HUTP-78/A004

THE USES OF INSTANTONS

Sidney Coleman*

Lyman Laboratory of Physics
Harvard University
Cambridge, Massachusetts 02138

Lectures delivered at the 1977 International School of Subnuclear Physics, *Ettore Majorana*

*Research supported in part by the National Science Foundation under Grant No. PHY77-22864.

2/78

TABLE OF CONTENTS

THE USES OF INSTANTONS

Sidney Coleman

Lyman Laboratory of Physics, Harvard University

Cambridge, Massachusetts 02138

I. INTRODUCTION

In the last two years there have been astonishing developments in quantum field theory. We have obtained control over problems previously believed to be of insuperable difficulty and we have obtained deep and surprising (at least to me) insights into the structure of the leading candidate for the field theory of the strong interactions, quantum chromodynamics. These goodies have come from a family of computational methods that are the subject of these lectures.

These methods are all based on semiclassical approximations, and, before I can go further, I must tell you what this means in the context of quantum field theory.

To be definite, let us consider the theory of a single scalar field in four-dimensional Minkowski space, with dynamics defined by the Lagrangian density

$$\mathcal{L} = \tfrac{1}{2}\,\partial_\mu \phi\, \partial^\mu \phi - \tfrac{1}{2} m^2 \phi^2 - g^2 \phi^4 . \tag{1.1}$$

For classical physics, g is an irrelevant parameter. The easiest way to see this is to define

$$\phi' = g\phi . \tag{1.2}$$

In terms of ϕ',

$$\mathcal{L} = \frac{1}{g^2}\left(\tfrac{1}{2}\,\partial_\mu \phi'\, \partial^\mu \phi' - \tfrac{1}{2} m^2 \phi'^2 - \phi'^4\right) . \tag{1.3}$$

Thus, g does not appear in the field equations; if one can solve the theory for any positive g, one can solve it for any other positive g; g is irrelevant. Another way of seeing the same thing is to observe that, in classical physics, g is a dimensionful parameter and can always be scaled to one.

Of course, g *is* relevant in quantum physics. The reason is that quantum physics contains a new constant, $\hbar$, and the important object (for example, in Feynman's path-integral formula) is

$$\mathcal{L}/\hbar = \frac{1}{g^2\hbar}\left(\tfrac{1}{2}\,\partial_\mu \phi'\, \partial^\mu \phi' + \ldots\right) . \tag{1.4}$$

As we see from thie expression, the relevant (dimensionless) parameter is $g^2\hbar$, and thus, semiclassical approximations, small-$\hbar$ approximations, are tantamount to weak-coupling approximations, small-g approximations.

At this point you must be puzzled by the trumpets and banners of my opening paragraph. Do we not have a perfectly adequate small-coupling approximation in perturbation theory? No, we do not; there is a host of interesting phenomena which occur for small coupling constant and for which perturbation theory is inadequate.

The easiest way to see this is to descend from field theory to particle mechanics. Consider the theory of a particle of unit mass moving in a one-dimensional potential,

$$L = \tfrac{1}{2}\dot{x}^2 - V(x;g) , \tag{1.5}$$

where

$$V(x;g) = \frac{1}{g^2} F(g\,x) , \tag{1.6}$$

and F is some function whose Taylor expansion begins with terms of order x^2. Everything I have said about the field theory defined by Eq. (1.1) goes through for this theory. However, let us consider the phenomenon of transmission through a potential barrier (Fig. 1). Every child knows that the amplitude for transmission

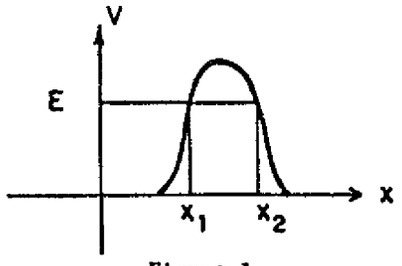


Figure 1

obeys the WKB formula,

$$|T(E)| = e^{-\frac{1}{\hbar}\int_{x_1}^{x_2} dx\sqrt{2(V-E)}}\,[1 + O(\hbar)]\ , \qquad (1.7)$$

where x_1 and x_2 are the classical turning points at energy E. This is a semiclassical approximation. Nevertheless, transmission, barrier penetration, is not seen in any order of perturbation theory, because Eq. (1.7) vanishes more rapidly than any power of $\hbar$, and therefore of g.

I can now make my first paragraph more explicit. There are phenomena in quantum field theory, and in particular in quantum chromodynamics, analogous to barrier penetration in quantum particle mechanics. In the last two years a method has been developed for handling these phenomena. This method is the subject of these lectures.

The organization of these lectures is as follows: In Section 2 I describe the new method in the context of particle mechanics, where we already know the answer by an old method (the WKB approximation). Here the instantons which play a central role in the new method and which have given these lectures their title first appear. In Section 3 I derive some interesting properties of gauge field theories. In Section 4 I discuss a two-dimensional model in which instantons lead to something like quark confinement and explain why a similar mechanism has (unfortunately) no chance of working in four dimensions. In Section 5 I explain 't Hooft's resolution of the U(1) problem. In Section 6 I apply instanton methods to vacuum decay. Only this last section reports on my own research; all the rest is the work of other hands.[1]

I thank C. Callan, R. Dashen, D. Gross, R. Jackiw, M. Peskin, C. Rebbi, G. 't Hooft, and E. Witten for patiently explaining large portions of this subject to me. Although I have never met A. M. Polyakov, his influence pervades these lectures, as it does the whole subject.[2]

A Note on Notation: In these lectures we will work in both Minkowski space and in four-dimensional Euclidean space. A point in Minkowski space is labeled x^μ, where $\mu = 0, 1, 2, 3$, and x^0 is the time coordinate. In Minkowski space I will distinguish between covariant and contravariant vectors, $x_\mu = g_{\mu\nu}x^\nu$, where the metric tensor has signature (+---). Euclidean space is obtained from Minkowski space by formal analytic continuation in the time coordinate, $x^4 = -ix^0$. A point in Euclidean space is labeled x^μ, where $\mu = 1, 2, 3, 4$. The signature of the metric tensor is (++++). Thus covariant and contravariant vectors are component-by-component identical, and I will not bother to distinguish between them. Note that x·y in Minkowski space continues to -x·y in Euclidean space. The Euclidean action is defined as -i times the continuation of the Minkowskian action. When discussing particle problems, I will use t for both Euclidean and Minkowskian time; which is meant will always be clear from the context. In Section 2 explicit factors of $\hbar$ are retained; elsewhere, $\hbar$ is set equal to one.

II. INSTANTONS AND BOUNCES IN PARTICLE MECHANICS

2.1 Euclidean Functional Integrals

In this section we will deal exclusively with the theory of a spinless particle of unit mass moving in a potential in one dimension:

$$H = \frac{p^2}{2} + V(x)\ , \qquad (2.1)$$

We will rederive some familiar properties of this much-studied system by unfamiliar methods. For the problem at hand, these methods are *much* more awkward than the standard methods of one-dimensional

quantum mechanics; however, they have the great advantage of being immediately generalizable to quantum field theory.

Our fundamental tool will be the Euclidean (imaginary time) version of Feynman's[3] sum over histories:

$$\langle x_f | e^{-HT/\hbar} | x_i \rangle = N \int [dx] e^{-S/\hbar} \quad . \tag{2.2}$$

Both sides of this equation require explanation:

On the left-hand side, $|x_i\rangle$ and $|x_f\rangle$ are position eigenstates, H is the Hamiltonian, and T is a positive number. The left-hand side of Eq. (2.2) is of interest because, if we expand in a complete set of energy eigenstates,

$$H|n\rangle = E_n |n\rangle \ , \tag{2.3}$$

then

$$\langle x_f | e^{-HT/\hbar} | x_i \rangle = \sum_n e^{-E_n T/\hbar} \langle x_f | n \rangle \langle n | x_i \rangle \ . \tag{2.4}$$

Thus, the leading term in this expression for large T tells us the energy and wave-function of the lowest-lying energy eigenstate.

On the right-hand side, N is a normalization factor, S is the Euclidean action[4]

$$S = \int_{-T/2}^{T/2} dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + V \right] \ , \tag{2.5}$$

and [dx] denotes integration over all functions x(t), obeying the boundary conditions, $x(-T/2) = x_i$ and $x(T/2) = x_f$. To be more specific, if $\bar{x}$ is any function obeying the boundary condition, then a general function obeying the boundary conditions can be written as

$$x(t) = \bar{x}(t) + \sum_n c_n x_n(t) \ , \tag{2.6}$$

where the x_n's are a complete set of real orthonormal functions vanishing at the boundaries,

$$\int_{-T/2}^{T/2} dt \, x_n(t) x_m(t) = \delta_{nm} \ , \tag{2.7a}$$

$$x_n(\pm T/2) = 0 \ . \tag{2.7b}$$

Then, the measure [dx] is defined by

$$[dx] = \prod_n (2\pi\hbar)^{-\frac{1}{2}} dc_n \ . \tag{2.8}$$

(This measure differs in normalization from the measure defined by Feynman;[3] this is why we need the normalization constant N. However, as we shall see, we shall never need an explicit formula for N.)

The right-hand side of Eq. (2.2) is of interest because it can readily be evaluated in the semiclassical (small-$\hbar$) limit. In this case the functional integral is dominated by the stationary points of S. For simplicity, let us assume for the moment that there is only one such stationary point, which we denote by $\bar{x}$,

$$\frac{\delta S}{\delta \bar{x}} = -\frac{d^2 \bar{x}}{dt^2} + V'(\bar{x}) = 0 \tag{2.9}$$

where the prime denotes differentiation with respect to x. Further, let us choose the x_n's to be eigenfunctions of the second variational derivative of S at $\bar{x}$,

$$-\frac{d^2 x_n}{dt^2} + V''(\bar{x}) x_n = \lambda_n x_n \ . \tag{2.10}$$

Then, in the small-$\hbar$ limit, the integral becomes a product of Gaussians, and we find

$$\begin{aligned} \langle x_f | e^{-HT/\hbar} | x_i \rangle &= N e^{-S(\bar{x})/\hbar} \prod_n \lambda_n^{-\frac{1}{2}} \left[1 + O(\hbar) \right] \\ &= N e^{-S(\bar{x})/\hbar} \left[\det \left(-\partial_t^2 + V''(\bar{x}) \right) \right]^{-\frac{1}{2}} \left[1 + O(\hbar) \right] \end{aligned} \tag{2.11}$$

(Of course, we are tacitly assuming here that all the eigenvalues are positive. We shall shortly see what to do when this is not the case.) If there are several stationary points, in general one has to sum over all of them.

Equation (2.9) is the equation of motion for a particle of unit mass moving in a potential *minus* V. Thus,

$$E = \frac{1}{2} \left(\frac{d\bar{x}}{dt} \right)^2 - V(\bar{x}) \tag{2.12}$$

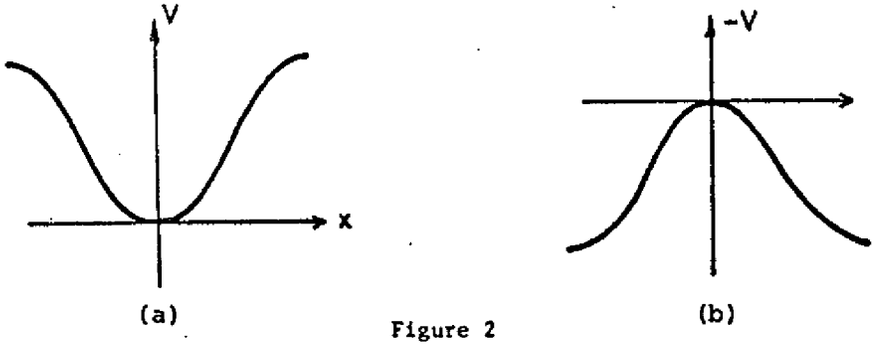


Figure 2

is a constant of the motion. This can be used to determine the qualitative features of the solutions of Eq. (2.9) by inspection.

As a simple example, consider the potential shown in Fig. 2a. Let us choose $x_i = x_f = 0$. Figure 2b shows the inverted potential, -V. It is obvious from the figure that the only solution of Eq. (2.9) which obeys the boundary conditions is

$$\bar{x} = 0 \ . \tag{2.13}$$

For this solution, S = 0. Thus, from Eq. (2.11),

$$\langle 0|e^{-HT/\hbar}|0\rangle = N[\det(-\partial_t^2 + \omega^2)]^{-\frac{1}{2}}[1 + O(\hbar)] \ , \tag{2.14}$$

where

$$\omega^2 = V''(0) \ . \tag{2.15}$$

In Appendix A, I show that, for large T,

$$N[\det(-\partial_t^2 + \omega^2)]^{-\frac{1}{2}} = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \ . \tag{2.16}$$

Thus, the ground-state energy is given by

$$E_0 = \tfrac{1}{2}\omega\hbar[1 + O(\hbar)] \ . \tag{2.17}$$

Also, the probability of the particle being at the origin when it is in its ground state is

$$|\langle x = 0|n = 0\rangle|^2 = (\omega/\pi\hbar)^{\frac{1}{2}}[1 + O(\hbar)] \ . \tag{2.18}$$

These are, of course, the correct semiclassical results. In the small-$\hbar$ limit, the particle is in a harmonic-oscillator ground-

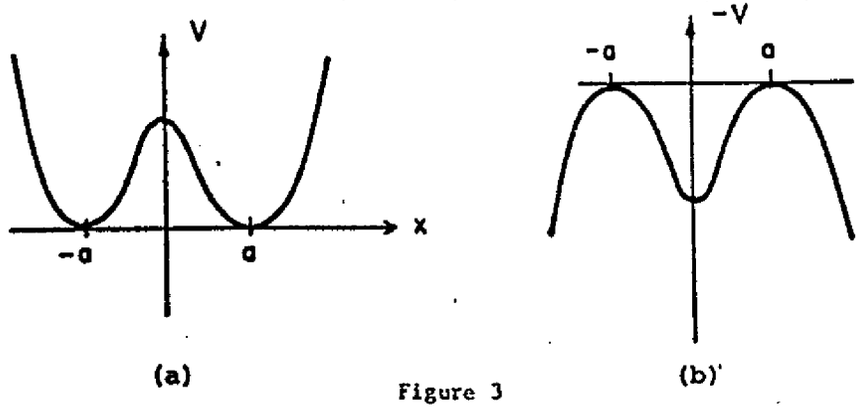


Figure 3

state concentrated at the origin and its energy is the ground-state energy of a harmonic oscillator.

2.2 The Double Well and Instantons

We now turn to a less trivial problem,[5] the double well of Fig. 3a. I will assume the potential is even, V(x) = V(-x), and will denote its minima by ±a. As before, I will add a constant to V, if necessary, to make V vanish at its minima, and I will denote $V''(\pm a)$ by ω^2.

We will attempt to compute both

$$\langle -a|e^{-HT}|-a\rangle = \langle a|e^{-HT}|a\rangle \ , \tag{2.19a}$$

and

$$\langle a|e^{-HT}|-a\rangle = \langle -a|e^{-HT}|a\rangle \ , \tag{2.19b}$$

by approximating the functional integral by its semiclassical limit, Eq. (2.11). Just as before, the first step is to find solutions of the classical Euclidean equation of motion, (2.9), consistent with our boundary conditions.

Of course, two such solutions are those in which the particle stays fixed on top of one or the other of the two hills in Fig. 3b. However, there is another potentially interesting solution, one where the particle begins at the top of one hill (say the left one) at time -T/2, and moves to the top of the right hill at time T/2.

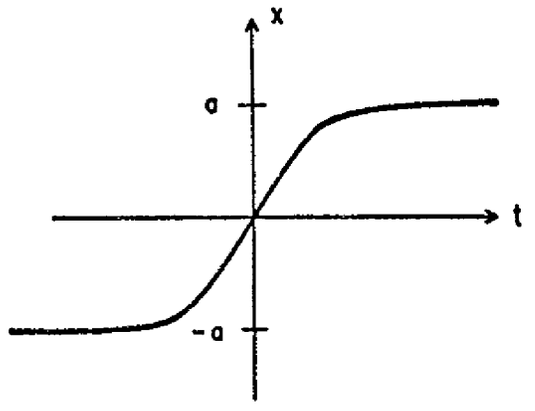


Figure 4

Since we plan eventually to take T to infinity, we will focus on the form of the solution in this limit, where the particle attains the tops of the hills at times plus and minus infinity. In this case, we are dealing with a solution of the equation of motion with vanishing E; whence

$$dx/dt = \sqrt{2V} \ . \tag{2.20}$$

Equivalently,

$$t = t_1 + \int_0^x dx'(2V)^{-\frac{1}{2}} \ , \tag{2.21}$$

where t_1 is an integration constant, the time at which x vanishes.

This solution is sketched in Fig. 4; it is called "an instanton with center at t_1". The name "instanton" was invented by 't Hooft. The idea is that these objects are very similar in their mathematical structure to what are called solitons or lumps,[6] particle-like solutions of classical field theories: thus the "-on". However, unlike lumps, they are structures in time (albeit Euclidean time): thus the "instant-". For the same reason, Polyakov suggested the name "pseudoparticle", also used in the literature.

Of course, we can also construct solutions that go from a to -a, simply by replacing t by -t in Eq. (2.21); these are called "anti-instantons".

Two properties of these solutions will be important to us:

(1) From Eq. (2.20), it is easy to derive a simple expression for S_0, the action of an instanton (or anti-instanton)

$$S_0 = \int dt[\tfrac{1}{2}(dx/dt)^2 + V] = \int dt(dx/dt)^2 = \int_{-a}^{a} dx\sqrt{2V} \ . \tag{2.22}$$

Note that this is the same as the integral that appears in the barrier-penetration formula, Eq. (1.7). We shall see shortly that this is no coincidence.

(2) For large t, x approaches a, and Eq. (2.20) can be approximated by

$$dx/dt = \omega(a - x) \ . \tag{2.23}$$

Thus, for large t,

$$(a - x) \propto e^{-\omega t} \ . \tag{2.24}$$

Thus, instantons are, roughly speaking, well-localized objects, having a size on the order of $1/\omega$.

This is of critical importance, because it means that, for large T, the instanton and the anti-instanton are not the only approximate solutions of the equation of motion; there are also approximate solutions consisting of strings of widely separated instantons and anti-instantons. (You may be troubled by the sudden appearance in the argument of approximate solutions, approximate stationary points of S. If so, bear with me; I'll give a fuller explanation of this point later.)

I shall evaluate the functional integral by summing over all such configurations, with n objects (instantons or anti-instantons) centered at $t_1 \ldots t_n$, where

$$T/2 > t_1 > t_2 \ldots > t_n > -T/2 \ . \tag{2.25}$$

Figure 5 (next page) shows one such configuration. T is assumed to be huge on the scale of the size of an instanton; thus the smooth curves of Fig. 4 appear as sharp jumps on the scale of Fig. 5. (The vertical marks on the time axis will be explained shortly.)

Now for the evaluation:

(1) For n widely separated objects, S is nS_0. This takes care of the exponential of the action.

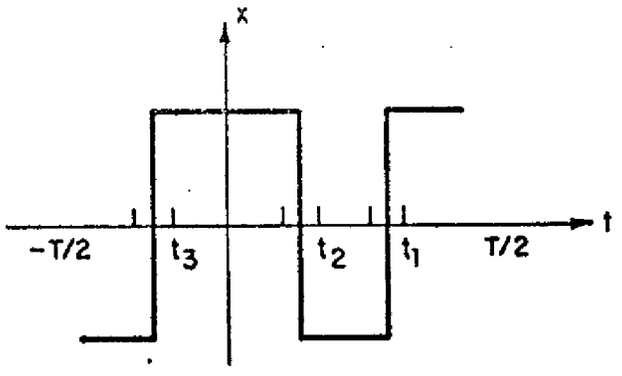


Figure 5

(2) The evaluation of the determinant is a bit trickier. Let us consider the time evolution operator, e^{-HT}, as a product of operators associated with evolution between the points indicated by the vertical marks on the time axis in Fig. 5. If it were not for the small intervals containing the instantons and anti-instantons, V'' would equal ω^2 over the entire time axis, and thus we would obtain the same result we obtained for a single-well potential in Section 2.1,

$$\left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} . \tag{2.26}$$

The small intervals containing the instantons and anti-instantons correct this formula. Thus we obtain

$$\left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} K^n , \tag{2.27}$$

where K is defined by demanding that this formula give the right answer for one instanton. Later we shall obtain a more explicit expression for K.

(3) We must integrate over the locations of the centers:

$$\int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{t} dt_2 \cdots \int_{-T/2}^{t_{n-1}} dt_n = T^n/n! \tag{2.28}$$

(4) We are not free to distribute instantons and anti-instantons arbitrarily. For example, if we start out at -a, the first object we encounter must be an instanton, the next one must be an anti-instanton, etc. Furthermore, if we are to end up back at -a, n must be even. Likewise, if we wish to end up at a, n must be odd.

Thus,

$$\langle -a|e^{-HT/\hbar}|-a\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \sum_{\text{even } n} \frac{(Ke^{-S_0/\hbar}T)^n}{n!} [1+O(\hbar)] , \tag{2.29}$$

while $\langle a|e^{-HT/\hbar}|-a\rangle$ is given by the same expression, summed over odd n's. These sums are trivial:

$$\langle \pm a|e^{-HT/\hbar}|-a\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \tfrac{1}{2}[\exp(Ke^{-S_0/\hbar}T) \mp \exp(-Ke^{-S_0/\hbar}T)] . \tag{2.30}$$

(From now on, to keep the page from getting cluttered, I will drop the factors of $[1+O(\hbar)]$; remember that they're omnipresent though unwritten.)

Comparing this to Eq. (2.4), we see that we have two low-lying energy eigenstates, with energies

$$E_\pm = \tfrac{1}{2}\hbar\omega \pm \hbar Ke^{-S_0/\hbar} . \tag{2.31}$$

If we call these eigenstates $|+\rangle$ and $|-\rangle$, we also see that

$$|\langle +|\pm a\rangle|^2 = |\langle -|\pm a\rangle|^2 = \langle a|->\langle -|-a\rangle = -\langle a|+\rangle\langle +|-a\rangle = \tfrac{1}{2}\left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} . \tag{2.32}$$

Of course, these are the expected results: the energy eigenstates are the spatially even and odd combinations of harmonic oscillator states centered at the bottoms of the two wells; the degeneracy of the two energy eigenvalues is broken only by barrier penetration (and thus the difference of the energies is proportional to the barrier-penetration factor, $e^{-S_0/\hbar}$), and the state of lower energy, which we have denoted by $|->$, is the spatially even combination.

Our next task is to evaluate K. Before we do this, though, some comments should be made about what we have done so far:

(1) Really we have no right to retain the second term in Eq. (2.31). It is not only exponentially small compared to the first term, it is exponentially small compared to the uncomputed $O(\hbar^2)$

corrections to the first term. However, it is the leading contribution to the difference of the energies, $E_+ - E_-$; a purist would retain it only in the expression for this difference and not in the expressions for the individual energies.

(2) Our approximation has been based on the assumption that the instantons and anti-instantons are all widely separated. As a consistency check, we should verify that the major portion of our final result comes from configurations where this is indeed the case.

This check is easy to carry out. For any fixed x, the terms in the exponential series, $\sum x^n/n!$, grow with n until n is on the order of x; after this point, they begin to decrease rapidly. Applying this to the sum in Eq. (2.29), we see the important terms are those for which

$$n \lesssim KTe^{-S_0/\hbar} \quad . \tag{2.33}$$

That is to say, for small $\hbar$, the important terms in the sum are those for which n/T, the density of instantons and anti-instantons, is exponentially small, and thus the average separation is enormous. Note that this average separation is independent of T; our approximation is indeed a small-$\hbar$ approximation; the conditions for its validity are independent of T, as long as T is sufficiently large.

This approximation of summing over widely-separated instantons is called the dilute-gas approximation, because of its similarity to the approximation of that name in statistical mechanics.

(3) Finally, I want to deliver the promised fuller explanation of the idea of an approximate stationary point of S. Let us begin by studying an integral over a single variable,

$$I = \int_0^T dt\, e^{-S(t)/\hbar} \quad , \tag{2.34}$$

where S is a function of t monotonically decreasing to some asymptotic value, $S(\infty)$. Thus the integrand has no stationary points in the region of integration. Nevertheless, it is easy to find the approximate form of the integral for small $\hbar$ and large T:

$$I \simeq Te^{-S(\infty)/\hbar} \tag{2.35}$$

Speaking loosely, the integral is dominated by the stationary point at infinity. It's straightforward to generalize this phenomenon to multi-dimensional integrals: We assume an integrand whose graph has a sort of trough in it; the line along the bottom of the trough flattens out only as we go to infinity. Speaking less pictorally, there is a line in the multi-dimensional space such that the integrand is a minimum with respect to variations perpendicular to the line and approaches some limiting value as one goes to infinity along the line. Of course, the line could itself be generalized to a hyperplane, a generalized "bottom of the trough". This is in fact the situation for our "approximate stationary points"; the locations of the instantons and anti-instantons are the variables along the bottom of the trough; S becomes stationary (and equal to nS_0) only when they all go to infinity.

This concludes the comments; we now turn to the evaluation of K.

