SLAC – PUB – 3909
March 1986
(T/E)

Superworlds / Hyperworlds

(The Proposition that Space-Time Has More than 4 Dimensions,
and What it Means to You)

MICHAEL E. PESKIN*

Stanford Linear Accelerator Center
Stanford University, Stanford, California, 94305

Lectures presented at the SLAC Summer Institute on Particle Physics
Stanford, California, July 29 – August 9, 1985

⋆ Work supported by the Department of Energy, contract DE – AC03 – 76SF00515.

It is an old topic for speculation that space-time contains more dimensions than the four familiar from our experience. Already in 1921, just after Einstein's development of general relativity had given a precise basis for investigations of the fabric of space-time, Kaluza[1] proposed an unseen fifth dimension as the origin of electromagnetism. In the past few years, however, this line of speculation has come to be a major element in the search for a fundamental theory underlying the known interactions of elementary particles.

An outsider to field theory, even one educated in quantum mechanics, might well be puzzled by this. If he is open-minded, he will admit that the number of space-time dimensions is an experimental question, and he might well grant that the answer to this question has not yet been decided. We see 4 macroscopic dimensions, but perhaps there exist others tightly curled, perhaps into rings of radius R. But then the minimum nonzero momentum that can be excited in these dimensions is of order $1/R$. If R is large, this momentum is very high, and we cannot probe directly for the signs of such extra dimensions until we reach comparably high energies. In principle, it is possible that R^{-1} is just out of reach at TeV energies, but theorists seriously contemplate values of R^{-1} of order the Planck mass $m_P \sim 10^{19}$ GeV. Such speculations would seem to be completely irrelevant to our current scientific concerns. But they are not, and that is the issue I wish to explain in these lectures. Even though the most direct manifestations of higher dimensions appear only at energies of order R^{-1}, there are consequences of this structure which can be felt at much lower energies, perhaps even at energies now accessible to experiment. The purpose of these lectures is to explain how this can be true, and what insights we might obtain from higher dimensions on the issues in elementary-particle physics which we puzzle over in today's experiments.

These lectures were prepared for an audience of experimentalists. They provide an introduction to the subject of physics in higher dimensions, but only at ~~the~~ most basic level. The student of theoretical physics who wishes to do research in this field might find a quick reading of these lectures useful, but he should then begin working through one of the more serious introductions to the subject, for example, Refs. 2–4.

These lectures will proceed as follows. In Section 1, I will carry out some simple exercises which clarify the physics of a 5-dimensional world in which one dimension is curled up to a radius R. In Section 2, I will discuss a technical problem necessary to generalize this discussion to d dimensions, the determination of the sizes of spinors in higher dimensions. One feature which we will see emerging in the simple analyses of these first two sections is the presence of eigenmodes of a higher-dimensional field with exactly zero energy. These *zero modes* will be of central importance for our discussion, since it is the existence of zero modes that allows the physics of the large scale R^{-1} to become visible at much smaller energies. In Section 3, then, I would like to explore the origin of such zero modes, using as examples conventional models of field theory. Section 4 gives the final bit of background material, a review of general relativity, formulated as a gauge theory. Finally, in Sections 5 and 6, I will come to the central results which I wish to describe. Section 5 will discuss, in general terms, the conditions for the appearance of zero modes in space-time geometries with compactified dimensions. Section 6 will illustrate the physics of these zero modes in a series of examples, from the original construct of Kaluza and Klein[1,5] to the currently fashionable superstring theory.[6,7]

1. The Cylinder World

To begin our discussion, consider space-time in the shape of a cylinder, with four extended dimensions $x^0 - x^3$ and a fifth spacelike dimension x^4 bound up to a size $2\pi R$, as indicated in Fig. 1. To be more precise, consider a space-time with coordinates

$$x^M = \left(x^0, x^1, x^2, x^3, x^4\right) , \tag{1.1}$$

on which all functions are periodic in x^4 with periodicity $2\pi R$. Let us examine the

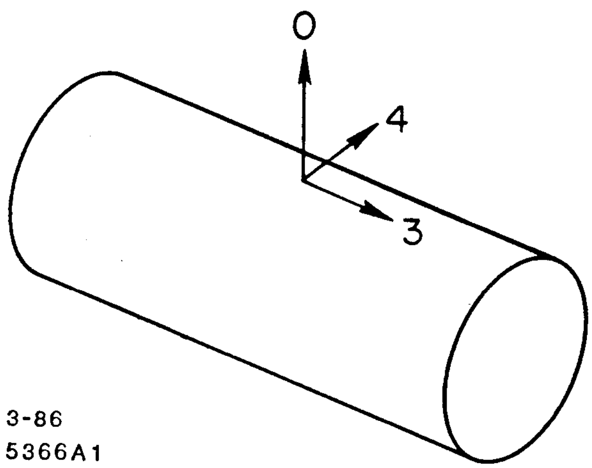

Fig. 1. A model cylindrical space-time.

solutions to the wave equation on such a space for a variety of different particles, from simple scalars to gravitons and gravitinos.

Our first example is a scalar field. The solutions to the wave equation $\partial^2\phi = 0$ for a scalar field on this cylindrical space have the form of plane waves,

$$\phi(x) = e^{-ik\cdot x} , \tag{1.2}$$

with $k^2 = 0$. The periodicity condition in x^4 implies that k^4 is quantized: $k^4 = n/R$; then the energy spectrum of scalar field modes takes the form

$$k^0 = \left[(\vec{k})^2 + (\frac{n}{R})^2\right]^{\frac{1}{2}} . \tag{1.3}$$

Apparently, we can interpret the various quantized values of k^4 as the squared masses of 4-dimensional propagating particles. The mass spectrum is shown in Fig. 2.

To impress you with the generality of this structure, let me digress to consider a more complicated background space. Let space-time have 6 dimensions, with two of them curled into a sphere of radius R, as is indicated in Fig. 3. The wave equation on such a space is given explicitly by

$$\partial^2\phi = \left\{\left[\frac{\partial^2}{\partial x^{02}} - (\vec{\nabla})^2\right] - \frac{1}{R^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\right\}\phi = 0 . \tag{1.4}$$

The solutions to this equation have the form

$$\phi = e^{-ik_\mu x^\mu} \cdot Y_{\ell m}(\theta, \varphi) . \tag{1.5}$$

The plane wave appears because the system is translation-invariant with respect to $x^\mu = (x^0, x^1, x^2, x^3)$; the spherical harmonics are of course the natural eigenfunctions on the 2-dimensional sphere. Inserting (1.5) into (1.4), we find that k^0

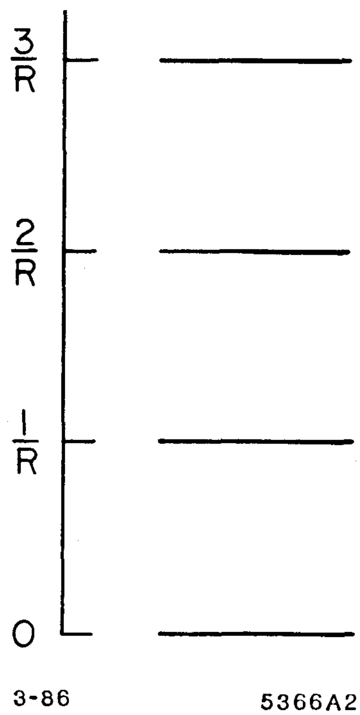

Fig. 2. Four-dimensional mass spectrum of modes of a scalar field on a 5-dimensional cylindrical space-time.

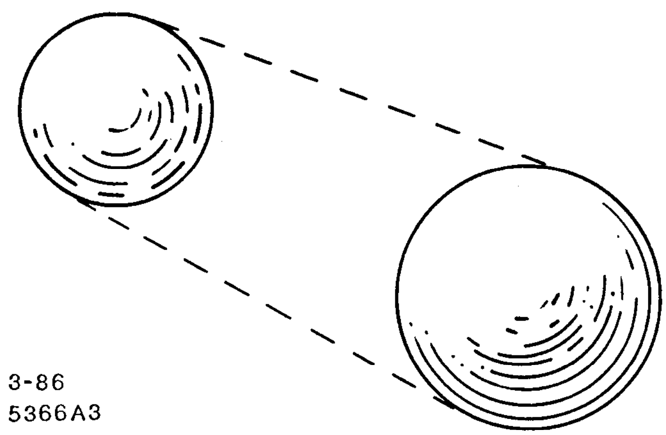

Fig. 3. A model space-time with two spherical dimensions.

must satisfy

$$k^0 = \left[(\vec{k})^2 + (\frac{\ell(\ell+1)}{R^2})\right]^{\frac{1}{2}} ; \tag{1.6}$$

that is, the eigenvalues of the spherical harmonics provide the values of m^2 for a series of the 4-dimensional scalar particles. The 4-dimensional mass spectrum emerging from this theory is shown in Fig. 4. It should be clear that this result generalizes to any space in which the wave equation can be solved by separation of variables.

