Sphalerons, spectral flow, and anomalies

Frans R. Klinkhamer and Christian Rupp

Institut für Theoretische Physik,
Universität Karlsruhe (TH),
76128 Karlsruhe, Germany

Abstract

The topology of configuration space may be responsible in part for the existence of sphalerons. Here, sphalerons are defined to be static but unstable finite-energy solutions of the classical field equations. Another manifestation of the nontrivial topology of configuration space is the phenomenon of spectral flow for the eigenvalues of the Dirac Hamiltonian. The spectral flow, in turn, is related to the possible existence of anomalies. In this review, the interconnection of these topics is illustrated for three particular sphalerons of SU(2) Yang–Mills–Higgs theory.

Invited paper for the special issue of the Journal of Mathematical Physics on “Integrability, topological solitons and beyond” edited by T. Fokas and N.S. Manton.
I. INTRODUCTION

One of the main themes of the present special issue concerns the so-called topological solitons. The field configurations of these classical solutions are characterized by a topologically nontrivial map of the space manifold (or part of it) into some internal space of the model considered. A well-known example is the Skyrme soliton \(^1\), for which the space manifold \(S^3\) (i.e., the compactified Euclidean space \(\mathbb{R}^3\)) is mapped into the internal space \(SU(2)\). Another example is the magnetic monopole \(^2\), for which the “sphere at infinity” \(S^2\) is mapped into the Higgs vacuum manifold \(SO(3)/SO(2)\).

There exist, however, other classical solutions, the so-called sphalerons, which themselves have trivial topology but trace back to nontrivial topology in the configuration space of the fields \(^3,\,4\). In this contribution, we intend to give an elementary discussion of sphaleron solutions in Yang–Mills–Higgs theory and the underlying topology. In order to get a clear picture of what goes on, we focus on a single Yang–Mills–Higgs theory and three specific sphalerons \(^5,\,6,\,7\).

Physically, the topological solitons and the sphalerons play a different role. Solitons are primarily relevant to the equilibrium properties of the theory (e.g., the existence of certain stable asymptotic states), whereas sphalerons are of importance to the dynamics of the theory. The sphaleron \(^8\) of the electroweak Standard Model \(^8\), for example, is believed to play a crucial role for baryon-number-violating processes in the early universe (see, e.g., Refs. \(^9,\,10\) for two reviews).

The outline of this article is as follows. In Section II, we present the theory considered, to wit, \(SU(2)\) Yang–Mills theory with a single complex isodoublet of Higgs fields. This particular Yang–Mills–Higgs theory forms the core of the electroweak Standard Model of elementary particle physics. In Section III, we recall some basic facts about the mapping of spheres into spheres, in particular their homotopy classes. In Section IV, we describe three sphaleron solutions and their topological \textit{raison d’être}. In Section V, we discuss another manifestation of the nontrivial topology of configuration space, namely the spectral flow of the eigenvalues of the Dirac Hamiltonian. The word “spectral flow” is used in a generalized sense, meaning \textit{any} type of rearrangement of the energy levels. Loosely speaking, the spectral flow makes it possible for a sphaleron to acquire a fermion zero-mode. In Section VI, we link the spectral flow to the possible occurrence of anomalies (which signal the loss of one or more classical symmetries). In Section VII, finally, we present some concluding remarks.
II. SU(2) YANG–MILLS–HIGGGS THEORY

In this article, we consider a simplified version of the electroweak Standard Model without the hypercharge $U(1)$ gauge field. This means, basically, that we set the weak mixing angle $\theta_w \equiv \arctan(g'/g)$ to zero, where $g'$ and $g$ are the coupling constants of the $U(1)$ and $SU(2)$ gauge groups, respectively. Also, we take only one family of quarks and leptons instead of the three known experimentally.

In general, the fields are considered to propagate in Minkowski spacetime with coordinates $x^\mu \in \mathbb{R}$, $\mu = 0, 1, 2, 3$, and metric $g_{\mu\nu}(x) = \text{diag}(+1, -1, -1, -1)$. But occasionally we go over to Euclidean spacetime with metric $g_{\mu\nu}(x) = \delta_{\mu\nu}$. Natural units with $\hbar = c = 1$ are used throughout.

The $SU(2)$ Yang–Mills gauge field is denoted by $A_\mu(x) \equiv A_\mu^a(x) \tau^a/(2i)$, where the $\tau^a$ are the three Pauli matrices acting on weak isospin space and the component fields $A_\mu^a(x)$ are real. (Repeated indices are summed over, unless stated otherwise.) The complex Higgs field transforms as an isodoublet under the $SU(2)$ gauge group and is given by $\Phi(x) = (\Phi_1(x), \Phi_2(x))^t$, where the suffix $t$ stands for transpose [cf. Eq. (5.5) below]. The fermion fields will be discussed in Section V.

The classical action of the gauge and Higgs fields reads

$$\Gamma_{YMH} = \int_{\mathbb{R}^4} d^4x \left\{ \frac{1}{2} \text{tr} F_{\mu\nu} F^{\mu\nu} + (D_\mu \Phi)^\dagger(D^\mu \Phi) - \lambda \left( \Phi^\dagger \Phi - \eta^2 \right)^2 \right\}, \quad (2.1)$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu + g[A_\mu, A_\nu]$ is the $SU(2)$ Yang–Mills field strength and $D_\mu \Phi \equiv (\partial_\mu + gA_\mu)\Phi$ the covariant derivative of the Higgs field. The theory has Yang–Mills coupling constant $g$ and quartic Higgs coupling constant $\lambda$, but the classical dynamics depends only on the ratio $\lambda/g^2$. The parameter $\eta$ has the dimension of mass and sets the scale of the Higgs expectation value. The three $W$ vector bosons then have equal mass, $M_W = g \eta/\sqrt{2}$. The single Higgs scalar boson has a mass $M_H = 2 \sqrt{\lambda} \eta$.

The action (2.1) is invariant under a local gauge transformation

$$\Phi'(x) = \Lambda(x) \Phi(x), \quad gA'_\mu(x) = \Lambda(x) \left( gA_\mu(x) + \partial_\mu \right) \Lambda(x)^{-1}, \quad (2.2)$$

for an arbitrary gauge function $\Lambda(x) \in SU(2)$. In addition, there are certain global $SU(2)$ and $U(1)$ symmetry transformations which operate solely on the Higgs field.

III. MAPS OF SPHERES INTO SPHERES

Let us consider continuous maps from a connected manifold $M$ to a connected manifold $N$. Two such maps, $f_1$ and $f_2$, are called homotopic if the one can be
obtained from the other by continuous deformation. More specifically, \( f_1 \) and \( f_2 \) are homotopic if there exists a continuous map \( h : [0,1] \times M \to N \) such that \( h(0,m) = f_1(m) \) and \( h(1,m) = f_2(m) \) for all \( m \in M \). All maps \( M \to N \) can be divided into equivalence classes, where two maps are equivalent if they are homotopic (see, e.g., Ref. [11]).

We are particularly interested in the case where \( M \) and \( N \) are the spheres \( S^m \) and \( S^n \), respectively. The set of homotopy classes is called the homotopy group and is denoted by \( \pi_m(S^n) \). Figure 1 shows two maps \( S^1 \to S^1 \) which are not homotopic. It is clear that in this particular case the homotopy classes can be labeled by integer numbers which describe how often the original circle \( S^1 \) is wrapped around the target circle \( S^1 \). This explains the result \( \pi_1(S^1) = \mathbb{Z} \), where \( \mathbb{Z} \) denotes the group of integers under addition. The two maps shown in Fig. 1 have winding numbers 1 and 0.

The homotopy classes of \( S^n \to S^n \), for \( n > 1 \), can be pictured analogously, since the representation of a sphere \( S^n \) in spherical coordinates contains exactly one azimuthal angle \( \phi \in [0,2\pi] \). The result is \( \pi_n(S^n) = \mathbb{Z} \). Further homotopy groups are: \( \pi_m(S^n) = \{0\} \) for \( m < n \), \( \pi_3(S^2) = \mathbb{Z} \), and \( \pi_4(S^3) = \mathbb{Z}_2 \), where \( \mathbb{Z}_2 \) denotes the group of integers \( \{0,1\} \) under addition modulo 2.

Next, consider families of maps \( S^m \to S^n \), where the family parameters themselves form a sphere \( S^p \). In short, consider \( S^p \times S^m \to S^n \). Imposing certain constraints, these families of maps can be viewed as maps \( S^{p+m} \to S^n \) and classified according to the homotopy groups of spheres.