We must study the eigenvalue equation, Eq. (2.10), with $\bar{x}$ a single instanton. Because of time translation invariance, this equation necessarily possesses an eigenfunction of eigenvalue zero,

$$x_1 = S_0^{-\frac{1}{2}} d\bar{x}/dt \quad . \tag{2.36}$$

[The normalization factor comes from Eq. (2.22).] Were we to integrate over the corresponding expansion coefficient, c_1, in Eq. (2.6), we would obtain a disastrous infinity. Fortunately, we have already done this integration, in the guise of integrating over the location of the center of the instanton in Eq. (2.28). The change of x(t) induced by a small change in the location of the center, t_1, is

$$dx = (d\bar{x}/dt)dt_1 \quad . \tag{2.37}$$

The change induced by a small change in the expansion coefficient, c_1, is

$$dx = x_1 dc_1 \quad . \tag{2.38}$$

Hence,

$$(2\pi\hbar)^{-\frac{1}{2}}\,dc_1 = (S_0/2\pi\hbar)^{\frac{1}{2}}\,dt_1 \ . \qquad (2.39)$$

Thus, in evaluating the determinant, we should not include the zero eigenvalue, but we should include in K a factor[7] of $(S_0/2\pi\hbar)^{\frac{1}{2}}$. Hence, the one-instanton contribution to the transition matrix element is given by

$$\langle a|e^{-HT}|-a\rangle_{\text{one inst.}} = NT(S_0/2\pi\hbar)^{\frac{1}{2}} e^{-S_0/\hbar}\,(\det'[-\partial_t^2 + V''(\bar{x})])^{-\frac{1}{2}} \ , \qquad (2.40)$$

where det' indicates that the zero eigenvalue is to be omitted when computing the determinant. Comparing this to the one-instanton term in Eq. (2.29), we find

$$K = (S_0/2\pi\hbar)^{\frac{1}{2}} \left| \frac{\det(-\partial_t^2 + \omega^2)}{\det'(-\partial_t^2 + V''(\bar{x}))} \right|^{\frac{1}{2}} \ . \qquad (2.41)$$

This completes the computation.

Some remarks:

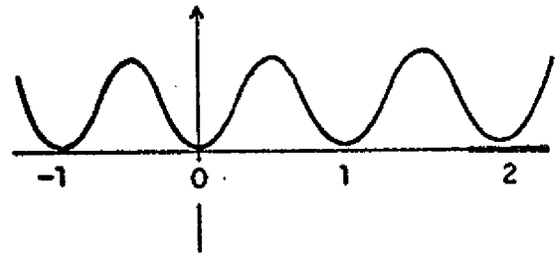
(1) To really sew things up, I should show that the formula we have obtained for the energy splitting is the same as that obtained by the traditional methods of wave mechanics. I do this in Appendix B.

(2) I have been tacitly assuming that all the eigenvalues in Eq. (2.10) are positive, other than the zero eigenvalue associated with x_1. It is easy to prove that this is indeed the case: It is well-known that the eigenfunction of a one-dimensional Schrödinger equation [like Eq. (2.10)] of lowest eigenvalue has no nodes, the next-lowest eigenfunction has one node, etc. Because the instanton is a monotone increasing function of t, x_1, proportional to the time derivative of the instanton, has no nodes. Thus zero is the lowest eigenvalue and all the other eigenvalues are positive.

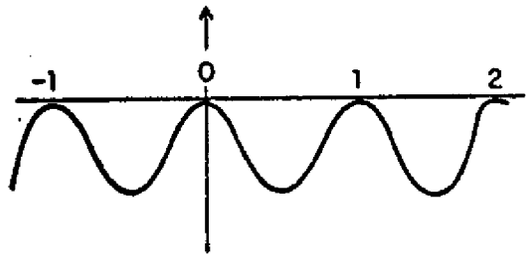
(3) K is proportional to $\hbar^{-\frac{1}{2}}$. This factor came from the zero eigenvalue associated with time-translation invariance. Later in these lectures we will be analyzing theories that have larger invariance groups and for which the instantons have more than one zero eigenvalue associated with them. Clearly, for every zero eigenvalue there will be a factor of $\hbar^{-\frac{1}{2}}$. This rule for counting powers of $\hbar$ will be very important to us, for, as I explained in Section 1, counting powers of $\hbar$ is equivalent to counting powers of coupling constants.

2.3 Periodic Potentials

Let us consider a periodic potential, like the one sketched in Fig. 6a. (For simplicity, I have chosen the minima of V to be the integers.) If we ignore barrier penetration, the energy eigenstates are an infinitely degenerate set of states, each concentrated at the bottom of one of the wells. Barrier penetration changes this single eigenvalue into a continuous band of eigenvalues; the true energy eigenstates are the eigenstates of unit translations, the Bloch waves. Let's see how this old result can be obtained by instanton methods.



(a)



(b)

Figure 6

As we see from Fig. 6b, the instantons are much the same as in the preceding problem. The only novelty is that the instantons can begin at any initial position, $x = j$, and go to the next one, $x = j+1$. Likewise, the anti-instantons can go from $x = j$ to $x = j-1$. Otherwise, everything is as before.

Thus, when doing the dilute-gas sum, we can sprinkle instantons and anti-instantons freely about the real axis; there is no constraint that instantons and anti-instantons must alternate. Of course, as we go along the line, each instanton or anti-instanton must begin where its predecessor ended. Furthermore, the total number of instantons minus the total number of anti-instantons must equal the change in x between the initial and final position eigenstates.

Thus we obtain

$$\langle j_+|e^{-HT/\hbar}|j_-\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \sum_{n=0}^{\infty}\sum_{\bar{n}=0}^{\infty} \frac{1}{n!\bar{n}!}\left(Ke^{-S_0/\hbar}T\right)^{n+\bar{n}} \delta_{n-\bar{n}-j_++j_-} \tag{2.42}$$

where n is the number of instantons and $\bar{n}$ the number of anti-instantons. If we use the identity

$$\delta_{ab} = \int_0^{2\pi} d\theta\, e^{i\theta(a-b)}/2\pi \ , \tag{2.43}$$

the sum becomes two independent exponential series, and we find

$$\langle j_+|e^{-HT/\hbar}|j_-\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \int_0^{2\pi} e^{i(j_--j_+)\theta}\,\frac{d\theta}{2\pi}\exp[2KT\cos\theta\, e^{-S_0/\hbar}] \ . \tag{2.44}$$

Thus we find a continuum of energy eigenstates labeled by the angle θ. The energy eigenvalues are given by

$$E(\theta) = \tfrac{1}{2}\hbar\omega + 2\hbar K\cos\theta\, e^{-S_0/\hbar} \ . \tag{2.45}$$

Also,

$$\langle\theta|j\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{4}}\left(2\pi\right)^{-\frac{1}{2}} e^{ij\theta} \ . \tag{2.46}$$

Heartening1y, this is just the right answer.

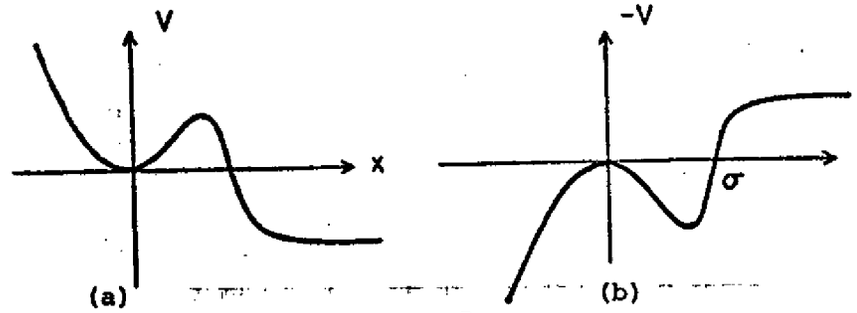


Figure 7

2.4 Unstable States and Bounces[8]

Galilean pastiche:

SAGREDO: Let me test my understanding of these instanton methods by studying the potential of Fig. 7a. If I neglect barrier penetration, in the semiclassical limit, this potential has an energy eigenstate sitting in the bottom of the well. I wish to compute the corrections to the energy of this state due to barrier penetration. If I turn the potential upside down (Fig. 7b), I observe that the classical equation of motion has a solution in which the particle begins at the top of the hill at $x = 0$, bounces off the classical turning point σ, and returns to the top of the hill (Fig. 8). I will call this motion "the bounce". I will compute the transition matrix element between $x = 0$ and $x = 0$ by summing over configurations consisting of widely separated bounces, just as one sums over instantons and anti-instantons in the study of the double well. Indeed, the sum is the same as that for the double well (with the obvious redefinitions of S_0, ω^2, etc.), save that there is no restriction to an even or odd number of bounces. Thus

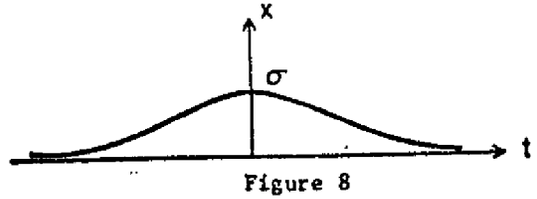


Figure 8

I obtain the complete exponential series, rather than just the odd or even terms, and I find that

$$\langle 0|e^{-HT/\hbar}|0\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{\frac{1}{2}} e^{-\omega T/2} \exp[KTe^{-S_0/\hbar}] , \qquad (2.47)$$

and the energy eigenvalue is given by

$$E_0 = \tfrac{1}{2}\omega\hbar + \hbar K e^{-S_0/\hbar} . \qquad (2.48)$$

SALVIATO: Alas, Sagredo, I fear you have erred in three ways. Firstly, the term you have computed is small compared to terms of order $\hbar^2$ which you have neglected, and thus you have no right to retain it. Secondly, I see by your sketch that the bounce has a maximum; therefore the eigenfunction x_1, which is proportional to the time derivative of the bounce, has a node. Thus it is not the eigenfunction of lowest eigenvalue, and there must be a nodeless eigenfunction, x_0, of a lower eigenvalue, that is to say, there must be a negative eigenvalue. Thus K, which is inversely proportional to the product of the square roots of the eigenvalues, is imaginary. Thirdly, the eigenvalue you attempt to compute is nowhere to be found in the spectrum of the Hamiltonian, because the state you are studying is rendered unstable by barrier penetration.

SAGREDO: Everything you say is correct, but I believe your criticisms show how to save the computation. An unstable state is one whose energy has an imaginary part; thus it is only to be expected that K should be imaginary. Furthermore, the term I have computed, though indeed small compared to neglected contributions to the real part of E_0, is the leading contribution to the imaginary part of E_0. Thus the correct version of Eq. (2.48) is

$$\mathrm{Im}E_0 = \Gamma/2 = \hbar|K|e^{-S_0/\hbar} , \qquad (2.49)$$

where Γ is, as usual, the width of the unstable state.

As you can see, the Tuscan twosome are as quick-witted as ever, although (also as ever) their arguments are sometimes a bit sloppy. Sagredo has missed a factor of $\frac{1}{2}$; the correct answer is

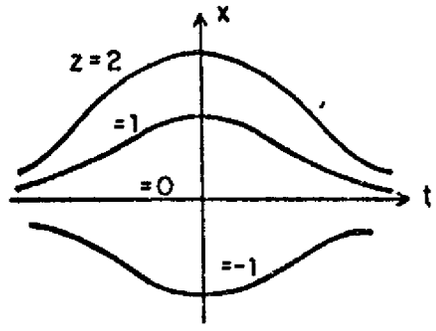


Figure 9

$$\Gamma = \hbar|K|e^{-S_0/\hbar} . \qquad (2.50)$$

To show that this is the case requires a more careful argument than Sagredo's. The essential point is Salviato's observation that the energy of an unstable state is not an eigenvalue of H; in fact, it's an object that can only be defined by a process of analytic continuation. I will now perform such a continuation.

To keep things as simple as possible, let us consider not an integral over all function space, but an integral over some path in function space parameterized by a real variable, z,

$$J = \int dz(2\pi\hbar)^{-\frac{1}{2}} e^{-S(z)/\hbar} , \qquad (2.51)$$

where S(z) is the action along the path. In particular, let us choose the path sketched in Fig. 9. This path includes two important functions that occur in the real problem: x(t) = 0, at z = 0, and the bounce, at z = 1. Furthermore, the path is such that the tangent vector to the path at z = 1 is x_0. Thus the path goes through the bounce in the "most dangerous direction", that direction with which the negative eigenvalue is associated, and z = 1 is a maximum of S, as shown in Fig. 10. S goes to minus infinity as z goes to infinity because the functions spend more and more time in the region beyond the turning point, where V is negative; note that this implies that Eq. (2.51) is hopelessly divergent.

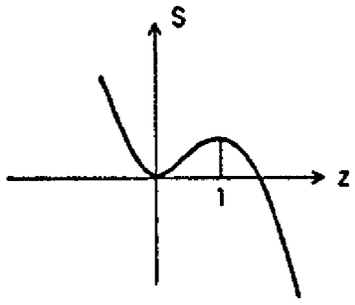


Figure 10

If $x = 0$ were the absolute minimum of V, that is to say, if V were as shown in Fig. 11a, we would have, for the same path, the situation shown in Fig. 11b, and there would be no divergence in Eq. (2.48). Now let us suppose we analytically change V in some way such that we go from this situation back to the one of interest. To keep the integral convergent, we must distort the right-hand portion of the contour of integration into the complex plane. How we distort it depends on the details of the analytic passage from one potential to the other. In Fig. 12, I have assumed that it is distorted into the upper half plane. Following the standard procedure of the method of steepest descents, I have led the contour along the real axis to $z = 1$, the saddle point, and then out along a line of constant imaginary part of S. The integral thus acquires

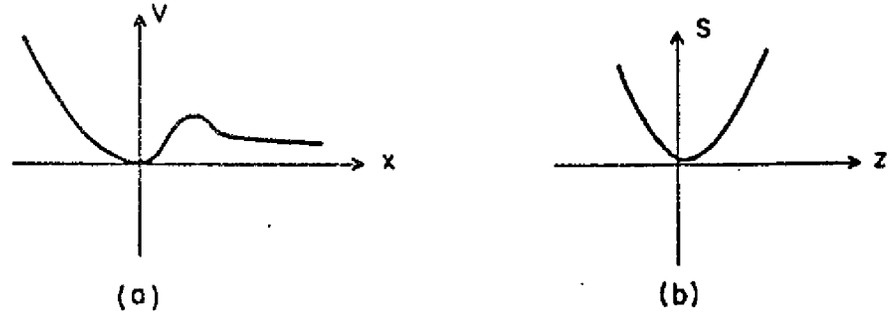


Figure 11

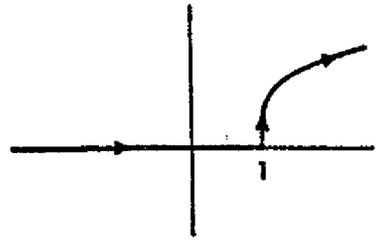


Figure 12

an imaginary part; in the steepest-descent approximation,

$$\text{im}J = \text{Im}\int_1^{1+i\infty} dz\,(2\pi\hbar)^{-\frac{1}{2}}\, e^{-S(1)\hbar}\, e^{-\frac{1}{2}S''(1)(z-1)^2/\hbar}$$
$$= \tfrac{1}{2} e^{-S(1)/\hbar} |S''(1)|^{-\frac{1}{2}} . \qquad (2.52)$$

Note the factor of $\frac{1}{2}$; this arises because the integration is over only half of the Gaussian peak.

(If we had passed from one potential to the other in the conjugate manner, the contour would have been distorted into the lower half plane, and we would have obtained the opposite sign for the imaginary part. This is just a reflection of the well-known fact that what sign you get for the imaginary part of the energy of an unstable state depends on how you do your analytic continuation.)

Now, we have studied a one-dimensional integral, but we can always reduce our functional integral to a one-dimensional integral simply by integrating (in the Gaussian approximation) over all the variables orthogonal to our path. These directions involve only positive or zero eigenvalues near the stationary point and give us no trouble. In this manner we obtain Sagredo's answer, Eq. (2.48), except that the negative eigenvalue carries a factor of $\frac{1}{2}$ with it; that is to say, we obtain Eq. (2.49).

III. THE VACUUM STRUCTURE OF GAUGE FIELD THEORIES[9]

3.1 Old Stuff

This subsection is a telegraphic compendium of formulas from gauge field theories. Its purpose is to establish notational conventions and possibly to jog your memory. If you don't already know the fundamentals of gauge field theory, you won't learn them here.[10]

Lie Algebras. A representation of Lie algegra is a set of N anti-Hermitian matrices, T^a, a = 1 ... N, obeying the equations

$$[T^a,T^b] = c^{abc}\,T^c\ , \tag{3.1}$$

where the c's are the structure constants of some compact Lie group, G. It is always possible to choose the T's such that $\mathrm{Tr}(T^aT^b)$ is proportional to δ^{ab}, although the constant of proportionality may depend on the representation. The Cartan inner product is defined by

$$(T^a,T^b) = \delta^{ab}\ . \tag{3.2}$$

Thus this is proportional to the trace of the product of the matrices.

So far I have not stated a convention that gives a scale to the structure constants and thus to the T's. For SU(2), the case I will spend most time discussing, I will choose c^{abc} to be equal to ϵ^{abc}. Thus, for the isospinor representation,

$$T^a = -i\sigma^a/2\ , \tag{3.3}$$

where the σ's are the Pauli spin matrices. In this case,

$$(T^a,T^b) = -2\,\mathrm{Tr}(T^a\,T^b)\ . \tag{3.4}$$

Occasionally I will discuss SU(n), in particular SU(3). In this case I will choose the structure constants to agree with the preceding convention for the SU(2) subgroup composed of unitary unimodular transformations on two variables only. Thus, for SU(3), T^a is $-i\lambda^a/2$, where the λ's are Gell-Mann's matrices.

Gauge Fields. The gauge potentials are a set of vector fields, $A^a_\mu(x)$. It is convenient to define a matrix-valued vector field, $A_\mu(x)$, by

$$A_\mu = g\,A^a_\mu\,T^a\ , \tag{3.5}$$

where g is a constant called the gauge coupling constant. The field-strength tensor, $F_{\mu\nu}(x)$, is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu,A_\nu]\ . \tag{3.6}$$

Pure gauge field theory is defined by the Euclidean action,

$$S = \frac{1}{4g^2}\int d^4x\,(F_{\mu\nu},F_{\mu\nu})\ . \tag{3.7}$$

Sometimes I will write this in a shorthand form,

$$S = \frac{1}{4g^2}\int (F^2)\ . \tag{3.8}$$

Gauge Transformations. A gauge transformation is a function, g(x), from Euclidean space into the gauge group, G. In equations,

$$g(x) = \exp\lambda^a(x)T^a\ , \tag{3.9}$$

where the λ's are arbitrary functions. (Please do not confuse g(x) with the coupling constant, g.) Under such a transformation,

$$A_\mu \to g\,A_\mu\,g^{-1} + g\,\partial_\mu\,g^{-1}\ , \tag{3.10}$$

and

$$F_{\mu\nu} \to g\,F_{\mu\nu}\,g^{-1}\ . \tag{3.11}$$

Thus, S is gauge-invariant. If $F_{\mu\nu}$ vanishes, then A_μ is a gauge-transform of zero; that is to say,

$$A_\mu = g\,\partial_\mu\,g^{-1}\ , \tag{3.12}$$

for some g(x).

Covariant Derivatives. The covariant derivative of the field strength tensor is defined by

$$D_\lambda\,F_{\mu\nu} = \partial_\lambda\,F_{\mu\nu} + [A_\lambda,F_{\mu\nu}]\ . \tag{3.13}$$

Equation (3.7) leads to the Euclidean equations of motion

$$D_\mu\,F_{\mu\nu} = 0\ . \tag{3.14}$$

Given a field ψ that gauge-transforms according to

$$\psi \to g(x)\psi \tag{3.15}$$

then the covariant derivative of ψ,

$$D_\mu \psi = \partial_\mu \psi + A_\mu \psi \ , \tag{3.16}$$

transforms in the same way.

3.2 The Winding Number

I propose to study Euclidean gauge field configurations of finite action (not necessarily solutions of the equations of motion).

Why?

The naive answer, sometimes given in the literature,[11] is that configurations of infinite action are unimportant in the functional integral, since, for such configurations, $e^{-S/\hbar}$ is zero. *This is wrong.* In fact, it is configurations of finite action that are unimportant; to be precise, they form a set of measure zero in function space. This has nothing to do with the divergences of quantum field theory; it is true even for the ordinary harmonic oscillator. (For a proof, see Appendix C.) The only reason we are interested in configurations of finite action is that we are interested in doing semiclassical approximations, and a configuration of infinite action does indeed give zero if it is used as the center point of a Gaussian integral.

The convergence of the action integral is controlled by the behavior of A_μ for large r, where r is the radial variable in Euclidean four-space. To keep my arguments as simple as possible, I will assume that, for large r, A_μ can be expanded in an asymptotic series in inverse powers of r. (This assumption can be relaxed considerably without altering the conclusions.[12]) Thus, for the action to be finite, $F_{\mu\nu}$ must fall off faster than $1/r^2$ as r goes to infinity; that is to say, $F_{\mu\nu}$ must be $O(1/r^3)$. One's first thought is that this implies that A_μ is $O(1/r^2)$, but this is wrong: vanishing $F_{\mu\nu}$ does not imply vanishing A_μ, but merely that A_μ is a gauge transform of zero. Thus A_μ can be of the form

$$A_\mu = g\,\partial_\mu g^{-1} + O(1/r^2) \ , \tag{3.17}$$

where g is a function from four-space to G of order one, that is to say, a function of angular variables only.

Thus, with every finite-action field configuration there is associated a group-element-valued function of angular variables, that is to say, a mapping of a three-dimensional hypersphere, S^3, into the gauge group, G. Of course, this assignment is not gauge invariant. Under a gauge transformation, h(x)

$$A_\mu \to h A_\mu h^{-1} + h\,\partial_\mu h^{-1} \ . \tag{3.18}$$

Thus,

$$g \to h g + O(1/r^2) \ . \tag{3.19}$$

If one could choose h to equal g^{-1} at infinity, one could transform g to one and eliminate it from Eq. (3.17). In general, though, this is not possible. The reason is that h must be a continuous function not just on the hypersphere at infinity, but throughout all four-space, that is to say, on a nested family of hyperspheres going all the way from r equals zero to r equals infinity. In particular, at the origin, h must be a constant, independent of angles. Thus, h at infinity can not be a general function on S^3, but must be one that can be obtained by continuous deformation from a constant function. Since any constant gauge transformation can trivially be obtained by continuous deformation from the identity transformation (all gauge groups are connected), we might as well say that h at infinity must be obtainable from h = 1 by a continuous deformation.

Given two mappings of one topological space into another, such that one mapping is continuously deformable into another, mathematicians say the two functions are "homotopic" or "in the same homotopy class". What we have shown is that by a gauge transformation we can transform g(x) into any mapping homotopic to g(x), but we can not transform it into a function in another homotopy class. Thus, the gauge-invariant quantity associated with a finite-action field configuration is not a mapping of S^3 to G but a homotopy class

of such mappings. Our task is to find these homotopy classes for physically interesting G's.

To warm up for this task, let me consider a baby version of the problem for which the geometry is somewhat easier to visualize. I will work with the simplest of all gauge groups, U(1), the group of complex numbers of unit modulus. Thus the gauge field theory is ordinary electromagnetism. (However, I will still keep to the notational conventions established in Sec. 3.1; in particular, A_μ will be an imaginary quantity, i times the usual vector potential.) Also, I will work not in Euclidean four-space but in Euclidean two-space. I will still study fields obeying Eq. (3.17), although, of course, in two-space this condition is not a consequence of finiteness of the action. Because we are working in two-space, we have, instead of a hypersphere, S^3, an ordinary circle, S^1.

Now to work:

(1) G is the unit circle in the complex plane; thus, topologically, G is also S^1, and we have to study homotopy classes of mappings of S^1 into S^1. We will label the circle in space, the domain of our functions, in the standard way, by an angle θ ranging from 0 to 2π.

(2) It will be useful to define some standard mappings from S^1 to S^1. One is the trivial mapping,

$$g^{(0)}(\theta) = 1 \ . \tag{3.20a}$$

Another is the identity mapping,

$$g^{(1)}(\theta) = e^{i\theta} \ . \tag{3.20b}$$

These are both part of a family of mappings,

$$g^{(\nu)}(\theta) = [g^{(1)}(\theta)]^\nu = e^{i\nu\theta} \ , \tag{3.20c}$$

where ν is an integer (positive, negative, or zero). ν is called "the winding number", because it is the number of times we wind around G when we go once around the circle at infinity in two-space. (By convention, winding around minus once means winding around once in the negative direction.)

(3) Every mapping from S^1 to S^1 is homotopic to one of the mappings (3.20c). We do not have the mathematical machinery to prove this rigorously, but I hope I can made it plausible. Imagine taking a rubber band and marking on it in ink a sequence of values of θ running from 0 to 2π. We then wrap the band about a circle representing G, such that each value of θ lies above the point into which it is mapped. (Figure 13 shows such a construction.) We can continuously deform the band, first to eliminate any folds, like the one on the top of the figure, and second to stretch the band so it lies uniformly on the circle. In this way we obtain some $g^{(\nu)}(\theta)$. (In the case shown, we obtain $g^{(1)}$.) Thus we can associate a winding number with every mapping. (Note that I have not yet shown that this number is uniquely defined.)