Let us now return to the cylinder world and examine the wave equation for fields of higher spin. In an open space of 5 dimensions, the wave equation for a vector field $A_M(x)$ would have solutions

$$A_M = \epsilon_M e^{-ik\cdot x} , \tag{1.7}$$

where $k^2 = 0$, and the polarization vector is spacelike and transverse:

$$\epsilon^0 = 0 , \quad \vec{k}\cdot\vec{\epsilon} + k^4\epsilon^4 = 0. \tag{1.8}$$

In 5 dimensions, there are, of course, three transverse directions of polarization. As in the scalar case, k^4 is quantized; k^0 again satisfies (1.3). Thus, we find a quantized mass spectrum of 4-dimensional particles. Consider first the massive particles, $n \neq 0$. The three polarization states satisfying (1.8) are:

$$A_M^{(i)} = (0, \vec{\epsilon}_i, 0)\, e^{-ik\cdot x} , \quad A_M = (0, \frac{\vec{k}k^4}{(\vec{k})^2}, 1)\, e^{-ik\cdot x} , \tag{1.9}$$

where $\vec{\epsilon}_i$, $i = 1, 2$, are the two vectors orthogonal to $\vec{k}$ in ordinary space. These three states naturally form a massive vector boson. For $n = 0$, we have massless modes. The possible polarization states are:

$$A_M^{(i)} = (0, \vec{\epsilon}_i, 0)\, e^{-ik\cdot x} , \quad A_M = (0, \vec{0}, 1)\, e^{-ik\cdot x} . \tag{1.10}$$

The two states involving the $\vec{\epsilon}_i$ give the two transversely polarized states of a massless vector particle. The third state is a new massless scalar. In general, in

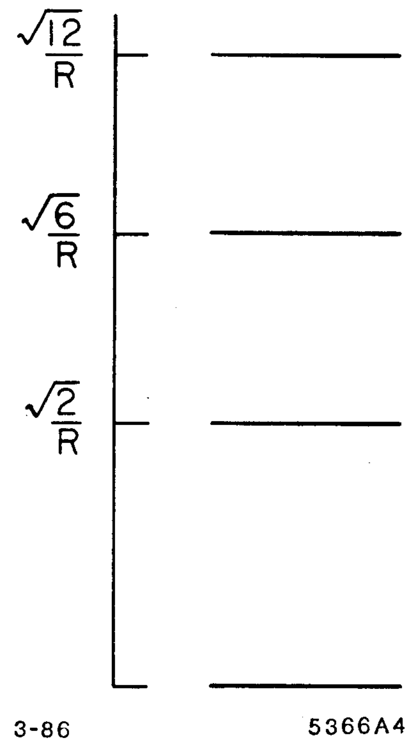

Fig. 4. Four-dimensional mass spectrum of modes of a scalar field on the 6-dimensional space-time of Fig. 3.

a d-dimensional theory with all dimensions but 4 curled into rings, the massless modes comprise a 4-dimensional transverse vector plus $(d-4)$ scalars.

A similar analysis can be done for the spin-2 field. Propagating gravitons have wavefunctions of the form

$$h_{MN}(x) \;=\; \epsilon_{MN}\, e^{-ik\cdot x} \;, \tag{1.11}$$

where ϵ is a symmetric matrix and k and ϵ satisfy the physical-state conditions

$$k^2 = 0 \;, \quad \epsilon_{0N} = \epsilon_{M0} = 0 \;, \quad \epsilon_M{}^M = 0 \;, \quad k^M \epsilon_{MN} = 0 \;. \tag{1.12}$$

I will write the solution to these constraints only for the sector corresponding to 4-dimensional massless particles: $k^M = (k, \vec{k}, 0)$. For simiplicity, take $\vec{k} \parallel \hat{1}$. Then there are two solutions for ϵ which involve only the dimensions 0–3:

$$\begin{pmatrix} & & & & \\ & & & & \\ & & 0 & \epsilon_s & \\ & & \epsilon_s & 0 & \\ & & & & \end{pmatrix} \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} & & & & \\ & & & & \\ & & \epsilon_c & 0 & \\ & & 0 & -\epsilon_c & \\ & & & & \end{pmatrix} . \tag{1.13}$$

These are the conventional gravitational waves of polarization $\uparrow\rightleftarrows\downarrow$ and $\rightarrow\updownarrow\leftarrow$. In addition, there are three solutions which involve the new fifth dimension:

$$\begin{pmatrix} & & & & \\ & & & & \\ & & & & \epsilon_2 \\ & & & & \epsilon_3 \\ & & \epsilon_2 & \epsilon_3 & \end{pmatrix} \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} & & & & \\ & & & & \\ & & -\frac{1}{2}\epsilon_5 & & \\ & & & -\frac{1}{2}\epsilon_5 & \\ & & & & \epsilon_5 \end{pmatrix} . \tag{1.14}$$

The first matrix displayed here represents the two components of a transverse vector. The second gives a massless scalar particle. This is the reduction first noticed by Kaluza and Klein: the content of 5-dimensional gravity, with one dimension compactified, includes 4-dimensional gravity plus a photon.

2. Fermions in Higher Dimensions

To complete our discussion of the cylinder world, we should investigate the spectra of spin-$\frac{1}{2}$ and gravitino fields. To do this, however, we must determine the correct higher-dimensional equations of motion for these fields. This generalization, immediate for integer-spin fields, requires an extra bit of analysis in the half-integer case.

To begin this analysis, recall the original basis of the Dirac equation. Dirac's theory tells us that if we can find a set of d matrices satisfying the algebra

$$\{\gamma^M, \gamma^N\} \;=\; 2g^{MN} \;, \tag{2.1}$$

then we can write an equation which implies the Klein-Gordon equation but has half the number of solutions. For, if we write

$$\left(i\gamma^M \partial_M \;+\; m\right)\psi \;=\; 0 \;, \tag{2.2}$$

acting on this equation with the operator $(i\gamma^M\partial_M \;-\; m)$ gives

$$0 \;=\; \left(\gamma^M\gamma^N\partial_M\partial_N \;+\; m^2\right)\psi \;=\; (\partial^2 + m^2)\psi \;, \tag{2.3}$$

if we use (2.1) to remove the γ's. Equation (2.2) is the Dirac equation, written in a form that applies to any dimensionality. The solutions to this equation are of the form

$$\psi \;=\; \xi\, e^{-ik\cdot x} \;, \quad k^2 = m^2 \;, \tag{2.4}$$

where ξ is a vector in the space on which the γ's act.

It is a standard result that, in 4 dimensions, the γ matrices must be 4×4. We can build up the rule for more general dimensions by examining a few more cases, beginning with 2-dimensional space-time. In 2 dimensions, the following set of 2×2 matrices satisfy the Dirac algebra (2.1):

$$\gamma^0_{(2)} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \gamma^1_{(2)} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} . \tag{2.5}$$

It is useful to define the *chirality* matrix $\overline{\gamma}$ by

$$\overline{\gamma} = \gamma^0_{(2)}\gamma^1_{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.6}$$

Here $\overline{\gamma}$ satisfies $\overline{\gamma}^2 = 1$ and anticommutes with the γ^M's. If we now take ψ to be a 2-component column vector, we can write the Dirac equation (2.2) in components as

$$\begin{aligned} (\partial_0 - \partial_1)\psi_2 + m\psi_1 &= 0 , \\ (-\partial_0 - \partial_1)\psi_1 + m\psi_2 &= 0 . \end{aligned} \tag{2.7}$$

In general, the Dirac field ψ will be complex-valued. However, it is clear from (2.7) that, in the specific case of 2 dimensions, the real and imaginary parts of ψ do not couple to one another, and we are free to insist that ψ is purely real. The imposition of such a reality condition is called a *Majorana reduction*. For a Dirac field of zero mass, a further reduction is possible. Setting $m = 0$ in (2.7) produces decoupled equations for ψ_1 and ψ_2, the eigenvectors of the chirality operator $\overline{\gamma}$. We are free to keep only one of these components. Then the content of (2.7) becomes

$$\begin{aligned} &\psi_1 = \psi_1(x^0 - x^1) \quad \text{(a right-moving fermion)}, \\ \text{or} \quad &\psi_2 = \psi_2(x^0 + x^1) \quad \text{(a left-moving fermion)}. \end{aligned} \tag{2.8}$$

The imposition of such a chirality condition is called a *Weyl reduction*.

Given a representation of the Dirac algebra in 2 dimensions, it is easy to construct one in 3 dimensions: Simply append to the algebra $\gamma^2_{(3)} = i\overline{\gamma}_{(2)}$. A similar trick allows one to construct Dirac matrices in any odd dimensionality $2n+1$ from those in dimensionality $2n$.

To reach 4 dimensions, however, we must increase the size of the matrices, since there do not exist 4 mutually anticommuting 2×2 matrices. It is possible, however, to build up the 4-dimensional matrices by using the 3-dimensional matrices as components. $\gamma^\mu_{(4)}$ may be given as 2×2 matrices of 2×2 matrices in the following way:

$$\gamma^\mu_{(4)} = \begin{pmatrix} & \gamma^\mu_{(3)} \\ \gamma^\mu_{(3)} & \end{pmatrix}, \ \mu = 0,1,2; \qquad \gamma^3_{(4)} = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}. \tag{2.9}$$

The matrix

$$\overline{\gamma} = i\gamma^0_{(4)}\gamma^1_{(4)}\gamma^2_{(4)}\gamma^3_{(4)} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \tag{2.10}$$

(written in 2×2 blocks) anticommutes with each member of (2.9) and so defines the 4-dimensional chirality. This matrix is, of course, just γ^5. This construction in fact gives a general procedure for finding a representation of the Dirac algebra in $2n+2$ dimensions, given a representation in $2n+1$.

