,\,\ldots\,. \tag{33.37}$$

33.5 Topological aspect

The topological properties are mentioned here not by chance. It is very instructive to discuss the topological aspect of the theoretical construction under consideration in more detail, in parallel with a similar discussion in Chapter 5, where we exploited the path integral formulation of Yang–Mills theory based on the Lagrangian formalism. At the same time, in the Schwinger model, so far, (Section 33.4) we used the Hamiltonian language to establish the existence of the θ vacuum.

The Schwinger model possesses the U(1) gauge invariance. An element of the U(1) group, as it is well-known, can be written as $e^{i\alpha}$. Using the gauge freedom one can reduce the fields $A_1(t,x)$ or $\psi(t,x)$ at a given moment of time to a standard form by choosing an appropriate gauge function $\alpha(t,x)$. The standard form of A_1 is $A_1 =$ const., which varies between, say, zero and $2\pi/L$. Gauge-equivalent points $A_1 = 0$, $\pm 2\pi/L$, $\pm 4\pi/L$, , ... are connected by "large" (topologically nontrivial) gauge transformations.

Moreover, under our boundary conditions the variable x represents a circle of length L, and, consequently, we deal here with the (continuous) mappings of the circle in the configuration space into the gauge group U(1). The set of the mappings can be divided in classes. The mathematical formula expressing the fact that the mappings are decomposed into classes is

Same topology as in the case of ANO strings

$$\pi_1\left(\mathrm{U}(1)\right) = \mathbb{Z}\,. \tag{33.38}$$

The meaning of Eq. (33.38) is very simple. Inside each class, all mappings, by definition, can be reduced to each other by continuous deformations. On the other hand, no continuous deformations transform mappings from one class into those belonging to another class.

When the mappings of a circle onto U(1) are considered, the difference between the classes is especially transparent (see Fig. 33.3). Assume that we started from a certain point, went around the circle **a** once, and returned

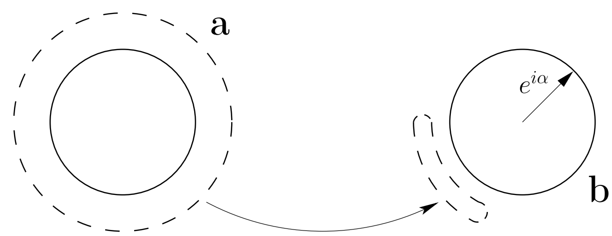

Fig. 33.3. Mapping of a circle in the coordinate space into U(1). The dashed contour near the circle **b** shows a topologically trivial mapping.

to the starting point. In doing so, we simultaneously went around the circle **b** $0\,,\ \pm 1\,,\ \pm 2\,,$ etc. times. (The negative sign corresponds to circulation in the opposite direction.) The number of windings around the circle **b** is just a class of the mapping. It is quite clear that all mappings with a given winding number are continuously deformable into each other. On the contrary, different winding numbers guarantee that a continuous deformation is impossible. The letter $\mathbb{Z}$ in Eq. (33.38) denotes the set of integers and shows that the set of different mapping classes is isomorphic to the set of integers; each class is characterized by an integer having the meaning of the winding number. The mappings corresponding to the zero winding number are called topologically trivial, the others are topologically non-trivial.

This information is sufficient to establish the existence of the vacuum sectors labeled by n ($n = 0\,,\ \pm 1\,,\ \pm 2\,,\ \dots$), for which $(A_\mu)_{\rm vac} \sim \partial_\mu \alpha_{(n)}$, without any explicit construction like (33.34) ($\alpha_{(n)}$ belongs to the $n^{\rm th}$ class). The necessity of introducing the vacuum angle θ also stems from this information.

33.6 The necessity of the θ vacuum

Finally, the last issue to be discussed in connection with the Schwinger model is as follows. Sometimes the question is raised as to why the vacuum wave function cannot be chosen in the form (33.34) with fixed n. The gauge invariance under "small" (topologically trivial) transformations is preserved which automatically implies the electric charge conservation. What is lost is only the invariance under the "large" (topologically non-trivial) transformations; it seems that there is nothing bad in that.† Then, why is it necessary to pass to $\Psi_{\theta\,\mathrm{vac}} = \sum_n e^{in\theta}\Psi_n$?