(4) I will now show that the winding number defined above is given by the integral formula

$$\nu = \frac{i}{2\pi}\int_0^{2\pi} d\theta g dg^{-1}/d\theta \ . \tag{3.21}$$

Firstly, by direct calculation, this gives the right answer for the standard mappings, Eq. (3.20c). Secondly, this quantity is invariant

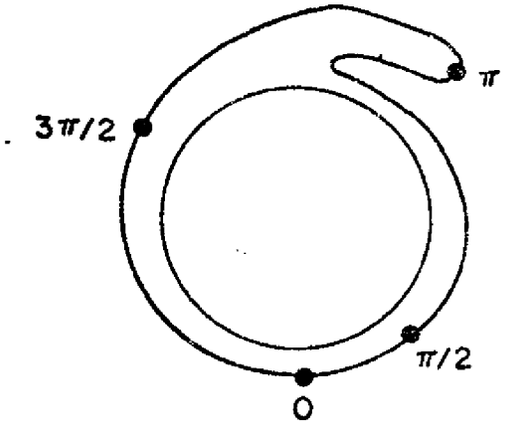


Figure 13

under continuous deformations. To prove this assertion it suffices to demonstrate invariance under infinitesimal deformations. A general infinitesimal deformation is of the form

$$\delta g = i(\delta\lambda)g \ , \tag{3.22}$$

where $\delta\lambda$ is some infinitesimal real function on the circle. Thus

$$\delta(gdg^{-1}/d\theta) = -id(\delta\lambda)/d\theta \ , \tag{3.23}$$

and the change in ν vanishes upon integration. (We now know that all of our standard mappings are in different homotopy classes and that the winding number is uniquely defined.)

(5) If

$$g(\theta) = g_1(\theta)g_2(\theta) \ , \tag{3.24a}$$

then

$$\nu = \nu_1 + \nu_2 \ . \tag{3.24b}$$

The proof is simple: The winding number is unchanged by continuous deformations. We can deform g_1 such that it is equal to one on the upper half of the circle ($0 \le \theta \le \pi$) and g_2 such that it is equal to one on the lower half of the circle ($\pi \le \theta \le 2\pi$). The integrand in Eq. (3.21) is then the sum of a part due to g_1 (vanishing on the upper semicircle) and a part due to g_2 (vanishing on the lower semicircle).

(6) Let us define

$$C_\mu = \frac{i}{2\pi}\cdot\varepsilon_{\mu\nu}A_\nu \ . \tag{3.25}$$

By Eqs. (3.17) and (3.21),

$$\nu = \lim_{r\to\infty}\int_0^{2\pi} rd\theta\hat{r}_\mu C_\mu \ , \tag{3.26}$$

where $\hat{r}_\mu$ is the radial unit vector. Thus, by Gauss's theorem,

$$\nu = \int d^2x\partial_\mu C_\mu \ . \tag{3.27}$$

Hence,

$$\nu = \frac{i}{4\pi}\int d^2x\varepsilon_{\mu\nu}F_{\mu\nu} \ . \tag{3.28}$$

I will now return to four-space, and take G to be SU(2). As we shall see, every argument will be a (mild) generalization of the arguments I have given for the baby problem.

(1) SU(2) is the group of unitary unimodular two-by-two matrices. It is well known that any such matrix can be uniquely written in the form

$$g = a + i\vec{b}\cdot\vec{\sigma} \ , \tag{3.29}$$

where $a^2 + |b|^2 = 1$. Thus, topologically, SU(2) is S^3, and we have to study homotopy classes of mappings from S^3 to S^3.

(2) It will be useful to define some standard mappings from S^3 to S^3. One is the trivial mapping,

$$g^{(0)}(x) = 1 \ . \tag{3.30a}$$

Another is the identity mapping,

$$g^{(1)}(x) = (x_4 + i\vec{x}\cdot\vec{\sigma})/r \ . \tag{3.30b}$$

These are both part of a family of mappings,

$$g^{(\nu)}(x) = [g^{(1)}(x)]^\nu \ . \tag{3.30c}$$

where ν is an integer, called the winding number. (It is also sometimes called the Pontryagin index.) It measures the number of times the hypersphere at infinity is wrapped around G. (By convention, we say the hypersphere is wrapped around G in a negative sense if a right-handed triad of tangent vectors is mapped into a left-handed triad.)

(3) Every mapping from S^3 to S^3 is homotopic to one of our standard mappings (3.30c). We do not have the mathematical machinery to prove this assertion rigorously, but a plausibility argument can be constructed just as in the baby problem, with hyperspheres replacing circles. (If you have problems envisioning hyperspheres wrapped around hyperspheres, just accept the assertion on faith.) In this way we can associate a winding number with every mapping. (Note that I have not yet shown that this number is uniquely defined.)

(4) Let us define

$$\nu = \frac{1}{48\pi^2}\int d\theta_1 d\theta_2 d\theta_3 \varepsilon^{ijk}(g\partial_i g^{-1},\ g\partial_j g^{-1} g\partial_k g^{-1}) \ . \tag{3.31}$$

where θ_1, θ_2 and θ_3 are three angles that parameterize S^3. How these angles are chosen is irrelevant to Eq. (3.31); the Jacobian determinant that comes from changing the angles is canceled by the Jacobian determinant from the ε-symbol. Equation (3.31) is written using the Cartan inner product, that is to say, in a representation-independent way. Of course, for any particular representation of SU(2), we can rewrite Eq. (3.31) in terms of traces; for example, for the two dimensional representation, by Eq. (3.4),

$$\nu = -\frac{i}{24\pi^2}\int d\theta_1 d\theta_2 d\theta_3 \mathrm{Tr}\varepsilon^{ijk} g\partial_i g^{-1} g\partial_j g^{-1} g\partial_k g^{-1} \ . \qquad (3.32)$$

I will show that this quantity is, firstly, a homotopy invariant, and secondly, agrees with the winding number as defined for our standard mappings. As before, a corollary of this proof will be that all of our standard mappings are in different homotopy classes and that the winding number is uniquely defined.

To show invariance under continuous deformations it suffices to show invariance under infinitesimal deformations. For any Lie group, a general infinitesimal transformation can be written as an infinitesimal right multiplication:

$$\delta g = g\delta\lambda^a(x)T^a \equiv g\delta T \ . \qquad (3.33)$$

Under this transformation,

$$\delta(g\partial_k g^{-1}) = -g(\partial_k \delta T)g^{-1} \ . \qquad (3.34)$$

The three derivatives in Eq. (3.32) make equal contributions to $\delta\nu$; thus,

$$\delta\nu \propto \int d\theta_1 d\theta_2 d\theta_3 \varepsilon^{ijk}\mathrm{Tr} g\partial_i g^{-1} g\partial_j g^{-1} g(\partial_k \delta T)g^{-1} \ . \qquad (3.35)$$

If we use the identity,

$$0 = \partial_i(gg^{-1}) = g\partial_i g^{-1} + (\partial_i g)g^{-1} \ , \qquad (3.36)$$

this becomes

$$\delta\nu \propto \int d\theta_1 d\theta_2 d\theta_3 \varepsilon^{ijk}\mathrm{Tr}\partial_i g^{-1}\partial_j g\partial_k \delta T \ , \qquad (3.37)$$

which vanishes upon integration by parts, because of the antisymmetry of the ε-symbol. This completes the proof of invariance under continuous deformations.

(5) Now to evaluate Eq. (3.32) for our standard mappings. The task is easiest for $g^{(1)}$, for the integrand is here obviously a constant, and we need evaluate it only at the north pole of the unit hypersphere, $x_4 = 1$, $x_i = 0$. At this point we might as well choose θ_i to equal x_i. Thus, from Eq. (3.30b),

$$g\partial_i g^{-1} = -i\sigma_i \ , \qquad (3.38)$$

and

$$\mathrm{Tr}\varepsilon^{ijk} g\partial_i g^{-1} g\partial_j g^{-1} g\partial_k g^{-1} = -12 \ . \qquad (3.39)$$

Since the area of a unit hypersphere is $2\pi^2$, we obtain the desired result, $\nu = 1$.

For the other standard mappings, the simplest way to proceed is to observe that if

$$g = g_1 g_2 \ , \qquad (3.40a)$$

then

$$\nu = \nu_1 + \nu_2 \ . \qquad (3.40b)$$

The argument is the same as for the baby problem, with semihyperspheres replacing semicircles.

(6) Let us define

$$G_\mu = 2\varepsilon_{\mu\nu\lambda\sigma}(A_\nu, \partial_\lambda A_\sigma + \tfrac{2}{3}A_\lambda A_\sigma) \ . \qquad (3.41)$$

A straightforward computation shows that

$$\partial_\mu G_\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}(F_{\mu\nu}, F_{\lambda\sigma}) \ . \qquad (3.42)$$

The dual of an antisymmetric tensor (denoted by a tilde) is conventionally defined by

$$\tilde{F}_{\mu\nu} \equiv \tfrac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}F_{\lambda\sigma} \ . \qquad (3.43)$$

(The factor of $\frac{1}{2}$ is inserted in the definition so that $\tilde{\tilde{F}} = F$.) Equation (3.42) can thus be rewritten as

$$\partial_\mu G_\mu = (F_{\mu\nu}, \tilde{F}_{\mu\nu}) \equiv (F, \tilde{F}) \ . \qquad (3.44)$$

From the definition of $F_{\mu\nu}$,

$$G_\mu = \varepsilon_{\mu\nu\lambda\sigma}(A_\nu, F_{\lambda\sigma} - \tfrac{2}{3}A_\lambda A_\sigma) \ . \qquad (3.45)$$

This expression is useful in evaluating

$$\int d^4x(F,\tilde{F}) = \int d^3S\hat{r}_\mu G_\mu \ , \qquad (3.46)$$

where d^3S is the element of area on a large hypersphere. The first term in Eq. (3.45) is $O(1/r^3)$ and makes no contribution to the integral; the second term simply gives (up to a multiplicative constant) the integral formula for the winding number, Eq. (3.31). Thus we obtain

$$\int d^4x(F,\tilde{F}) = 32\pi^2\nu \ . \qquad (3.47)$$

Summary and Generalizations. This has been a long analysis, and you may have lost track of what we were doing, so let me summarize the main results of this subsection: For a gauge field theory based on the group SU(2), every field configuration of finite action in four-dimensional Euclidean space has an integer associated with it, the Pontryagin index or winding number, ν. It is not possible to continuously deform a configuration of one winding number into one of a different winding number while maintaining the finiteness of the action. We have two integral formulas for the winding number, one in terms of a surface integral over a large sphere, Eq. (3.31), and one in terms of a volume integral over all four-space, Eq. (3.47).

How much of this depends on the gauge group being SU(2)? Firstly, if the gauge group is U(1), it is easy to see that every mapping of S^3 into U(1) is continuously deformable into the trivial mapping (all of S^3 mapped into a single point). *Thus, for an Abelian gauge field theory, there is no analog of the winding number.* Secondly, for a general simple Lie group, G, there is a remarkable theorem due to Raoul Bott[14] that states that any continuous mapping of S^3 into G can be continuously deformed into a mapping into an SU(2) subgroup of G. *Thus, everything we have discovered for SU(2) is true for an arbitrary simple Lie group; in particular, it is true for SU(n).* I stress that "everything" means *everything;* In particular, not a single numerical factor in the integral formulas for the winding number needs alteration, so long as we choose the normalization of the Cartan inner product appropriately (as we have). Finally, since a general compact Lie group is locally the direct product of an Abelian group and a string of simple groups, *for a general gauge field theory, there is an independent winding number for every simple factor group.*

3.3 Many Vacua

We have learned a lot about classical gauge field theories; now it is time to confront the quantum theory. In principle, the Euclidean functional integral tells how to go from the classical theory to the quantum theory. As I explained in Sec. 2, we can use the functional integral to study the energy eigenstates of the theory; also, by adding appropriate source terms to the Hamiltonian (equivalently, to the Euclidean action) and then differentiating with respect to the sources at the end of the computation, we can study the expectation values of strings of operators, Euclidean Green's functions. However, for gauge field theories, there is a famous complication: to make the functional integral well-defined, we must impose a gauge-fixing condition.[15]

I will choose to work in axial gauge, $A_3 = 0$. I have several reasons for this choice: (1) It is possible to show[16] that every non-singular gauge field configuration can be put in axial gauge by a non-singular gauge transformation. It is by no means clear whether this is true for covariant gauges, for example. (2) In axial gauge the functional integral is directly equivalent to a canonical formulation of the theory;[17] there is no need of the ghost terms that occur in covariant gauges, or of the subsidiary conditions on the space of states that are needed in such gauges as $A_0 = 0$. (3) Most of the treatment in the literature of the phenomena we are about to discuss is in the gauge $A_0 = 0$. It's nice to show explicitly that the answers don't depend on this gauge choice. (4) Although axial gauge is terribly awkward for specific computations, once we have obtained functional-integral expressions for quantities of interest, we can use the standard Fadeev-Popov methods to transform these into some more convenient gauge.

In field theory, we normally plunge directly into infinite space. However, I will here study gauge field theory in a finite box of three-volume V, with definite boundary conditions, which I shall specify shortly. Just as in Sec. 2, I will also restrict the theory to a finite range of Euclidean time, T, with appropriate boundary conditions at initial and final times. Thus we are integrating over a box in Euclidean four-space, with boundary conditions on the (three-dimensional) walls of the box. Of course, I will eventually send both V and T to infinity. I again have reasons for this choice: (1) Certainly nothing is lost by beginning in a finite box; if the transition to infinite space goes smoothly, at worst we will have wasted a little time. (2) In some theories, we can gain information about the structure of the theory by seeing how things depend on the boundary conditions imposed on the walls of the box. For example, in a scalar field theory with spontaneous symmetry breakdown, the expectation value of the scalar field in the center of the box depends on the boundary conditions on the walls, no matter how large the box; this is one of the easiest ways to see that the theory has many vacua. (3) In the canonical quantization of the theory, it is necessary to eliminate A_0 from the action. To do this, it is necessary to find A_0 from $\partial_3^2 A_0$. In infinite space, this problem has many solutions; this ambiguity is usually resolved by applying *ad hoc* conditions on the behavior of A_0 at infinity. In a box with appropriate boundary conditions, this problem always has a unique solution.

There are many possible types of boundary conditions we could impose: we could fix some components of A_μ, some components of $F_{\mu\nu}$, some combinations of these, etc. A clue to a wise choice of boundary conditions is given by the surface term in the expression for the variation of the action. For example, for a free scalar field theory,

$$\delta S = \int d^3S n^\mu \partial_\mu \phi \delta\phi + \cdots . \tag{3.48}$$

Here, d^3S is the element of surface area, n^μ is the normal vector to the surface, and the triple dots denote the usual volume integral of the Euler-Lagrange equations. From this expression we see that one way to make the surface terms vanish is to fix the value of ϕ on the walls of the box. Likewise, for a gauge field theory,

$$\delta S = \frac{1}{g^2}\int d^3S n^\mu F_{\mu\nu}\delta A^\nu + \cdots . \tag{3.49}$$

From this expression we see that one way to make the surface term vanish is to fix the tangential components of A_μ on the surface. Note that there is no need to fix the normal component of A_μ; because $F_{\mu\nu}$ is antisymmetric, this makes no contribution to the surface integral.

We are not totally free to choose the tangential components of A_μ arbitrarily. Firstly, they must be chosen consistent with our gauge condition, $A_3 = 0$. Secondly, because we want to do semiclassical computations, we must choose our boundary conditions to be consistent with finiteness of the action, as the box goes to infinity. Equivalently, the boundary conditions must be consistent with the box being filled with a field configuration of a definite winding number. Furthermore, for fixed boundary conditions, this winding number is fixed, for only the tangential components of A_μ are needed to compute the normal component of G_μ. [See Eq. (3.41).]

Thus at least one relic of our boundary conditions remains no matter how large the box: we can not put an arbitrary finite-action field configuration in the box, but only one of a definite winding number. It turns out that the winding number is the *only* relic of the boundary conditions that survives as the box goes to infinity. The hand-waving argument for this is that the winding number is the *only* gauge-invariant quantity associated with the large-distance behavior of the fields. If you do not find this argument convincing, you will find a more careful one in Appendix D.

Thus, for large boxes, we can forget about the boundary conditions in the functional integral and simply integrate over all configurations where the winding number, ν, has some definite value, n. I will denote the result of such an integration by $F(V,T,n)$. In

equations,

$$F(V,T,n) = N\int [dA]e^{-S}\delta_{\nu n} \quad . \tag{3.50}$$

where $[dA]$ denotes $[dA_1][dA_2][dA_4]$. Also, I have set $\hbar$ to one; we can always keep track of the powers of $\hbar$ by keeping track of the powers of g, as explained in Sec. 1.

$F(V,T,n)$ is a transition matrix element from some initial state to some final state (determined by our boundary conditions). What these states are will not be important to us. What is important is that for large times, T_1 and T_2,

$$F(V,T_1+T_2,n) = \sum_{n_1+n_2=n} F(V,T_1,n_1)F(V,T_2,n_2) \quad . \tag{3.51}$$

This follows from Eq. (3.47), the expression for the winding number as the integral of a local density; this tells us that the way to put total winding number n in a large box is to put winding number n_1 in one part of the box and winding number n_2 in the remainder of the box, with $n = n_1 + n_2$. (Of course, such counting misses field configurations with significant action density on the boundary between the two sub-boxes, for there is no reason for the winding-number integral for each sub-box to be an integer for such configurations. However, we expect this to be a negligible surface effect for sufficiently large boxes.)

Pretty as it is, Eq. (3.51) is not what we would expect from a transition-matrix element that has a contribution from only a single energy eigenstate. Such an object would be a simple exponential, and would obey a multiplicative composition law for large times, not the convolutive composition law of Eq. (3.51). However, it is easy enough to turn convolutions into multiplications. The technique is called Fourier transformation:

$$\begin{aligned} F(V,T,\theta) &\equiv \sum_n e^{in\theta} F(V,T,n) \\ &= N\int [dA]e^{-S}e^{i\nu\theta} \quad . \end{aligned} \tag{3.52}$$

From Eq. (3.51),

$$F(V,T_1+T_2,\theta) = F(V,T_1,\theta)F(V,T_2,\theta) \quad . \tag{3.53}$$

This is the correct composition law for a simple exponential. Thus we identify $F(V,T,\theta)$ as being (up to a normalization constant) the expectation value of e^{-HT} in an energy eigenstate, which we denote by $|\theta\rangle$ and call the θ vacuum.

$$\begin{aligned} F(V,T,\theta) &\propto \langle\theta|e^{-HT}|\theta\rangle \\ &= N'\int [dA]e^{-S}e^{i\nu\theta} \quad . \end{aligned} \tag{3.54}$$

where N' is a new normalization constant.

Our analysis has been simple and straightforward (I hope), but we have been led to a very unintuitive conclusion. Our original gauge field theory seems to have split up into a family of disconnected sectors, labeled by the angle θ, each with its own vacuum. Furthermore, in each of these sectors, the computational rules are the same as those we would have naively written down if we had not gone through any of this analysis, except that an extra term, proportional to $(F,\tilde{F})$, has been added to the Lagrangian density. Probably half the people who have played with gauge field theories have thought, at one time or another, of adding such a term, and they have discarded the possibility, because the added term is a total divergence [see Eq. (3.44)] and thus has no effect on the equations of motion and therefore "obviously" has no effect on the physics of the theory. Of course, at this stage in our investigation, it is still possible that we have been fooling ourselves, that the extra term indeed has no effect on the physics, and that all the θ vacua we think we have discovered are simply duplicates of the same state. We shall eliminate this possibility immediately.

[I should remark that what we have done here closely parallels the treatment of a periodic potential in Sec. 2.3, except the arguments are somewhat more abstract and in a different order. The winding number is something like the total change in x (the difference between the number of instantons and the number of anti-instantons) in Sec. 2.3, and the θ vacua are something like the $|\theta\rangle$ eigenstates. The two big differences are that we found the analogs

of the $|\theta>$ states without pausing to talk about the analogs of the $|j>$ states, and that we did the Fourier transform that untangled the energy spectrum before we saturated the functional integral with instantons. The first difference is unimportant; if I had wanted to, I could have added two extra paragraphs when I was talking about $F(V,T,n)$ and discussed the analogs of the $|j>$ states. (They're called n vacua.) As for the instantons, they are the subject of the next subsection.]

3.4 Instantons: Generalities

In the next subsection I shall explicitly construct instantons, finite-action solutions of the Euclidean gauge-field equations with $\nu = 1$. Most of the qualitative consequences of these solutions are independent of their detailed structure and follow merely from the fact of their existence. Therefore, in this subsection, I will simply assume that instantons exist and draw some conclusions from this assumption.

I will denote the action of an instanton by S_0. Because S_0 is finite, the instanton can not be invariant under spatial translations. Thus there exists at least a four-parameter family of instanton solutions; I will call these parameters "the location of the center of the instanton". The winding number is parity-odd. Thus there must also exist at least a four-parameter family of solutions with $\nu = -1$, the parity transforms of the instanton solutions, which I will call anti-instantons. Just as in Sec. 2, we can build approximate solutions consisting of n instantons and $\bar{n}$ anti-instantons, with their centers at arbitrary widely-separated locations. These approximate solutions have $\nu = n - \bar{n}$.

Again as in Sec. 2, we approximate Eq. (3.54) by summing over all these configurations. Thus we obtain

$$<\theta|e^{-HT}|\theta> \propto \sum_{n,\bar{n}} (Ke^{-S_0})^{n+\bar{n}} (VT)^{n+\bar{n}} e^{i(n-\bar{n})\theta}/n!\bar{n}!$$
$$= \exp(2KVTe^{-S_0}\cos\theta) \ , \qquad (3.55)$$

where K is a determinental factor, defined as in Sec. 2. Thus, the

energy of a θ vacuum is given by

$$E(\theta)/V = -2K\cos\theta\, e^{-S_0} \ . \qquad (3.56)$$

Note that, as should be the case in a field theory, the different vacua are distinguished not by different energies, but by different energy densities. [Also note the similarity with the energy spectrum of a periodic potential, Eq. (2.45).]

We can go on and compute the expectation values of various operators. A particularly easy (and particularly instructive) computation is that of the expectation value of $(F,\tilde{F})$. By translational invariance,

$$<\theta|(F(x),\tilde{F}(x))|\theta> = \frac{1}{VT}\int d^4x <\theta|(F,\tilde{F})|\theta> \ . \qquad (3.57)$$

Thus, by Eq. (3.47),

$$<\theta|(F,\tilde{F})|\theta> = \frac{32\pi^2\int[dA]\nu e^{-S}e^{i\nu\theta}}{VT\int[dA]e^{-S}e^{i\nu\theta}}$$
$$= -\frac{32\pi^2 i}{VT}\frac{d}{d\theta}\ln\left(\int[dA]e^{-S}e^{i\nu\theta}\right) \ . \qquad (3.58)$$

Hence there is no need to do a fresh summation over a dilute instanton-anti-instanton gas, since we have just evaluated the quantity in parentheses in Eq. (3.55). Thus in our approximation,

$$<\theta|(F,\tilde{F})|\theta> = -64\pi^2 iKe^{-S_0}\sin\theta \ . \qquad (3.59)$$

Some comments:

(1) The expectation value is independent of V and T, as it should be.

(2) The expectation value is an imaginary number, again as it should be. The reason is that

$$(F,\tilde{F}) = (F_{12},F_{34}) + \text{permutations} \ . \qquad (3.60)$$

When we continue from Euclidean space to Minkowski space, F_{12} remains F_{12}, but, just as x_4 becomes ix_0, so does F_{34} become iF_{30}. Thus, if we had obtained a real answer, we would have found that in Minkowski space (the real world) a Hermitian operator would have had an imaginary vacuum expectation value, a disaster.

(3) Both the vacuum energy density and the vacuum expectation value depend non-trivially on θ. Thus the θ vacua are indeed all different from each other.

3.5 Instantons: Particulars

$$\int d^4x(F,F) = [\int d^4x(F\ F)\int d^4x(\tilde{F},\tilde{F})]^{\frac{1}{2}}$$
$$\geq \left|\int d^4x(F,\tilde{F})\right| , \qquad (3.61)$$

by the Schwartz inequality. Thus, for any winding number, we have an absolute lower bound on the action,

$$S \geq \frac{8\pi^2}{g^2}|\nu| . \qquad (3.62)$$

Furthermore, equality is attained if and only if

$$F = \pm\tilde{F} , \qquad (3.63)$$

where the positive (negative) sign holds for positive (negative) ν.

This inequality was first derived by Belavin, Polyakov, Schwarz, and Tyupkin,[9] who used it to search for instantons. Their idea was to look for solutions of Eq. (3.63). If such solutions exist, they are minima of the action for fixed winding number, and thus stationary points of the action under local variations, that is to say, solutions of the field equations. Furthermore, since they have lower action than any other solutions of the same winding number (if other solutions exist), they dominate the functional integral, and, for our purposes, are the only solutions we need worry about. Finally, as a bonus, Eq. (3.63) is a first-order differential equation and considerably more tractable than the second-order field equations.

Let us begin the search with $\nu = 1$. We know that any field configuration with $\nu = 1$ can be gauge-transformed such that

$$A_\mu = g^{(1)}\partial_\mu[g^{(1)}]^{-1} + O(1/r^2) , \qquad (3.64)$$

where

$$g^{(1)} = \frac{x_4 + i\vec{x}\cdot\vec{\sigma}}{r} . \qquad (3.65)$$

Equation (3.64) is rotationally invariant, in the sense that the effect of any four-dimensional rotation can be undone by an appropriate gauge transformation. This is a consequence of the statement that a rotation is a continuous deformation and thus does not change the winding number. There is also a short direct proof: Under a general rotation

$$g^{(1)} \to gg^{(1)}h^{-1} , \qquad (3.66)$$

where g and h are elements of SU(2) determined by the rotation. [This is a standard formula; it is the usual way of demonstrating the isomorphism between SO(4) and SU(2) ⊗ SU(2).] Thus,

$$A_\mu \to gA_\mu g^{-1} + O(1/r^2) . \qquad (3.67)$$

This, as promised, can be undone by a gauge transformation, indeed, by a gauge transformation of the first kind, a constant gauge transformation.