In 4 dimensions, all 4×4 matrix representations of the Dirac algebra are equivalent up to unitary transformations, so we are free to convert (2.9) to another, more convenient, form. One possible choice is the one made in the Bible.[8] A second is a representation in which the $\gamma^\mu_{(4)}$ are all pure imaginary. This representation allows us to define a Majorana reduction. A third choice is one which is manifestly amenable to a Weyl reduction:

$$\gamma^\mu_{(4)} = \begin{pmatrix} & \overline{\sigma}^\mu \\ \sigma^\mu & \end{pmatrix}, \tag{2.11}$$

where $\sigma^\mu = (1, \vec{\sigma})$, $\overline{\sigma}^\mu = (0, -\vec{\sigma})$, and $\vec{\sigma}$ are the standard Pauli sigma matrices.

In this basis, $\overline{\gamma}$, defined in (2.10), again takes the form

$$\overline{\gamma} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} . \tag{2.12}$$

The massless Dirac equation $i\gamma\cdot\partial\psi = 0$ has solutions which are plane waves with $k^2 = 0$, $k^\mu = (k, \vec{k})$. If we write

$$\psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} , \tag{2.13}$$

the two-component Weyl spinors ψ_R, ψ_L obey independent equations

$$(k - \vec{k}\cdot\vec{\sigma})\psi_R = 0 \ , \quad (k + \vec{k}\cdot\vec{\sigma})\psi_L = 0 \ . \tag{2.14}$$

These equations express the fact that ψ_R has right-handed spin polarization, and ψ_L has left-handed polarization, as shown in Fig. 5. It is worth recalling that Weyl spinors are physically very important: Since the weak gauge group $SU(2) \times U(1)$ has chiral couplings, it sees as fundamental objects Weyl, rather than Dirac, spinors.

It is not hard to trace the pattern of the representations of the Dirac algebra into higher dimensions. In $2n$ or $2n+1$ dimensions, $2^n \times 2^n$ matrices are required. The Weyl reduction is possible, if $m = 0$, in any even dimensionality. The Majorana reduction, however, is more subtle; this reduction is possible only in $8n + 2$ and $8n + 4$ dimensions, and the Majorana and Weyl reductions may be *simultaneously* applied only in $8n+2$ dimensions.[9] We saw that this simultaneous reduction is indeed possible in 2 dimensions, but not in 4. It is next possible in 10 dimensions.

With this introduction to spinors in higher dimensions, let us take up the question of spinor fields in the cylinder world. Consider first the case of spin $\frac{1}{2}$. Let us construct the plane wave solutions to the Dirac equation, imposing, as

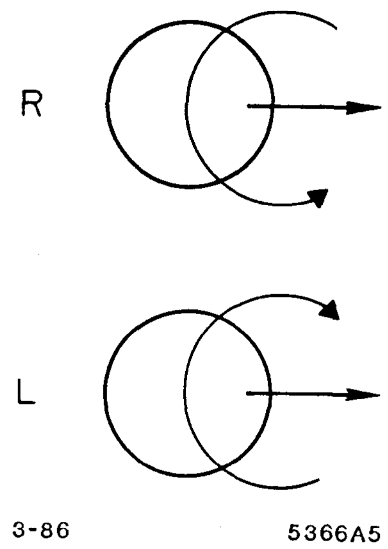

Fig. 5. Polarization of Weyl fermions in 4 dimensions.

before, $k^4 = n/R$. Since 5-dimensional spinors have 4 components and $\gamma^4_{(5)} = \overline{\gamma}_{(4)}$, we can write the Dirac equation as the following condition on the spinor ξ in (2.4):

$$\left(\gamma^{\mu}_{(4)} k_{\mu} - i\frac{n}{R}\overline{\gamma}_{(4)}\right)\xi \;=\; 0 \;. \tag{2.15}$$

Define $\overline{U} = \exp[-i\frac{\pi}{8}\overline{\gamma}]$. This matrix satisfies $\overline{U}\gamma^{\mu}_{(4)}\overline{U}^{\dagger} = \overline{U}^2\gamma^{\mu}_{(4)} = (-i\overline{\gamma})\cdot\gamma^{\mu}_{(4)}$. Thus, (2.15) is equivalent by a unitary transformation to an ordinary Dirac equation in 4 dimensions. The particle mass is n/R, as before. A crucial feature, though, is that we cannot obtain Weyl spinors, even for $n = 0$, except as members of R-L pairs (as in (2.14)). The same result would have been obtained if we had started from a 6-dimensional Weyl fermion, since this is also a 4-component object containing both left- and right-handed 4-dimensional spinors.

We have now encountered a severe problem with the idea that the world is in fact higher-dimensional. Higher-dimensional fermions necessarily contain components of both 4-dimensional chiralities. However, the weak interactions are observed to couple to fermions of definite chirality. How could this be possible? I will refer to this question as the *chirality problem*. One of the main goals of my lectures will be to explain the quite nontrivial mechanism which gives a solution to this problem.

Let us turn, finally, to our last example on the cylinder world, the spin-$\frac{3}{2}$ or gravitino field. Plane-wave gravitini have wavefunctions of the form

$$\psi_{Ma} \;=\; \xi_{Ma}\, e^{-ik\cdot x} \;, \tag{2.16}$$

where a is a spinor index, and the elements of (2.16) satisfy

$$k^2 = 0 \;, \quad \xi_{0a} = 0 \;, \quad \gamma^M_{ab}\,\xi_{Mb} = 0 \;, \quad k^M \xi_{Ma} = 0 \;. \tag{2.17}$$

The third condition removes the spin-$\frac{1}{2}$ piece of the polarization vector. I will write the solutions to these equations only for the massless sector—$n = 0$, $k =$

$(k, \vec{k}, 0)$—setting $\vec{k} \parallel \hat{1}$. We may represent the ξ's as matrices with five rows, corresponding to the possible values of M, and two columns, giving the $\overline{\gamma} = \pm 1$ pieces of the spinor. Two sets of solutions correspond to spin-$\frac{3}{2}$ particles in 4 dimensions:

$$\begin{pmatrix} & \\ & \\ \xi_{1R} & \\ \xi_{2R} & \\ & \end{pmatrix} \quad \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \quad \begin{pmatrix} & \\ & \\ & \xi_{1L} \\ & \xi_{2L} \\ & \end{pmatrix} . \tag{2.18}$$

The spinors satisfy $\gamma^\mu \xi_\mu = 0$. Since there are 2 gravitini, the 5-dimensional theory must reduce to $N = 2$ supergravity. In addition, there are two sets of solutions which correspond to spin-$\frac{1}{2}$ particles in 4-dimensions:

$$\begin{pmatrix} & \\ & \\ & \frac{1}{2}\sigma^2\eta_R \\ & \frac{1}{2}\sigma^3\eta_R \\ \eta_R & \end{pmatrix} \quad \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \quad \begin{pmatrix} & \\ & \\ -\frac{1}{2}\sigma^2\eta_L & \\ -\frac{1}{2}\sigma^3\eta_L & \\ & \eta_L \end{pmatrix} . \tag{2.19}$$

All four of these states come in paired chiralities, a reflection of the generality of the chirality problem we have just posed.

3. The Theory of Zero Modes

Having now gained some experience with multidimensional physics by considering the cylinder world, we are ready to discuss the features of the higher-dimensional spectrum of states in greater generality. Let us take as our starting point the observed fact that, whatever the dimension of space-time might be, 4 of its dimensions are extended, while the rest are extremely tightly curled. The extended dimensions curve slowly over distances of order 10^{27}cm; for the purposes of elementary particle physics, we may regard them as flat. The curled dimensions have $R < 10^{-16}$cm, based on the fact that we observe no massive counterparts of the photon such as the ones which appeared in the analysis of the previous section. I encourage you to imagine, however, that R might be much smaller, of order the Planck scale, 10^{-33}cm.

The solutions to the wave equation on such a space have the form

$$\Phi \sim e^{-ik_\mu x^\mu} \cdot F_\lambda(x^4, x^5, \ldots) . \tag{3.1}$$

The plane wave in the familiar dimensions reflects their simple Minkowski character. The functions $F_\lambda(x^a)$, by analogy to the discussion of the previous sections, are eigenfunctions of an appropriate wave operator on the compact $(d-4)$-dimensional space. Their eigenvalues become the masses of the 4-dimensional particles which this theory produces. Most of these masses are very large: by dimensional analysis, we expect $m_\lambda^2 \sim 1/R^2$. Only eigenvalues which turn out to be small compared to the characteristic scale of $1/R^2$ will correspond to particles which we can readily observe.

If we are really envisioning values of $1/R$ of order m_P, the only relevant eigenvalues will be those which are extremely small on the natural scale. It is most natural to require that the relevant eigenvalues be exactly equal to zero, at least in a first approximation. For quarks and leptons, this is the requirement that these particles have zero bare mass, that is, that all of their mass arises from $SU(2) \times U(1)$ symmetry breaking. We must ask, then, under what circumstances wave operators have eigenvalues exactly zero. This may happen by accident, but presumably there are no accidents in the Grand Design. It may happen also for reasons of physics. In fact, two rather elegant mechanisms for generating zero eigenvalue modes are known from the study of model field theories—one relies on symmetry, the other on topology. In the remainder of this section, I would like to explain these two mechanisms for generating zero-eigenvalue states (*zero modes*), by discussing some simple (nongravitational) examples.