To this end, we introduce the smash product [11] of two spheres \( S^p \) and \( S^m \). The smash product \( S^p \wedge S^m \) is defined as the Cartesian product \( S^p \times S^m \) with the set \((\{x_0\} \times S^m) \cup (S^p \times \{y_0\})\) considered as a single point, for some arbitrarily chosen \( x_0 \in S^p \) and \( y_0 \in S^m \). It can be shown that \( S^p \wedge S^m \) is homeomorphic to the sphere \( S^{p+m} \) (see Fig. 2 for a sketch of the proof).
FIG. 2: Left: Cartesian product space $S^1 \times S^1$ drawn as a torus, with two circles representing the sets $\{x_0\} \times S^1$ and $S^1 \times \{y_0\}$ for some arbitrarily chosen points $x_0$ and $y_0$. Middle: surface after shrinking the first of these sets to a point. Right: surface after shrinking the second set to a point, which gives the smash product $S^1 \wedge S^1 \sim S^2$.

A simple corollary will be important in the following. Any map $f : S^p \times S^m \rightarrow S^n$ can effectively be considered as a map defined on $S^p \wedge S^m$ if $f(x_0, y)$ is independent of $y$ and $f(x, y_0)$ is independent of $x$, for an appropriate choice of $x_0 \in S^p$ and $y_0 \in S^m$.

IV. SPHALERONS

The word sphaleron is of Greek origin and means “ready to fall” (see Ref. [5] for the etymology). It is used to denote a static but unstable solution of the classical field equations with finite total energy of the fields.

In this article, only finite-energy configurations of the fields will be considered. By analogy to Morse theory [12], sphalerons can then be looked for by a minimax procedure [3] if the configuration space of the underlying field theory is multiply connected.

The procedure runs as follows: first, construct a noncontractible $p$-dimensional sphere $S^p$ in configuration space, then determine its maximal energy configuration, and, finally, “shrink” the sphere to minimize this maximal energy. If the configuration space were compact, this procedure would be guaranteed to give a saddle point. But configuration space is infinite-dimensional and noncompact, so that the minimax procedure produces at best only a candidate solution. It has to be checked explicitly that the appropriate minimax-configuration solves the classical field equations. If this is the case, the minimax-configuration is a genuine sphaleron.
A. Sphaleron S

For the sphaleron S of the SU(2) Yang–Mills–Higgs theory, we consider three-space to be compactified by adding the “sphere at infinity.” Configuration space is then the space of all static three-dimensional gauge and Higgs field configurations in a particular gauge which have finite energy. The static gauge field can be written as a Lie-algebra-valued one-form,

\[ A(x^1, x^2, x^3) \equiv A^a_m(x^1, x^2, x^3) \tau^a/(2\pi) \, dx^m, \]

with implicit sums of \( a \) and \( m \) over 1, 2, 3. Furthermore, we use spherical coordinates \((r, \theta, \phi)\) over \(\mathbb{R}^3\) and employ the radial gauge condition \(A_r = 0\), together with \(A_0 = 0\).

Since the energy has to be finite, only those configurations are admissible for which the gauge field tends towards a pure-gauge configuration as \(r \to \infty\) and the Higgs field towards its associated vacuum value,

\[ g \, A^\infty = -d\omega \, \omega^{-1}, \quad \Phi^\infty = \eta \, \omega \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \]

for a map \(\omega\) of the “sphere at infinity” \(S^2\) into the gauge group \(SU(2)\).

Any loop in configuration space induces a loop in the space of these mappings \(\omega\). The corresponding map is denoted by

\[ U : S^1 \times S^2 \to SU(2), \quad (\mu, \theta, \phi) \mapsto U(\mu, \theta, \phi), \]

where \(\mu \in [0, \pi]\) is the parameter of the loop of configurations and \(\theta \in [0, \pi]\) and \(\phi \in [0, 2\pi]\) are spherical coordinates in three-space.

By imposing certain constraints on this map, we may effectively reduce the set of allowed loops, so that \(U\) becomes a map \(S^3 \to SU(2) \sim S^3\) which falls into homotopy classes according to \(\pi_3(SU(2)) = \pi_3(S^3) = \mathbb{Z}\). To be specific, the map \(U\) for \(\mu = 0\) and \(\mu = \pi\) must not depend on \(\theta\) and \(\phi\), and the map \(U\) for \(\theta = 0\) has to be independent of \(\mu\). Then \(U\) is effectively defined on the smash product \(S^1 \wedge S^2 \sim S^3\), as explained in the last paragraph of Section III. Now there exist noncontractible loops of field configurations for which the minimax procedure can be performed.

An appropriate expression for the map (4.3) is given by

\[ U(\mu, \theta, \phi) = y^1 \, (-1\tau_1) + y^2 \, (-1\tau_2) + y^3 \, (-1\tau_3) + y^4 \, \Pi_2, \]

with

\[ \begin{pmatrix} y^1 \\ y^2 \\ y^3 \\ y^4 \end{pmatrix} = \begin{pmatrix} -\sin \mu \sin \theta \sin \phi \\ -\sin \mu \sin \theta \cos \phi \\ \sin \mu \cos \mu \, (\cos \theta - 1) \\ \cos^2 \mu + \sin^2 \mu \cos \theta \end{pmatrix}. \]
In order to calculate the winding number of this particular map $U$, we examine its relation to the standard spherical coordinates on $S^3$, 

$$\begin{pmatrix} z^1 \\ z^2 \\ z^3 \\ z^4 \end{pmatrix} = \begin{pmatrix} \cos \theta_2 \\ \sin \theta_2 \cos \theta_1 \\ \sin \theta_2 \sin \theta_1 \sin \phi_1 \\ \sin \theta_2 \sin \theta_1 \cos \phi_1 \end{pmatrix},$$

(4.6) 

with polar angles $\theta_1, \theta_2 \in [0, \pi]$ and azimuthal angle $\phi_1 \in [0, 2\pi]$.

We first observe that the two-vector

$$\vec{w} = \begin{pmatrix} \cos \mu \\ \sin \mu \cos \theta \end{pmatrix}$$

(4.7) 

sweeps over the unit disk if $\mu$ and $\theta$ run from 0 to $\pi$. Since rotations map the unit disk onto itself and leave the length of $\vec{w}$ invariant,

$$|\vec{w}|^2 = 1 - \sin^2 \mu \sin^2 \theta,$$

(4.8) 

the relation

$$\begin{pmatrix} \cos \theta_2 \\ \sin \theta_2 \cos \theta_1 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \mu \\ \sin \mu \cos \theta \end{pmatrix}$$

(4.9) 

describes an admissible reparametrization of the disc. By choosing $\alpha = -\mu$, we find $y^4 = z^1$ and $y^3 = z^2$. With $\sin \theta_1 \sin \theta_2 = \sin \mu \sin \theta$ and $\phi_1 = \phi$, we have also $y^1 = -z^3$ and $y^2 = -z^4$.

The conclusion is that the map $U$ as defined by Eqs. (4.4)–(4.5) covers the target sphere $S^3$ exactly once. The map $U$ has winding number one (or minus one, depending on the definition of the winding number) and corresponds to a nontrivial element of the homotopy group $\pi_3(S^3) = \mathbb{Z}$.

For the static $SU(2)$ gauge and Higgs fields of the noncontractible loop (NCL), we make the Ansatz

$$g A(\mu, r, \theta, \phi) = -f(r) \, dU(\mu, \theta, \phi) \, U(\mu, \theta, \phi)^{-1},$$

(4.10a) 

$$\Phi(\mu, r, \theta, \phi) = \eta \, h(r) \, U(\mu, \theta, \phi) \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \eta \left(1 - h(r)\right) \begin{pmatrix} 0 \\ \exp(-1 \mu \cos \mu) \end{pmatrix},$$

(4.10b) 

with the following boundary conditions for the radial functions $f$ and $h$:

$$\lim_{r \to 0} f(r)/r = 0, \quad \lim_{r \to \infty} f(r) = 1,$$

(4.11a) 

$$\lim_{r \to 0} h(r) = 0, \quad \lim_{r \to \infty} h(r) = 1.$$  

(4.11b)
FIG. 3: Sphaleron S on top of a noncontractible loop in configuration space. The field energy is zero for the vacuum and positive for the sphaleron S.

The energy density of the fields (4.10) turns out to be spherically symmetric. Indeed, the fields of the NCL can also be written in a manifestly spherically symmetric form [14].