This suggests that we search for a solution of Eq. (3.63) that is rotationally invariant in the same sense. That is to say, we make the ansatz,

$$A_\mu = f(r^2)g^{(1)}\partial_\mu[g^{(1)}]^{-1} , \qquad (3.68)$$

where, to avoid a singularity, f must vanish at the origin. From here on it's straightforward plug-in-and-crank, which I will spare you. It turns out that we do indeed obtain a solution in this way, if

$$f = \frac{r^2}{r^2+\rho^2} , \qquad (3.69)$$

where ρ is an arbitrary constant, called "the size of the instanton". The existence of solutions of arbitrary sizes is a necessary consequence of the scale invariance of the classical field theory. (This fact will occasion some embarrassment shortly.)

Once we have a solution to any field theory, we can obtain new solutions by applying the invariances of the theory. In the case at hand, these are generated by (1) scale transformations, (2) rotations, (3) the four-parameter group of spatial translations, (4) the four-parameter group of special conformal transformations, and (5) gauge transformations. Scale transformations simply change the size of the instanton: thus they just shift around

the members of our one-parameter family of solutions but generate no new solutions. Rotations, as I have shown, can always be undone by gauge transformations. Spatial translations generate genuinely new solutions, and give us four more parameters, "the location of the center of the instanton". Although I do not have time to demonstrate it here, it turns out[18] that special conformal transformations can be undone by gauge transformations and translations.

Gauge transformations, as usual, require special consideration. It is easy to see that any non-trivial gauge transformation changes (3.68): Because $g^{(1)}$ is a function of angles only, the radial component of A_μ, A_r, vanishes. Thus, under a general non-singular gauge transformation, $g(x)$,

$$A_r \to gA_rg^{-1} + g\partial_r g^{-1} = g\partial_r g^{-1} \ . \qquad (3.70)$$

Hence, if the gauge transformation is not to change A_μ, g must be independent of r. That is to say, its value everywhere must be its value at the origin; g must be a constant gauge transformation. But the only constant gauge transformation that leaves A_μ unchanged is the identity. (Remember, the effect of a constant gauge transformation is the same as that of a rotation.)

You might think that this discussion of gauge transformations is irrelevant. After all, when we do the quantum theory, we must work in a fixed gauge, such as axial gauge, and it is commonly said that once we have fixed the gauge we have no freedom to make gauge transformations. However, although commonly said, this is not strictly true; all standard gauges still allow constant gauge transformations.[19] This is as it should be. Constant gauge transformations act like ordinary symmetries; they put particles into multiplets (if there is no spontaneous symmetry breakdown), impose selection rules on scattering processes, etc. Thus, in a sensible formulation of the theory, they should remain as manifest symmetries of the Hamiltonian. Whether you accept this philosophy or not, the fact remains that constant gauge transformation applied to an instanton solution (transformed to obey the gauge conditions) will

generate a different solution still obeying the gauge conditions. Thus we have found an eight-parameter family of solutions, one parameter from scale transformations, four from translations, and three from constant gauge transformations.

Are there other solutions with unit winding number? Atiyah and Ward[20] state that there are none. I can not give their proof here because I do not understand it. Nevertheless, mathematicians I trust say that their argument is not only legitimate but brilliant, so let us assume they are right and continue.

Solutions of higher winding number (if they exist) are of no interest to us. We have used approximate solutions consisting of n widely separated objects (instantons or anti-instantons) to evaluate the functional integral. These approximate solutions depend on 8n parameters, 8 for each object. Now suppose there are exact solutions that can be interpreted as n objects; that is to say, they depend on 8n (or fewer) parameters and become our approximate solutions when some of the parameters (the separations between the objects) become large. In this case, all we learn by knowing these exact solutions exist is that the dilute-gas approximation is better than we think it is - but we already know that it is good enough for our purposes. There might also be exact solutions that can not be interpreted in this way. To have a definite example, let me suppose there were a "binstanton", a brand-new solution of winding number two. Then in evaluating the functional integral, we would have to sum over a dilute gas of instantons, anti-instantons, binstantons, and anti-binstantons. Thus, Eq. (3.56) would be replaced by

$$E(\theta)/V = -2K\cos\theta e^{-S_0} - 2K'\cos 2\theta e^{-S_0'} \qquad (3.71)$$

where the primed quantities are the action and determinantal factor for a binstanton. But S_0' is twice S_0, so the new term is exponentially small compared to the old one and should be neglected.[21]

3.6 The Evaluation of the Determinant and an Infrared Embarrassment

We now know enough to go a long way towards explicitly evalu-

ating the right-hand side of Eq. (3.56).

(1) S_0 is $8\pi^2/g^2$.

(2) We have an eight-parameter family of solutions and thus eight eigenmodes of eigenvalue zero in the small-vibration problem. Thus K contains a factor of $(1/\sqrt{K})^8$, or, equivalently $1/g^8$. Everything else in K is independent of $\mathcal{K}$, and thus independent of g.

(3) We have already done the integral over instanton locations. The integral over constant gauge transformations is an integral over a compact group and thus gives only a constant numerical factor, the volume of SU(2). The integral over instanton sizes is potentially troublesome, since ρ can be anywhere between zero and infinity, so we will, for the moment, keep it as an explicit integral.

(4) Thus we obtain

$$E(\theta)/V = -\cos\theta\, e^{-2\pi^2/g^2}\, g^{-8}\int_0^\infty \frac{d\rho}{\rho^5}\, f(\rho M)\ , \qquad (3.72)$$

where f is an unknown function and M is the arbitrary mass (more properly, arbitrary inverse wavelength) that is needed to define the renormalization prescription in a massless field theory. (I have avoided mentioning renormalization until now, but renormalization is essential in any computation that involves an infinite number of eigenmodes, as does this one. In Sec. 5 I will give a more detailed discussion of the ultraviolet divergences in determinantal factors and their removal by the usual one-loop renormalization counterterms.) The form of the integral is determined by dimensional analysis; an energy density has dimensions of $1/(\text{length})^4$.

(5) However, M and g are not independent parameters. Renormalization-group analysis[22] tells us that they must enter expressions for observable quantities only in the combination

$$\frac{1}{g^2} \sim \beta_1 \ln M + O(g^2)\ , \qquad (3.73)$$

where β_1 is a coefficient which can be computed from one-loop perturbation theory. In the case at hand, β_1 is $11/12\pi^2$.

(6) This fixes the form of f. Thus,

$$E(\theta)/V = -A\cos\theta\, e^{-8\pi^2/g^2}\, g^{-8}\int_0^\infty \frac{d\rho}{\rho^5}\, (\rho M)^{8\pi^2\beta_1}[1+O(g^2)]\ , \qquad (3.74)$$

where A is a constant independent of g, ρ, and M.

(7) To determine A requires a lot of hard work,[23] so I'll stop the calculation here. Even though we haven't been able to carry things out to the end, it's remarkable how far we have been able to go with so little effort.

No doubt you have noticed that the integral we have derived is infrared divergent. The origin of the divergence is clear from the derivation of the integral: the effective coupling constant (in the sense of the renormalization group) becomes large for large instantons, and this makes the integrand blow up. Thus the divergence is an embarrassment but not a catastrophe. It would be a catastrophe if we obtained a divergent answer in a regime in which we trusted our approximations. This is not the situation here; the divergence arises in the regime of large effective coupling constant, where all small-coupling approximations are certainly wrong. Phrased another way, the fact that the integrand has the wrong behavior for large ρ is overshadowed by the fact that it is the wrong integrand. Thus we are free to hope that strong-coupling effects (which we can not at the moment compute) introduce some sort of effective infrared cutoff in the integrand. This hope might be wrong, but it is not ruled out by anything we have done so far.

I admit that this argument is blatant hand-waving. However, it is not some new hand-waving special to instanton calculations, but the same old hand-waving that accompanies any discussion of the large-scale behavior of non-Abelian gauge field theories. For example,[23] there is evidence that the observed hadrons are made of weakly-coupled quarks. But if the quarks are weakly coupled, why can't we knock them out of the hadron? Well, in a gauge field theory the effective coupling constant grows at large distances, etc.,

etc., much hand-waving, infrared slavery and quark confinement.

Everything that we have done for SU(2) can be extended straightforwardly to SU(3). To begin with, an SU(2) instanton solution can trivially be made into an SU(3) instanton solution; all that needs to be done is to say that three of the gauge fields, those associated with an SU(2) subgroup, are of the form given, while the other five vanish. It is believed that these exhaust the set of solutions of Eq. (3.63) with unit winding number, although, unlike the SU(2) case, there is, to my knowledge, no rigorous proof of this statement. If this is indeed the case, there are only two minor differences between the SU(3) computation and the SU(2) one:
(1) Instead of three parameters associated with constant gauge transformations, we have seven. [One of the eight SU(3) generators commutes with the SU(2) subgroup and does not change the solution.) Thus the factor of g^{-8} in Eq. (3.74) is replaced by one of g^{-12}.
(2) β_1 has the proper value for an SU(3) gauge theory, $11/8\pi^2$.

IV. THE ABELIAN HIGGS MODEL IN 1+1 DIMENSIONS[24]

In this section I will discuss a field theory in which instanton effects drastically change the particle spectrum, the Abelian Higgs model in two-dimensional space-time.

In any number of dimensions, this is the theory of a complex scalar field with quartic self-interactions, minimally coupled to an Abelian gauge field with gauge coupling constant e, called the electric charge. In our notation, the theory is defined by the Euclidean Lagrangian density,

$$\mathcal{L} = \frac{1}{4e^2}(F,F) + D_\mu\psi^* D_\mu\psi + \frac{\lambda}{4}(\psi^*\psi)^2 + \frac{\mu^2}{2}\psi^*\psi \ , \qquad (4.1)$$

where λ is a positive number and μ^2 may be either positive or negative. To this must be added renormalization counterterms; however, renormalization will play no part in our computations, and, to keep things as simple as possible, I will not distinguish between bare and renormalized parameters.

Perturbation theory tells us that for weak coupling the qualitative properties of the theory depend critically on the sign of μ^2:

(1) If μ^2 is positive, the theory is simply the electrodynamics of a charged scalar meson. The mass spectrum consists of the charged meson, its antiparticle, and a massless vector meson, the photon. The force between widely separated external charges is the ordinary Coulomb force. These statements require some modification in two dimensions. Firstly, because there are no transverse directions, there is no photon. Secondly, because the Coulomb force is independent of distance, it is impossible to separate a meson and an antimeson; in contemporary argot, the charged particles are confined. The spectrum of the theory consists of a sequence of meson-antimeson bound states, rather like the spectrum of positronium, except that these states are all stable, since they can not decay through the emission of (nonexistent) photons.

(2) If μ^2 is negative, the Higgs phenomenon takes place. In the ground state of the theory,

$$|\langle\psi\rangle|^2 = -\mu^2/\lambda \equiv a^2 \ . \qquad (4.2)$$

The particle spectrum consists of a massive neutral scalar meson and a massive neutral vector meson. The force between widely separated external charges falls off exponentially rapidly. These statements require no modification in two dimensions.

In the remainder of this section, I will argue that the preceding sentence is a lie; contrary to the predictions of perturbation theory, the qualitative properties of the model for negative μ^2 are the same as those for positive μ^2; the two-dimensional Abelian Higgs model does not display the Higgs phenomenon. To be precise, I will show that, for negative μ^2, the theory admits instantons, and, when the effects of these instantons are taken into account, the long-range force between external charges is independent of their separation. Also, I will be able to argue, from the behavior of the long-range force, that the theory contains (confined) charged particles. There is a quantitative difference between positive and

negative μ^2, though: For positive μ^2, the strength of the long-range force is independent of $\hbar$; for negative μ^2, the strength of the long-range force is exponentially small in $\hbar$, the mark of an instanton effect.

Just as in Sec. 3, we must begin the analysis by classifying classical field configurations of finite action. Of course, before doing this, we must add a constant to the Lagrangian density so the minimum of the action is zero. Thus we write

$$\mathcal{L} = \frac{1}{4e^2}(F,F) + |D_\mu \psi|^2 + \frac{\lambda}{4}(|\psi|^2 - a^2)^2 \ . \qquad (4.3)$$

This is the sum of three positive terms. In order that the third term not make a divergent contribution to the action, it is necessary that $|\psi|$ approach a as r goes to infinity. However, there is no restriction on the phase of ψ. In equations,

$$\lim_{r\to\infty} \psi(r,\theta) = g(\theta)a \ , \qquad (4.4)$$

where g is a complex number of unit modulus, an element of U(1). In order that the second term not make a divergent contribution to the action, it is necessary that

$$A_\mu = g\partial_\mu g^{-1} + O(1/r^2) \ . \qquad (4.5)$$

(Remember, in our conventions, A_μ is an imaginary field.) The first term now automatically makes a finite contribution to the action.

The lovely thing about Eq. (4.5) is that it is identical to Eq. (3.17); that is to say, the problem of classifying finite-action configurations is the baby problem of Sec. 3.2. Thus the finite-action configurations are characterized by an integer, ν, the winding number, just as they are for four-dimensional gauge field theories. By Eq. (3.28), the integral expression for the winding number is

$$\nu = \frac{i}{4\pi}\int d^2x\, \epsilon_{\mu\nu} F_{\mu\nu} \ . \qquad (4.6)$$

Equivalently,

$$\nu = \frac{i}{2\pi}\oint A_\mu \, dx_\mu \ , \qquad (4.7)$$

where the integral is over the circle at infinity.

Although I won't bother to explicitly display them here, it turns out that the Euclidean field equations have solutions with unit winding number, instantons, again just like four-dimensional gauge theories.[25] The only relevant difference, for our purposes, is that the Higgs model is not scale invariant; thus the instantons have a fixed size and the problems associated with integrating over scale transformations don't arise. Otherwise, though, everything is much the same as it was before, and we can copy step-by-step our earlier analysis and uncover the vacuum structure of the theory.

Thus, just as before, we have a family of θ-vacua, with energy densities given by

$$E(\theta)/L = -2Ke^{-S_0}\cos\theta \ . \qquad (4.8)$$

Here L is the volume of (one-dimensional) space, S_0 is the action of an instanton, and K is a determinantal factor. Also, by copying the derivation of Eq. (3.59), we find that

$$\langle\theta|\epsilon_{\mu\nu} F_{\mu\nu}|\theta\rangle = 8\pi K e^{-S_0}\sin\theta \ . \qquad (4.9)$$

As before, this has the right reality properties; when we continue to Minkowski space, we pick up a factor of i that cancels the factor of i in our definition of A_μ. We see from this equation that the θ-vacua are characterized by a constant expectation value of the electric field F_{01}. In two dimensions, unlike four, such a constant "background field" is not in conflict with Lorentz invariance.[26]

Now that we understand the vacuum structure, let's compute the force between widely separated external charges. To be more precise, let us introduce into the system two static charges of equal magnitude, q, and opposite sign, separated by a distance L', and let us compute (for large L') Δ, the change in the energy of a θ-vacuum caused by these charges. The standard method of computing Δ uses Wilson's loop integral,[27]

$$W = \exp\left(-\frac{q}{e}\oint A_\mu \, dx_\mu\right) \ , \qquad (4.10)$$

where the integration is over the rectangular path shown in Fig. 14. According to Wilson, the vacuum energy shift is given by

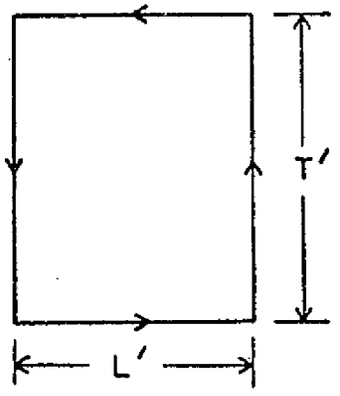


Figure 14

$$\Delta = -\lim_{T'\to\infty} \frac{1}{T'} \ell n\langle\theta|W|\theta\rangle \ . \tag{4.11}$$

In our case,

$$\langle\theta|W|\theta\rangle = \frac{\int [dA][d\psi^*][d\psi] W e^{-S} e^{i\nu\theta}}{\int [dA][d\psi^*][d\psi]\, e^{-S} e^{i\nu\theta}} \tag{4.12}$$

and our task is to compute these two functional integrals in our standard dilute-gas approximation, for large L' and T' (and, of course, for even larger L and T, the spatial and temporal extent of the universe). In Eq. (4.8) we have already calculated the denominator. To calculate the numerator, let us divide the sum over instantons and anti-instantons into two independent sums: one over objects lying inside the loop and one over objects lying outside the loop. By this division we neglect contributions coming from configurations in which instantons and anti-instantons overlap the loop, but, for large L, T, L', and T', this is a very small portion of the available configurations and can reasonably be neglected. (Of course, if our calculation gives zero for its answer, then these configurations will be the most important ones and we will have to go back and compute them.) The functional integrand splits neatly into the product of an "outside" term and an "inside" term: $S = S^{\text{outside}} + S^{\text{inside}}$, $\nu = \nu^{\text{outside}} + \nu^{\text{inside}}$, while

$$W = \exp(2\pi i q\, \nu^{\text{inside}}/e) \ . \tag{4.13}$$

Thus, for the outside objects, we have the same sum as for the denominator, except that the available volume of Euclidean two-space is not LT but LT - L'T'. For the inside objects, we also have the same sum, except that the available volume is L'T', and θ is replaced by $\theta + 2\pi q/e$.

Thus,

$$\begin{aligned}\ell n\langle\theta|W|\theta\rangle = 2Ke^{-S_0}[&(LT - L'T')\cos\theta \\ &+ L'T'\cos(\theta + 2\pi q/e) \\ &- LT\cos\theta] \ ,\end{aligned} \tag{4.14}$$

where the first term comes from the outside sum, the second from the inside sum, and the third from the denominator. Hence,

$$\Delta = 2L'Ke^{-S_0}[\cos\theta - \cos(\theta + 2\pi q/e)] \ . \tag{4.15}$$

This is proportional to L', the separation between the external charges; thus there is a constant force between external charges at large separation. As announced, there is no quantitative difference between positive and negative μ^2. However, there is a qualitative difference. For positive μ^2, the strength of the force is proportional to q^2 for small $\hbar$; for negative μ^2, it is exponentially small in $\hbar$. (Remember, if we had not chosen our units so $\hbar$ was one, S_0 would have been $S_0/\hbar$.)

There is a simple physical interpretation of this result. For small θ and small q/e,

$$\Delta = L'Ke^{-S_0}[(\theta + q/e)^2 - \theta^2] \ , \tag{4.16}$$

$$E(\theta) = LKe^{-S_0}\theta^2 + \text{constant} \ , \tag{4.17}$$

and,

$$\langle\theta|F_{12}|\theta\rangle = 4\pi Ke^{-S_0}\theta \ . \tag{4.18}$$

These expressions have an obvious interpretation: In a θ-vacuum, there is a background electric field, and an energy density proportional to the square of this field. Because we are in one

spatial dimension, the external charges act like condenser plates in three dimensions; they induce a constant field proportional to their charge in the region between them, which is added to the pre-existing background field. Thus the energy shift is the separation multiplied by the difference of the energy density of the new field and that of the old. Equation (4.14) is just this trivial picture complicated by non-linear terms in the expression for the energy density as a function of the field.

One aspect of these non-linear complications is of physical import: Eq. (4.14) is periodic in q with period e. This is explicable if the theory contains charged particles of charge e. If this is the case, there is a process that can change the charge on our condenser plates by ±e: a particle-antiparticle pair can materialize in the region between the plates, and the particle can fly to one plate and the antiparticle to the other. This process will occur whenever it is energetically favorable. For sufficiently large L', this is equivalent to saying that it will occur whenever it lowers the energy density, because the energetic cost of making a pair is independent of L', and the energetic gain of lowering the energy density is proportional to L'. Thus q's that are equal modulo e lead to identical physics; no matter which one you start out with, pairs are made until the charge on the plates reaches its optimum value, the one that gives minimum energy density.

What if we were to do a parallel computation in a four-dimensional gauge field theory, with non-Abelian external charges? Would we also obtain a force independent of separation? Alas, we would not. There is an L' in Eq. (4.15) because there is an L'T' in Eq. (4.14), that is to say, because even an instanton deep within the loop has a non-negligible effect on the loop integral. This is precisely what does not happen in four dimensions. At large distances from an instanton, A_μ is $g\partial_\mu g^{-1}$, plus terms that fall off far too rapidly to affect the loop integral. However, the loop integral is gauge-invariant, and we can always gauge-transform g such

that it is constant everywhere except within a small cone emerging from the instanton perpendicular to the plane of the loop. Whatever confines quarks, it's not instantons.

V. 'T HOOFT'S SOLUTION OF THE U(1) PROBLEM

5.1 The Mystery of the Missing Meson

The U(1) problem is an apparent contradiction between two pieces of accepted wisdom. One is wisdom of the '70's, that hadronic physics is quantum chromodynamics. The other is wisdom of the '60's, that hadronic physics is approximately invariant under chiral SU(2) ⊗ SU(2). Let me remind you of the meaning of these two propositions.

Quantum chromodynamics is a field theory whose dynamical variables are an octet of SU(3) gauge fields and a family of SU(3) triplet Dirac bispinor fields, called quarks. In Minkowski space, the Lagrangian density is

$$\mathcal{L} = -\frac{1}{4g^2}(F_{\mu\nu}\cdot F^{\mu\nu}) + \sum_f \bar{\psi}_f (i D_\mu \gamma^\mu - m_f)\psi_f \ . \qquad (5.1)$$

where f, called the flavor index, labels the various triplets. The usual exact and approximate symmetries of hadron physics [charge, isospin, Gell-Mann's SU(3), etc.] act only on the flavor indices; all physical hadrons are supposed to be singlets under the gauge group. (This last statement is sometimes called quark confinement; it is still far from proved, although there are some suggestive arguments.) ψ_1 and ψ_2 form an isodoublet, the non-strange quarks; ψ_3 is the strange quark; ψ_4 is the charmed quark; there may or may not be additional flavors.

Chiral SU(2) ⊗ SU(2) is the group generated by the strangeness-conserving weak-interaction currents and their parity transforms. Its diagonal subgroup is conventional isospin. This group is very close to being an exact symmetry of the strong interactions; it is a much better symmetry than SU(3) and roughly as good a symmetry as isospin. However, were this symmetry to be exact, only the isospin

subgroup would be a manifest symmetry; the remainder of the group would be a Nambu-Goldstone symmetry, with three massless Goldstone bosons, the pions. The smallness of the pion mass (on a hadronic mass scale) is a measure of the goodness of the symmetry. This is the picture that stands in back of all the stunningly successful soft-pion computations of the mid '60's.

Now for the apparent contradiction: In quantum chromodynamics, the limit of perfect SU(2) ⊗ SU(2) symmetry is the limit in which the non-strange quarks are massless. In this limit, the Lagrangian (4.1) obviously has a further chiral U(1) symmetry; it is invariant under

$$\psi_f \to e^{-i\alpha\gamma_5}\psi_f , \qquad (f = 1,2) \qquad (5.2)$$

where α is a real number. The associated conserved current is

$$j^5_\mu = \sum_{f=1}^{2} \bar{\psi}_f \gamma_\mu \gamma_5 \psi_f . \qquad (5.3)$$

I emphasize that the appearance of this additional chiral symmetry is very special to quantum chromodynamics; for example, the σ model has no such additional symmetry in the chiral limit.

Now, either this additional symmetry is manifest or it is spontaneously broken. If it were manifest, all non-massless hadrons would occur in parity doublets. This is not the case; thus it must be spontaneously broken. But if it is spontaneously broken, Goldstone's theorem tells us there must be an associated isoscalar pseudoscalar Goldstone boson. This is the U(1) problem: *What happened to the fourth Goldstone boson?*

One's first thought is that the missing meson is the eta, but this is wrong. The chiral U(1) symmetry is broken by the same mass term that breaks chiral SU(2) ⊗ SU(2), and thus the fourth Goldstone boson should have roughly the same mass as the pions. The eta is far too heavy. This can be made more precise: Using conventional soft-pion methods, Weinberg[28] has shown that a U(1) Goldstone boson must have a mass less than $\sqrt{3}\,m_\pi$. The eta grossly disobeys this inequality. Also, if we consider the approximation in which the

strange quark mass also vanishes, and in which we have perfect chiral SU(3) ⊗ SU(3) symmetry, the eta takes its place with the pions in an octet of Goldstone bosons. But in this limit we still have an additional U(1) symmetry and we still have a missing meson.

[This should be all that I need to say about the eta. However, there is some confusion abroad on this point, and thus I emphasize that *there is no connection between the eta and the U(1) problem.* The eta is a red herring; it is just another hadron; it is no more a relic of a U(1) Goldstone boson than is the N^{**}.]

It may seem that I have posed an insoluble problem; this is because I have lied to you. In fact, j^5_μ is not a conserved current; it is afflicted with the famous Adler-Bell-Jackiw anomaly.[29] In the limit of N massless quarks,

$$\partial^\mu j^5_\mu = \frac{N}{32\pi^2}\,\epsilon^{\mu\nu\lambda\sigma}(F_{\mu\nu}, F_{\lambda\sigma}) . \qquad (5.4)$$

(Note the similarity between the right-hand side of this equation and the Pontryagin density. This will be important to us later.)