3.1 SYMMETRY ZERO MODES

Let us first consider some examples of zero modes generated by symmetry. The simplest example arises in considering a domain wall between two regions of spontaneously broken symmetry. Such a system would be described by the Hamiltonian

$$H = \int dx \left[\frac{1}{2}(\nabla\phi)^2 + V(\phi)\right] , \tag{3.2}$$

where $V(\phi)$ is a double-well potential of the form shown in Fig. 6(a). The states of lowest energy are those for which $\phi(x)$ remains always at a minimum of V: $\phi(x) = \pm\phi_0$. However, if we insist that $\phi(x) \to +\phi_0$ as $x \to \infty$ but that $\phi(x) \to -\phi_0$ as $x \to -\infty$, we find instead a solution of the form of Fig. 6(b). In particular, the variational equation which follows from (3.2) is

$$-\nabla^2\phi(x) + \frac{\partial V}{\partial\phi}(\phi(x)) = 0. \tag{3.3}$$

Let $\hat{\phi}(x)$ be the solution to this equation of the form of Fig. 6(b).

Consider the small oscillations about this solution:

$$\phi(x) = \hat{\phi}(x) + \delta\phi(x). \tag{3.4}$$

Inserting (3.4) into (3.2), we find

$$\begin{aligned} H = \int dx \Big[&\{\frac{1}{2}(\nabla\hat{\phi})^2 + V(\hat{\phi})\} + \delta\phi\{-\nabla^2\hat{\phi} + \frac{\partial V}{\partial\phi}(\hat{\phi})\} \\ &+ \frac{1}{2}\{(\nabla\delta\phi)^2 + \frac{\partial^2 V}{\partial\phi^2}(\hat{\phi})(\delta\phi)^2\} + O(\delta\phi^3)\Big] . \end{aligned} \tag{3.5}$$

The term linear in $\delta\phi$ vanishes by virtue of (3.3). To analyze the stability of the

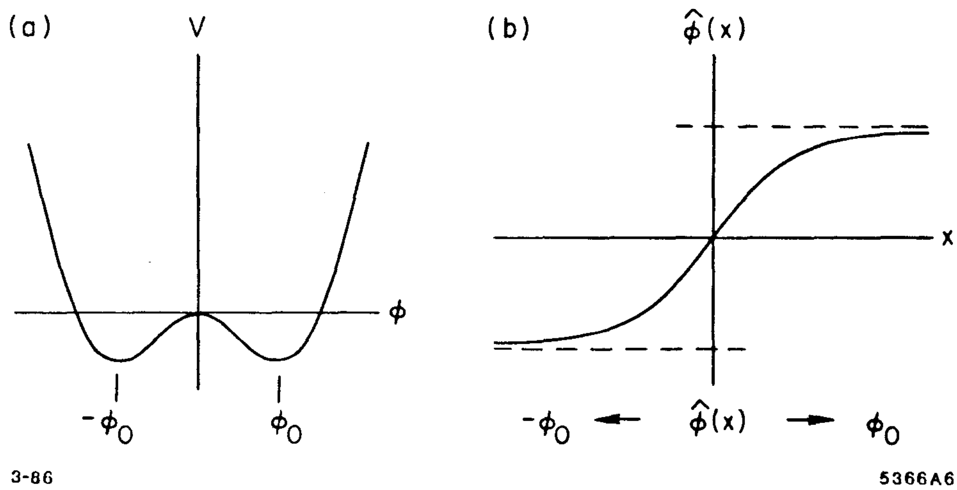

Fig. 6. Appearance of domain walls in a model system with spontaneously broken symmetry: (a) the potential energy $V(\phi)$; (b) the form of the domain wall solution $\hat{\phi}(x)$.

domain wall solution, we must treat the $(\delta\phi)^2$ term as an eigenvalue problem

$$-\nabla^2\delta\phi + \frac{\partial^2 V}{\partial\phi^2}(\hat{\phi})\delta\phi = \epsilon^2\delta\phi . \tag{3.6}$$

Try, in this equation, the solution

$$\delta\phi = \frac{d}{dx}\hat{\phi} . \tag{3.7}$$

The left-hand side of (3.6) becomes

$$\frac{d}{dx}(-\nabla^2\hat{\phi} + \frac{\partial V}{\partial\phi}(\hat{\phi})) = 0 ! \tag{3.8}$$

Thus, (3.7) is a *zero mode* of the differential operator (3.6).

There is a good physical reason to expect a zero mode to appear in this situation: the original problem was invariant under translations, but the domain wall solution singles out a preferred position, its center. Any translation of the wall will give a new solution of the original variational equation (3.3) with the same energy. Then the infinitesimal translation must be a mode of neutral stability. This is precisely the statement that (3.7) should be a zero mode of the stability operator (3.6). (It should be noted that the higher order terms in the expansion of H in small fluctuations do induce an energy cost for translations. If we go to eigenstates of P, the generator of translations of a domain wall, terms of higher order induce a kinetic energy term $E = P^2/2M$.)

Let us now consider a second example of symmetry zero modes, to show that such modes can potentially form a multiplet under a non-Abelian symmetry group. This example involves the model of nucleon structure invented by Skyrme[10] and recently revived by Balachandran, Nair, Rajeev, and Stern[11] and Witten.[12] This model imagines the nucleon to be a condensate of pions, in the following way: in the strong interactions, the pions form an isospin triplet and thus are in 1-to-1 correspondence with the generators of isospin. Construct, then, the $SU(2)$ unitary matrix field

$$U(x) \;=\; e^{i\vec{\pi}(x)\cdot\vec{\sigma}/f_\pi} \,. \tag{3.9}$$

To make the exponent dimensionless, I have supplied a factor f_π with the dimensions of mass. (Using current algebra, one can in fact argue that this factor must be the pion decay constant.) Equation (3.9) is of the form

$$e^{i\vec{\alpha}\cdot\vec{\sigma}} \;=\; \cos|\vec{\alpha}| \;+\; i\frac{\vec{\alpha}\cdot\vec{\sigma}}{|\vec{\alpha}|}\sin|\vec{\alpha}| \;=\; n^0 \;+\; i\vec{n}\cdot\vec{\sigma}\,, \tag{3.10}$$

where n defined by this equation is a general unit vector in 4 dimensions. The set of all vectors n sweeps out the unit sphere in 4 dimensions. Since this sphere is a 3-dimensional surface, we may imagine cutting it open at the north pole and stretching it out to cover 3-dimensional space. This produces a field configuration topologically trapped in 3 dimensions. Just as was true for the domain wall, this configuration cannot be continuously deformed to a trivial configuration in which $\vec{\pi}(x)$ is constant. The form of the pion field corresponding to this solution is shown graphically in Fig. 7, its explicit form is

$$\hat{U}(x) \;=\; e^{if(r)\hat{r}\cdot\vec{\sigma}}\,, \quad \frac{\vec{\pi}(x)}{f_\pi} \;=\; f(r)\hat{r}\,, \tag{3.11}$$

where $f(r)$ tends to 0 as $r \to \infty$ and equals π (the south pole, according to (3.10)) at $r = 0$. More generally, we may imagine placing on 3-dimensional space N bumps of the form of Fig. 7, so that the n vectors associated with $\vec{\pi}/f_\pi$ cover the unit sphere N times. The number N is unchanged by small deformations; it is a conserved quantum number. Skyrme identified the solution (3.11) with a baryon and the conserved number N with baryon number.

— Our main interest in this solution lies in the problem of the small oscillations about $\hat{U}(x)$. Because the form of $\hat{U}(x)$ involves a spatial vector $\hat{r}$ and an isovector

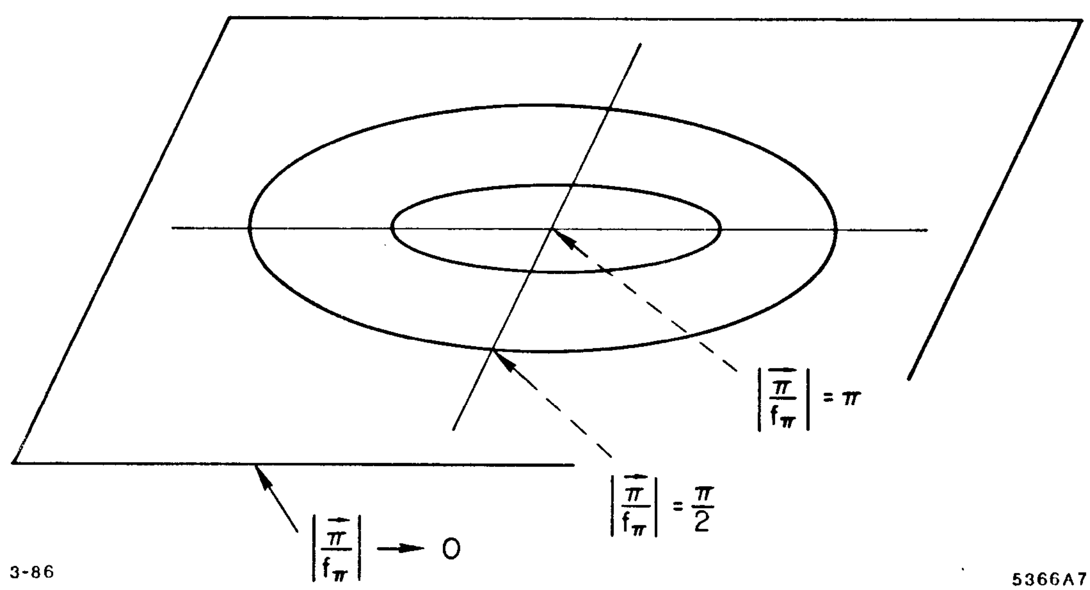

Fig. 7. The pion field configuration in Skryme's model of the baryon.