Cluster decomposition and stability with regards to e.g. mass deformations

The point is that Ψ_n taken as the vacuum wave function violates clusterization – one of the basic properties in the field theory which can be traced back to causality and unitarity of the theory. The following is understood by clusterization: the vacuum expectation value of the product of several local operators at causally independent points must be reducible to the product of vacuum expectation values for each operator, for example,

$$\langle O_1 O_2 \rangle = \langle O_1 \rangle \langle O_2 \rangle \,. \tag{33.39}$$

The violation of the clusterization can be demonstrated explicitly. Consider the two-point function

$$\mathcal{A}(t) = \langle \Psi_n | \, T\left\{ O^\dagger(t)\,, O(0) \right\} | \, \Psi_n \rangle \,,$$
$$O(t) = \int \bar{\psi}(t,x)(1+\gamma^5)\psi(t,x)\, dx \,. \tag{33.40}$$

The operator O changes the axial charge of the state by two units (adds a particle and a hole to the Dirac sea), $O^\dagger$ returns it back, and, as a result, $\mathcal{A}(t) \neq 0$. Moreover, if $t \to \infty$ in the Euclidean domain $\mathcal{A}(t) \to \text{const}$. (For a concrete calculation, see, e.g., [2] based on the bosonization method. In this work the limit $L \to \infty$ is considered but all relevant expressions can be readily rewritten for finite L.) The fact that $\mathcal{A}(t)$ tends to a non-vanishing constant at $t \to \infty$ means, according to clusterization, that the operators $\bar{\psi}(1 \pm \gamma^5)\psi$ acquire a nonvanishing vacuum expectation value.

On the other hand, if $|\mathrm{vac}\rangle = |\Psi_n\rangle$ then $\langle \bar{\psi}(1 \pm \gamma^5)\psi \rangle = 0$ for a trivial reason. Indeed, the operator $\bar{\psi}(1 \pm \gamma^5)\psi$ acting on Ψ_n produces an electron and a hole, and the corresponding state is obviously orthogonal to Ψ_n itself.

The clusterization property restores itself if one passes to the θ-vacuum

† The contents of this section should be compared to Section 18.2. For a discussion of subtle and contrived modifications which are possible, but will not concern us here, see [5, 6].

(33.35) . In this case there emerges a nondiagonal expectation value

$$\langle\Psi_{n+1}|\,\bar\psi\,(1\pm\gamma^5)\,\psi\,|\Psi_n\rangle\sim L^{-1}\exp\Big(-\frac{\pi^{3/2}}{e_0 L}\Big)\,. \qquad (33.41)$$

If the line of reasoning based on clusterization seems too academic to the reader, it might be instructive to consider another argument (connected with Eqs. (33.40) and the subsequent discussion). Let us ask the question: what will happen if instead of the massless Schwinger model we consider the model with a small mass, i.e. introduce an extra mass term $\Delta\mathcal{L}_m=-m\,\bar\psi\,\psi$ in the Lagrangian (33.1) ? Naturally, all physical quantities obtained in the massless model will be shifted. It is equally natural to require, however, the shifts to be small for small m, so that there would be no change in the limit $m\to 0$. Otherwise, we would encounter an unstable situation while we would like to have the mass term as a small perturbation.

But in the presence of the degenerate states (and the states Ψ_n with different n are degenerate) any perturbation is potentially dangerous and can lead to large effects. Just such a disaster occurs, in particular, if $\Delta\mathcal{L}_m$, acting on the vacuum, is nondiagonal.

If we prescribe the states like Ψ_n to be the vacuum, then $\Delta\mathcal{L}_m$ will by no means be diagonal, as follows from the discussion after Eqs. (33.40) . This we cannot accept. On the other hand, the mass term is certainly diagonalized in the basis of the wave functions (33.35) ,

$$\langle\Psi_{\theta'\,\rm vac}|\Delta\mathcal{L}_m|\Psi_{\theta\,\rm vac}\rangle=0 \qquad \text{if} \quad \theta'\neq\theta\,. \qquad (33.42)$$

33.7 Two faces of the anomaly

In conclusion, it will be extremely useful to discuss the connection between the picture presented above and the more standard derivation of the chiral anomaly in the Schwinger model. This discussion will represent a bridge between the physical picture described above and the standard approach to anomalies.