You might think that this is the end of the story; if the current is not conserved, there is no U(1) symmetry to worry about. Alas, life is not so simple. In Sec. 3, we showed that the Euclidean counterpart of the right-hand side of Eq. (5.4) could be written as the divergence of a (gauge-variant) function of A_μ and $F_{\mu\nu}$. It is easy to see that the same construction works in Minkowski space. Thus, if we define

$$J^5_\mu = j^5_\mu - \frac{N}{16\pi^2}\,\epsilon_{\mu\nu\lambda\sigma}(A^\nu, F^{\lambda\sigma} - \tfrac{2}{3}A^\lambda A^\sigma) , \qquad (5.5)$$

this current is gauge-variant but conserved.

If we work in a covariant gauge (and why shouldn't we?), the added term commutes with the quark fields at equal times. Thus we can derive, for Green's functions made of one J^5_μ and a string of gauge-invariant quark multilinears, chiral U(1) Ward identities of the usual form. And since these are of the usual form, they lead to the usual conclusion: Chiral U(1) is a symmetry; either Green's

functions made of quark multilinears alone are U(1) symmetric, or there are Goldstone poles in Green's functions for one J^5_μ and a string of quark multilinears.

Is there no way out? Well, there is one. The Hilbert space of a gauge field theory quantitized in a covariant gauge is notoriously full of negative-norm timelike photons and similar gauge phantoms, states that never couple to gauge-invariant operators. Could it be that the Goldstone boson is such a phantom? No, this is not possible; the formulation of the question is wrong. If the Goldstone boson does not couple at all to gauge-invariant operators, it con't produce a pole in a Green's function for one J^5_μ and a string of gauge-invariant operators.

The proper formulation of the question was found by Kogut and Susskind,[30] who had the bright idea of looking at the Schwinger model, massless spinor electrodynamics, in 1+1 dimensions in a covariant gauge. The Schwinger model is an exactly soluble theory that has properties very close to those we have been discussing. In particular, there is a gauge-invariant axial current with an anomalous divergence and a gauge-variant conserved axial current, and, most important, there is chiral symmetry breakdown without Goldstone poles in gauge-invariant Green's functions. What Kogut and Susskind found in the covariant-gauge Schwinger model were two free massless fields, ϕ_+ and ϕ_-. ϕ_+ creates quanta of positive norm and has the usual propagator; ϕ_- creates quanta of negative norm and has minus the usual propagator. (Remember, a covariant gauge is full of negative-norm states from the very beginning.) All gauge-invariant quantities couple to the sum of these fields, $\phi_+ + \phi_-$; this has zero propagator and produces no singularities. Thus gauge-invariant Green's functions are free of Goldstone poles. However, the gauge-variant conserved current couples to the gradient of the difference, $\partial_\mu(\phi_+ - \phi_-)$. Thus, when one considers a Green's function for one gauge-variant current and a string of gauge-invariant fields, the relative minus sign in the coupling cancels

the relative minus sign in the propagators, and Goldstone poles appear where they should. This set-up is called a Goldstone dipole. (The terminology is a bit misleading, because there are only single poles in Green's functions, but I'll stick with it anyway.)

Thus according to Kogut and Susskind, the proper formulation of our question is, is the U(1) symmetry of quantum chromodynamics spontaneously broken via a Goldstone dipole? You might think that this is a question that could be asked seriously only by a field theorist driven mad by spending too many years in too few dimensions. Nevertheless, as 't Hooft[9] brilliantly showed, the answer is yes. The remainder of this section is an explanation of his computation.

5.2 Preliminaries: Euclidean Fermi Fields

Before we can treat quantum chromodynamics by functional integration, we must know how to integrate over Euclidean Fermi fields. This section is a description of the theory of such integration, with all mathematical fine points ruthlessly suppressed.[31] I will develop the theory by defining Fermi integration as a "natural" generalization of Bose integration. At the end, I will justify my definitions by showing that they lead to formulas equivalent to those obtained by conventional canonical quantization.

Let us begin by defining our integration variables. For Bose theories, we integrate over c-number Euclidean fields. These are objects that commute with each other at arbitrary separations; they can be thought of as the classical (vanishing $\hbar$) limit of quantum Bose fields. This suggests that the proper variables for a Fermi theory should be classical Fermi fields, objects which *anticommute* with each other at arbitrary separations. Thus, for example, for the theory of a single Dirac field, we would expect our integration variables to be two Euclidean bispinors, $\bar\psi$ and ψ, obeying

$$\{\psi(x),\psi(y)\} = \{\bar\psi(x),\bar\psi(y)\} = \{\psi(x),\bar\psi(y)\} = 0\,, \qquad (5.6)$$

for all Euclidean points x and y.

The last of these relations is crucial, for it implies that $\bar\psi$ can not be in any sense the adjoint of ψ times some matrix. For if

this were so, the last relation (multiplied by the inverse matrix) would state that the sum of two positive semi-definite objects, $\psi\psi^+$ and $\psi^+\psi$, was zero. This would only be possible if ψ vanished, not a happy situation for a prospective integration variable. Thus, if we are to have any hope of founding a sensible integration theory, we must treat ψ and $\bar{\psi}$ as *totally independent variables*.

This independence is the main novelty of Euclidean Fermi fields; the rest of the construction is straightforward. We define the Euclidean γ-matrices to be four Hermitian matrices obeying

$$\{\gamma_\mu,\gamma_\nu\} = 2\delta_{\mu\nu} \quad . \tag{5.7}$$

We use these to define the O(4) transformation law for ψ in the usual way, and define $\bar{\psi}$ to transform like the adjoint of ψ. We define γ_5, a Hermitian matrix, by

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4 \quad . \tag{5.8}$$

Thus, $\bar{\psi}\psi$ is a scalar, $\bar{\psi}\gamma_5\psi$ a pseudoscalar, $\bar{\psi}\gamma_\mu\psi$ a vector, etc.

The Euclidean action for a free Dirac field is

$$S = -\int d^4x\bar{\psi}(i\partial_\mu\gamma_\mu - im)\psi \quad . \tag{5.9}$$

The minus sign is pure convention; we could always absorb it into ψ if we wanted to. (Remember, we are free to transform ψ without touching $\bar{\psi}$.) The i in front of the mass term is not conventional. It is there to insure that the Euclidean propagator is proportional to $(\not{p}+im)/(p^2+m^2)$; if it were not for the i, we would have tachyon poles. If m vanishes, Eq. (5.9) is invariant under chiral transformations,

$$\psi \to e^{-i\alpha\gamma_5}\psi, \quad \bar{\psi} \to \bar{\psi}e^{-i\alpha\gamma_5} \quad . \tag{5.10}$$

The quark part of the Euclidean action for quantum chromodynamics is obtained from Eq. (5.9) by replacing ordinary derivatives by covariant derivatives.

So much for the integrand; now for the integration. For Bose fields, we defined functional integration as iterated integration over ordinary numbers. Therefore, let us begin by defining

integration for a function of a single anticommuting quantity, a. (Of course, for a single quantity, the anticommutation algebra degenerates to a single equation, $a^2 = 0$.)

We want to define $\int da f(a)$, (5.11)

for an arbitrary function, f. We want this to have the usual linearity property: the integral of a linear combination of two functions should be the linear combination of the integrals. In addition, we would like the integral to be translation-invariant

$$\int da f(a+b) = \int da f(a) \tag{5.12}$$

where b is an arbitrary anticommuting quantity. I will now show that these conditions determine the integral, up to a normalization factor.

The reason is that there are only two linearly independent functions of a, 1 and a; all higher powers vanish. We will choose our normalization such that $\int da\, a = 1$. (5.13)

From this, and Eq. (5.12), $\int da\, 1 = 0$. (5.14)

For functions of many anticommuting variables, we define multiple integrals as iterated single integrals. Thus, for example, a complete integration table for the four linearly independent functions of two anticommuting variables, a and $\bar{a}$, is

$$\int da d\bar{a} \begin{Bmatrix} \bar{a}a \\ \bar{a} \\ a \\ 1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \quad . \tag{5.15}$$

As an application of this table, I will evaluate

$$\int da d\bar{a}\, e^{\lambda\bar{a}a} = \int da d\bar{a}(1+\lambda\bar{a}a) = \lambda \quad . \tag{5.16}$$

We can now define integration over Fermi fields exactly as we

defined integration over Bose fields in Sec. 2. We introduce two arbitrary complete orthonormal sets of c-number functions, ψ_r and $\overline{\psi}_r$,

$$\int d^4x\, \psi_r^+ \psi_s = \int d^4x\, \overline{\psi}_r \overline{\psi}_s^+ = \delta_{rs} \ . \tag{5.17}$$

We expand the Fermi fields in terms of these functions,

$$\psi = \sum_r a_r \psi_r \ , \qquad \overline{\psi} = \sum_r \overline{a}_r \overline{\psi}_r \ , \tag{5.18}$$

and define

$$[d\psi][d\overline{\psi}] = \prod_r da_r\, d\overline{a}_r \ . \tag{5.19}$$

As an application let me evaluate

$$\int [d\psi][d\overline{\psi}] e^{-S} \ , \tag{5.20a}$$

where

$$S = -\int d^4x\, \overline{\psi} A \psi \ , \tag{5.20b}$$

and A is some linear operator, possibly depending on external c-number fields. For simplicity, let me assume that A commutes with A^+. (This is the case for a quark in an external gauge field.) Then we can choose the ψ_r's to be the eigenfunctions of A,

$$A\psi_r = \lambda_r \psi_r \ , \tag{5.21}$$

and we can choose $\overline{\psi}_r$ to be ψ_r^+. Thus

$$S = -\sum_r \lambda_r \overline{a}_r a_r \ , \tag{5.22}$$

and

$$\int [d\psi][d\overline{\psi}] e^{-S} = \prod_r \lambda_r = \det A \ . \tag{5.23}$$

Note that this is the inverse of the answer we would have obtained had we done the identical integral with ψ and $\overline{\psi}$ complex Bose fields.

I will now show that Eq. (5.22) is the correct answer, that it is identical to the normal field-theoretic expression for the vacuum-to-vacuum transition amplitude in a theory of a quantized Dirac field interacting with external c-number fields. In this theory, this amplitude is the sum of all Feynman graphs with no external Fermi lines. This in turn is the exponential of the sum of

all connected (that is to say, one-loop) graphs. Now, if ψ were a Bose field, we know that the amplitude would be the inverse determinant, because we trust functional integration for Bose fields. But the only effect of replacing bosons by fermions is to multiply the one-loop graphs by minus one. This inverts the exponential of the one-loop graphs, that is to say, it turns the inverse determinant into the determinant.

In any theory in which the Fermi fields enter the action at most bilinearly, we can always integrate over the Fermi fields, using Eq. (5.23), before we integrate over the Bose fields. In diagrammatic language, we can always sum the Fermi loops before we integrate over virtual bosons. Thus, because our definition of Fermi integration gives the right answer for a Dirac field in an external c-number fields, it also gives the right answer for a Dirac field interacting with a quantum Bose field. In particular, it gives the right answer for quantum chromodynamics.

5.3 Preliminaries: Chiral Ward Identities

In this section is a discussion of the chiral Ward identities for a theory of a set of quantum Dirac fields interacting with c-number gauge fields. In the sequel, we shall use these identities in several different cases; thus it's useful to have them written down in their most general form, at hand when we need them.

Let ψ be a set of Euclidean Dirac fields, assembled into a big vector, which transforms according to some representation of SU(n), not necessarily irreducible, generated by a set of matrices, T^a. Let us define the constant C by

$$\mathrm{Tr}\, T^a T^b = -C\delta^{ab} \ . \tag{5.24}$$

Thus, for example, for a set of N fields each transforming according to the n-dimensional representation of SU(n),

$$C = N/2 \ . \tag{5.25}$$

We wish to study the theory of these fields interacting with given c-number gauge fields,

$$S = -i\int d^4x\,\bar\psi(\gamma_\mu D_\mu - M)\psi\ , \tag{5.26}$$

where D_μ is the covariant derivative defined by Eq. (3.16), and M is the mass matrix for the Dirac fields, assumed to be SU(n)-invariant. Let $\phi^{(r)}$, r = 1...m, be a set of local multilinear functions of the Dirac fields. The Euclidean Green's functions for these objects are defined by

$$\langle\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A = \frac{\int[d\psi][d\bar\psi]e^{-S}\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)}{\int[d\psi][d\bar\psi]e^{-S}} \tag{5.27}$$

where I have inserted the superscript A to remind you that we are working in an external gauge field.

Now let us perform an infinitesimal change of variables in the numerator of Eq. (5.27),

$$\delta\psi = -i\gamma_5\psi\delta\alpha,\quad \delta\bar\psi = -i\bar\psi\gamma_5\delta\alpha\ , \tag{5.28a}$$

where $\delta\alpha$ is an infinitesimal function of Euclidean space. Since the ϕ's are functions of the Dirac fields, they will change under the change of variables; we define $\partial\phi^{(r)}/\partial\alpha$ by

$$\delta\phi^{(r)} = (\partial\phi^{(r)}/\partial\alpha)\delta\alpha\ . \tag{5.28b}$$

Thus, for example, $\partial\bar\psi\psi/\partial\alpha$ is $-2i\bar\psi\gamma_5\psi$. A change of variables does not change the integral; thus, taking the variational derivative with respect to $\delta\alpha$, we find

$$\begin{aligned}&\partial^\mu\langle j^5_\mu(y)\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A\\ &+\langle\bar\psi M\gamma_5\psi(y)\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A\\ &+\delta^{(4)}(y-x_1)\langle\partial\phi^{(1)}(x_1)/\partial\alpha\ldots\phi^{(m)}(x_m)\rangle^A\\ &+\ldots\\ &+\delta^{(4)}(y-x_m)\langle\phi^{(1)}(x_1)\ldots\partial\phi^{(m)}(x_m)/\partial\alpha\rangle^A = 0\ ,\end{aligned} \tag{5.29}$$

where j^5_μ is $\bar\psi\gamma_\mu\gamma_5\psi$.

These are, of course, just the Euclidean version of the Ward identities we would have obtained in Minkowski space by studying the divergence of j^5_μ, and, of course, they are wrong, for they take no account of the Adler-Bell-Jackiw anomaly. I don't have the time here to recapitulate the theory of the anomaly, and I will simply state the correct version of Eq. (5.29): The zero on the right-hand side is replaced by

$$-\frac{iC}{8\pi^2}\,(F(y),\tilde F(y))\langle\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A\ . \tag{5.30}$$

We can obtain a very useful equation by integrating the corrected Ward identity over y. The first term on the left vanishes by integration by parts; the theory contains no massless particles that could give a non-vanishing surface term. Also, on the right we can use

$$\int d^4y(F,\tilde F) = 32\pi^2\nu\ . \tag{3.47}$$

Thus we obtain

$$\begin{aligned}&2\langle\int d^4y\,\bar\psi M\gamma_5\psi(y)\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A + \frac{\partial}{\partial\alpha}\langle\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A\\ &\qquad = -4iC\nu\langle\phi^{(1)}(x_1)\ldots\phi^{(m)}(x_m)\rangle^A\ .\end{aligned} \tag{5.31}$$

Now all our artillery is at the ready; we can begin our assault on quantum chromodynamics.

5.4 QCD (Baby Version)

I will begin by analyzing a baby version of quantum chromodynamics, in which the gauge group is SU(2), and in which there is only a single isodoublet quark, of mass zero. In equations,

$$S = \int d^4x\left[\frac{1}{4g^2}\,(F,F) - i\bar\psi D_\mu\gamma_\mu\psi\right]\ . \tag{5.32}$$

After we have worked out the baby theory, we will go on to the real thing.

Most of the analysis of Sec. 3 is essentially unaltered by the presence of a quark. In particular, all of our old instanton solutions are still solutions of the Euclidean equations of motion (with the quark fields set equal to zero). Thus we still have all the θ-vacua, and formulas like

$$E(\theta)/V = -2K\cos\theta\, e^{-S_0}\ , \tag{3.56}$$

and $$\langle\theta|(F,\tilde{F})|\theta\rangle = -64\pi^2 i K e^{-S_0}\sin\theta\,, \tag{3.59}$$

remain unaltered. The only effect of the quarks is to insert into the definition of K a term proportional to

$$\det\left[\frac{i\not{D}}{i\not{\partial}}\right] = \det\left[\frac{i(\partial_\mu + A_\mu)\gamma_\mu}{i\partial_\mu\gamma_\mu}\right]\,, \tag{5.33}$$

where A_μ is the field of an instanton.

This is a trifling alteration, but it is a tremendous trifle, for, as we shall see, $i\not{D}$ has a vanishing eigenvalue. Thus the determinant vanishes, as does $E(\theta)/V$ and $\langle\theta|(F,\tilde{F})|\theta\rangle$!

The vanishing eigenvalue can be demonstrated either by a short explicit computation or by a long indirect argument. I will choose the second method. Despite what you might think, this is not a perverse choice. (Well, not totally perverse.) The indirect argument will have some byproducts that will be very useful to us later.

For simplicity, I will assume (falsely) that $i\not{D}$ has a purely discrete spectrum,[12]

$$i\not{D}\psi_r = \lambda_r\psi_r\,. \tag{5.34}$$

Because $i\not{D}$ is Hermitian, all the λ's are real. Because γ_5 anticommutes with γ_μ,

$$i\not{D}\gamma_5\psi_r = -\lambda_r\gamma_5\psi_r\,. \tag{5.35}$$

Thus nonvanishing eigenvalues always occur in pairs of opposite sign. Eigenfunctions of vanishing eigenvalue, on the other hand, can always be chosen to be eigenfunctions of γ_5,

$$\gamma_5\psi_r = \chi_r\psi_r\,, \qquad (\lambda_r = 0) \tag{5.36}$$

Because $\gamma_5^2 = 1$, $\chi_r = \pm 1$. I will denote the number of eigenfunctions of these two types by $n_\pm$.

I will now prove the remarkable sum rule,[13]

$$n_- - n_+ = \nu\,. \tag{5.37}$$

Thus, not only is there a zero eigenvalue in the field of an instanton, there is a zero eigenvalue in any gauge field of non-zero winding number, whether or not it is a solution of the Euclidean equations of motion.

The proof rests on the chiral Ward identities for the quantum theory of a *massive* quark interacting with an external gauge field.

$$S = -i\int d^4x\,\bar{\psi}(\not{\partial} - m)\psi\,. \tag{5.38}$$

If we take the case of *no* ϕ's, Eq. (5.31) becomes

$$-2i\nu = 2\langle\int d^4y\,\bar{\psi}m\gamma_5\psi\rangle^A = \frac{2\int[d\psi][d\bar{\psi}]e^{-S}\int d^4y\,\bar{\psi}m\gamma_5\psi}{\int[d\psi][d\bar{\psi}]e^{-S}}\,. \tag{5.39}$$

(Remember, in the case at hand, $C = \frac{1}{2}$.) To evaluate the functional integrals, we need the eigenfunctions and eigenvalues of $i(\not{\partial} - m)$. The eigenfunctions are those of $\not{\partial}$, and the eigenvalues are simply shifted by $-im$,

$$i(\not{\partial} - m)\psi_r = (\lambda_r - im)\psi_r\,. \tag{5.40}$$

If we expand the fields in the ψ_r's, the functional integrals become trivial, and we obtain

$$-2i\nu = \frac{2m\sum_r\int d^4y\,\psi_r^+\gamma_5\psi_r\prod_{s\neq r}(\lambda_s - im)}{\prod_r(\lambda_r - im)} = 2m\sum\int d^4y\,\psi_r^+\gamma_5\psi_r(\lambda_r - im)^{-1}\,. \tag{5.41}$$

Because eigenfunctions of a Hermitian operator with different eigenvalues are orthogonal,

$$\int d^4y\,\psi_r^+\gamma_5\psi_r = 0 \quad \text{if } \lambda_r \neq 0\,, \tag{5.42}$$

while

$$\int d^4y\,\psi_r^+\gamma_5\psi_r = \chi_r \quad \text{if } \lambda_r = 0\,. \tag{5.43}$$

Thus,

$$-2i\nu = 2i(n_+ - n_-)\,. \tag{5.44}$$

This is the desired result.

It turns out that the instanton obeys the sum rule by having one eigenfunction of vanishing eigenvalue with $\chi = -1$ and none with $\chi = +1$. (This also can be seen indirectly, without dirtying one's hands with explicit computations; see Appendix E.) We'll never

need the explicit form of the eigenfunction, but, just for completeness, I'll write it down here. For an instanton with center at X and size ρ,

$$\psi_0(x-X,\rho) = \rho[\rho^2 + (x-X)^2]^{-\frac{3}{2}}u , \qquad (5.45)$$

where u is a constant spinor. Likewise, for an anti-instanton, there is one eigenfunction of vanishing eigenvalue with X = +1, the parity transform of Eq. (5.45). For n widely separated instantons and anti-instantons, there are n such eigenfunctions, one centered about each object. (More properly, I should say that there are n approximate eigenfunctions with approximately vanishing eigenvalues, but, for the dilute-gas approximation, the qualifications are irrelevant.)

What is important for our purposes is that the sum rule implies that any field configuration with nonvanishing winding number has at least one eigenfunction of vanishing eigenvalue and thus a vanishing Fermi determinant. Thus, not just in the dilute gas approximation, but to *all* orders in the semiclassical expansion, all the θ vacua have the same energy and they all have a vanishing expectation value for $(F,\tilde{F})$.

A phenomenon this general must have a deep cause. We can discover this cause if we consider the chiral Ward identities for vanishing quark mass. There is a technical obstacle to this; for vanishing quark mass, the denominator in Eq. (5.27) vanishes, at least for fields with $\nu \neq 0$. This is easily surmounted; we define denominator-free Green's functions,

$$\langle\langle\phi^{(1)}(x_1)\ldots\rangle\rangle^A \equiv \int[d\psi][d\bar{\psi}]e^{-S}\phi^{(1)}(x_1)\ldots \qquad (5.46)$$

By the same reasoning as before, these obey the Ward identities,

$$[\frac{\partial}{\partial\alpha} + 2i\nu]\langle\langle\phi_1(x_1)\ldots\rangle\rangle^A = 0 , \qquad (5.47)$$

i.e., Eq. (5.31) without the mass term. The Green's functions of our baby version of chromodynamics are given by

$$\langle\theta|\phi^{(1)}(x_1)\ldots|\theta\rangle = \frac{\int[dA]e^{-S_g}e^{i\nu\theta}\langle\langle\phi^{(1)}(x_1)\ldots\rangle\rangle^A}{\int[dA]e^{-S_g}e^{i\nu\theta}\langle\langle 1\rangle\rangle^A} , \qquad (5.48)$$

where S_g is the gauge-field part of the action. By Eq. (5.47),

$$\left[\frac{\partial}{\partial\alpha} + 2\frac{\partial}{\partial\theta}\right]\langle\theta|\phi^{(1)}(x_1)\ldots|\theta\rangle = 0 . \qquad (5.49)$$

Thus, the effect of a chiral U(1) transformation can be undone by a change of θ. That is to say, chiral U(1) transformations turn one θ vacuum into another; chiral U(1) symmetry is spontaneously broken, and the θ vacua are the many vacua that appear when a symmetry suffers spontaneous breakdown. This is startling; after all, when we first met the θ vacua in Sec. 3, they had no connection with chiral symmetry -- there was no chiral symmetry for them to be connected with! Nevertheless, it is an inevitable result of our analysis, and it explains why all the θ vacua have the same energy density and the same expectation value of $(F,\tilde{F})$; it is because these quantities are chiral U(1) invariants.

[Parenthetical remark: The factor of 2 in Eq. (5.49) is worth comment. It tells us that when we make a chiral rotation by π we return to the same θ vacuum. This is as it should be.

$$e^{-i\pi\gamma_5} = -1 . \qquad (5.50)$$

Thus a chiral rotation by π has the same effect on the fields as a spatial rotation by 2π; we would be very unhappy if this symmetry suffered spontaneous breakdown.]

There is one possible loophole in the argument I have given: It remains a logical possibility that, for every Green's function, the derivative with respect to α and the derivative with respect to θ both vanish. If this happened, we would have, not spontaneous symmetry breakdown, but manifest symmetry, and the θ vacua would be mathematical artifacts, superfluous duplicates of a single vacuum.

I will now eliminate this possibility by computing, in the dilute-gas approximation

$$\langle\theta|\sigma_\pm(x)|\theta\rangle = \frac{\int[dA][d\psi][d\bar{\psi}]e^{-S}e^{i\nu\theta}\sigma_\pm(x)}{\int[dA][d\psi][d\bar{\psi}]e^{-S}e^{i\nu\theta}} \qquad (5.51)$$

where
$$\sigma_{\pm} = \tfrac{1}{2}\bar{\psi}(1 \pm \gamma_5)\psi \ . \tag{5.52}$$

These are chiral eigenfields,

$$\partial\sigma_{\pm}/\partial\alpha = \mp 2i\sigma_{\pm} \ . \tag{5.53}$$

Thus, if we obtain a non-zero answer, we will know that spontaneous symmetry breakdown has occurred.

The computation will parallel closely that of the vacuum energy of a pure gauge field theory in Sec. 3. Indeed, as the calculation proceeds, we will accumulate all the terms that led to our earlier expression for the determinantal factor, K, as an integral over instanton size, ρ,

$$K = 2g^{-8}\int_0^{\infty}\frac{d\rho}{\rho^5}\, f\,(\rho M) \ , \tag{5.54}$$

where M is the renormalization mass. As these old terms come up, I'll call them to your attention, but I won't bother to write them down; I'll keep explicit track only of new terms that modify the integrand in Eq. (5.54).