$\vec{\sigma}$, $\hat{U}(x)$ picks out a preferred orientation in space and in isospace. Rotations in space or isospace generally change the form of $\hat{U}(x)$:

$$\begin{aligned} \hat{U}(x) &\to e^{if(r)(R_J\hat{r})\cdot\vec{\sigma}} \quad \text{(rotation in space)} \,, \\ \hat{U}(x) &\to e^{if(r)\hat{r}\cdot(R_I\vec{\sigma})} \quad \text{(rotation in isospace)} \,, \end{aligned} \tag{3.12}$$

where R_J, R_I are rotation matrices. Note that these two motions duplicate one another, so that combined rotation generated by $\vec{I}+\vec{J}$ leaves the Skyrme solution invariant:

$$\vec{I}+\vec{J}\,|\text{Skyrmion}\rangle = 0. \tag{3.13}$$

Either can be expressed as an isospin rotation of the pion field. Since space and isospace rotations are symmetries of the strong interactions, the rotated and unrotated pion fields must have the same energy. Then the infinitesimal rotation

$$\frac{\partial}{\partial\theta^i} R^{ij}(\theta)\,\pi^i(x)\big|_{\theta=0} \tag{3.14}$$

must be a zero mode of the stability problem associated with the Skyrme solution. The three independent rotations give three zero modes

$$\delta^{(j)}\pi^i = \epsilon^{ijk}\left[\hat{r}^k f(r)\right] ; \tag{3.15}$$

these three states form an spin triplet under the symmetry $(\vec{I}+\vec{J})$ preserved by the Skyrme solution. (In a higher order analysis, one finds that states of definite I, J (with $I = J$, because of (3.13)) receive energy $E = I(I+1)/2\mathcal{I}$. It is permissible to quantize I, J as half-integer angular momenta[13] ; in this case, the two lowest states have $I = J = \frac{1}{2}$ (the nucleon) and $I = J = \frac{3}{2}$ (the Δ).)

3.2 TOPOLOGY ZERO MODES

The second physical mechanism for the appearance of zero modes is a more subtle one which is easier to illustrate than to explain. The simplest example in which it appears is in the motion of Dirac fermions about a magnetic vortex in a superconductor. Let me pose this problem carefully and show you that a zero mode does appear.

Let us first set up a magnetic flux quantum in a 2-dimensional superconductor. Particle physicists view a superconductor as a system with a complex-valued Higgs field (physically, this is the electron pair condensate) which acquires a vacuum expectation value $|\langle\phi\rangle| = \phi_0$. Deep inside the superconductor, there can be no electric or magnetic fields; this is the Meissner effect. It is possible, though, that some magnetic flux lines might thread through the superconductor, driving it normal or partially normal in a localized region. This situation is illustrated in Fig. 8. We should try to determine $\langle\phi(r)\rangle$ and the magnetic vector potential $\vec{A}(r)$ in the vicinity of such a flux tube. These functions are constrained by the integral

$$\int d^2s\, B \;=\; \oint d\vec{\ell}\cdot\vec{A}\,, \tag{3.16}$$

which gives the flux of B passing through the contour C in Fig. 8. If C is far from the flux tube, $B = 0$ there. But this does not imply that $\vec{A} = 0$, since if $\vec{A} = \vec{\nabla}\lambda$ for some λ, we still have $B = \vec{\nabla}\times\vec{A} = 0$. This statement reflects the fact that $\vec{A}$ can be changed by a gauge transformation: a gauge motion carries

$$\vec{A} = 0\,,\; \phi = \phi_0 \;\rightarrow\; \vec{A} = \vec{\nabla}\lambda\,,\; \phi = e^{i\frac{e}{\hbar c}\lambda}\phi_0\;. \tag{3.17}$$

Using this form of $\vec{A}$, we can evaluate (3.16) as

$$\oint d\vec{\ell}\cdot\vec{A} \;=\; \oint d\vec{\ell}\cdot\vec{\nabla}\lambda \;=\; \Delta\lambda|_{\text{circuit}}\;. \tag{3.18}$$

This integral is not necessarily zero: by (3.17), λ is proportional to the phase of $\langle\phi\rangle$ along the contour C. If one circles this contour, $\langle\phi\rangle$ must come back to the

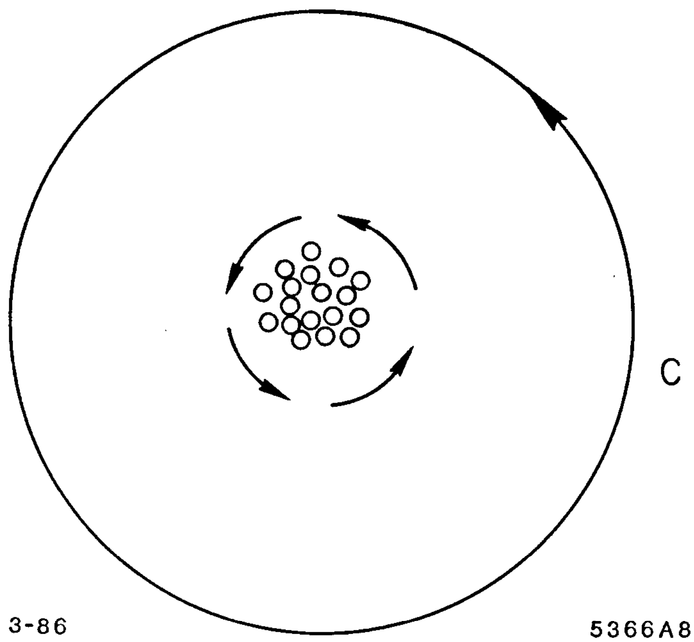

Fig. 8. Configuration of magnetic flux and supercurrents when a bit of magnetic flux traverses a superconductor.

same value, but its phase may advance by $2\pi n$. In that case,

$$\int d^2s\, B = \oint d\vec{\ell}\cdot\vec{A} = \frac{\hbar c}{e}\cdot 2\pi n\ . \tag{3.19}$$

The total magnetic flux flowing though C is then quantized! A simple, regular form for $\vec{A}$ which satisfies the constraint (3.19) (with $n=1$) for a suffiiciently large contour is

$$\vec{A} = \frac{f(r)}{er}\hat{\theta} = \frac{f(r)}{er}\left(\frac{x_2}{r}, -\frac{x_1}{r}\right)\ , \tag{3.20}$$

where $r = (x_1^2+x_2^2)^{\frac{1}{2}}$ and $f(r)$ is a function which vanishes as $r\to 0$ and tends to 1 as $r\to\infty$ (and I have returned to $\hbar c = 1$).

Let us use this $\vec{A}$ in the massless 2-dimensional Dirac operator

$$i\gamma^\mu(\partial_\mu - ieA_\mu)\psi\ , \tag{3.21}$$

and look for zero modes. To write (3.21) explicitly, we need two spacelike γ^μ's:

$$\gamma^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad \gamma^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}\ . \tag{3.22}$$

Inserting (3.22) and (3.20) into (3.21), we find that the condition for a zero mode in ψ_1 is

$$\left\{-(\partial_1 - i\partial_2) - \frac{f(r)}{r}\left(\frac{x_1 - ix_2}{r}\right)\right\}\psi_1 = 0\ . \tag{3.23}$$

Try a solution of the form $\psi_1 = g(r)$; this equation becomes

$$-\left(\frac{x_1-ix_2}{r}\right)\left\{\frac{d}{dr}g + \frac{f(r)}{r}g\right\} = 0\ . \tag{3.24}$$

We find, then, that

$$\psi_1 = \exp\left[-\int_0^r dr'\frac{f(r')}{r'}\right] \tag{3.25}$$

is a zero mode of (3.21). Note that the conditions on $f(r)$ given at the end of the previous paragraph imply that this form for ψ_1 is regular at $r=0$ and falls to zero as $r \to \infty$. An analogous argument can be made for ψ_2; this leads to a solution of the form of (3.25), but with the crucial change of $-$ to $+$ in the exponent. This solution blows up as $r \to \infty$ and so does not belong to the physical Hilbert space. Thus, in this example, ψ_1 has a zero mode, but ψ_2 does not.