The fields (4.10) of the NCL at $\mu = 0$ or $\pi$ correspond to the Higgs vacuum with $A(x) = 0$ and $\Phi(x) = (0, \eta)^t$. This particular configuration is independent of the radial functions $f$ and $h$ and has zero energy according to Eq. (2.1). The NCL configuration at $\mu = \pi/2$ is distinguished by having parity reflection symmetry (the only other configuration of the NCL with this property is the vacuum at $\mu = 0$). For given functions $f$ and $h$, this $\mu = \pi/2$ configuration is also the maximum energy configuration over the NCL. The minimax procedure now consists of adjusting the radial functions $f$ and $h$ while maintaining the boundary conditions (4.11), so that the energy at $\mu = \pi/2$ is minimized. The resulting configuration is the sphaleron S, as sketched in Fig. 3.

Using numerical methods, one finds for the sphaleron energy the value

$$E_S \approx 3.04 \times (4\pi/g^2) M_W ,$$

(4.12)

which holds for the case of vanishing quartic Higgs coupling constant ($\lambda/g^2 = 0$). [The sphaleron energy $E_S$ has also been calculated for the full $SU(2) \times U(1)$ Yang–Mills–Higgs theory of the electroweak Standard Model. The energy $E_S$ is found to be weakly dependent on the mixing angle $\theta_w$, at least near $\theta_w = 0$. The emergence of a large magnetic dipole moment $\mu_S \propto (4\pi/g^2) g \tan \theta_w/M_W$ is perhaps more interesting. See Refs. [5,15,16] for details.] For large enough values of $\lambda/g^2$, additional solutions appear, the so-called “deformed sphalerons” [17, 18]. The appearance of these extra sphalerons can be explained [14] by a simple deformation of the energy surface in Fig. 3.

The sphaleron S by itself has trivial topology, with the “sphere at infinity” $S^2$ mapped into the Higgs vacuum manifold $SU(2) \sim S^3$; cf. Section [11]. (As mentioned
in the Introduction, the magnetic monopole in \( SU(2) \) Yang–Mills theory with a real isotriplet of Higgs is based on the nontrivial map \( S^2 \to SO(3)/SO(2) \sim S^2 \).) Note that the original S Ansatz, with the so-called hedgehog structure, was discovered ten years before the construction of S via the NCL.

In the radial gauge, the vacuum configuration of the \( SU(2) \) gauge field is uniquely fixed, \( gA^\text{vac}_m(x^1, x^2, x^3) = 0 \). If this gauge condition is abandoned, any pure-gauge configuration \( gA_m = -\partial_m U U^{-1} \), for arbitrary time-independent \( SU(2) \)-valued field \( U \), is a possible vacuum configuration. Depending on the topology of three-space, these vacuum configurations may or may not fall into different unconnected classes. This does not happen for our compactification \( \mathbb{R}^3 \cup S^2_\infty \). But the situation changes if, instead, we choose a one-point compactification \( \mathbb{R}^3 \cup \{\infty\} \), with all fields approaching a single direction-independent value as \( r \to \infty \). Each vacuum configuration then corresponds to a map \( S^3 \to S^3 \) and there are topologically distinct vacuum classes, since \( \pi_3(S^3) = \mathbb{Z} \).

In fact, it is possible to perform a gauge transformation on the NCL which changes the asymptotic behavior of the gauge fields, so that they can be considered to live on \( S^3 = \mathbb{R}^3 \cup \{\infty\} \). Let \( \omega(\mu, r, \theta, \phi) \) be an \( SU(2) \)-valued map which approaches \( U(\mu, \theta, \phi) \) for \( r \to \infty \) and \( \mathbb{1}_2 \) for \( r \to 0 \). The radial dependence of \( \omega(\mu, r, \theta, \phi) \) implements a path which connects the map \( U(\mu, \theta, \phi) = \omega(\mu, \infty, \theta, \phi) \) to the constant map \( \mathbb{1}_2 = \omega(\mu, 0, \theta, \phi) \). [Note that \( U(\mu, \theta, \phi) \) for fixed \( \mu \) is a map \( S^3 \to S^3 \) and therefore contractible.]

The crucial point, now, is that the map \( \omega(0, r, \theta, \phi) \) is homotopically different from the map \( \omega(\pi, r, \theta, \phi) \). [Otherwise, the radial dependence of \( \omega(\mu, r, \theta, \phi) \) would yield a contraction of \( \omega(\mu, \infty, \theta, \phi) \), considered as a \( \mu \)-dependent map \( S^3 \to S^3 \), which is impossible.] Both maps \( \omega(0, r, \theta, \phi) \) and \( \omega(\pi, r, \theta, \phi) \) can also be viewed as maps \( S^3 \to S^3 \), since \( \omega(\mu, \infty, \theta, \phi) = U(\mu, \theta, \phi) = \mathbb{1}_2 \) for \( \mu = 0 \) and \( \mu = \pi \). The conclusion is then that the corresponding vacuum configurations \( A_m(x^1, x^2, x^3) \) at \( \mu = 0 \) and \( \mu = \pi \) belong to different homotopy classes. This result will be discussed further in Section VI A.

### B. Sphaleron \( S^* \)

The three-space of our \( SU(2) \) Yang–Mills–Higgs theory is again compactified by adding the “sphere at infinity.” This time, however, we do not consider one-parameter families (loops) of static finite-energy configurations but two-parameter
families (spheres). At spatial infinity, these families are characterized by the map

\[ U : S^2 \times S^2 \to SU(2), \quad (\mu, \nu, \theta, \phi) \mapsto U(\mu, \nu, \theta, \phi), \quad (4.13) \]

where \((\mu, \nu)\) are the parameters of the sphere of configurations and \((\theta, \phi)\) are the polar and azimuthal angles of the spherical coordinates in three-space. The parameters \(\mu\) and \(\nu\) run from \(-\pi/2\) to \(+\pi/2\) and the boundary of the \((\mu, \nu)\)-square at \(|\mu| = \pi/2\) or \(|\nu| = \pi/2\) is mapped to the same element of \(SU(2)\).

Next, restrict the class of mappings \(U\) by requiring that \(U(\mu, \nu, 0, \phi)\) is independent of \((\mu, \nu, \phi)\) and \(U(-\pi/2, -\pi/2, \theta, \phi)\) independent of \((\theta, \phi)\). Then \(U\) is effectively a mapping from \(S^4\) to \(S^3\), which has a nontrivial homotopy structure, \(\pi_4(S^3) = \mathbb{Z}_2\).

The general idea, now, is to construct a noncontractible sphere, to determine the maximal energy configuration on that sphere and to continuously deform the sphere so that its maximal energy is minimized [19].

The construction of the required nontrivial map \(S^4 \to S^3\) is done in two steps. First, a nontrivial map \(S^3 \to S^2\) is found and, second, an operation is performed to increase the dimension of both spheres.

The relevant map \(S^3 \to S^2\) is given by the well-known Hopf fibration [11], which can be explained as follows. Consider the three-sphere \(S^3\) to be a subset of \(\mathbb{C}^2\), namely \(\{(z_1, z_2) \mid |z_1|^2 + |z_2|^2 = 1\}\). Each \(\mathbb{C}\)-line through the origin in \(\mathbb{C}^2\) then intersects with this three-sphere in a great circle \(S^1\). These great circles \(S^1\) form a pairwise disjoint covering of \(S^3\). Two points of \(S^3\) are defined to be equivalent \((\simeq)\), if they lie on the same great circle \(S^1\). The corresponding projection,

\[ S^3 \to S^3 / \simeq, \quad (4.14) \]

is the desired Hopf map, since the topological space \(S^3 / \simeq\) is homeomorphic to \(S^2\).

The topological equivalence of \(S^3 / \simeq\) and \(S^2\) can be shown by considering the \(\mathbb{C}\)-lines which label the great circles \(S^1\) discussed in the previous paragraph. All but one of these \(\mathbb{C}\)-lines can be parametrized by complex numbers \(c \in \mathbb{C}\). Specifically, the coordinates of such a line read

\[ (z_1, z_2) = (w, c w), \quad \text{for } w \in \mathbb{C}. \quad (4.15) \]

In addition, there is the single \(\mathbb{C}\)-line given by

\[ (z_1, z_2) = (0, w), \quad \text{for } w \in \mathbb{C}. \quad (4.16) \]

Hence, the total parameter space of \(S^3 / \simeq\) is given by the one-point-compactified plane, i.e., the Riemann sphere \(S^2\).
The suspension of a sphere $S^n$ is essentially the same as the smash product $S^1 \wedge S^n$. It can be used to increase the dimension of the spheres appearing in the above discussion. The resulting suspension of the Hopf map corresponds to a nontrivial element of the homotopy group $\pi_4(S^3) = \mathbb{Z}_2$.