There is one important novelty in the dilute-gas approximation. For n widely separated instantons and anti-instantons, $i\not{D}$ has n vanishing eigenvalues. Thus the integral over Fermi fields will vanish unless the integrand contains

$$\prod_{\lambda_r=0} \bar{a}_r a_r \ . \tag{5.55}$$

Such a term can appear only if we are computing a Green's function involving at least 2n Dirac fields. Hence, for any fixed Green's function, the potentially infinite sum over instantons and anti-instantons terminates.

I will first do the σ_- computation:

In the denominator of Eq. (5.51), the only configuration that does not have a surplus of vanishing eigenvalues is one of no instantons and no anti-instantons, that is to say, the classical vacuum, $A_\mu = 0$. Thus the denominator is simply the product of a Bose determinant and a Fermi determinant. The same Bose determinant appeared in the denominator in our earlier computation. The Fermi determinant, $\det(i\not{\partial})$, is a new factor.

In the numerator, we need a configuration with $\nu = 1$, by Eq. (5.47). The only one that does not have a surplus of vanishing eigenvalues is one instanton and no anti-instantons. Let us do the Fermi integral first; this gives

$$\tfrac{1}{2}\psi_0^{\dagger}(x-X,\rho)(1-\gamma_5)\psi_0(x-X,\rho)\prod_{\lambda_r\neq 0}\lambda_r = \psi_0^{\dagger}\psi_0(x-X,\rho)\det'(i\not{D}) \ , \tag{5.56}$$

where det', as always, denotes a determinant with vanishing eigenvalues removed. The Bose integral gives a determinant and a bunch of collective-coordinate factors identical to those that go into K. Because $\det'(i\not{D})$ does not depend on X, the integration over the instanton location is trivial,

$$\int d^4X\, \psi_0^{\dagger}\psi_0(x-X) = 1 \ . \tag{5.57}$$

Finally, we have a factor of $e^{-8\pi^2/g^2}$ from the instanton action, and a factor of $e^{i\theta}$ from the $e^{i\nu\theta}$.

The σ_+ computation is almost identical to the σ_- one; the only difference is that the relevant configuration is one anti-instanton, and thus, instead of a factor of $e^{i\theta}$, we have one of $e^{-i\theta}$.

Putting all this together, we find

$$\langle\theta|\sigma_{\pm}(x)|\theta\rangle = e^{-8\pi^2/g^2}\, e^{\mp i\theta}\, g^{-8}\, 2\int_0^{\infty}\frac{d\rho}{\rho^5}\, f(\rho M)\frac{\det'(i\not{D})}{\det\,(i\not{\partial})} \ . \tag{5.58}$$

(In case you've lost track of the meaning of my symbols, I remind you that $i\not{D}$ is the Dirac operator in the field of an instanton of size ρ.)

Just as before, we can use dimensional analysis to study the integrand in this formula. The eigenvalues of $i\not{D}$ have the dimensions of 1/length. One eigenvalue has been removed from the primed determinant; thus the ratio det'/det has dimensions of length, and must be of the form

$$\frac{\det'(i\not{D})}{\det\,(i\not{\partial})} = \rho\, h(\rho M) \ , \tag{5.59}$$

where h is an unknown function. Note that this gives the right dimensions for the expectation values of $\sigma_\pm$, $1/(\text{length})^3$.

From here on the argument is a rerun of that of Sec. 3: We can use the renormalization group to determine the form of the integrand up to an arbitrary multiplicative constant, be embarrassed in the infrared, wave our hands about new physics giving an effective infrared cutoff, etc.

We now know spontaneous symmetry breakdown occurs. Are there Goldstone bosons? Let's look for them in

$$\langle\theta|\sigma_+(x)\sigma_-(0)|\theta\rangle \ . \tag{5.60}$$

By reasoning which should now be familiar to you, only two field configurations are relevant: $A_\mu = 0$, and one instanton plus one anti-instanton. The first of these just gives the usual one-loop perturbation theory expression; this has a two-quark cut, but no Goldstone pole. The second just gives the product $\langle\theta|\sigma_+|\theta\rangle\langle\theta|\sigma_-|\theta\rangle$. This also has no Goldstone pole. By similar methods one can investigate other gauge-invariant Green's functions, such as $\langle\theta|j^5_\mu\sigma_\pm|\theta\rangle$ or $\langle\theta|j^5_\mu j^5_\nu|\theta\rangle$, and again find no Goldstone poles, but really there is no need to do these computations. If Goldstone bosons appear anywhere, they should appear in (5.60), and they don't.

In the last sentence, I should have said not "appear anywhere", but "appear among the physical states", that is to say, as singularities in gauge-invariant Green's functions. The situation is very different if we study a gauge-variant Green's function such as

$$\langle\theta|J^5_\mu(x)\sigma_-(0)|\theta\rangle = \langle\theta|j^5_\mu(x)\sigma_-(0)|\theta\rangle + \frac{1}{16\pi^2}\langle\theta|G_\mu(x)\sigma_-(0)|\theta\rangle \ , \tag{5.61}$$

where G_μ is defined in Eq. (3.41). As I have said, the first of the terms on the right has no Goldstone pole, but, as I will show, the second does. The argument is simple: In a covariant gauge, there is a Goldstone pole if and only if

$$\int d^4x\, \partial_\mu \langle\theta|G_\mu(x)\sigma_-(0)|\theta\rangle \neq 0 \ . \tag{5.62}$$

If we use the identity,

$$\int d^4x\, \partial_\mu G_\mu = 32\pi^2\nu \ , \tag{5.63}$$

and the fact that the only configurations that contribute to (5.62) have $\nu = 1$, we find

$$\int d^4x\, \partial_\mu \langle\theta|G_\mu(x)\sigma_-(0)|\theta\rangle = 32\pi^2\langle\theta|\sigma_-|\theta\rangle \neq 0 \ . \tag{5.64}$$

On the other hand, for $\langle\theta|J_\mu J_\nu|\theta\rangle$, the contributing configurations have vanishing ν, and thus there is no Goldstone pole.

To summarize, we have found in the dilute-gas approximation: spontaneous breakdown of chiral U(1) symmetry, no Goldstone poles in gauge-invariant Green's functions, no Goldstone poles in the propagator of a gauge-variant conserved current, and a Goldstone pole in the Green's function for one gauge-variant current and one gauge-invariant operator. This is the Goldstone dipole of Kogut and Susskind.

5.5 QCD (The Real Thing)

Real quantum chromodynamics in the chiral SU(2) ⊗ SU(2) limit differs from our baby version in two respects. Firstly, we have triplet quarks with gauge group SU(3) rather than doublet quarks with gauge group SU(2). Secondly, we have two massless quarks, rather than one. [I will ignore the massive quarks; they are irrelevant to the U(1) problem.]

Replacing an SU(2) doublet by an SU(3) triplet makes hardly any change. If this were the only difference, we would still have instantons, and the constant C of Eq. (5.25) would still be ½; the only thing we would need to change in Sec. 5.4 would be the integral over instanton size, where g^{-8} would become g^{-12}.

In contrast, replacing one massless triplet by two makes a profound change. C is doubled, and thus the sum rule (5.37) is changed to

$$n_- - n_+ = 2\nu \ . \tag{5.65}$$

Hence, $i\not{D}$ in an instanton field has two vanishing eigenvalues rather than one. (We don't really need a fancy sum rule to see this; we

have two independent quark fields, so every eigenvalue occurs twice, once for ψ_1 and once for ψ_2.) Thus, two fields no longer suffice to take care of all the vanishing eigenvalues, and all quark bilinears have zero expectation values.

This is no obstacle to demonstrating the spontaneous breakdown of chiral U(1) symmetry; we just have to study quadrilinears rather than bilinears. For example, the same computation that before gave a non-vanishing expectation value for $\bar{\psi}_1(1-\gamma_5)\psi_1$ will now give a non-vanishing expectation value for $\bar{\psi}_1(1-\gamma_5)\psi_1\bar{\psi}_2(1-\gamma_5)\psi_2$.

There is a reason for this. We have found spontaneous breakdown of chiral U(1), but not of chiral SU(2) ⊗ SU(2); the θ vacua are all invariant under chiral SU(2) ⊗ SU(2). [There are two ways to see this: (1) There are too few θ vacua for them to be anything but invariant; for spontaneous breakdown of chiral SU(2) ⊗ SU(2) we need at least a three-parameter family of vacua. (2) Chiral U(1) transformations are connected to θ by the anomalous divergence of the isosinglet axial current; the isotriplet axial current is anomaly-free.] All Lorentz-invariant quark bilinears transform according to the representation $(\frac{1}{2},\frac{1}{2})$ of SU(2) ⊗ SU(2), and must have vanishing expectation values. However, there are quadrilinear SU(2) ⊗ SU(2) singlets, such as

$$\tfrac{1}{4}\varepsilon_{ij}\varepsilon_{k\ell}\bar{\psi}_i(1-\gamma_5)\psi_k\bar{\psi}_j(1-\gamma_5)\psi_\ell = \bar{\psi}_1(1-\gamma_5)\psi_1\bar{\psi}_2(1-\gamma_5)\psi_2 - \bar{\psi}_1(1-\gamma_5)\psi_2\bar{\psi}_2(1-\gamma_5)\psi_1 . \tag{5.66}$$

These operators can have non-vanishing expectation values.

The doubling of C also changes Eq. (5.49) to

$$\left[\frac{\partial}{\partial\alpha} + 4\,\frac{\partial}{\partial\theta}\right]\langle\theta|\phi_1(x_1)\ldots|\theta\rangle = 0 . \tag{5.67}$$

Thus a chiral rotation by π/2, rather than π,

$$\psi_{1,2} \to -i\gamma_5\psi_{1,2} , \tag{5.68}$$

returns us to the same θ vacuum. Again, this is an effect of unbroken SU(2) ⊗ SU(2). If we multiply this by the SU(2) ⊗ SU(2) transformation,

$$\psi_1 \to -i\gamma_5\psi_1 , \quad \psi_2 \to i\gamma_5\psi_2 , \tag{5.69}$$

we obtain

$$\psi_1 \to \psi_1 , \quad \psi_2 \to -\psi_2 , \tag{5.70}$$

which should not be spontaneously broken.

Of course, we don't want unbroken SU(2) ⊗ SU(2) in quantum chromodynamics; we want spontaneous breakdown; we want pions. However, there is no reason to be disturbed that pions have not emerged from our computations. Our methods are semiclassical, valid in the limit of vanishing ħ, in principle capable only of revealing those phenomenon that occur for arbitrarily weak coupling. We have learned that the breakdown of chiral SU(2) ⊗ SU(2) is not such a phenomenon. This is no surprise. What is a surprise (and a wonderful surprise) is that the breakdown of chiral U(1) *is* such a phenomenon.

5.6 Miscellany[36]

There are some topics that I do not have the time to discuss in the detail they deserve but which I can not resist mentioning:

(1) For most theories with spontaneous symmetry breakdown, symmetry is restored at sufficiently high temperatures. Is this true here? This is an easy question to answer. Finite-temperature Green's functions are given by functional integrals over a Euclidean time inversely proportional to the temperature, with periodic time boundary conditions for Bose fields and antiperiodic ones for Fermi fields. Thus, as the temperature goes up, instantons of any given size eventually get squeezed out; there's no way to fit them into the available region of Euclidean space. However, no matter how high the temperature, there are always instantons so small that they barely notice the time boundary conditions. Thus, although assymmetries go to zero as a (calculable) power of the inverse temperature, symmetry is never fully restored. For extremely high temperatures, the only relevant instantons are so small that the effective coupling constant is extremely weak; thus we could make numerical computations of extreme accuracy, but only in a regime that is

totally inaccessible to experiment. I stress that this persistence of symmetry breakdown is a reflection of the scale invariance of classical chromodynamics, not of any property of instanton effects in general. For example, in the model of Sec. 4, there is a definite instanton size, and thus, at sufficiently high temperatures, all instanton effects disappear.

(2) Callan, Dashen, and Gross[1] have recently proposed a detailed picture of the dynamic structure of quantum chromodynamics. To explain their ideas, let me restrict myself to chromodynamics with two massless quarks, and let me imagine the universe cooling down from a very high temperature. Then, according to Callan, Dashen, and Gross:

(a) At very high temperatures, when the effective coupling constant is very small, chiral U(1) is spontaneously broken by instantons, but chiral SU(2) ⊗ SU(2) is still a good symmetry, and quarks are still unconfined. (Of course, this part is the standard picture which I have described in detail.)

(b) At somewhat lower temperatures, the effective coupling constant grows larger, and chiral SU(2) ⊗ SU(2) suffers spontaneous breakdown.[35] This is also an instanton effect, but an indirect one that can not be seen in the dilute gas approximation. Nevertheless, the effective coupling constant, although not tiny, is still small enough so that weak-coupling approximations are fairly reliable. (This part looks good to me.) Quarks are still unconfined.

(c) As still lower temperatures, and still larger effective couplings, new field configurations, called "merons", become important in the functional integral. These produce a long-range force that confines the quarks.[36] (I can see nothing wrong with this idea in principle, but the details of the argument involve a stupendous amount of hand-waving. This part is just a suggestion (although a very clever suggestion) that may or may not someday become a theory of confinement.)

If you will excuse me for beating a dead horse one more time, this picture shows very sharply how misleading it is to say that "instantons give the U(1) Goldstone boson a mass". This implies that quarks get their masses through spontaneous symmetry breakdown, with the appearance of four Goldstone bosons, and then instantons come to the rescue. This is not what happens.

(To be fair, I should modify the last sentence and say, "This is not what happens in the picture of Callan, Dashen, and Gross." A skeptic might imagine replacing paragraph (c) above by, "At still lower temperatures, and still larger effective couplings, new field configurations become important which restore chiral U(1) invariance. At a yet later stage, this suffers spontaneous breakdown and a Goldstone boson appears." To my knowledge, there is no chromodynamic computation that offers the slightest evidence for this disgusting alternative, but it is not logically excluded.[37])

(3) I have stressed several times that spontaneous breakdown of U(1) (without Goldstone bosons) is independent of spontaneous breakdown of SU(2) ⊗ SU(2) (with Goldstone bosons). In a recent paper, Crewther[38] has argued ingeniously that these phenomena are not just independent; they are inconsistent. This would be bad news if it were true, but I do not believe that it is; I think Crewther's arguments are invalid. However, since Crewther and I are at this moment entering our fourth month of correspondence on this matter, and since neither of us has yet convinced the other of the error of his ways, I will say no more about this.

(4) In all the θ vacua, except for $\theta = 0$ or π, CP-noninvariant operators have nonvanishing expectation values. Thus it seems that in most of the θ vacua we have observable strong CP violation. Of course, this is an illusion; the θ vacua are transformed into each other by the U(1) group, and thus all experiments must yield the same results in any vacuum. Phrased more explicitly, for every θ vacuum there is a discrete symmetry under which the vacuum is invariant, the product of CP and an appropriate U(1) transformation,

and we are free to redefine CP to be this transformation.

All this is for massless quarks. The situation changes drastically when the quarks have masses, either because we have put them in by hand, or because they have Yukawa couplings to weak-interaction Higgs mesons. Now we no longer have U(1) symmetry; there is a potential clash between the definition of CP selected by θ and that selected by the quark mass operator, and there is the disastrous possibility of strong CP violation.

[Let me dispose of a red herring. You might think that all this might be said of a theory in which U(1) breaks down in the ordinary way, with Goldstone bosons, as in the U(1) σ model. In this case, there is no problem; as soon as we add a U(1) violating interaction, no matter how weak, the order parameter, the analog of θ, automatically aligns itself with the perturbation. *This is not what happens here.* The easiest (and unfortunately also the least convincing) way of seeing this is to remember that when all the dust of Sec. 3 settled, θ emerged as effectively a coupling constant, the coefficient of a term in the action. Thus we would no more expect θ to change discontinuously in response to an external perturbation than we would expect g to.]

Several mechanisms have been suggested for avoiding this disaster.[39] At the moment I favor an up quark with vanishing bare mass, that is to say, with vanishing coupling to the Higgs fields. In this case, we still have a U(1) symmetry, chiral U(1) acting on the up quark only, and thus we have no CP problem. Unfortunately, this conflicts with current-algebra estimates of the up mass; these all agree that it is somewhere between $\frac{1}{2}$ and $\frac{2}{3}$ of the down mass. However, all these estimates are based on soft-kaon and soft-eta computations, and these are notoriously less accurate than soft-pion computations. For example, only soft-pion methods are needed to compute the slope of $\eta \to 3\pi$, in good agreement with experiment; soft-eta methods are needed to compute the rate, off by a factor of three.[40] So perhaps a massless up quark is not such a silly idea.

Still, I would be happier if I had a more elegant solution, and one with more predictive power.

VI. THE FATE OF THE FALSE VACUUM[41]

6.1 Unstable Vacua

In Sec. 2.4 I explained how to use instanton methods to study a particle theory with a false (that is to say, unstable) ground state. In this section I will apply these methods to a field theory with a false ground state, that is to say, a false vacuum.

For simplicity, I will restrict myself to the theory of a single scalar field in four-dimensional space-time, with dynamics defined by the Euclidean action

$$S = \int d^4x [\tfrac{1}{2}(\partial_\mu \phi)^2 + U(\phi)] \ , \qquad (6.1)$$

where U is a function of the form shown in Fig. 15. Note that U possesses two relative minima, ϕ_+ and ϕ_-, but only ϕ_- is an absolute minimum. In analogy to Sec. 2.4, I have used my freedom to add a constant to U to insure that $U(\phi_+) = 0$. The state of the classical field theory for which $\phi = \phi_-$ is the unique classical state of lowest energy, and, at least for weak coupling, corresponds to the unique vacuum state of the quantum theory. The state of the classical field theory for which $\phi = \phi_+$ is also a stable classical equilibrium state. However, in the quantum theory it is rendered unstable by barrier penetration; it is a false vacuum.

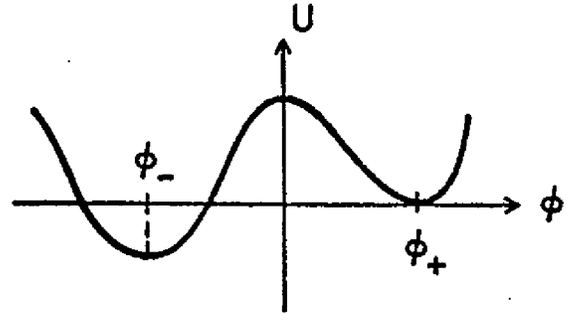


Figure 15

Even without any knowledge of instantons and bounces, it is easy to understand the qualitative features of the decay of the false vacuum. The decay closely parallels the nucleation processes of statistical physics, like the crystallization of a supersaturated solution or the boiling of a superheated fluid. Imagine Fig. 15 to be a plot of the free energy of a fluid as a function of density. The false vacuum corresponds to the superheated fluid phase and the true vacuum to the vapor phase. Thermodynamic fluctuations are continually causing bubbles of vapor to materialize in the fluid. If the bubble is too small, the gain in volume energy caused by the materialization of the bubble is more than compensated for by the loss in surface energy, and the bubble shrinks to nothing. However, once in a while a bubble is formed large enough so that it is energetically favorable for the bubble to grow. Once this occurs, there is no need to worry about fluctuations anymore; the bubble expands until it converts the available fluid to vapor (or coalesces with another bubble).

An identical picture describes the decay of the false vacuum, with quantum fluctuations replacing thermodynamic ones. Once in a while a bubble of true vacuum will form large enough so that it is energetically favorable for the bubble to grow. Once this happens, the bubble spreads throughout the universe, a cancer of space, converting false vacuum to true.

Thus the thing to compute is not a decay probability per unit time, Γ, but a decay probability per unit time per unit volume, Γ/V, for the probability per unit time that in a given volume a critical bubble will form is proportional to the volume (at least if the volume is much bigger than the bubble).

Of course, such a computation would be bootless were it not for cosmology. An infinitely old universe must be in a true vacuum, no matter how slowly the false vacuum decays. However, the universe is not infinitely old, and, at the time of the big bang, the universe might well have been in the false vacuum. For example, in

the Weinberg-Salam model, if the mass of the Higgs meson exceeds Weinberg's lower bound, the asymmetric vacuum, in which we live, has a lower energy than the symmetric vacuum. However, if the Higgs mass is less than $\sqrt{2}$ times the lower bound, the symmetric vacuum is a local minimum of the potential, a possible false vacuum. Now we know that at high temperatures (i.e., in the early universe), symmetry breaking disappears in this model; the symmetric vacuum is the true ground state. Thus it is possible to envision a situation in which the universe gets into the false vacuum early in its history and is stuck there as it cools off; in such a situation, knowledge of Γ/V is essential if we wish to describe the future of the universe.

[I stress that I'm just using the Weinberg-Salam model as an example. I've chosen it because it's familiar and concrete, but in some ways it's a bad choice for our purposes. Firstly, the model involves, not one scalar field, but many scalar and vector fields. Secondly, the vacuum stability features I have described are not properties of the classical potential, $U(\phi)$, but require consideration of one-loop corrections. Thus the formalism I am going to develop is not applicable to this case. As long as we're talking about this model, though, you might be tempted to consider the possibility that the Higgs mass is less than Weinberg's lower bound, that we are living in the false vacuum. As Linde[41] has pointed out, this is silly; if this were the case, there would be no way for the universe to get into the false vacuum in the first place.]

The relevant parameter for cosmology is that cosmic time for which the product of Γ/V and the volume of the past light cone is of order unity. If this time is on the order of microseconds, the universe is still hot when the false vacuum decays, even on the scale of high-energy physics, and a zero-temperature computation of Γ/V is inapplicable. If this time is on the order of years, the decay of the false vacuum will lead to a sort of secondary big bang, with interesting cosmological consequences. If this time is on the

order of billions of years, we have occasion for anxiety.

6.2 The Bounce

We know from Sec. 2.4 how to compute Γ/V. We must find the bounce, $\bar{\phi}$, a solution of the Euclidean equations of motion,

$$\partial_\mu \partial_\mu \bar{\phi} = U'(\bar{\phi}) \ , \tag{6.2}$$

that goes from the false ground state at time minus infinity to the false ground state at time plus infinity,

$$\lim_{x_4 \to \pm\infty} \bar{\phi}(\vec{x}, x_4) = \phi_+ \ . \tag{6.3}$$

To these boundary conditions we can add another. It is easy to see that if the action of the bounce is to be finite,

$$\lim_{|\vec{x}| \to \infty} \bar{\phi}(\vec{x}, x_4) = \phi_+ \ . \tag{6.4}$$

Once we have found the bounce, it is trivial to compute Γ/V. To leading order in $\hbar$,

$$\Gamma/V = K e^{-S_0} \ . \tag{6.5}$$

where S_0 is $S(\bar{\phi})$ and K is a determinantal factor, defined as in Sec. 2.4.

I will shortly construct the bounce. Before I do so, though, I want to make some comments:

(1) We already see the power of our method. The problem of barrier penetration in a system with an infinite number of degrees of freedom has been reduced to a study of the properties of a single classical partial differential equation.

(2) The factor of V in the expression for Γ arises automatically in our method. No non-trivial solution of Eqs. (6.2)-(6.4) is translation invariant. Thus we must integrate over the location of the bounce. This gives us a factor of V, just as did the integration over instanton location in Sec. 3.

(3) It might be that there are many solutions to Eqs. (6.2)-(6.4). We are only interested in the solutions of minimum action, for these make the dominant contribution to the functional integral.

(4) We are not interested in the trivial solution, $\phi = \phi_+$. For this solution, $\delta^2 S/\delta\phi^2$ has no negative eigenvalues, and thus makes no contribution to the vacuum decay probability.

(5) If we imbed $\bar{\phi}$ in a one-parameter family of functions,

$$\phi_\lambda(x) = \bar{\phi}(x/\lambda) \ , \tag{6.6}$$

then,

$$S(\phi_\lambda) = \tfrac{1}{2}\lambda^2 \int d^4x (\partial_\mu \bar{\phi})^2 + \lambda^4 \int d^4x U(\bar{\phi}) \ . \tag{6.7}$$

Because $\bar{\phi}$ is a solution of the equations of motion, this must be stationary at $\lambda = 1$. Thus,

$$\int d^4x (\partial_\mu \bar{\phi})^2 = -4 \int d^4x U(\bar{\phi}) \ , \tag{6.8}$$

and

$$S_0 = \tfrac{1}{2} \int d^4x (\partial_\mu \bar{\phi})^2 > 0 \ . \tag{6.9}$$

This is reassuring. Since U is somewhere negative, one might worry about the possibility that S_0 was negative, which would lead to a very strange dependence of the decay probability on $\hbar$. This possibility has now been eliminated. Also,

$$d^2S/d\lambda^2 = -\tfrac{1}{2} \int d^4x (\partial_\mu \bar{\phi})^2 < 0 \ . \tag{6.10}$$

Thus, at $\bar{\phi}$, $\delta^2 S/\delta\phi^2$ has at least one negative eigenvalue, and $\bar{\phi}$ does contribute to the decay probability. Of course, if there were more than one negative eigenvalue, we would have to rethink the analysis of Sec. 2.4. However, as I shall show eventually, this does not happen; there is only one negative eigenvalue.

Now for the construction of the bounce: Eqs. (6.2)-(6.4) are O(4) invariant. Thus it is not unreasonable to guess that the bounce might also be O(4) invariant, that is to say, that $\bar{\phi}$ might depend only on the distance from some point in Euclidean space. Recently, Glaser, Martin, and I were able to show that this guess is right, under mild conditions on U; there always exists an O(4)-invariant bounce and it always has strictly lower action than any O(4)-noninvariant bounce.[42] The rigor of our proof is matched only

by its tedium; I wouldn't lecture on it to my worst enemy. However, it is possible to give a sloppy argument for the first part (existence) although, unfortunately, not for the second (action minimization).