We have already emphasized the double nature of the anomaly which shows up as the infrared effect in the current and the ultraviolet effect in the divergence of the current. The line of reasoning accepted thus far put more emphasis on the infrared aspect of the problem – the finite "box" served as a natural infrared regularization. The same result for $\partial_\mu j^{\mu 5}$ as in Eqs. (33.16) could be obtained with no reference to the infrared regularization.

A conventional treatment of the issue is based on the standard Feynman diagram technique. The usual explanation one can find in numerous textbooks connects the anomalies to the ultraviolet divergence of certain

Feynman graphs. The assertion of the ultraviolet divergence is valid if one deals directly with $\partial_\mu j^{\mu 5}$. Thus, the emphasis is shifted to the ultraviolet aspect of the anomaly.

Below, first of all, we will sketch the standard derivation. Then we will show that the diagrammatic language used, as a rule, for the analysis of $\partial_\mu j^{\mu 5}$ from the point of view of the ultraviolet regularization, can be successfully used for the "infrared" derivation of the anomaly. The fact that the anomalies reveal themselves in the infrared behavior of Feynman graphs is rarely mentioned in the literature, and, hence, deserves a more detailed discussion. The pragmatically oriented reader *can omit this subsection in the first reading.*

Thus, we would like to demonstrate that

$$\partial_\mu j^{\mu 5} = -\frac{1}{2\pi}\,\varepsilon^{\mu\nu} F_{\mu\nu}\,, \tag{33.43}$$

by considering directly $\partial_\mu j^{\mu 5}$, not $j^{\mu 5}$ as previously. Then we need to bother only about the ultraviolet regularization, and, in particular, the theory can be considered in the infinite space since the finiteness of L does not affect the result coming from the short distances.

One of the convenient methods of the ultraviolet regularization is due to Pauli and Villars. In the model at hand it reduces to the following. In addition to the original massless fermions in the Lagrangian, heavy regulator fermions are introduced with the mass M_0 ($M_0 \to \infty$) and the opposite metric. The latter means that each loop of the regulator fermions is supplied by an extra minus sign relatively to the normal fermion loop. The interaction of the regulator fermions with the photons is assumed to be just the same as for the original fermions, and the only difference is the mass. Then the role of the Pauli–Villars fermions in the low-energy processes ($E \ll M_0$) is to provide the ultraviolet cut-off in the formally divergent integrals with the fermion loops. Such a regularization procedure, clearly, automatically guarantees gauge invariance and the electromagnetic current conservation.

In the model regularized according to Pauli and Villars the axial current has the form

$$j^{\mu 5} = \bar\psi\gamma^\mu\gamma^5\psi + \bar R\gamma^\mu\gamma^5 R\,, \tag{33.44}$$

where R is the fermion regulator. In calculating the divergence of the regularized current the naive equations of motion can be used. Then

$$\partial_\mu j^{\mu 5} = 2iM_0\bar R\gamma^5 R\,.$$

The divergence does not vanish (the axial current is not conserved!), but, as was expected, $\partial_\mu j^{\mu 5}$ contains only the regulator anomalous term.

The last step is contracting the regulator fields in the loop in order to convert $M_0\bar{R}\gamma^5 R$ in the "normal" light fields in the limit $M_0 \to \infty$. The relevant diagrams are displayed in Fig. 33.4 where the solid line denotes the standard heavy fermion propagator $i(\not{p} - M_0)^{-1}$. The graph **(a)** does