To understand the significance of this result, we should think about the more general eigenvalue problem

$$i\gamma^\mu(\partial_\mu - ieA_\mu)\psi = \lambda \cdot \psi \,. \tag{3.26}$$

Defining $\overline{\gamma}$ for this system by $\overline{\gamma} = -i\gamma^1\gamma^2$, we can see that, if we have an eigenvector ψ_λ with eigenvalue λ, then $\psi = \overline{\gamma}\psi_\lambda$ is another eigenvector with eigenvalue $-\lambda$. The spectrum of eigenvalues of (3.26) thus consists of pairs $\pm\lambda$, as indicated in Fig. 9. The corresponding eigenvectors must be states of mixed chirality. The only states that need not be paired are the zero modes, and these may have definite chirality. Let the number of zero modes of positive and negative chirality be N_+, N_-. Consider now the effect of making a small change in the vector potential $\vec{A}$ used to define (3.26). The nonzero eigenvalues may move slightly, but their pairing is preserved. In principle, zero modes may move to nonzero values. But they must move away from zero in pairs, each pair being composed of the two possible linear combinations of a positive and a negative chirality eigenfunction. Pairs of nonzero modes may move to zero by reversing this process. In all cases, however, the value of $N_+ - N_-$ remains unchanged. The difference $N_+ - N_-$ must then be determined by an expression which is insensitive to local changes in $\vec{A}$. Indeed, Atiyah and Singer[14] have proved a theorem on the zero modes of the Dirac operator (the *Index Theorem*) which reads, in this case,

$$\begin{aligned} N_+ - N_- &= \int d^2x \, \frac{e}{4\pi} \epsilon^{\mu\nu} F_{\mu\nu} \\ &= \frac{e}{2\pi} \oint d\vec{\ell} \cdot \vec{A} \,, \end{aligned} \tag{3.27}$$

where the integral in the last line is taken around a contour at infinity. Our

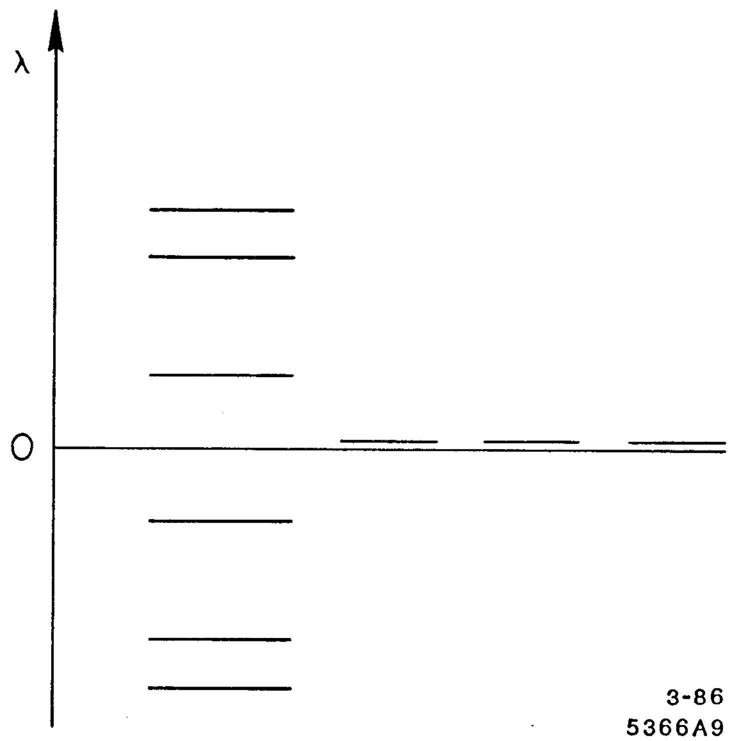

Fig. 9. Spectrum of eigenvalues of the Dirac operator in a typical electromagnetic field.

explicit result in the previous paragraph, $N_+ = 1, N_- = 0$ for a situation with (3.19), $n = 1$, is a special case of this general formula.

As was first noted by 't Hooft,[15] the Atiyah-Singer Index Theorem is closely related to a well-known result in field theory called the *axial vector anomaly*. By considering a process of the form shown in Fig. 10 in 1+1-dimensional space-time, 't Hooft argued that a zero mode of the 2-dimensional Dirac equation could be interpreted as a chirality-flip process in 1+1-dimensions. Let $j^{\mu 5} = \overline{\psi}\gamma^\mu \overline{\gamma}\psi$ be the chirality current and $Q^5 = \int dx\, j^{05}$ the associated charge. Then 't Hooft's arguments related (3.27) to the equation

$$\int dt\, \frac{d}{dt} Q^5 = \int d^2x\, \frac{e}{2\pi} \epsilon^{\mu\nu} F_{\mu\nu} \,, \tag{3.28}$$

or, in local form,

$$\partial_\mu j^{\mu 5} = \frac{e}{2\pi} \epsilon^{\mu\nu} F_{\mu\nu} \,. \tag{3.29}$$

This last equation also follows directly from the Feynman diagram of Fig. 11(a), as was first shown by Schwinger.[16]

The whole set of results we have just discussed have a natural generalization to 4 dimensions. The analogue of (3.29) is the Adler-Bell-Jackiw anomaly equation[17]

$$\partial_\mu j^{\mu 5} = \frac{e^2}{16\pi^2} \epsilon^{\mu\nu\lambda\sigma} F_{\mu\nu} F_{\lambda\sigma} \,. \tag{3.30}$$

(The right-hand side of this equation is proportional to $\vec{E} \cdot \vec{B}$.) Equation (3.30) follows from the Feynman diagram of Fig. 11(b). It is well-known to field theorists because it constrains the whole set of diagrams shown in Fig. 11(c) and thus allows one to compute the $\pi^0 \to 2\gamma$ decay width. The integral of (3.30) over 4-dimensional space, connected to zero modes by 't Hooft's argument, gives the 4-dimensional Atiyah-Singer Index Theorem.

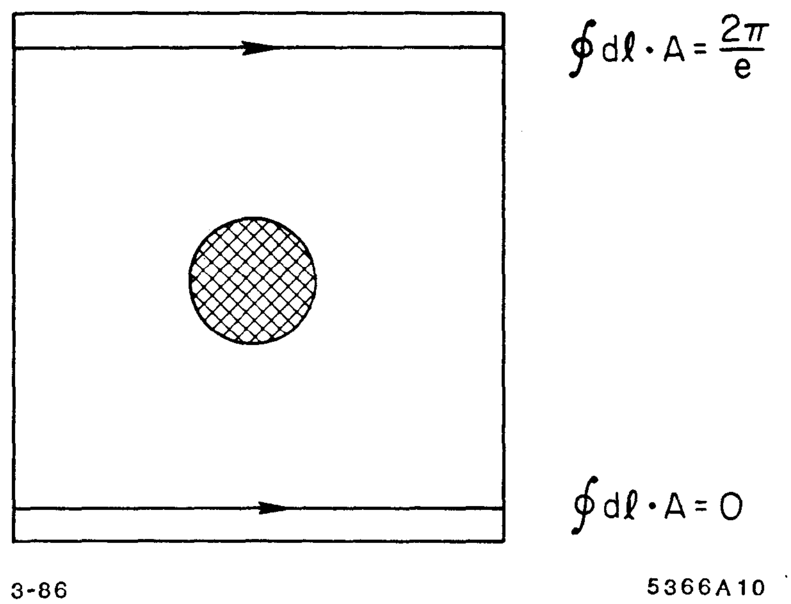

Fig. 10. A process in 1+1 dimensions which leads to a chirality flip in the vicinity of a magnetic flux quantum.

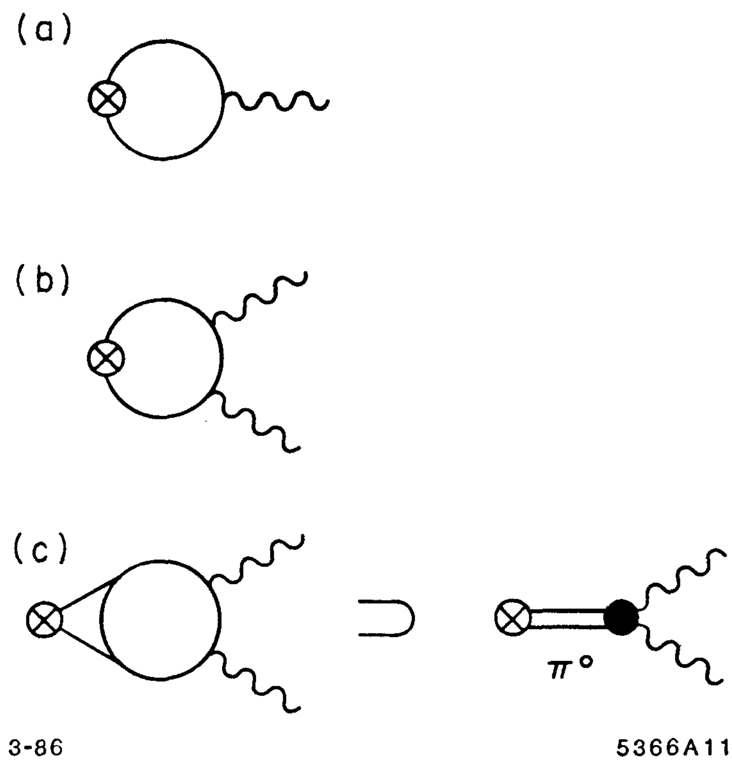

Fig. 11. Feynman diagrams which violate chiral current conservation (a) in 1+1 dimensions, (b) in 3+1 dimensions. This latter diagram enters the computation of the class of diagrams shown in (c) and actually dominates the evaluation of the $\pi^0 \to 2\gamma$ coupling.

The Atiyah-Singer theorem actually applies in general higher (even) dimensions. The its general form is

$$N_{+} - N_{-} \;=\; (\text{constant}) \int d^{d}x\, \epsilon^{\mu\nu\lambda\sigma\cdots} F_{\mu\nu} F_{\lambda\sigma} \cdots \;. \tag{3.31}$$

The quantity on the right-hand side can, quite generally, be written as a surface integral at infinity.

We have now discussed two physical mechanisms for obtaining zero modes of wave equations. The first arose from considerations of symmetry. The second arose from the nontrivial topology of the background fields. It is reasonable that both of these mechanisms have analogues in the study of wave equations in higher-dimensional spaces. The topological mechanism is particularly promising, because it can produce fermion zero modes of definite chirality; thus, it could potentially solve the chirality problem noted at the end of the previous section. Let us now investigate how exactly these mechanisms can work in gravitational contexts.