In a particular parametrization, the required map (4.13) takes the form

$$U(\mu, \nu, \hat{x}) = \left( \sin \mu + i \cos \mu \exp[+i(\nu + \pi/2) \tau_3] \hat{x} \cdot \tau \exp[-i(\nu + \pi/2) \tau_3] \right) \times \left( \sin \mu - i \cos \mu \hat{x} \cdot \tau \right),$$

(4.17)

where $\mu$ and $\nu$ range over $[-\pi/2, \pi/2]$ and describe a two-sphere, as does the unit three-vector $\hat{x} \equiv (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$. The map $U(\mu, \nu, \hat{x})$ is effectively defined on the smash product $S^2 \wedge S^2$, since $U$ is independent of $\hat{x}$ for $\mu = \pm \pi/2$ or $\nu = \pm \pi/2$ and independent of $\mu$ and $\nu$ for $\hat{x} = (0, 0, 1)$. (Note that the suspended Hopf map also plays a role in the physics of Skyrme solitons [20].)

With the map (4.17) in hand, it is possible to construct a noncontractible sphere (NCS) of static Yang–Mills–Higgs configurations and to obtain the corresponding nontrivial classical solution, the sphaleron $S^*$, just as for the NCL and the sphaleron $S$ of the previous subsection. The construction of $S^*$ is, however, rather subtle. Here, we only describe the four basic steps and refer the reader to Ref. [6] for more information.

First, we observe that the map (4.17) singles out the $x^3$ axis, which suggests the use of the cylindrical coordinates $\rho$, $\phi$, and $z$, defined by $(x^1, x^2, x^3) = (\rho \cos \phi, \rho \sin \phi, z)$. Then, it is not difficult to construct a NCS of static $SU(2)$ Yang–Mills–Higgs configurations, whose behavior at infinity is governed by the $SU(2)$ matrix (4.17). Specifically, the NCS configurations can be written in terms of six axial functions $f_i(\rho, z)$ and $h_j(\rho, z)$, for $i = 0, 1, 2, 3$ and $j = 1, 2$, with appropriate boundary conditions (for example, $f_0(\rho, z) \to 1$ and $h_1(\rho, z) \to 1$ as $\rho^2 + z^2 \to \infty$). The $SU(2)$ gauge and Higgs field configuration of the NCS are by construction axially symmetric.

Second, the configuration at $\mu = \nu = 0$ is also invariant under parity reflection and gives the maximum energy of the NCS. Moreover, it can be verified that this $\mu = \nu = 0$ configuration, in terms of the six axial functions $f_i(\rho, z)$ and $h_j(\rho, z)$, provides a self-consistent Ansatz for the $SU(2)$ Yang–Mills–Higgs field equations. Concretely, this means that the Ansatz reduces the field equations to precisely six partial differential equations (PDEs) for the six functions $f_i(\rho, z)$ and $h_j(\rho, z)$, with appropriate boundary conditions which trace back in part to the finite-energy condition. (This result agrees with the so-called principle of symmetric criticality [21], which states that, under certain conditions, it suffices to consider variations that
FIG. 4: Sphaleron S∗ on top of a noncontractible sphere in configuration space.

respect the symmetry of the Ansatz.) The solution of these PDEs then determines the field configurations of the sphaleron S∗. See Fig. 4 for a sketch of configuration space.

Third, the reduced field equations for the sphaleron S∗ can be solved numerically. For approximately vanishing quartic Higgs coupling constant \( \lambda/g^2 = 1/800 \), the numerical solution of the six PDEs with the correct boundary conditions gives the following value for the energy:

\[
E_{S^*} \approx 1.91 \times E_S,
\]

where \( E_S \) denotes the corresponding energy of the sphaleron S [cf. Eq. (4.12) above]. In fact, the sphaleron S∗ is found to have the structure of a di-atomic molecule, binding together a sphaleron S and an “anti-sphaleron” \( \bar{S} \). See Ref. [6] for a plot of the energy density and further discussion.

Fourth, the construction of S∗ via the NCS can be extended to the full \( SU(2) \times U(1) \) theory of the electroweak Standard Model by the introduction of one more axial function, \( f_4(\rho, z) \), with trivial boundary conditions at infinity. But for nonvanishing weak mixing angle \( \theta_w \), there are only preliminary numerical results for the sphaleron S∗ and it would be worthwhile to obtain accurate results over the full range of values of \( \lambda/g^2 \) and \( \theta_w \).

C. Z-string

Now consider static field configurations of the \( SU(2) \) Yang–Mills–Higgs theory which are independent of one spatial coordinate, the \( z \)-coordinate, and have vanishing gauge potential in that direction, \( A_z = 0 \). In order to have finite total
energy, the z-direction has to be compact and three-space is taken to be \( \mathbb{R}^2 \times S^1 \) instead of \( \mathbb{R}^3 \). Also, choose cylindrical polar coordinates \((\rho, \phi, z)\) and work in the polar gauge for which \( A_\rho = 0 \).

Since the energy density in a plane with fixed \( z \) has to be finite, the remaining gauge field component \( A_\phi \) reduces to a pure-gauge configuration asymptotically,

\[
g A_\phi \rightarrow - (\partial_\phi \omega) \omega^{-1}, \quad \text{as} \quad \rho \rightarrow \infty,
\]

(4.19)

for a map \( \omega : S^1 \rightarrow SU(2) \). Basically, this means that the plane \( \mathbb{R}^2 \) is compactified by adding the “circle at infinity.”

It is possible to construct a noncontractible sphere of these field configurations by restricting the corresponding maps

\[
U : S^2 \times S^1 \rightarrow SU(2), \quad (\mu, \nu, \phi) \mapsto U(\mu, \nu, \phi),
\]

(4.20)

in such a way that they are effectively defined on the smash product \( S^2 \wedge S^1 \sim S^3 \). Specifically, the sphere is parametrized by \( \mu \) and \( \nu \) which take values in \([-\pi/2, +\pi/2]\). The rim of the \((\mu, \nu)\)-square is identified and corresponds to a single point on \( S^2 \). The map \( U \) is restricted to be independent of \( \phi \) if \((\mu, \nu)\) lies on this rim and independent of \((\mu, \nu)\) if \( \phi = 0 \).

An appropriate expression for the map (4.20) is given by

\[
U(\mu, \nu, \phi) = V(\mu, \nu, 0)^{-1} V(\mu, \nu, \phi),
\]

(4.21a)

\[
V(\mu, \nu, \phi) = z^1 (-i \tau_1) + z^2 (-i \tau_2) + z^3 (-i \tau_3) + z^4 \mathbb{1}_2,
\]

(4.21b)

\[
\begin{pmatrix}
  z^1 \\
  z^2 \\
  z^3 \\
  z^4
\end{pmatrix}
= 
\begin{pmatrix}
  \sin \mu \\
  \cos \mu \sin \nu \\
  \cos \mu \cos \nu \sin \phi \\
  \cos \mu \cos \nu \cos \phi
\end{pmatrix}.
\]

(4.21c)

Note that the factor \( V(\mu, \nu, 0)^{-1} \) in (4.21a) serves a dual purpose. First, it assures that the rim of the \((\mu, \nu)\)-square is mapped to a single element, since \( V(\mu, \nu, \phi) \) is independent of \( \phi \) on this boundary. Second, it makes \( U \) independent of \( \mu \) and \( \nu \) if \( \phi = 0 \).

For the two-dimensional \( SU(2) \) gauge and Higgs fields of the noncontractible sphere (NCS), we make the \textit{Ansatz}

\[
g A(\mu, \nu, \rho, \phi) = - f(\rho) \, dU(\mu, \nu, \phi) U(\mu, \nu, \phi)^{-1},
\]

(4.22a)

\[
\Phi(\mu, \nu, \rho, \phi) = \eta h(\rho) \, U(\mu, \nu, \phi) \begin{pmatrix}
  0 \\
  1
\end{pmatrix},
\]

(4.22b)
FIG. 5: Z-string on top of a noncontractible sphere (balloon) in configuration space.

with parameters \(|\mu|, |\nu| \leq \pi/2\). The polar functions \(f\) and \(h\) have the following boundary conditions:

\[
\begin{align*}
\lim_{\rho \to 0} f(\rho)/\rho &= 0, & \lim_{\rho \to \infty} f(\rho) &= 1, \\
\lim_{\rho \to 0} h(\rho) &= 0, & \lim_{\rho \to \infty} h(\rho) &= 1.
\end{align*}
\]

But no point on this NCS corresponds to a vacuum configuration, since the Higgs field in Eq. (4.22b) vanishes at \(\rho = 0\) for all values of \((\mu, \nu)\). Therefore, the point of the NCS at the boundary of the \((\mu, \nu)\)-square must be connected to the vacuum by an additional line segment. The corresponding Ansatz is simply

\[
\begin{align*}
g A(\mu, \nu, \rho, \phi) &= 0, \\
\Phi(\mu, \nu, \rho, \phi) &= \eta \left(1 - \sin[\mu\nu] + h(\rho) \sin[\mu\nu]\right) \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\end{align*}
\]

for \([\mu\nu] \equiv \max\{|\mu|, |\nu|\} > \pi/2\) with the parameter range of \(\mu\) and \(\nu\) extended to \([-\pi, +\pi]\). The set of configurations (4.21)–(4.24) is like a “balloon” which is tied to the ground by a rope.