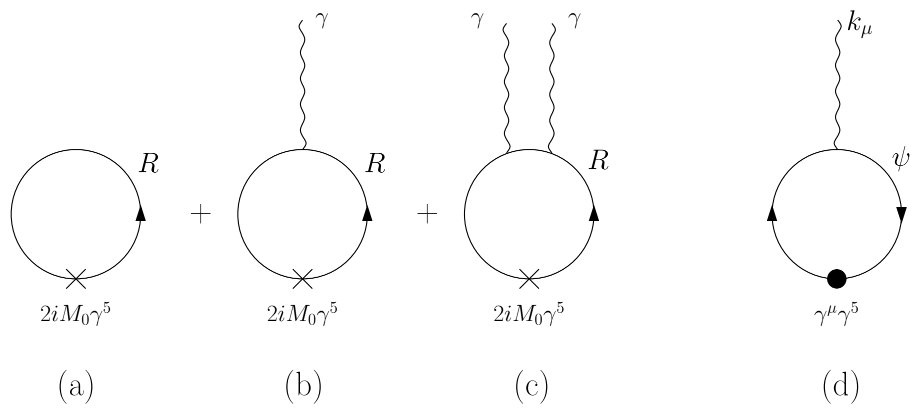

Fig. 33.4. Diagrammatic representation of the anomaly in the axial current in the Schwinger model. (a), (b), (c): Heavy regulator fields in the divergence of the current. (d): Infrared anomalous contribution in $\bar{\psi}\gamma^\mu\gamma^5\psi$.

not depend on the external field. The corresponding contribution to $\partial_\mu j^{\mu 5}$ represents a number which can be set equal to zero. The graph **(c)**, with two photon legs, and all others having more legs die off in the limit $M_0 \to \infty$. The only surviving is the graph **(b)**. Calculation of this diagram is trivial,

$$2iM_0\bar{R}\gamma^5 R \to -\frac{1}{2\pi}\,\varepsilon^{\mu\nu}F_{\mu\nu}\,. \tag{33.45}$$

(Do not forget an extra minus sign in Pauli–Villars fermion loop.) As a result, we reproduce the anomalous relation (33.43) obtained previously by a different method.

The easiest method allowing one to check Eq. (33.45) in another way is, probably, the so-called background field technique. I will not dwell here on its detail because the corresponding explanation would lead us far astray. The interested reader is referred to the review [7] where all relevant nuances are fully discussed. I will limit myself to intuitively obvious features and use self-evident notation. Then

Background field formula

$$2iM_0\bar{R}\gamma^5 R = -2M_0\,\mathrm{Tr}\left[\gamma^5\left(\not{P} - M_0\right)^{-1}\right], \tag{33.46}$$

where $P_\mu = i\mathcal{D}_\mu = i\partial_\mu + A_\mu$ is the generalized momentum operator, and we have taken into account the fact that the minus sign in the fermion loop does not appear for the regulator fields.

Moreover,

$$(\not{P} - M_0)^{-1} = (\not{P} + M_0)\left(P^2 + \frac{1}{2}\, i\,\varepsilon^{\mu\nu} F_{\mu\nu}\gamma^5 - M_0^2\right)^{-1}. \qquad (33.47)$$

Now, since $M_0 \to \infty$ the trace in Eq. (33.46), can be expanded in the inverse powers of M_0,

$$\mathrm{Tr}\left[\gamma^5\,(\not{P} - M_0)^{-1}\right] = \qquad (33.48)$$

$$\mathrm{Tr}\left[\gamma^5\,(\not{P} + M_0)\left(\frac{1}{P^2 - M_0^2} - \frac{1}{P^2 - M_0^2}\,\frac{1}{2}\, i\,\varepsilon^{\mu\nu} F_{\mu\nu}\gamma^5\,\frac{1}{P^2 - M_0^2} + \ldots\right)\right].$$

The first term in the expansion vanishes after taking the trace of the γ matrices. The third and all other terms are irrelevant because they vanish in the limit $M_0 \to \infty$. The only relevant term is the second one where we can substitute the operator P_μ by the momentum p_μ since the result is explicitly proportional to the background field $F_{\mu\nu}$, and the chiral anomaly in the Schwinger model is linear in $F_{\mu\nu}$. Then

$$2\, i M_0 \bar{R}\gamma_5 R = -2M_0^2 \int \frac{d^2p}{(2\pi)^2}\,\frac{i}{(p^2 - M_0^2)^2}\,\varepsilon^{\mu\nu} F_{\mu\nu}.$$

Upon performing the Wick rotation and integrating over p we arrive at Eq. (33.45).