4. General Relativity as a Gauge Theory

Before beginning this discussion, however, we should review one further bit of formalism. We would like to generalize results obtained in ordinary gauge theories to theories involving gravity. This generalization will be made easier if we can cast the theory of gravity into a form which looks as much like a gauge theory as possible. In the process, we will gain some useful insight into the gauge structure implicit in the theory of gravity.

Einstein's theory of gravity has as its basic dynamical variables parameters of the geometry of space. Consider, then, some general curved space, for example, the one shown in Fig. 12. One way to describe distances on this space is to define a set of coordinates x^{μ} and then to summarize local measurements of distance in

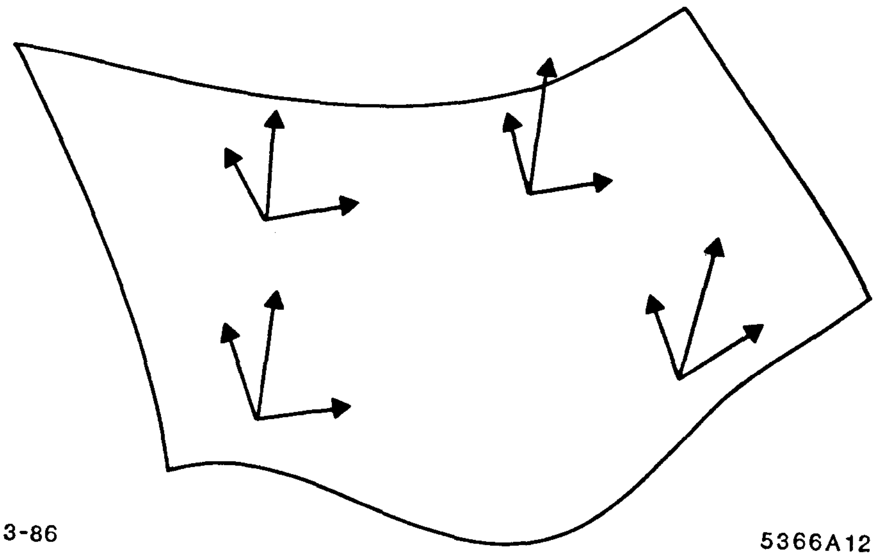

Fig. 12. A piece of curved space, and some orthonormal frames to use in measuring it.

a metric tensor $g_{\mu\nu}(x)$:

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu dx^\nu \ . \tag{4.1}$$

Let V^μ be the components of a vector referred to this coordinate system. An alternative approach, which we will find preferable, is to set-up at each point in the space an arbitrary orthonormal frame. Let $e_\alpha{}^\mu(x)$ be the expression in the coordinate representation of the αth vector of the frame at x. Then the components of a vector with respect to the orthonormal basis obey

$$V^\mu = e_\alpha{}^\mu V^\alpha \ . \tag{4.2}$$

In 4-dimensional space-time, $e_\alpha{}^\mu$ represents a set of 4 vectors and so is called a *vierbein* or *tetrad*. (Its higher-dimensional analogue is, of course, the *vielbein*.) Let us define the matrix $e^\alpha{}_\mu$ to be the inverse of $e_\alpha{}^\mu$; this matrix inverts the relation (4.2), converting components V^μ to V^α. The invariant square of a vector is given by

$$V^2 = \eta_{\alpha\beta} V^\alpha V^\beta = g_{\mu\nu} V^\mu V^\nu \ , \tag{4.3}$$

where η_{ab} is the usual metric tensor on Minkowski space: $\eta_{aa} = (1, -1, -1, -1)$. Inserting into (4.3) the inverse of (4.2), and comparing terms, we find:

$$g_{\mu\nu} = \eta_{\alpha\beta}\, e^\alpha{}_\mu e^\beta{}_\nu \ . \tag{4.4}$$

The orthonormal frames we have constructed can have any orientations, and these orientations can vary arbitrarily from point to point. This arbitrariness in the way we may construct the frames is precisely a local gauge invariance, in which the gauge group is the group of rotations or Lorentz transformations. To make this gauge symmetry an invariance of the equations of motion of fields, we should introduce a covariant derivative, including a gauge field which I will label

$\omega_\mu{}^A$. The superscript A labels a group transformation; the elementary transformations are rotations in planes, so we may replace A by a pair of antisymmetrized indices $\alpha\beta$. We may now write the covariant derivative more explicitly, on vectors

$$D_\mu V^\alpha = \left(\partial_\mu V^\alpha + \omega_\mu{}^\alpha{}_\beta V^\beta\right), \tag{4.5}$$

and on spinors

$$D_\mu \psi = \left(\partial_\mu \psi + \frac{1}{2}\omega_\mu{}^{\alpha\beta}\Sigma_{\alpha\beta}\psi\right), \tag{4.6}$$

using $\Sigma^{\alpha\beta} = \frac{i}{2}[\gamma^\alpha, \gamma^\beta]$, the generators of Lorentz transformations on spinors. A vector (or spinor) is *parallel-transported* along a curve in the direction of the vector n^μ if

$$n^\mu D_\mu V^\alpha = 0 . \tag{4.7}$$

A vector which is parallel-transported around a closed loop on a curved surface will generally return rotated from its original orientation. This is illustrated, for the case of a sphere, in Fig. 13. Around an infinitesimal loop, we find an infinitesimal rotation, which depends on the plane (α, β) of the rotation and on the plane $(\mu\nu)$ of the loop:

$$\delta V^\alpha = \int d^2\sigma^{\mu\nu} R_{\mu\nu}{}^\alpha{}_\beta V^\beta . \tag{4.8}$$

This equation defines the *Riemann curvature tensor*. This tensor is given explicitly by

$$R_{\mu\nu}{}^{\alpha\beta} = \partial_\mu \omega_\nu{}^{\alpha\beta} - \partial_\nu \omega_\mu{}^{\alpha\beta} + \omega_\mu{}^\alpha{}_\gamma \omega_\nu{}^{\gamma\beta} - \omega_\mu{}^\alpha{}_\gamma \omega_\nu{}^{\gamma\beta} . \tag{4.9}$$

This is exactly the form of a field strength tensor $F_{\mu\nu}$ corresponding to the non-Abelian gauge symmetry of Lorentz transformations.

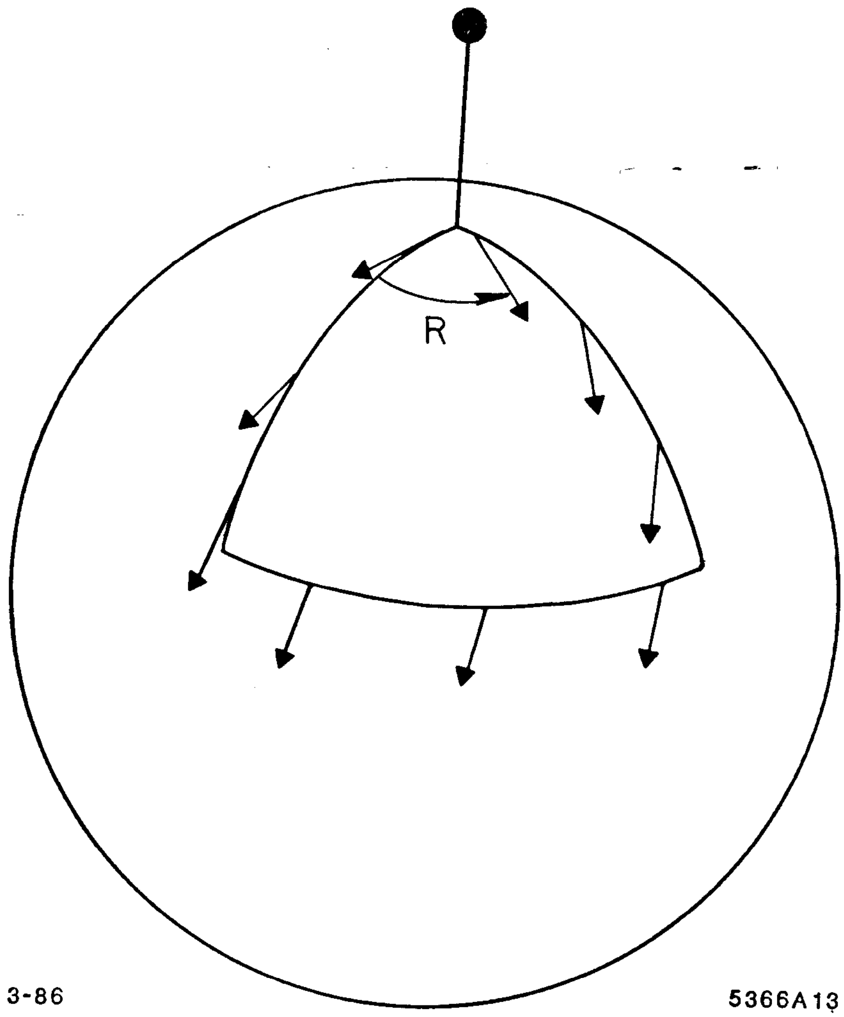

Fig. 13. Parallel transport of a vector around a triangle drawn on a sphere, beginning and ending at the north pole.