The energy of the NCS has a global maximum at \(\mu = \nu = 0\). By minimizing this maximal energy with respect to the functions \(f\) and \(h\), one finds the coupled differential equations

\[
\begin{align*}
\rho f'' - f' &= \frac{1}{2} g^2 \eta^2 \rho h^2 (f - 1), \\
\rho^2 h'' + \rho h' &= h (1 - f)^2 + 2\lambda \eta^2 \rho^2 h (h^2 - 1),
\end{align*}
\]

for (4.25a), (4.25b).
where the prime indicates a derivative with respect to \( \rho \). The same differential
equations, with boundary conditions (4.23), hold for the so-called Z-string [22, 23, 24],
which excites the Z boson and Higgs fields of the electroweak Standard Model.

The Z-string is thus the sphaleron on the NCS given by Eqs. (4.21)–(4.24); see
Fig. 5. This particular sphere (balloon) in configuration space will be discussed
further in Section VI C. Note, finally, that the configurations of the NCS can also
be embedded in the full \( SU(2) \times U(1) \) gauge theory of the electroweak Standard
Model; see Ref. [7] for details and numerical results.

V. SPECTRAL FLOW

The classical field configurations of the previous section may serve as background
fields for massless Dirac fermions, whose left-handed components form an isodou-
blet under the \( SU(2) \) gauge group and whose right-handed components are gauge
singlets. The Dirac equation for the spinor \( \Psi(x) \) reads in this case

\[
\left( \gamma^\mu D_{\mu} - k (\Phi_M P_L + \Phi_M P_R) \right) \Psi = 0, \tag{5.1}
\]

with the Yukawa coupling constant \( k \), the Feynman slash notation \( D_{\mu} \equiv \gamma^\mu D_{\mu} \), and
the covariant derivative

\[
D_{\mu} \Psi(x) \equiv [\partial_{\mu} + g A_{\mu}(x) P_L] \Psi(x), \tag{5.2}
\]

which shows that only the left-handed fermions interact with the \( SU(2) \) gauge field.

The left- and right-handed projectors are, as usual, defined by \( P_L \equiv \frac{1}{2} (I - \gamma_5) \)
and \( P_R \equiv \frac{1}{2} (I + \gamma_5) \). With the Minkowski metric of Section III the Dirac matrices
obey the following Clifford algebra and Hermiticity conditions:

\[
\gamma^{\mu} \gamma^{\nu} + \gamma^{\nu} \gamma^{\mu} = 2 g^{\mu\nu}, \quad \gamma^{\mu \dagger} = \gamma^{0} \gamma^{\mu} \gamma^{0}, \quad \gamma_5 \equiv 1 \gamma^0 \gamma^1 \gamma^2 \gamma^3 = \gamma_5^\dagger. \tag{5.3}
\]

(For the Euclidean metric, all Dirac matrices are chosen Hermitian, \( \gamma^{\mu \dagger} = \gamma^{\mu} \).) The spacetime manifold considered in this article is flat and there is no need to use the vierbeins (tetrads) explicitly.

The matrix \( \Phi_M \) in (5.1) contains the two Higgs field components \( \Phi_1 \) and \( \Phi_2 \),

\[
\Phi_M = \begin{pmatrix} \Phi_2 & \Phi_1 \\ -\Phi_1^* & \Phi_2 \end{pmatrix}, \tag{5.4}
\]

so that

\[
\Phi_M \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \Phi. \tag{5.5}
\]
For the Higgs vacuum with \( A_\mu(x) = 0 \) and \( \Phi_M(x) = \text{diag}(\eta, \eta) \), the effective fermion mass is given by \( m = k \eta \).

The model considered may serve as the starting point for a consistently renormalized quantum field theory with gauge group \( SU(2) \times U(1) \) if we include three colors of left-handed quark isodoublets for each left-handed lepton isodoublet, so that the perturbative gauge anomalies \(^{25, 26, 27, 28}\) cancel between the quarks and the leptons \(^{29, 30}\). (A similar cancellation occurs for the nonperturbative \( SU(2) \) anomaly \(^{31}\) to be discussed in Section VI B.) But, for our purpose, it suffices to consider a single isodoublet of left-handed fermions, since the fermion isodoublets of the full theory behave identically.

The time-dependent solutions of the Dirac equation (5.1) are, however, not our main interest. Rather, we are interested in the eigenvalues \( E \) of the corresponding Dirac Hamiltonian,

\[
H = -\gamma^0 \gamma^m D_m + k \gamma^0 \left( \Phi^*_M P_L + \Phi_M P_R \right),
\]

(5.6)

where use has been made of the fact that \( A_0 = 0 \) for our gauge field configurations and the covariant derivative \( D_m \), for \( m = 1, 2, 3 \), has already been given in Eq. (5.2). The eigenvalues \( E \) are real, since the Dirac Hamiltonian \( H \) is Hermitian.

Now consider periodic one-parameter families (loops) of static background fields. The spectral flow invariant \(^{32}\) is then defined as the number of eigenvalues that cross zero from below minus the number of eigenvalues that cross zero from above as the loop parameter varies over its range (in a prescribed direction). See, e.g., Ref. \(^{33}\) for an elementary introduction to the concept of spectral flow.

Even if the spectral flow invariant vanishes, there may still be a nontrivial rearrangement (permutation) of the energy levels. We speak about “spectral flow” also in this case. (Mathematicians would perhaps say that there is no spectral flow if the spectral flow invariant is zero.) In addition, we will look for “spectral flow” in two-parameter families of background fields (which may be characterized by a different topological invariant).

**A. Spectral flow for the sphaleron S**

Consider the noncontractible loop (NCL) used in Section IV A to construct the sphaleron S, with parameter \( \mu \) running from 0 to \( \pi \). At the beginning of the loop (\( \mu = 0 \)) and at the end (\( \mu = \pi \)), the static background field (4.10) is the same vacuum configuration and the spectrum of the Dirac Hamiltonian (5.6) is purely continuous with a mass gap \( 2m \) according to the Higgs mechanism \( (m \propto \eta) \). For the
sphaleron $S$ at $\mu = \pi/2$, on the other hand, it has been shown [34, 35, 36, 37, 38] that the Dirac Hamiltonian $H$ has a single normalizable eigenfunction with eigenvalue zero.

The overall picture, starting from $\mu = 0$, is that a negative eigenvalue $E(\mu)$ of the Dirac Hamiltonian $H$ rises above the negative continuous spectrum, crosses zero when the background fields pass the sphaleron barrier ($\mu = \pi/2$), and finally reaches the positive continuous spectrum for $\mu = \pi$. See Fig. 6 for a sketch and Ref. 38 for numerical results.

The nonvanishing spectral flow over the NCL is guaranteed by the Atiyah–Singer index theorem [39], which relates the analytic index of the four-dimensional Dirac operator (the loop parameter $\mu$ playing the role of an imaginary time) to the topological charge associated with the NCL. Further details will be given in Section VI A. Here, we only remark that the NCL gauge field (4.10), defined in Minkowski space, has essentially the same topology as the BPST instanton solution of Euclidean Yang–Mills theory [40, 41].

**B. Spectral flow for the sphaleron $S^*$**

For the fermion behavior over the noncontractible sphere (NCS) through $S^*$, we need to resort to more abstract reasoning, since no complete numerical or analytic solution has been obtained up till now.

First, consider massless Dirac fermions with equal gauge couplings for the right- and left-handed components. It has then been shown that there exist two fermion zero-modes of the four-dimensional Euclidean Dirac operator $i\not{\partial}_4$, one of each chirality, if the fermions are placed in the background of the constrained instanton
I* [42, 43, 44]. (Note that a particular time slice through I* corresponds to the
three-dimensional configuration of the S* sphaleron. For practical purposes, one
may consider I* as a bound state of a BPST instanton I and an anti-instanton ¯I,
just as the sphaleron S* may be viewed as a composite of a sphaleron S and an
anti-sphaleron ¯S; see Eq. (4.18) and the lines below.)