This computation completes the standard derivation of the anomaly. One should have a very rich imagination to be able to see in these formal manipulations the simple physical nature of the phenomenon which has been described above (restructuring of the Fermion sea and the level crossing). Nevertheless, this is the same phenomenon viewed from a different angle – less transparent but more economic since we can get the final result very quickly using the well-developed machinery of the diagram technique, familiar to everybody.

Let us ask the following question: “What is the infrared connection (or infrared face, if you wish) of the anomaly in the diagram language?” To extract the infrared aspect from the Feynman graphs it is necessary to turn back to consideration of the current $j^{\mu 5}$. Our aim is to calculate the matrix element of the current $j^{\mu 5}$ in the background photon field. Unlike $\partial_\mu j^{\mu 5}$ the matrix element $\langle j^{\mu 5}\rangle$ contains an infrared contribution. Because of this it is impossible to consider $\langle j^{\mu 5}\rangle$ for the on-mass-shell photon, with the momentum $k^2 = 0$. We are forced to introduce “off-shellness” to ensure the infrared regularization (a substitute for finite L, see above). Thus, we will consider the photon field A_μ which does not obey the equations of motion.

General arguments (such as gauge invariance) imply the following expression for the matrix element $\langle j^{\mu 5}\rangle$ stemming from the diagram **(d)** of Fig. 33.4 :

$$\langle j^{\mu 5}\rangle = \text{const}\,\frac{k^{\mu}}{k^2}\,\varepsilon^{\alpha\beta}F_{\alpha\beta}\ , \tag{33.49}$$

where the constant in the right-hand side can be determined by an explicit computation of the graph. In principle, there is one more structure with the appropriate dimension and quantum numbers, namely $\varepsilon^{\mu\nu}A_\nu$, but it cannot appear by itself if gauge invariance is maintained. In other words, one can say that the *local* structure $\varepsilon^{\mu\nu}A_\nu$ can always be eliminated by subtraction of an ultraviolet counterterm.

It is worth noting that, purely kinematically,

$$k^{\mu}\varepsilon^{\alpha\beta}F_{\alpha\beta} = -2i\,\varepsilon^{\mu\nu}\left[k^2A_\nu - k_\nu(k^{\rho}A_\rho)\right]\,. \tag{33.50}$$

It is seen that in order to distinguish an infrared singular term proportional to k^{-2} from the local term depending on the ultraviolet regularization it is necessary to assume that $k^\rho A_\rho \neq 0$. The infrared singular term is fixed unambiguously by the diagram **(d)** of Fig. 33.4 . The easiest way to get it is just to compute this graph in a straightforward way,

$$\langle j^{\mu 5}\rangle = (-1)\int\frac{d^2p}{(2\pi)^2}\,\text{Tr}\left[\gamma^\mu\gamma^5\,\frac{i\not{p}}{p^2}\,i\,\gamma^\rho\,\frac{i(\not{p}+\not{k})}{(p+k)^2}\right]A_\rho\,. \tag{33.51}$$

Performing the p integration and disregarding terms non-singular in k^2 we get

$$\int\frac{p^\alpha}{p^2}\,\frac{(p+k)^\beta}{(p+k)^2}\,\frac{d^2p}{(2\pi)^2}\to\frac{i}{4\pi}\,\frac{k^\alpha k^\beta}{k^2}\ ,$$

which implies, in turn,

$$\langle j^{\mu 5}\rangle_{\text{sing.}} = -\frac{1}{4\pi k^2}\,\text{Tr}\left[\gamma^\mu\gamma^5\not{k}\,\gamma^\rho\not{k}\right]A_\rho\to\frac{1}{\pi k^2}\,\varepsilon^{\mu\nu}k_\nu(k^\rho A_\rho)\,.$$

Anomaly from the IR side

Now, inserting the local term in order to restore the gauge invariance and using Eq. (33.50) we arrive at

$$\langle j^{\mu 5}\rangle = -\frac{i}{2\pi}\,\frac{k^{\mu}}{k^2}\,\varepsilon^{\alpha\beta}F_{\alpha\beta}\ . \tag{33.52}$$

Taking the divergence is equivalent to multiplication of the right-hand side by $-ik_\mu$, and we reproduce, now for the third time, the anomalous relations (33.43) .