I will now give this argument.

If we choose the center of symmetry to be the origin of coordinates, then O(4) symmetry is the statement that $\bar{\phi}$ is a function only of the radial variable, r. Thus Eq. (6.2) becomes

$$\frac{d^2\bar{\phi}}{dr^2} + \frac{3}{r}\frac{d\bar{\phi}}{dr} = U'(\bar{\phi}) \ , \tag{6.11}$$

while Eqs. (6.3) and (6.4) both become

$$\lim_{r\to\infty} \bar{\phi}(r) = \phi_+ \ . \tag{6.12}$$

Also,

$$\left.\frac{d\bar{\phi}}{dr}\right|_{r=0} = 0 \ . \tag{6.13}$$

Otherwise, $\bar{\phi}$ would be singular at the origin.

The key to the argument is the observation that if we interpret $\bar{\phi}$ as a particle position and r as time, Eq. (3.9) is the mechanical equation for a particle moving in a potential *minus* U and subject to a somewhat peculiar viscous damping force with Stoke's law coefficient inversely proportional to the time. The particle is released at rest at time zero, Eq. (6.13); we wish to show that if the initial position is properly chosen, the particle will come to rest at time infinity at ϕ_+, that is to say, on top of the right-hand hill in Fig. 16.

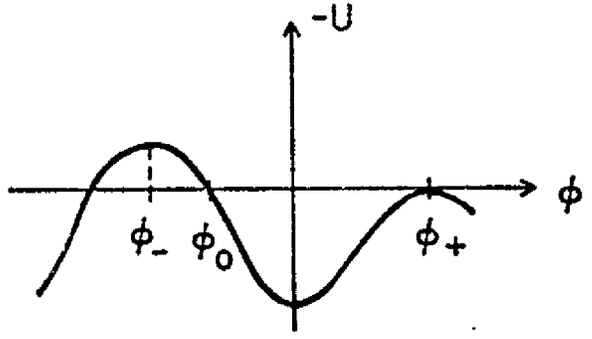


Figure 16

I shall demonstrate this by showing that if the particle is released to the right of ϕ_-, and is sufficiently close to ϕ_-, it will overshoot and pass ϕ_+ at some finite time. On the other hand, if it is released sufficiently far to the right of ϕ_-, it will undershoot and never reach ϕ_+. Thus (arguing in the worst tradition of nineteenth century British mathematics) by continuity there must be an intermediate initial position for which it just comes to rest at ϕ_+.

To demonstrate undershoot is trivial. If the particle is released to the right of ϕ_0, it does not have enough energy to climb the hill to ϕ_+. The damping force does not affect this argument, because viscous damping always diminishes the energy.

To demonstrate overshoot requires a little more work. For ϕ very close to ϕ_-, we may safely linearize Eq. (6.11),

$$(\frac{d^2}{dr^2} + \frac{3}{r} - \mu^2)(\bar{\phi} - \phi_-) = 0 \ , \tag{6.14}$$

where μ^2 is $U''(\phi_-)$. The solution to Eq. (6.14) is

$$\bar{\phi} - \phi_- = 2[\bar{\phi}(0) - \phi_-]I_1(\mu r)/\mu r \ . \tag{6.15}$$

Thus, if we choose ϕ to be initially sufficiently close to ϕ_-, we can arrange for it to stay arbitrarily close to ϕ_- for arbitrarily large r. But for sufficiently large r, the viscous damping force can be neglected, since it is inversely proportional to r. But if we neglect viscous damping, the particle overshoots. Q.E.D.

We have made great progress. We have reduced the partial differential equation for the bounce to an ordinary differential equation. But we can go even farther; in the limit of small energy-density difference between the true and false vacuum, we can obtain an explicit expression for the bounce and for S_0, as I shall now show.

6.3 The Thin-Wall Approximation

Let $U_+(\phi)$ be an even function of ϕ,

$$U_+(\phi) = U_+(-\phi) \ , \tag{6.16}$$

with minima at some points $\pm a$,

$$U_+'(\pm a) = 0 \ . \tag{6.17}$$

Also, let us define

$$\mu^2 = U_+''(\pm a) \ . \tag{6.18}$$

Now let us add to U_+ a small term that breaks the symmetry,

$$U = U_+ + \varepsilon(\phi - a)/2a \ , \tag{6.19}$$

where ε is a positive number. This defines a theory of the sort we have been discussing. To lowest non-trivial order in ε,

$$\phi_\pm = \pm a \ . \tag{6.20}$$

and ε is the energy-density difference between the true and the false vacuums.

It is easy to see the qualitative form of the bounce in the limit of small ε from the mechanical analogy of Sec. 6.2. In order not to lose too much energy, we must choose $\phi(0)$, the initial position of the particle, very close to ϕ_-. The particle then stays close to ϕ_- until some very large time, $r = R$. Near time R, the particle moves quickly through the valley in Fig. 16, and slowly comes to rest at ϕ_+ at time infinity. Translating from the mechanical analogy back into field theory, the bounce looks like a large four-dimensional spherical bubble of radius R, with a thin wall separating the false vacuum without from the true vacuum within.

To go on, we need more information about the wall of the bubble. For r near R, we can neglect the viscous damping term and we can also neglect the ε-dependent term in U. We thus obtain

$$d^2\bar{\phi}/dr^2 = U'_+(\bar{\phi}) \ . \tag{6.21}$$

This is the classical equation of motion for a particle in a symmetric double-welled potential, the equation we studied in Sec. 2.2, the equation that had one-dimensional instantons for its solutions. Indeed, a one-dimensional instanton centered at R is the solution we need here, for such a function goes from -a to a as r increases through R, just what we want. This is our approximate description of the bounce.

The only thing missing from this description is the value of R. This is easily obtained by a variational computation:

$$S = 2\pi^2 \int_0^\infty r^3 dr [\tfrac{1}{2}(d\bar{\phi}/dr)^2 + U] \ . \tag{6.22}$$

We can divide this integral into three regions: the outside of the bubble, the skin of the bubble, and the inside of the bubble. Within the accuracy of our approximation, in the outside region, $\phi = \phi_+$ and $U = 0$; thus we get no contribution from this part of the integral. In the inside region, $\phi = \phi_-$ and $U = -\varepsilon$; thus from this part of the integral we get

$$-\tfrac{1}{2}\pi^2 R^4 \varepsilon \ . \tag{6.23}$$

Over the skin, r is approximately R, and, over this small region, the ε-dependent terms in U are negligible; thus from this part of the integral we get

$$2\pi^2 R^3 \int dr [\tfrac{1}{2}(d\bar{\phi}/dr)^2 + U_+] = 2\pi^2 R^3 S_1 \ , \tag{6.24}$$

where S_1 is the action of a one-dimensional instanton,

$$S_1 = \int_{-a}^{a} \sqrt{2U_+}\, d\phi \ . \tag{6.25}$$

Putting all this together, we find

$$S = -\tfrac{1}{2}\pi^2 R^4 \varepsilon + 2\pi^2 R^3 S_1 \ . \tag{6.26}$$

Varying with respect to R, we find

$$dS/dR = 0 = -2\pi^2 R^3 \varepsilon + 6\pi^2 R^2 S_1 \ . \tag{6.27}$$

Hence,

$$R = 3S_1/\varepsilon \tag{6.28}$$

This completes the approximate description of the bounce. We also know S_0:

$$S_0 = 27\pi^2 S_1^4/2\varepsilon^3 \ . \tag{6.29}$$

I have described what we have done as an approximation that is valid in the limit of small ε. Now that we have gone through the computation, we can phrase the condition for the validity of the approximation more precisely: the approximation is good if the radius of the bubble is much larger than the thickness of the bubble wall; R must be much larger than $1/\mu$, or, equivalently,

$$3S_1\mu \gg \varepsilon \ . \tag{6.30}$$

6.4 The Fate of the False Vacuum

In a particle problem like that of Sec. 2.4, we can describe the decay process in the language of the old quantum theory. The particle sits at the bottom of the potential well until, at some random time, it makes a quantum jump to the other side of the barrier, materializing at the point labeled σ in Fig. 7. At this point, the potential energy of the particle is the same as it was at the bottom of the well; thus its kinetic energy must vanish; equivalently, it has zero velocity. These conditions give the initial-value data for the subsequent motion of the particle, which is totally governed by classical mechanics. Like all descriptions of quantum-mechanical processes in the language of the old quantum theory, this one must be taken with a large grain of salt; it will certainly lead us astray if we try to use it to describe meaurements made just outside the potential barrier. Nevertheless, it is very useful as an asymptotic description, for discussing what happens far from the barrier and long after the time the system decays. For example, this is the description we all use when we discuss the macroscopic detection of an alpha particle emitted by an unstable nucleus.

This description can readily be extended to a system with many degrees of freedom. The point σ becomes the point in multi-dimensional configuration space where all velocities vanish; that is to say, it is the midpoint of the bounce. Thus, for the field theory we have been studying, the description of the vacuum decay process in the language of the old quantum theory is: The classical field makes a quantum jump (say at time zero) to the state defined by

$$\phi(x_0=0,\vec{x}) = \bar{\phi}(\vec{x},x_4=0) \ , \tag{6.31a}$$

and

$$\partial_0\phi(x_0=0,\vec{x}) = 0 \ . \tag{6.31b}$$

Afterwards, it evolves according to the classical Minkowskian field equation,

$$(\nabla^2-\partial_0^2)\phi = U'(\phi) \ . \tag{6.32}$$

The first of these equations implies that the same function, $\bar{\phi}(r)$, that gives the shape of the bounce in four-dimensional Euclidean space also gives the shape of the bubble at the moment of its materialization in ordinary three-space. Indeed, it does more; because the Minkowskian field equation is simply the analytic continuation of the Euclidean field equation back to real time, the desired solution of Eqs. (6.31) and (6.32) is simply the analytic continuation of the bounce:

$$\phi(x_0,\vec{x}) = \bar{\phi}(r=\sqrt{|\vec{x}|^2-x_0^2}\ ) \ . \tag{6.33}$$

[As a consequence of Eq. (6.13), $\bar{\phi}$ is an even function of r, so we need not worry about which branch of the square root to take.]

We can immediately draw some very interesting consequences of Eq. (6.33):

(1) O(4) invariance of the bounce becomes O(3,1) invariance of the solution of the classical field equations. In other words, the growth of the bubble, after its materialization, looks the same to any Lorentz observer.

(2) In the case of small ε, discussed in Sec. 6.3, there is a thin wall, localized at r=R, separating true vacuum from false. As the bubble expands, this wall traces out the hyperboloid

$$|\vec{x}|^2 - x_0^2 = R^2 \ . \tag{6.34}$$

Typically, we would expect R to be a microphysical number, on the order of a fermi, give or take ten orders of magnitude. This means that by macrophysical standards, once the bubble materializes it begins to expand almost instantly with almost the velocity of light.

(3) As a consequence of this rapid expansion, if a bubble were expanding toward us at this moment, we would have essentially no warning of its approach until its arrival. This is shown graphically in Fig. 17. The heavy curve is the bubble wall, Eq. (6.34). A stationary observer, O, cannot tell a bubble has formed until he intercepts the future light cone, H, projected from the wall at the time of its formation. A time R later, that is to say, on the order

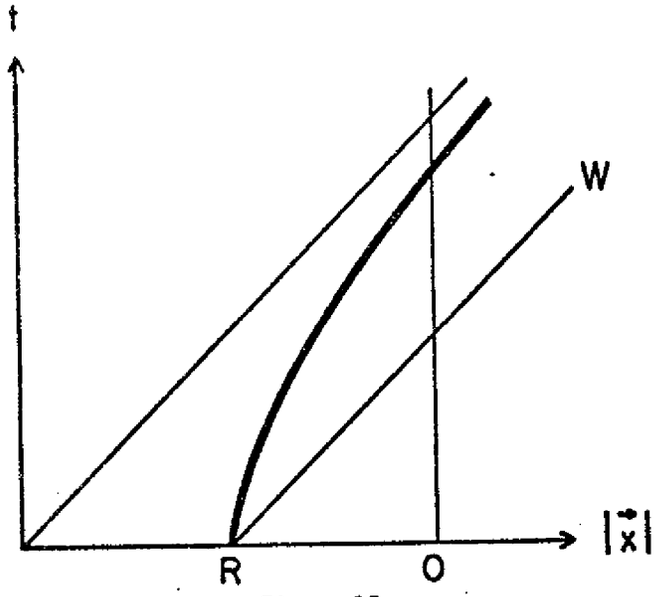


Figure 17

of 10^{-10}-10^{-30} sec. later, he is inside the bubble and dead. (In the true vacuum, the constants of nature, the masses and couplings of the elementary particles, are all different from what they were in the false vacuum, and thus the observer is no longer capable of functioning biologically, or even chemically.) Since even 10^{-10} sec. is considerably less than the response time of a single neuron, there is literally nothing to worry about; if a bubble is coming toward us, we'll never know what hit us.

(4) The rapidly expanding bubble wall obviously carries a lot of energy. How much? A section of bubble wall at rest carries energy S_1 per unit area. Because any part of the bubble wall at any time is obtained from any other part by a Lorentz transformation, a section of wall expanding with velocity v carries energy $S_1/\sqrt{1-v^2}$ per unit area. Thus, at a time when the radius of the bubble is $|\vec{x}|$, the energy of the wall is

$$E_{wall} = 4\pi|\vec{x}|^2 S_1/\sqrt{1-v^2} \ . \qquad (6.35)$$

By Eq. (6.34),

$$v = d|\vec{x}|/dt = \sqrt{1-R^2/|\vec{x}|^2} \ . \qquad (6.36)$$

Thus,

$$E_{wall} = 4\pi|\vec{x}|^3 S_1/R = 4\pi\epsilon|\vec{x}|^3/3 \ . \qquad (6.37)$$

Thus, in the thin-wall approximation, all the energy released by converting false vacuum to true goes to accelerate the bubble wall. This refutes the naive expectation that the decay of the false vacuum would leave behind it a roiling sea of mesons. In fact, the expansion of the bubble leaves behind only the true vacuum.

6.5 Determinants and Renormalization

I said earlier that the determinantal factor K in Eq. (6.5) was defined as in the particle problem of Sec. 2.4. This is basically true, but there are three technical differences: (1) In particle mechanics, we had only one infinitesimal translation, and thus one zero eigenvalue, to worry about; here we have four. (2) It was critical in the analysis of Sec. 2.4 that the second variational derivative of the action at the bounce had one and only one negative eigenvalue. Is the same true here? (3) Whenever we study a relativistic field theory, we must deal with ultraviolet divergences and renormalization. Of course, this last remark also applied to the gauge field theories of Sec. 3, where I swept renormalization problems under the rug. However, we now have a problem with a much simpler renormalization structure (only a single scalar field to worry about, no problems with gauge invariance and gauge-fixing terms, etc.), so it's worth confronting renormalization head-on.

I will deal with these three problems in the order in which I have stated them.

(1) Vanishing Eigenvalues. Because we have four infinitesimal translations, we have four eigenfunctions with eigenvalue zero, proportional to $\partial_\mu\bar{\phi}$. We must determine the constant of proportionality, that is to say, the normalization of the eigenfunctions. This is easy to do. By the spherical symmetry of the bounce,

$$\int d^4x\, \partial_\mu \bar\phi\, \partial_\nu \bar\phi = \tfrac{1}{4}\delta_{\mu\nu}\int \partial_\lambda \bar\phi\, \partial_\lambda \bar\phi$$

$$= \delta_{\mu\nu} S_0 , \tag{6.38}$$

by Eq. (6.9).

Thus, as far as zero eigenvalues go, the only difference between the problem at hand and the particle problem of Sec. 2.4 is that we have four factors of $(S_0/2\pi)^{\frac{1}{2}}$ rather than one. Hence,

$$K = \frac{S_0^2}{4\pi^2}\left|\frac{\det'[-\partial_\mu\partial_\mu + U''(\bar\phi)]}{\det[-\partial_\mu\partial_\mu + U''(\phi_+)]}\right|^{-\frac{1}{2}} , \tag{6.39}$$

assuming we have no problems with negative eigenvalues or renormalization.

(2) Negative Eigenvalues. We already know that $\delta^2 S/\delta\phi^2$ evaluated at the bounce has at least one negative eigenvalue. Can there be more than one? To answer this question I will have to steal some information from the paper by Glaser, Martin, and I that I referred to earlier.[42] There we showed that the bounce could be characterized as the absolute minimum of S for fixed

$$V = \int d^4x\, U . \tag{6.40}$$

This implies that there can not be two independent eigenvectors with negative eigenvalues; for, if there were, we could form a linear combination of the eigenvectors tangent to the surface of constant V, and the bounce would not even be a local minimum of S with fixed V, let alone an absolute minimum.

(3) Renormalization. Until now all of our dynamics has been expressed in terms of unrenormalized quantities. We must now recast our formulas in terms of renormalized quantities. We begin with S itself,

$$S = S_R + \sum_{n=1}^{\infty} S^{(n)} . \tag{6.41}$$

Here S_R is the renormalized action, a functional of exactly the same form as S, but with all unrenormalized quantities replaced by their renormalized counterparts, and $S^{(n)}$ is the action induced by standard renormalization counterterms computed from the sum of all

n-loop graphs. To avoid excessive clutter in my equations, I will redefine ϕ to be the renormalized field, U to be the polynomial that occurs in S_R, $\bar\phi$ to be the bounce as computed from S_R, and S_0 to be $S_R(\bar\phi)$.

The renormalization counterterms serve to remove all ultraviolet divergences from all one-particle irreducible Green's functions. Equivalently, they serve to remove all ultraviolet divergences from the effective action, $\Upsilon(\phi)$, the generating functional of these Green's functions. To one-loop order,[43]

$$\exp\Upsilon(\phi) = \exp[S_R(\phi) + S^{(1)}(\phi)]$$

$$\times \det[-\partial_\mu\partial_\mu + U''(\phi)]^{-\frac{1}{2}} . \tag{6.42}$$

It will be important to us shortly that (for renormalizable U's) the right-hand side of this equation is free of ultraviolet divergences for arbitrary ϕ.

Now let us imagine computing Γ/V iteratively, first treating S_R as if it were the total action, and then taking account of the renormalization counterterms perturbatively. If we had not set $\hbar$ equal to one, $S^{(n)}$ would have been proportional to $\hbar^n$. Thus, to the order in which we are working, the only counterterm we need consider is $S^{(1)}$.

The first thing we must realize is that the counterterms may destroy our convention that $S(\phi_+)$ vanishes. We can take care of this trivially by replacing S_0 in Eq. (6.5) by the difference $S_0 - S(\phi_+)$.

Secondly, adding new terms to S_R will change the stationary points of S. In particular, it will change the bounce. Let us write

$$\bar\phi \to \bar\phi + \Delta\bar\phi . \tag{6.43}$$

Then

$$S(\bar\phi) \to S_0 + \int d^4x \frac{\delta S_R}{\delta\bar\phi}\Delta\bar\phi + S^{(1)}(\bar\phi) + \ldots , \tag{6.44}$$

where the triple dots indicate terms that are negligible in the order in which we are working. The second term vanishes because

the bounce is a stationary point of S_R. Thus, for our purposes,

$$S(\bar{\phi}) = S_0 + S^{(1)}(\bar{\phi}) \ . \tag{6.45}$$

By the same reasoning,

$$S(\phi_+) = S^{(1)}(\phi_+) \ . \tag{6.46}$$

Putting all this together, we find

$$\Gamma/V = \frac{S_0^2}{4\pi^2} \exp[-S_0 - S^{(1)}(\bar{\phi}) + S^{(1)}(\phi_+)] \times \left| \frac{\det'[-\partial_\mu\partial_\mu + U''(\bar{\phi})]}{\det[-\partial_\mu\partial_\mu + U''(\phi_+)]} \right|^{-\frac{1}{2}} \tag{6.47}$$

The point of this exercise is not the simplicity of this formula. Equation (6.47) is an ugly mess, and I know no way of evaluating it for even the simplest theories without using a computer. Rather, the point is that ordinary renormalization works for instanton computations: As a good renormalized expression should be, Eq. (6.47) is free of ultraviolet divergences; each determinant is paired with an exponential of $S^{(1)}$, just as in Eq. (6.42). (That one of the factors is a primed determinant is irrelevant; omitting any finite number of eigenvalues has no effect on the ultraviolet divergence.)

6.6 Unanswered Questions

This concludes what I know about the fate of the false vacuum. There remain many interesting unanswered questions:

(1) I have discussed the expansion of a bubble of true vacuum into false vacuum. What if the initial state of the world is not the false vacuum, but some state of nonzero particle density built on the false vacuum? What happens when a bubble wall encounters a particle?

(2) I have discussed spontaneous decay of the false vacuum. However, there is also the possibility of induced decay. In particular, in a collision of two particles of very high energy, there might be a non-negligible cross section for the production of a bubble. How can one estimate this cross section?

(3) If we assume that the universe starts out in a false vacuum, at some time in its expansion bubbles begin to form. Because the formation of bubbles is totally Lorentz invariant, the average distance between bubbles at their time of formation must be of the same order of magnitude as the time at which bubbles begin to appear. Because bubble walls expand with the speed of light, after a time interval of the same order of magnitude, bubble walls begin to collide. What happens then? Can such events be accommodated in the history of the early universe?

The preceding paragraphs are taken verbatim from a paper I wrote at the end of 1976. I still don't know the answers to any of these questions; maybe you'll be able to do better than I.

Appendix A

HOW TO COMPUTE DETERMINANTS[44]

We wish to study the equation

$$(-\partial_t^2 + W)\psi = \lambda\psi \ , \tag{A.1}$$

where W is some bounded function of t. Let us define $\psi_\lambda(t)$ as the solution of this equation obeying the boundary conditions

$$\psi_\lambda(-T/2) = 0 \ , \qquad \partial_t\psi_\lambda(-T/2) = 1 \ . \tag{A.2}$$

The operator $-\partial_t^2 + W$ (acting on the space of functions vanishing at $\pm T/2$) has an eigenvalue, λ_n, if and only if

$$\psi_{\lambda_n}(T/2) = 0 \ . \tag{A.3}$$

As in the text, we define

$$\det(-\partial_t^2 + W) = \prod_n \lambda_n \ . \tag{A.4}$$

Now, let $W^{(1)}$ and $W^{(2)}$ be two functions of t, and let $\psi_\lambda^{(1,2)}$ be the associated solutions of Eq. (A.1). I will prove that

$$\det\left[\frac{-\partial_t^2 + W^{(1)} - \lambda}{-\partial_t^2 + W^{(2)} - \lambda}\right] = \frac{\psi_\lambda^{(1)}(T/2)}{\psi_\lambda^{(2)}(T/2)} \ . \tag{A.5}$$

Proof: The left-hand side of this formula is a meromorphic function of λ, with a simple zero at each $\lambda_n^{(1)}$ and a simple pole at each $\lambda_n^{(2)}$. By elementary Fredholm theory, it goes to one as λ goes to infinity in any direction except along the positive real axis. The right-hand side is a meromorphic function with exactly the same zeroes and poles. By elementary differential-equation theory, it also goes to one in the same limit. Thus the ratio of the two sides is an analytic function of λ that goes to one as λ goes to infinity in any direction except along the positive real axis. That is to say, it is one. Q.E.D.

If we define a quantity N by

$$\frac{\det(-\partial_t^2 + W)}{\psi_0(T/2)} = \pi\hbar N^2 , \tag{A.6}$$

then, by Eq. (A.5), N is independent of W. I will use this expression to define the normalization constant N in the functional integral. (Note that no explicit definition of this quantity was given in the text, so I am perfectly free to define it as I wish here.) Thus we have the desired formula for evaluating Gaussian functional integrals,

$$N[\det(-\partial_t^2 + W)]^{-\frac{1}{2}} = [\pi\hbar\psi_0(T/2)]^{-\frac{1}{2}} \tag{A.7}$$

As a specific example, for the harmonic oscillator, $W = \omega^2$,

$$\psi_0 = \omega^{-1}\sinh\omega(t + T/2) , \tag{A.8}$$

from which Eq. (2.16) immediately follows.

Appendix B

THE DOUBLE WELL DONE DOUBLY WELL[45]

In this appendix I shall show that the formulas derived in the text for the splitting of the ground-state energies in a double-well potential, Eqs. (2.31) and (2.41), are equivalent to the results of ordinary wave mechanics. To do this, I will have to both evaluate the determinants that appear in Eq. (2.41) (using the method of Appendix A) and do the wave-mechanical computation. To keep my equations as simple as possible, I will choose my units such that $\omega = 1$.

Evaluating Determinants

We have to evaluate a primed determinant, one with the zero eigenvalue omitted. I will do this by evaluating the full determinant on a finite interval, $[-T/2, T/2]$, dividing this by its smallest eigenvalue, λ_0, and then letting T go to infinity.