Now the instanton I*, which depends on four Euclidean spacetime coordinates,
can be viewed as a path in configuration space which passes over the S* barrier.
(In other words, this path is homotopic to a particular closed loop on the S*-NCS
modulo gauge transformations; cf. Fig. 4.) The two zero-modes of \( i \mathcal{D}_4 \) in the I*
background, being time-dependent solutions of the Dirac equation (with imaginary
time), can be calculated in the adiabatic approximation, where the state at time
t \( \to +\infty \) and negative for \( t \to -\infty \).

With left- and right-handed chiralities, there are then two energy levels \( E^{(1,2)}(t) \)
crossing zero from below (these energy levels may, of course, be degenerate). In
addition, there are two eigenvalues \( E^{(3,4)}(t) \) which cross zero from above, so that the
total spectral flow invariant is zero (note that the loop through S* over the NCS is
contractible). For these last two eigenvalues, there are no zero-modes of \( i \mathcal{D}_4 \) because
the corresponding four-dimensional wave functions are not normalizable. Thus we
have two pairs of levels which cross at zero energy, one left-handed pair and one
right-handed pair. Returning to the Dirac Hamiltonian (5.6) with only left-handed
fermions interacting with the SU(2) gauge fields, we have the spectral flow of the
eigenvalues \( E^{(1)}(t) \) and \( E^{(3)}(t) \) as shown in Fig. 7.

Recently, numerical results [45] have been obtained for the eigenvalues of the \( i \mathcal{D}_4 \)
operator along a particular path over the I* barrier. It would be interesting to use
similar methods to calculate the spectral flow related to S* and also to consider
fermion representations other than isodoublet.

C. Spectral flow for the Z-string

Finally, we turn to the fermion behavior over the noncontractible sphere (NCS)
with the Z-string at the top [46]. First, we choose a path on the NCS [41, 21]–(1, 21),
which starts in the vacuum, passes over the Z-string and ends up in the vacuum.
FIG. 7: Spectral flow for a path through the sphaleron $S^*$ (cf. Fig. 4).

FIG. 8: Spectral flow over the noncontractible sphere (NCS) through the $Z$-string (cf. Fig. 5). In the picture on the left, the NCS parameter $\nu$ is held fixed at the value 0 and $\mu$ is varied. In the picture on the right, the Dirac eigenvalues are shown for the NCS patch $|\mu|, |\nu| \leq \pi/2$.

To be concrete, we put $\nu = 0$ in (4.21) and let $\mu$ run from $-\pi/2$ to $+\pi/2$. Such a loop is contractible and there is no net spectral flow to be expected (just as for the loop through $S^*$ considered in the previous subsection). What happens instead is that one negative eigenvalue $E^{(1)}(\mu, 0)$ is raised from the negative continuous spectrum and one positive eigenvalue $E^{(2)}(\mu, 0)$ is lowered from the positive continuous spectrum. Both eigenvalues meet at energy zero when the background fields pass the $Z$-string configuration ($\mu = \nu = 0$), cross and reach the opposite region of continuous eigenvalues (see the picture on the left in Fig. 8). The fermion zero-modes of the $Z$-string have been studied in Refs. [47, 48].

We can also consider the behavior of the Dirac eigenvalues over the whole two-
parameter family \(122\). Plotted over the \((\mu, \nu)\)-square, the eigenvalues \(E(\mu, \nu)\) form a double cone meeting at \(\mu = \nu = 0\) (see the picture on the right in Fig. 8).

VI. ANOMALIES

In this section, we review the relation between the sphalerons presented in Section IV and so-called anomalies. The connection between sphalerons and anomalies is precisely the spectral flow discussed in Section V.

A. Chiral anomaly and fermion number violation

The chiral \(U(1)\) anomaly, which turns out to be related to the sphaleron \(S\), eliminates a rigid \(U(1)\) symmetry of the classical action, viz. chiral invariance. This anomaly can be found in theories with massless fermions, for which there is a classical Ward identity

\[
\sum_f \left( \frac{\delta \Gamma_{cl}}{\delta \Psi_f} \delta \Psi_f - \delta \bar{\Psi}_f \frac{\delta \Gamma_{cl}}{\delta \bar{\Psi}_f} \right) = \partial^\mu j_5^\mu, \tag{6.1}
\]

where \(\Gamma_{cl}\) is the classical action and \(\Psi_f(x)\) denotes a Dirac fermion field, with \(f\) labeling the different flavors (fermion species). The rigid chiral transformation of the fermion fields is given by

\[
\delta \Psi_f(x) = \gamma_5 \Psi_f(x), \quad \delta \bar{\Psi}_f(x) = \bar{\Psi}_f(x) \gamma_5. \tag{6.2}
\]

Since the left-hand side of (6.1) vanishes for solutions of the classical equations of motion, the current \(j_5^\mu(x) \equiv \sum_f \bar{\Psi}_f(x) \gamma_5 \gamma_\mu \Psi_f(x)\) is conserved classically. This implies that the chiral charge \(Q_5 \equiv \int d^3x j_5^0(x)\) does not change with time \((t \equiv x^0)\),

\[
\left. \frac{dQ_5}{dt} \right|_{\text{classical}} = 0. \tag{6.3}
\]

Now suppose that the \(SU(2)\) gauge field couples equally to left- and right-handed fermions in the fundamental representation (as is the case for the \(SU(3)\) gauge field which is believed to be responsible for quark confinement in the Standard Model). Then the spectral flow for a path over the sphaleron \(S\) with unit winding number is as shown in Fig. 9 for each isodoublet of fermions a left-handed state crosses zero from below and a right-handed one crosses zero from above. (Essentially the
FIG. 9: Opposite spectral flow for left- and right-handed fermions, which leads to the creation of two units of chiral charge. Filled states are drawn black, empty states gray.

The same type of spectral flow has been found \cite{49} in the Schwinger model, i.e., two-dimensional quantum electrodynamics with a massless Dirac fermion. In the Dirac-sea picture of the second-quantized vacuum, this means that a pair of fermions is created from an initial vacuum state, namely one chiral fermion and one chiral antiparticle corresponding to a hole in the Dirac sea of antichiral negative-energy states. Hence, the total chiral charge $Q_5$ changes by two units per isodoublet, which contradicts the classical conservation equation (6.3).

This result is supported by the Atiyah–Singer index theorem \cite{39} for the four-dimensional chiral Dirac operator (see, e.g., Refs. \cite{50, 51, 52}). For $N$ isodoublets, the relation between the change of chiral charge and the appropriate characteristic of the background gauge field is simply the integrated version of the perturbative Ward identity for the chiral current containing the Adler–Bell–Jackiw anomaly \cite{25, 26},

$$
\sum_{f=1}^{N} \left( \frac{\delta \Gamma}{\delta \Psi_f} \delta \Psi_f - \frac{\delta \Psi_f}{\delta \Psi_f} \frac{\delta \Gamma}{\delta \Psi_f} \right) = \left[ \partial^\mu j_5^{\mu} \right] \Gamma + \frac{g^2 N}{8 \pi^2} \left[ \text{tr} \tilde{F}_{\mu \nu} F^{\mu \nu} \right] \Gamma ,
$$

where $\Gamma$ is now the fully quantized vertex functional and the bullet denotes an operator insertion.

The anomalous term in Eq. (6.4) includes the Pontryagin density

$$
q(x) \equiv - \frac{g^2}{16 \pi^2} \text{tr} \tilde{F}_{\mu \nu}(x) F^{\mu \nu}(x) ,
$$

with $\tilde{F}_{\mu \nu} \equiv (1/2) \epsilon_{\mu \nu \rho \sigma} F^{\rho \sigma}$. The integral of this density over the spacetime manifold $M$ is a topological invariant called the Pontryagin index,

$$
Q \equiv \int_M d^4 x \, q(x) .
$$
FIG. 10: Spectral flow in the electroweak Standard Model, where only left-handed fermions interact with the $SU(2)$ gauge fields. $S$ denotes the sphaleron, which has a single fermion zero-mode. The spectral flow leads to a change of fermion number between initial and final states (see text).

For compact spacetime manifolds $M$, the Pontryagin index is an integer number and is also called the winding number or topological “charge” (hence, the notation $Q$).

Next, turn to the simplified version of the electroweak Standard Model, as described in Sections II and V. Here, the fermion fields are fundamentally massless, even though they behave as massive particles in the Higgs vacuum. More importantly, the gauge field now couples only to the left-handed parts of the fermion fields; cf. Eqs. (5.1) and (5.2). Hence, the spectral flow for a single fermion flavor is made up of only one state which crosses zero from below. This implies that fermion number conservation is violated [53, 54]. See Fig. 10 and compare with Fig. 9. (It is, of course, important to define carefully what is meant by “the fermion number” of a given state [55, 56, 57]; see also the discussion in the last three paragraphs of this subsection.)