Thus we must construct solutions of

$$[-\partial_t^2 + U''(\bar{x})]\psi_\lambda = \lambda\psi . \tag{B.1}$$

We already know one solution with $\lambda = 0$,

$$x_1 = S_0^{-\frac{1}{2}}\, d\bar{x}/dt \to A e^{-|t|} , \quad t \to \pm\infty$$

The constant A is determined by the integral expression for the instanton, Eq. (2.21),

$$t = \int_0^{\bar{x}} dx(2V)^{-\frac{1}{2}} = -\ell n[S_0^{-\frac{1}{2}}A^{-1}(a - \bar{x})] + O(a - \bar{x}) . \tag{B.3}$$

Equation (B.1) must have a second solution with $\lambda = 0$, which I denote by y_1. It will be convenient to normalize y_1 such that its Wronskian with x_1 is given by

$$x_1\partial_t y_1 - y_1\partial_t x_1 = 2A^2 . \tag{B.4}$$

Thus,

$$y_1 \to \pm A e^{|t|} , \quad t \to \pm\infty . \tag{B.5}$$

We can now construct ψ_0 of Appendix A. For large T,

$$\psi_0(t) = (2A)^{-1}(e^{T/2}x_1 + e^{-T/2}y_1) . \tag{B.6}$$

Hence,

$$\psi_0(T/2) = 1 . \tag{B.7}$$

This takes care of the determinant. To find the lowest eigenvalue, we must find $\psi_\lambda(t)$ for small λ. This can be done by a standard method: we turn Eq. (B.1) into an integral equation and

iterate once. This can readily be seen to yield

$$\psi_\lambda(t) = \psi_0(t) - \lambda(2A^2)^{-1}\int_{-T/2}^{t} dt'[y_1(t)x_1(t') - x_1(t)y_1(t')]\psi_0(t') , \tag{B.8}$$

plus terms of order λ^2, which we neglect. By Eq. (B.6),

$$\psi_\lambda(T/2) = 1 - \lambda(4A^2)^{-1}\int_{-T/2}^{T/2} dt[e^T x_1^2 - e^{-T} y_1^2] . \tag{B.9}$$

For large T, the second term in this expression is bounded, and thus negligible compared to the first term. Thus, for large T,

$$\psi_\lambda(T/2) = 1 - \lambda(4A^2)^{-1}e^T , \tag{B.10}$$

because x_1 is properly normalized.

Thus the lowest eigenvalue is given by

$$\lambda_0 = (4A^2)e^{-T} , \tag{B.11}$$

and, for large T,

$$\frac{\det'[-\partial_t^2 + U''(\bar{x})]}{\det[-\partial_t^2 + \omega^2]} = \frac{\psi_0(T/2)}{\lambda_0 e^T/2} = \frac{1}{2A^2} . \tag{B.12}$$

Reassuringly, this is non-zero and T-independent.

Plugging this in to Eqs. (2.31) and (2.41), we find that the lowest energy levels are given by

$$E_\pm = \hbar/2 \pm A(\hbar S_0/\pi)^{1/2} e^{-S_0/\hbar} . \tag{B.13}$$

Solving the Schrödinger Equation

We wish to study the solutions of

$$-\tfrac{1}{2}\hbar^2 \partial_x^2 \psi + V\psi = E\psi . \tag{B.14}$$

As long as x is not near the bottoms of the wells, we can use standard WKB solutions. Near the bottom of each well, though, there are two turning points. These are *not* separated by many wavelengths, so we can not use the standard connection formulas for a linear turning point. Fortunately, near the bottom of a well, in a region that includes both turning points, we may safely approximate V by a harmonic-oscillator potential. Thus, for example, for x near a, we may write

$$-\tfrac{1}{2}\hbar^2 \partial_x^2 \psi + \tfrac{1}{2}(x-a)^2\psi = E\psi . \tag{B.15}$$

Our strategy will be to match WKB solutions of Eq. (B.14) outside the wells to solutions of Eq. (B.15) in the bottoms of the wells. Furthermore, since we know the solutions are either even or odd, we can restrict ourselves to positive x, and only have to do this awkward matching for the right-hand well.

I will begin by constructing the even and odd WKB solutions for $0 \le x < a$. If we define

$$k(x) = [2(V-E)]^{1/2} , \tag{B.16}$$

then these are

$$\psi_\pm = k^{-1/2}[\exp\hbar^{-1}\int_0^x k\,dx' \pm \exp -\hbar^{-1}\int_0^x k\,dx'] . \tag{B.17}$$

For the solutions we are interested in, E is itself of order $\hbar$. Thus we may ignore E in the factor of $k^{-1/2}$, and expand to first order in the exponential,

$$k = (2V)^{1/2} - E(2V)^{-1/2} . \tag{B.18}$$

To match on to the solutions of Eq. (B.15), we need the form of the WKB solutions as x enters the regime of validity of the quadratic approximation to V, $V = (a-x)^2/2$. In this regime k(x) is just (a-x), while we can compute the E-independent term in the integral by

$$\int_0^x dx(2V)^{1/2} = \int_0^a dx(2V)^{1/2} - \int_a^x dx(2V)^{1/2}$$

$$= \tfrac{1}{2}S_0 - \tfrac{1}{2}(a-x)^2 . \tag{B.19}$$

For the E-dependent term in the integral, we can use Eq. (B.3). Thus we obtain

$$\psi_\pm = (a-x)^{-1/2}\Big\{\exp\hbar^{-1}[\tfrac{1}{2}S_0 - \tfrac{1}{2}(a-x)^2 + E\,\ell n\, S_0^{-1/2}A^{-1}(a-x)]$$

$$\pm \exp -\hbar^{-1}[\tfrac{1}{2}S_0 - \tfrac{1}{2}(a-x)^2 + E\,\ell n\, S_0^{-1/2}A^{-1}(a-x)]\Big\} \tag{B.20}$$

If we write

$$E = \hbar(\tfrac{1}{2} + \varepsilon) , \tag{B.21}$$

then Eq. (B.20) becomes

$$\psi_\pm = \Big\{ e^{S_0/2\hbar} S_0^{-\frac{1}{4}} A^{-\frac{1}{2}} \exp[-(a-x)^2/2\hbar] \pm (a-x)^{-1} e^{-S_0/2\hbar} S_0^{\frac{1}{4}} A^{\frac{1}{2}} \exp[(a-x)^2/2\hbar] \Big\} \times [1+O(\varepsilon)] \ . \tag{B.22}$$

We will hold this expression in reserve while we go on to study the solutions of Eq. (B.15).

We already know one solution of Eq. (B.15), for $\varepsilon = 0$,

$$\psi_1 = \exp[-(a-x)^2/2\hbar] \ . \tag{B.23}$$

Of course, there is another (odd, increasing) solution, ϕ_1. This does not have a simple form in terms of elementary functions, but its asymptotic form, for $|x-a| \gg \hbar$, is easily computed by the WKB approximation, or just read off from Eq. (B.22),

$$\phi_1 = (a-x)^{-1}\exp[(a-x)^2/2\hbar] \ . \tag{B.24}$$

It will turn out that this is all that we need. Note that I have normalized ϕ_1 such that the Wronskian of the two solutions is

$$\phi_1 \partial_x \psi_1 - \psi_1 \partial_x \phi_1 = 2/\hbar \ . \tag{B.25}$$

We wish to solve Eq. (B.15) for small ε. By the same arguments as led to Eq. (B.8),

$$\psi = \psi_1 - \varepsilon \int_x^\infty dx' \psi_1(x')[\psi_1(x')\phi_1(x) - \phi_1(x')\psi_1(x)] \ . \tag{B.26}$$

I have chosen here the solution that vanishes as x goes to plus infinity. Thus, this is the appropriate solution for matching with the decreasing WKB solution in the region $(x-a) \gg \hbar$. Thus, the only matching left to do is in the region $(a-x) \gg \hbar$.

In this region, we can use

$$\int_{-\infty}^\infty dx\, \psi_1^2 = \sqrt{\pi\hbar} \ . \tag{B.27}$$

to write

$$\psi = \exp[-(a-x)^2/2\hbar][1+O(\varepsilon)] - \varepsilon(\pi\hbar)^{\frac{1}{2}}(a-x)^{-1}\exp[(a-x)^2/2\hbar] \ . \tag{B.28}$$

As it should be, this is proportional to Eq. (B.22), if we choose

$$\varepsilon = e^{-S_0/\hbar} A (S_0/\pi\hbar)^{\frac{1}{2}} \ . \tag{B.29}$$

This is the desired result, and it is identical to the result of the dilute-gas approximation, Eq. (B.13).

Almost identical methods to these can be used to check the dilute-gas formula for the width of an unstable state, Eq. (2.50). You might find it an instructive exercise to see that things work out in this case also.

Appendix C

FINITE ACTION IS ZERO MEASURE[46]

In this appendix I will show that, even for a one-dimensional harmonic oscillator, motions of finite action form a set of measure zero in function space.

If we define eigenvalues λ_n and expansion coefficients c_n as in Sec. 2.1, then, for a harmonic oscillator, the quadratic approximation to the action is exact,

$$S = \tfrac{1}{2} \sum_n \lambda_n c_n^2 \ . \tag{C.1}$$

If we introduce new variables, $b_n = c_n \sqrt{\lambda_n/\hbar}$, then

$$S = \hbar/2 \sum_n b_n^2 \ . \tag{C.2}$$

Let us define a slightly unconventional normalization constant, N', by

$$N'[dx] = \prod_n (2\pi)^{-\frac{1}{2}} d b_n \ . \tag{C.3}$$

This has been chosen such that

$$N' \int [dx] e^{-S/\hbar} = 1 \ . \tag{C.4}$$

How much of this integral comes from motions of finite action? The integrand is positive, and every motion of finite action lies in a cube of side L

$$|b_n| \le L \quad \text{for all } n \ , \tag{C.5}$$

for sufficiently large L. Thus, the finite-action contribution to

the integral must be less than

$$\lim_{L\to\infty} \prod_{n=1}^{\infty} (2\pi)^{-\frac{1}{2}} \int_{-L}^{L} d b_n \, e^{-b_n^2} = 0 \ . \qquad (C.6)$$

Q.E.D.

Appendix D

ONLY WINDING NUMBER SURVIVES

This appendix is the promised (in Sec. 3.3) demonstration that, for a sufficiently large box, the only relic of the boundary conditions imposed on the walls of the box is the winding number.

Consider a rectangular box in Euclidean four-space, with sides $L_1 \ldots L_4$. I will label the eight hyperplanes that bound the box by their normal vectors; thus I will refer to the upper 1-wall, the lower 1-wall, the upper 2-wall, etc. (Upper and lower here refer to greater and lesser values of the appropriate coordinate.)

On the walls of the box the tangential components of A_μ are given in a way consistent with finiteness of the action, that is to say, consistent with

$$A_\mu = g \partial_\mu g^{-1} \ . \qquad (D.1)$$

Thus, giving the tangential components of A_μ on the walls is equivalent to giving g on the walls (up to an irrelevant multiplicative constant). The gauge condition $A_3 = 0$ still allows arbitrary x_3-independent gauge transformations. I will use the freedom to make such a transformation to transform g to one on the lower 3-wall. Because the vanishing of A_3 implies the vanishing of $\partial_3 g$, g is automatically one on all walls except the upper 3-wall. On this wall, g is given as a function of three variables, $g(x_1, x_2, x_4)$, equal to one on the boundary of the wall. (I stress that the only function of this gauge transformation is to simplify my subsequent arguments. Since the functional integral is gauge-invariant, anything I can prove with this gauge convention I could prove without it; it's just that the arguments would be clumsier.)

Now let us imbed our original box, with boundary conditions

given by $g_1(x_1, x_2, x_4)$ in a larger box, with the same lowermost corner (chosen to be the origin of coordinates), and with the same sides L_1, L_2, and L_4, but with third side $L_3 + \Delta$. Let the boundary conditions on the larger box be given by some function $g_2(x_1, x_2, x_4)$.

Theorem: If g_1 and g_2 are in the same homotopy class, then any field configuration defined inside the original box consistent with its boundary conditions can be extended to a field configuration defined inside the larger box, consistent with its boundary conditions and the gauge condition $A_3 = 0$, at the cost of an increase in action of order $1/\Delta$.

Before I prove this theorem I will make some comments:

(1) The theorem would certainly not be true if g_1 and g_2 were in different homotopy classes. In this case, to get from g_1 to g_2, we would have to put at least one instanton in the new volume; this would increase the action by at least $8\pi^2/g^2$, independent of the value of Δ.

(2) We are free to choose Δ to be proportional to, say, $L_3^{\frac{1}{2}}$. Thus, for a very large box, the fractional change in the volume of the box is negligible, as is the change in the action. In the language of statistical physics, changing the boundary conditions while keeping the winding number fixed is just a surface effect, not a volume effect.

(3) There is an apparent paradox that may have bothered you: For any fixed configuration of instantons and anti-instantons, $g(x_1, x_2, x_4)$ is fixed. How then can we get all configurations consistent with a fixed winding number with a single set of boundary conditions? The theorem supplies the answer: We don't get all these configurations; we get only a small portion of them. However, we do get "close relatives" of all of them, configurations that differ only by a small distortion very close to the upper 3-wall. The difference caused by this small distortion is negligible for a sufficiently large box.

Now for the proof: By assumption, g and g' are in the same

homotopy class. Thus there is a continuous function of four variables, $g(x_1, x_2, s, x_4)$, with $0 \le s \le 1$, such that

$$g(x_1, x_2, 0, x_4) = g_1 , \qquad g(x_1, x_2, 1, x_4) = g_2 . \tag{D.2}$$

Let $g(x)$ be a function defined in the added volume by

$$g(x) = g(x_1, x_2, (x_3 - L_3)/\Delta, x_4) . \tag{D.3}$$

If we could choose

$$A_\mu = g \partial_\mu g^{-1} , \tag{D.4}$$

then we could effect the desired transition at *no* cost in added action. Unfortunately, this is impossible; Eq. (D.4) is inconsistent with the gauge condition $A_3 = 0$. However,

$$\begin{aligned} A_\mu &= g \partial_\mu g^{-1} , \qquad \mu \neq 3 , \\ &= 0 , \qquad \mu = 3 , \end{aligned} \tag{D.5}$$

is consistent with the gauge condition and will effect the transition.

We must compute the action associated with Eq. (D.5). If we make a gauge transformation by g^{-1}, Eq. (D.5) becomes

$$\begin{aligned} A_\mu &= 0 , \qquad \mu \neq 3 , \\ &= g^{-1} \partial_\mu g , \qquad \mu = 3 . \end{aligned} \tag{D.6}$$

(A gauge transformation does not change the action.) From Eq. (D.6), we see that A_3 is proportional to $1/\Delta$. The only non-vanishing components of $F_{\mu\nu}$ are $F_{\mu 3}$, also proportional to $1/\Delta$. Thus the Lagrangian density is proportional to $1/\Delta^2$. However, the volume of integration is only proportional to Δ. Q.E.D.

Appendix E

NO WRONG-CHIRALITY SOLUTIONS[47]

In this appendix I will show that, if

$$F_{\mu\nu} = \tilde{F}_{\mu\nu} , \tag{E.1}$$

then the only normalizable solution of both

$$D_\mu \gamma_\mu \psi = 0 , \tag{E.2}$$

and

$$\gamma_5 \psi = \psi , \tag{E.3}$$

is $\psi = 0$.

From Eq. (E.2),

$$D_\nu \gamma_\nu D_\mu \gamma_\mu \psi = D_\mu D_\mu \psi + F_{\mu\nu} \gamma_\mu \gamma_\nu \psi = 0. \tag{E.4}$$

Also,

$$F_{\mu\nu} \gamma_\mu \gamma_\nu \gamma_5 = -\tilde{F}_{\mu\nu} \gamma_\mu \gamma_\nu . \tag{E.5}$$

Thus,

$$D_\mu D_\mu \psi = 0 . \tag{E.6}$$

Multiplying by ψ^+ and integrating, we find

$$\int d^4x \, D_\mu \psi^+ D_\mu \psi = 0 . \tag{E.7}$$

Hence

$$D_\mu \psi = 0 , \tag{E.8}$$

for all μ. If we go to axial gauge, this implies, in particular, that ψ is independent of x_3. The only such normalizable function is $\psi = 0$. Q.E.D.

NOTES

1. These topics are all drawn from the classic part of the theory. "Classic", in this context, means work done more than six months ago. A good summary of the more recent research of one of the most active groups in this field is C. Callan, R. Dashen, and D. Gross, "Toward a Theory of the Strong Interactions" (to be published in Phys. Rev.).

2. Polyakov's early work is summarized in A. M. Polyakov, Nucl. Phys. B 121, 429 (1977).
3. See, for example, R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
4. See the note on notation at the end of Sec. 1.
5. It was Polyakov (Ref. 2) who recognized the double well as the prototypical instanton problem.
6. For a review of lumps, see my 1975 Erice lectures, "Classical Lumps and Their Quantum Descendants", in *New Phenomena in Subnuclear Physics*, ed. by A. Zichichi (Plenum Press, New York, 1977).
7. This is, of course, nothing but the standard prescription for handling collective coordinates in soliton problems. See J. L. Gervais and B. Sakita, Phys. Rev. D 11, 2943 (1975).
8. The treatment here follows that of C. Callan and S. Coleman, Phys. Rev. D 16, 1762 (1977). The idea of handling unstable states this way goes back to Langer's analysis of the droplet model in statistical mechanics [J. S. Langer, Ann. Phys.(N.Y.) 41, 108 (1967)]. The factor of $\frac{1}{2}$, of which much is made below, occurs in Langer's analysis and was explained to me by Michael Peskin.
9. The order of my exposition will not be the historical order of discovery. Here is the way it happened: The topological structure of finite-action Euclidean gauge-field configurations was uncovered and the instanton solutions discovered by A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, Phys. Lett. 59B, 85 (1975). The importance of the instantons was realized by G. 't Hooft [Phys. Rev. Lett. 37, 8 (1976); Phys. Rev. D 14, 3432 (1976)] who used them to solve the U(1) problem. (I won't get to this until Sec. 5.) 't Hooft's work was clarified and extended by R. Jackiw and C. Rebbi [Phys. Rev. Lett. 37, 172 (1976)] and by C. Callan, R. Dashen, and D. Gross [Phys. Lett. 63B, 334 (1976)], who discovered the properties of pure gauge field theories discussed in this section.
10. For a review of gauge field theories, see my 1973 Erice lectures, "Secret Symmetries", in *Laws of Hadronic Matter*, ed. by A. Zichichi (Academic Press, New York and London, 1975). See also Ref. 6.
11. And sometimes given by me. I thank Arthur Wightman for awakening me from my dogmatic slumbers.
12. It suffices to assume that the gauge field is without (gauge-invariant) singularities if we make a stereographic projection of four-space onto a four-sphere. I would love to found the analysis on finiteness of the action, without even this assumption about the behavior of the fields at infinity, but I have not been able to do so.
13. Sign convention: In n-space, $\epsilon_{1\ldots n} = 1$. Symbols with upper indices are defined by raising with the appropriate (Euclidean or Minkowskian) metric.
14. R. Bott, Bull. Soc. Math. France 84, 251 (1956).
15. See Ref. 10.
16. See Ref. 6.
17. At least in a box; see the next paragraph.
18. R. Jackiw and C. Rebbi, Phys. Rev. D 14, 517 (1976).
19. Even this is true for axial gauge (in infinite space) only if we add additional gauge conditions. (See Ref. 6.)
20. M. Atiyah and R. Ward, Comm. Math. Phys. 55, 117 (1977).
21. Although I have just argued that this knowledge is irrelevant to our immediate purposes, an enormous amount *has* been learned recently about solutions to the Euclidean gauge-field equations. In fact, "binstantons" don't exist, but $8|\nu|$-parameter families of solutions with winding number ν do. For a review (with references to the original literature) see R. Jackiw, C. Nohl, and C. Rebbi, "Classical and Semiclassical Solutions to Yang-Mills Theory" (to appear in the proceedings of the 1977 Banff School, to be published by Plenum Press).

22. For a review of the renormalization group applied to gauge theories, see Ref. 10.

23. Done by 't Hooft, a hard worker (second paper cited in Ref. 9). 't Hooft's computation has been somewhat simplified. See A. Belavin and A. M. Polyakov, Nucl. Phys. B 123, 429 (1977); F. Ore, Phys. Rev. D 16, 2577 (1977); S. Chahda, A. D'Adda, P. di Vecchia, and F. Nicodemi, Phys. Lett. 72B, 103 (1977).

24. The analysis reported here is based on C. Callan, R. Dashen, and D. Gross, Phys. Lett. 66B, 375 (1977). The fact that the Abelian Higgs model in two dimensions doesn't display the Higgs phenomenon was discovered independently by two of my graduate students, Frank De Luccia and Paul Steinhardt. They didn't write up their results because I didn't believe them. I take this occasion to apologize to them for my stupidity.

25. The problem is identical to that of constructing flux tubes in superconductors. See Ref. 6, and references cited therein.

26. Indeed, θ vacua, with precisely the same interpretation (but derived in a completely different way), occur in the massive Schwinger model, quantum electrodynamics of charged fermions in 1+1 dimensions. [See S. Coleman, R. Jackiw, and L. Susskind, Ann. Phys. (N.Y.) 93, 267 (1975), and S. Coleman, Ann. Phys. (N.Y.) 101, 239 (1976).] The arguments that work for the Schwinger model also work for the Higgs model when μ^2 is positive, so we also obtain θ vacua in this case despite the absence of instantons.

27. K. Wilson, Phys. Rev. D 10, 2445 (1974). The standard expression for W has a factor of -iq where I have one of -q/e; the source of the difference is the factor of ie hidden in my definition of A_μ.

28. S. Weinberg, Phys. Rev. D 11, 3583 (1975). This paper, titled "The U(1) Problem", gives a characteristically lucid description of the situation just before 't Hooft's breakthrough. (As a major unanswered question, Weinberg lists "How does the underlying gluon-gauge invariance enforce the equal coupling of the positive- and negative-metric Goldstone bosons to gauge-invariant operators?")

29. S. L. Adler, Phys. Rev. 177, 2426 (1969). J. S. Bell and R. Jackiw, Nuovo Cimento 60, 47 (1969). W. Bardeen, Phys. Rev. 184, 1848 (1969).

30. J. Kogut and L. Susskind, Phys. Rev. D 11, 3594 (1976).

31. For more details on Fermi integration, see F. A. Berezin, *The Method of Second Quantization* (Academic Press, New York and London, 1966).

32. The easiest way to fix this up is to stereographically project Euclidean four-space onto a four-sphere; $\not{D}$ is then projected into an operator with a pure discrete spectrum. This changes the determinant, but only by a factor that is independent of the gauge field. Since, as we shall see, our final results will only depend on ratios of determinants, this change is irrelevant.

33. To my knowledge, this sum rule was first derived by A. S. Schwarz, Phys. Lett. 67B, 172 (1977). The derivation in the literature closest to the one given here is that of L. Brown, R. Carlitz, and C. Lee, Phys. Rev. D 16, 417 (1977).

34. This section is mainly afterthoughts; I didn't know most of these things at the time these lectures were given.

35. A related picture of how instantons break SU(2) ⊗ SU(2) is advanced by D. Caldi, Phys. Rev. Lett. 39, 121 (1977).

36. An (apparently) very different picture of how merons effect confinement has been advanced by J. Glimm and A. Jaffe ["A Droplet Model for Quark Confinement" (unpublished)]. G. 't Hooft has advocated completely different configurations ["On the Phase Transition Towards Permanent Quark Confinement" (unpublished)].

37. The prededing paragraph is the product of conversations with Michael Peskin, who has observed that a group of two-dimensional

models analyzed by C. Callan, R. Dashen, and D. Gross [Phys. Rev. D16, 2526 (1977)] display (in a certain sense) a restoration of chiral symmetry at large scales, the first half of the above scenario.

38. R. Crewther, Phys. Lett. 70B, 349 (1977).
39. R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. 38, 1440 (1977); Phys. Rev. D16, 1791 (1977). F. Wilczek, Phys. Rev. Lett. 40, 279 (1978). S. Weinberg, Phys. Rev. Lett. 40, 223 (1978) and "Instantons Without Axions" (unpublished).
40. These ideas are the product of discussions with S. Glashow and D. Nanopoulos.
41. The key paper on this subject is M. B. Voloshin, I. Yu. Kobzarev, and L. B. Okun, Yad. Fiz. 20, 1229 (1974) [Sov. J. Nucl. Phys. 20, 644 (1975)]. The instanton approach to the problem was developed in S. Coleman, Phys. Rev. D15, 2929 (1977), and C. Callan and S. Coleman, Ref. 8. (Large portions of the text of this section are plagiarized from these two papers.) Similar ideas were developed independently by M. Stone, Phys. Rev. D14, 3568 (1976) and Phys. Lett. 67B, 186 (1977). P. Frampton was the first to study these phenomena in the Weinberg-Salam model [Phys. Rev. Lett. 37, 1378 (1976)]; however, Frampton's conclusions have been criticized severely (and, I think, correctly) by A. Linde (unpublished).
42. S. Coleman, V. Glaser, and A. Martin, Comm. Math. Phys. (in press).
43. See, for example, R. Jackiw, Phys. Rev. D9, 1686 (1974).
44. Formulas related to the one developed here can be found throughout the literature. Two references out of many: J. H. Van Vleck, Proc. Nat. Acad. Sci. 14, 178 (1928). R. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D10, 4114 (1974). The derivation given here was developed in conversations with Ian Affleck.
45. This appendix reports on computations done with C. Callan. A somewhat different attack on the problem (with the same conclusions) is E. Gildener and A. Patrascioiu, Phys. Rev. D16, 423 (1977).
46. I give no reference not because these results are novel but because they are a standard part of the theory of Weiner integrals.
47. This appendix is a transcription of an argument of Brown *et al.* (Ref. 33).

ACKNOWLEDGMENTS

After I gave these lectures at Erice, I repeated them at the Ecole Normale Superieure in August and the Lawrence Berkeley Laboratory in September. I thank these institutions for their hospitality.

During these reruns, I made improvements in parts of the lectures, in large measure in response to comments made at Erice, and I have incorporated these in my notes. Thus, if in the transcripts of the discussion sessions you find students asking questions already answered in detail in the notes, it's not that they didn't listen, but that I did.

This work was supported in part by the National Science Foundation under Grant No. PHY77-22864.