The map $U$ given in Section IV A essentially provides a map $S^3 \to S^3$, characterized by the topological charge $|Q| = 1$. The above considerations can be generalized to other (integer) values of $Q$ and to a model with $N_{\text{fam}}$ families of quarks and leptons. The sum of baryon number $B$ and lepton number $L$ is then found to be nonconserved [53],

$$\Delta(B - L) = 0, \quad \Delta(B + L) = 2 N_{\text{fam}} Q,$$

where $\Delta B$ denotes the change of baryon number between initial and final states and similarly for $\Delta L$.

As explained at the end of Section IV A, the NCL can be transformed into a path connecting two topologically distinct vacua in one-point-compactified three-space.
FIG. 11: Potential energy over a slice of configuration space, parametrized by the Chern–Simons number \( N_{CS} \). The height of the energy barrier between topologically different vacua \( (N_{CS} = n \in \mathbb{Z}) \) is set by the sphaleron \( S \), which appears in different gauge copies \( (N_{CS} = 1/2 + n, \text{ for } n \in \mathbb{Z}) \). This figure essentially “unwraps” the loop of Fig. 3.

The general form of such a vacuum is given by a static pure-gauge configuration,

\[
g A = -d\chi \chi^{-1}, \quad \Phi = \eta \chi \left( \begin{array}{c} 0 \\ 1 \end{array} \right),
\]

(6.8)

for a map \( \chi : \mathbb{R}^3 \to SU(2) \) which approaches \( 1_2 \) at spatial infinity. The homotopy class to which \( \chi \) belongs is characterized by the integer Chern–Simons number

\[
N_{CS}[\chi] = -\frac{1}{24 \pi^2} \int d^3 x \epsilon^{k l m} \text{tr} \left\{ (\partial_k \chi \chi^{-1})(\partial_l \chi \chi^{-1})(\partial_m \chi \chi^{-1}) \right\}.
\]

(6.9)

The topological charge \( Q \) of the map \( \omega(\mu, r, \theta, \phi) \), as discussed in the last two paragraphs of Section IV A, is then the difference of the Chern–Simons numbers of the vacua at the start and end of the associated path,

\[
Q = \Delta N_{CS} \equiv N_{CS}[\omega(\pi, r, \theta, \phi)] - N_{CS}[\omega(0, r, \theta, \phi)].
\]

(6.10)

Of course, it is also possible to map the \( \mu \)-interval \([0, \pi]\) on the time interval \([-\infty, +\infty]\).

The sphaleron now corresponds to an energy barrier between these vacua, as sketched in Fig. 11. The transition between different vacua can, for example, take place by tunneling through the sphaleron barrier \([53, 54]\) or by passing over the barrier due to a thermal fluctuation of the fields \([5, 58]\). Especially the latter mechanism is expected to contribute significantly to fermion-number-violating processes in the early universe (see, e.g., Refs. \([9, 10]\)).

The rate of fermion-number-violating processes at relatively low energies \( (E \ll E_S \approx 10 \text{ TeV}) \) can be calculated from the Euclidean path integral \([54, 59]\). But for a reliable discussion of these processes at high energies \( (E \gtrsim E_S) \) it is advisable to
remain in Minkowski spacetime. The problem, then, is that the compactification of four-space which we used as the starting point of our topological considerations is not really physically sensible for Minkowski spacetime. The topological charge $Q$, in particular, need not be an integer quantity in Minkowski spacetime. The crucial point here is the role of *energy conservation* for background fields that solve the equations of motion; see, e.g., Ref. [60]. The general question of which type of gauge field leads to nontrivial spectral flow remains unanswered for the moment.

There exists, however, a result for *strongly dissipative* $SU(2)$ gauge fields [55, 56]. In this case, the spectral flow is given by the difference in winding numbers of the asymptotic vacuum configurations for $t \to \pm \infty$. Roughly speaking, this coincides with the previous result in Euclidean spacetime, namely Eq. (6.10) inserted into Eq. (6.7).

For the case of spherically symmetric fields, there is also a result for *generic* (i.e., nondissipative) gauge fields,

$$\Delta(B - L) = 0, \quad \Delta(B + L) = 2N_{\text{fam}} (\Delta N_{\text{winding}} + \Delta N_{\text{twist}}) \bigg|_{\text{spher. symm.}}. \quad (6.11)$$

The change of $B + L$ is now determined by two integers. The first, $\Delta N_{\text{winding}}$, again corresponds to Eq. (6.10). But the second, $\Delta N_{\text{twist}}$, is an entirely new characteristic of spherically symmetric $SU(2)$ gauge fields, which is related to the asymptotic behavior of the solutions of a (nonlinear) Riccati equation embedded in the (linear) zero-energy Dirac equation [61]. The integer $\Delta N_{\text{twist}}$ is zero for strongly dissipative $SU(2)$ gauge fields. See Ref. [62] for further discussion of the issues involved.

### B. Witten’s global gauge anomaly

The global $SU(2)$ gauge anomaly, which turns out to be related to the sphaleron $S^*$, differs from the case discussed in the previous subsection in that not just a symmetry of the theory is eliminated but the theory itself.

As mentioned in Section 6.B, the crossing of energy levels for paths over the $S^*$ barrier is related to the existence of two normalizable zero-modes of the four-dimensional Euclidean Dirac operator $\slashed{D}_4$, one of each chirality. The noncontractible sphere of three-dimensional configurations can also be viewed as a noncontractible loop of four-dimensional configurations. Furthermore, as explained at the end of Section 6.A, it is possible to pass from a *loop* of gauge field configurations in the radial gauge to a *path* of gauge field configurations without the radial gauge condition. The resulting path has topologically inequivalent vacua at the start and at the end.
FIG. 12: Left: sketch of the eigenvalues of the Dirac Hamiltonian for a particular path over the $S^*$ noncontractible sphere (the theory considered has both left- and right-handed fermions interacting with the $SU(2)$ gauge fields, which is not the case in Fig. 7). Middle: eigenvalues of $\mathcal{D}_4$ for the corresponding noncontractible loop through the constrained instanton $I^*$. Right: eigenvalues of $\mathcal{D}_5$ for the corresponding five-dimensional configuration. Each path with an energy level crossing zero from below (thick line) corresponds to a normalizable fermion zero-mode one dimension higher. The sphaleron $S^*$ has four fermion zero-modes, the constrained instanton $I^*$ two, and the five-dimensional configuration one.

Now consider the change of eigenvalues of $\mathcal{D}_4$ along such a path. Since for one "point" of the path (i.e., the $I^*$-configuration) there are known zero-modes [44], it is to be expected that some level crossing is occurring also here.

That this is indeed the case has been shown in Ref. [31] by increasing the dimension once more. The one-parameter family of four-dimensional Dirac operators can also be considered as a single five-dimensional one. (In other words, the whole NCS serves as a single background configuration.) It then follows from the so-called mod–2 Atiyah–Singer index theorem [63] that the corresponding five-dimensional Dirac operator has a normalizable zero-mode. For $\mathcal{D}_4$, this implies that an eigenvalue is crossing zero from negative to positive values as the path is traversed. Simultaneously, there is a second eigenvalue which passes from positive to negative values. This discussion is summarized in Fig. 12, which also gives the corresponding spectral flow in three dimensions. (The mod–2 index theorem guarantees only an odd number of zero-modes for the five-dimensional configuration, but for simplicity we have assumed there is just one. See Ref. [45] for numerical results and further discussion.)

Witten also argued that the spectral flow of $\mathcal{D}_4$ leads to a global gauge anomaly [31]. In the Euclidean path integral of $SU(2)$ Yang–Mills theory with a single isodoublet of Weyl fermions, there effectively appears a square root of the Dirac
determinant,

\[ \int DA_\mu \ \sqrt{\det i D_\mu} \ \exp \left( \frac{1}{2} \int d^4x \ tr F_{\mu\nu} F^{\mu\nu} \right), \]  

(6.12)

if one recalls that two Weyl fermions of opposite chiralities make a single Dirac fermion.

The Dirac determinant in Eq. (6.12) depends on the background gauge fields \( A_\mu(x) \) and its square root can be defined as the product of the positive eigenvalues [starting from a given gauge field configuration, say \( A_\mu(x) = 0 \)]. The above considerations then show that for a particular continuous variation of the gauge fields we end up with gauge fields, which are related to the starting configuration by a large gauge transformation and which have a \( \sqrt{\det i D_\mu} \) of opposite sign (one positive eigenvalue having become negative; cf. the middle picture of Fig. 12). In the path integral, one has to integrate over all gauge fields (taking out the infinite factor due to gauge invariance afterwards). Hence, for every contribution \( \sqrt{\det i D_\mu} \) there is also a contribution \( -\sqrt{\det i D_\mu} \) arising from the gauge-transformed background fields. This implies that the path integral (6.12) vanishes.

More precisely, the path integral over the gauge fields is not well defined, because there is no satisfactory way to restrict the integration over the gauge fields so that a single Weyl isodoublet gives a continuous gauge-invariant contribution. This, then, is the Witten anomaly, which can also be proven without the mod–2 Atiyah–Singer index theorem but with the perturbative Bardeen anomaly instead [64, 65].

C. Z-string global gauge anomaly

Just as for the Witten anomaly and the \( S^\ast \) sphaleron of the previous subsection, there exists a global gauge anomaly related to the Z-string sphaleron [46]. In order to explain this anomaly, we need a modified noncontractible sphere (NCS), obtained by continuous deformation of the balloon as given in Section IV C. This modification has the advantage of being a real sphere, that is, without degenerate points. The modified NCS still has one point corresponding to the vacuum and one point corresponding to the Z-string (see the picture in the middle of Fig. 13).

For the modified NCS, the \( z \)-independent \( SU(2) \) gauge field is in the polar gauge \( A_\rho = 0 \). But like the case of the sphalerons \( S \) and \( S^\ast \), it is possible to relax the polar gauge condition and to demand instead that the vacuum reached for \( \rho \rightarrow \infty \) is the trivial one. Then one ends up with a disc of configurations with a loop of pure-gauge configurations on the boundary (see the picture on the right of Fig. 13).
Considering the compactified radial coordinate $\rho$ to be a polar angle $\theta$, the fields are effectively defined on a sphere $S^2$. The loop of vacuum configurations on this two-sphere, restricted to the smash product $S^1 \wedge S^2$, corresponds to a nontrivial element of the homotopy group $\pi_3(S^3)$.

We keep this in mind for later and turn to the eigenvalues of the four-dimensional Euclidean Dirac operator $i\mathcal{D}_4$, where the time-dependent background fields are taken to be paths over the $Z$-disc, with the start and end point (not necessarily the same) lying on the rim of vacuum configurations. For any such path passing through the $Z$-string, we know from Section \ref{sec:Z-String} that $i\mathcal{D}_4$ has a single normalizable zero-mode corresponding to the eigenvalue of the Dirac Hamiltonian which crosses zero from below.

Now consider a particular family of operators $i\mathcal{D}_4$ corresponding to a family of paths over the $Z$-disc, which starts from a constant path corresponding to a point on the rim of the disc, passes through a path via the $Z$-string, and ends up in a pure vacuum path formed by the boundary of the disc (see Fig. \ref{fig:Z-Disc} where the $Z$-disc of Fig. \ref{fig:Z-Balloon} has been flattened). This family of four-dimensional Dirac operators sweeps over the whole $Z$-disc and we expect that there is spectral flow corresponding to the winding number of the underlying map $S^3 \to S^3$. In our case, this means that a single eigenvalue of $i\mathcal{D}_4$ crosses zero. The zero crossing can be expected to occur for the path labeled (3) in Fig. \ref{fig:Z-Disc}.

Like the case of the Witten anomaly in Section \ref{sec:Witten-Anomaly}, this prevents us from defining the square root of the fermion determinant as a continuous gauge-invariant function of the bosonic background fields. Note that our path of four-dimensional configurations begins in a time-independent, topologically trivial vacuum and ends up in a gauge-transformed, time-dependent and topologically nontrivial one.

It is particularly interesting to see how this global gauge anomaly manifests itself in the space of time-independent fermion states. Since the Dirac Hamiltonian is a

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Z-Balloon.png}
\caption{Z-balloon, Z-sphere and Z-disc.}
\end{figure}
FIG. 14: Five closed loops on the \( Z \)-disc (cf. Fig. 13). Loop (1) corresponds to a single point on the boundary, loop (3) passes through the \( Z \)-string, and loop (5) consists of the whole boundary of the disc.

A real Hermitian operator, we may choose real energy-eigenstates. Concretely, look at the one-dimensional subspace spanned by the eigenstate which crosses zero from below in the picture on the left in Fig. 8. For the background fields, we use an arbitrary loop on the \( Z \)-disc, which circumnavigates the \( Z \)-string exactly once and which is parametrized by \( \alpha \in [0, 2\pi] \). Then the energy eigenstate defines a real line bundle over \( S^1 \).

It has been shown in Ref. 46 that this bundle is, in fact, the Möbius bundle. A normalized eigenstate \( |\Phi(0)\rangle \) transported around the loop ends up as \( |\Phi(2\pi)\rangle = -|\Phi(0)\rangle \); see Fig. 15. The phase factor found is determined by the Berry phase for adiabatic transport 66. The Berry phase \( \pi \) (mod \( 2\pi \)) is of topological origin and does not change under continuous deformation of the loop, as long as the loop of configurations does not touch the fermion degeneracy “point” corresponding to the \( Z \)-string. This observation also shows that the boundary of the \( Z \)-disc (Fig. 13) is a noncontractible loop of vacuum configurations, since the real Berry phase factor \( -1 \) on it cannot be continuously changed to \( +1 \).

The variation of the eigenstate along the rim of the \( Z \)-disc defines a projective action of the gauge group on the fermionic matter. There is then a global gauge anomaly, because it is impossible to define a real, continuous, and proper (i.e., non-projective) representation of the local gauge group on the fermion states. Since the vacuum of quantum field theory is the Dirac sea with all negative-energy eigenstates filled, this also means that the second-quantized vacuum state acquires the Berry phase factor \( -1 \). See Section 6 of Ref. 46 for further discussion. (We take the opportunity to correct a slip of the pen. In the last sentence of Footnote 6 in Ref. 46, the words “and vice versa” must be deleted.)

A similar interpretation of the Witten anomaly in terms of a Berry phase has been
FIG. 15: Möbius bundle structure of the gauge orbit, with loop parameter $\alpha \in [0, 2\pi]$. The line represents a normalized real eigenstate $|\Phi(\alpha)\rangle$ of the Dirac Hamiltonian, with $|\Phi(2\pi)\rangle = -|\Phi(0)\rangle$.

given in Ref. [67]. There is, however, a crucial difference between the $Z$-string global gauge anomaly and the Witten anomaly. For the $Z$-string anomaly, namely, there does exist a local counterterm in the action which restores gauge invariance, but at the price of violating Lorentz and CPT invariance [68]. More generally, if gauge invariance is enforced, there appears a new anomaly, the so-called CPT anomaly (see Refs. [69, 70] for the main result and Ref. [71] for a review).

VII. CONCLUSION

The space of finite-energy three-dimensional field configurations of $SU(2)$ Yang–Mills–Higgs theory (in short, configuration space) has nontrivial topology [3, 4], which leads to the existence of a new type of classical solutions, the so-called sphalerons. Sphalerons are unstable static finite-energy solutions of the classical field equations, whereas solitons are stable solutions.

In Section VI, we have explained the topology behind the $S$, $S^*$, and $Z$-string sphalerons [5, 6, 7] of the $SU(2)$ Yang–Mills–Higgs theory (2.1). Precisely this theory appears in the electroweak Standard Model of elementary particle interactions [8]. Knowledge of these classical solutions may, therefore, be of great importance to physics.

Adding chiral fermions to the $SU(2)$ Yang–Mills–Higgs theory, the nontrivial topology of configuration space makes itself felt by the occurrence of spectral flow [32], as discussed in Section V. In turn, the general phenomenon of spectral flow is related to the possible existence of anomalies which invalidate certain properties of the classical theory, as discussed in Section VI.

The spectral flow over a noncontractible loop through the sphaleron $S$ is related
to the chiral $U(1)$ anomaly $^{25,26}$, which corresponds to a breakdown of baryon and lepton number conservation in the electroweak Standard Model $^{53}$. The spectral flow through the $S^*$ and $Z$-string sphalerons does not lead to a global $SU(2)$ gauge anomaly $^{31,46}$, because the electroweak Standard Model has an even number of chiral isodoublets. Still, there is nontrivial spectral flow (more precisely, spectral rearrangement) over configuration space, but its physical implications remain to be clarified (cf. Refs. $^{44,69}$). Indeed, we need a better understanding of the role of configuration space topology in concrete physical problems, such as the behavior of elementary particle fields at high energies or temperatures.


J.S. Bell and R. Jackiw, “A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the sigma model,” Nuovo Cimento A 60 (1969) 47.