Since $R_{\mu\nu}{}^{\alpha\beta}$ is a field strength, it is transforms covariantly under gauge transformations. It therefore can reflect directly the symmetries of the space it describes. Consider, for example, the curvature tensor of the 3-dimensional sphere, which is parametrized by 3 angles ω, θ, ϕ. The nonzero components of $R_{\mu\nu}{}^{\alpha\beta}$ are

$$R_{\theta\phi}{}^{\theta\phi} = \frac{1}{R^2}\,, \quad R_{\omega\theta}{}^{\omega\theta} = \frac{1}{R^2}\,, \quad R_{\omega\phi}{}^{\omega\phi} = \frac{1}{R}\,, \tag{4.10}$$

with R on the right-hand side being the radius of the sphere. The equality of these three elements reflects the spherical symmetry. For completeness, let us define two contractions of the curvature tensor, the Ricci tensor $R_{\alpha\gamma}$ and the curvative scalar R:

$$R_\alpha{}^\gamma = e_\alpha{}^\mu e_\beta{}^\nu R_{\mu\nu}{}^{\gamma\beta}\,, \qquad R = R_\alpha{}^\alpha\,. \tag{4.11}$$

For the 3-dimensional sphere, the Ricci tensor is forced by symmetry to be proportional to the unit matrix

$$R_\alpha{}^\gamma = \begin{pmatrix} \frac{2}{R^2} & & 0 \\ & \frac{2}{R^2} & \\ 0 & & \frac{2}{R^2} \end{pmatrix}\,. \tag{4.12}$$

We may now write Einstein's equation for the gravitational field:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G_N T_{\mu\nu}\,. \tag{4.13}$$

So far, we have discussed only the gauge invariance of the theory of gravity under local Lorentz rotations. But this theory has a second set of local invariances, the freedom to make local changes of coordinates. This freedom may be viewed as a local translation invariance, as is illustrated in Fig. 14. To understand how this second local gauge invariance is implemented, it will be useful

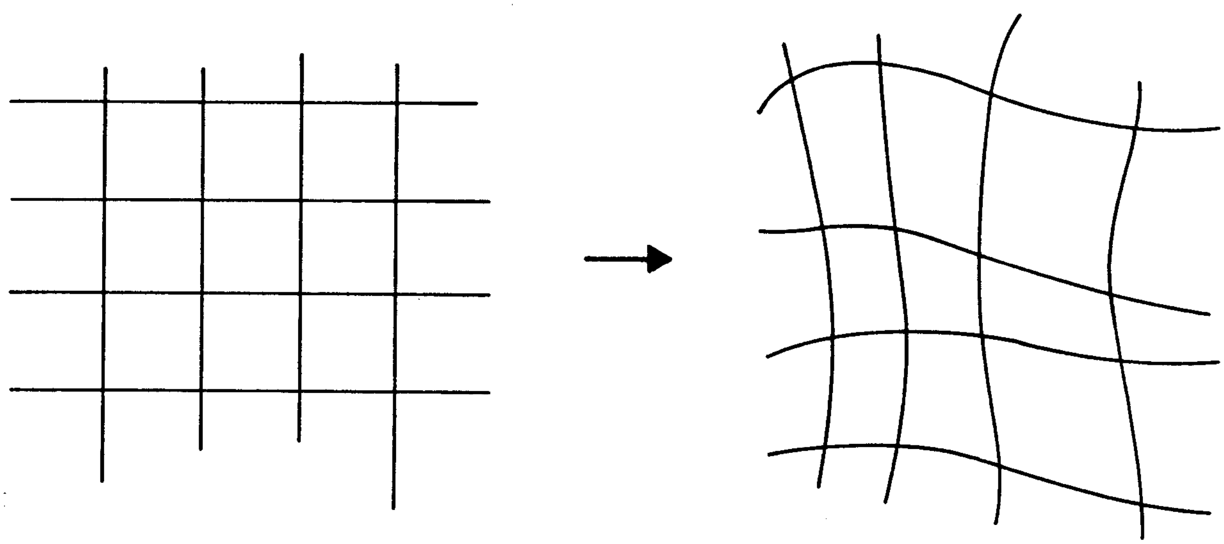

Fig. 14. A general change of the coordinate system, accomplished by making local translations of the coordinate points.

to work out the variation of the vierbein associated with a change of coordinates $x^\mu \to x'^\mu$. Let us write the expression for an infinitesimal vector of fixed orientation located at the point x':

$$\begin{aligned} dx'^\mu \, e'^\alpha{}_\mu(x) &= dx^\nu \, e^\alpha{}_\nu(x(x')) \\ &= dx'^\mu \frac{\partial x^\nu}{\partial x'^\mu} \, e^\alpha{}_\nu(x(x')) \; . \end{aligned} \tag{4.14}$$

Let x' differ from x only infinitesimally: $x^\mu = x'^\mu + \xi^\mu(x)$. Inserting this formula into (4.14), we find the transformation law:

$$e'^\alpha{}_\mu \; - \; e^\alpha{}_\mu \; = \; \frac{\partial}{\partial x^\mu}\xi^\nu e^\alpha{}_\nu \; + \; \xi^\nu \frac{\partial}{\partial x^\nu} e^\alpha{}_\mu \; . \tag{4.15}$$

If we consider an approximately flat space and take $e^\alpha{}_\mu$ to be a natural uniform choice of the orientation of frames, $e^\alpha{}_\mu = \delta^\alpha{}_\mu$, the transformation law (4.15) takes the form

$$\delta e^\alpha{}_\mu \; = \; \partial_\mu \xi^\alpha \; , \tag{4.16}$$

a transformation of a set of gauge fields by the gauge parameters ξ^α. At least in this linearized approximation, then, the vierbein provides the gauge fields for a local translation invariance.

The field strength constructed from this gauge field,

$$T_{\mu\nu}{}^\alpha \; = \; D_\mu e^\alpha{}_\nu \; - \; D_\nu e^\alpha{}_\mu \tag{4.17}$$

is called the *torsion*. In Einstein's theory of gravity, the torsion has no physical content. The Principle of Equivalence implies that all invariant information about the structure of space is contained in the curvature tensor. It requires us to set $T_{\mu\nu}{}^\alpha = 0$; this constraint is a set of equations connecting $e^\alpha{}_\mu$ and $\omega_\mu{}^{\alpha\beta}$ which fixes the ω_μ's in terms of the vierbein. Note, however, that the constraint is covariant with respect to both of the gauge invariances that we have discussed and leaves them both in force.

5. Zero Modes of Geometrical Equations

Now we are ready to address the basic issue toward which we have been working: given a space of some higher number of dimensions, with some dimensions curled into a compact space, what zero modes will we find for wave equations on this space? To answer this question, we will need to generalize the discussion given in Section 3. I will discuss this generalization in some detail for symmetry zero modes, and then more quickly for topology zero modes. In this section, I will refer to the decomposition of a vector x^M into components lying along the extended and compact directions by writing

$$x^M = \left(x^\mu, y^\eta\right) \tag{5.1}$$

with $\mu = 0, \ldots, 3$, $\eta = 4, \ldots$.

Let us first ask how to find symmetry zero modes of the basic wave equations. These are produced in several ways. The first possibility is quite trivial. The scalar field wave operator on a compact space

$$-D^2 \, \phi(y) \tag{5.2}$$

always has a zero mode given simply by the constant function. Thus, scalar fields on a space with extended and compactified dimensions will always have low energy modes of the form of (3.1), with F in this equation a function independent of the y^η. The components of vector and tensor fields in which all of the indices point into the extended directions 0–3 also may be decomposed into terms of the form (3.1) with $F(y)$ an eigenfunction of the scalar operator (5.2). These fields thus have low-energy modes

$$A_M(x,y) = \begin{pmatrix} A_\mu(x) & \quad \end{pmatrix} , \quad g_{MN}(x,y) = \begin{pmatrix} g_{\mu\nu}(x) & \\ & \end{pmatrix} , \tag{5.3}$$

corresponding to one 4-dimensional photon or graviton field for each such field in the d-dimensional theory. Note, however, that this trick does not work for spinor fields, either spin-$\frac{1}{2}$ or spin-$\frac{3}{2}$. The reason is that, while a vector in d-dimensions can be written as the sum of a 4-dimensional and a $(d-4)$-dimensional vector, a spinor in d-dimensions is, by the considerations of Section 2, the product of a 4-dimensional spinor and a $(d-4)$-dimensional spinor.

A second mechanism for producing zero modes is more closely connected to the symmetries of the compactified space. To introduce this mechanism, let me consider the simplest example, the cylinder world of Section 1. We have already analyzed directly the light particle content of this theory, but it is instructive to reconsider this theory in order to count its gauge invariances. Despite its compact structure, the cylinder has its full translation symmetry in 5 dimensions. A 4-dimensional local translation symmetry has its gauge fields contained in the 4-dimensional gravitational field shown in (5.3). But we have one more gauge degree of freedom which requires a corresponding gauge field. Where is it? To see, consider the effect of making a local translation into the 5th dimension

$$\xi^M = (0, \vec{0}, \lambda(x)) . \tag{5.4}$$

Inserting this into (4.15), we find

$$\delta e^A{}_M = \partial_M \lambda \, e^A{}_4 + \lambda \, \partial_4 e^A{}_M , \tag{5.5}$$

or, using (4.4),

$$\delta g_{MN} = \partial_M \lambda \, g_{4N} + \partial_N \lambda \, g_{4M} + \lambda \, \partial_4 g_{MN} . \tag{5.6}$$

If g_{MN} is independent of $y = x^4$, the last term vanishes. Then we find

$$\delta g_{\mu 4} = \partial_\mu \lambda . \tag{5.7}$$

Thus, $g_{\mu 4}(x)$ must be the gauge field corresponding to this last local translation symmetry. By the usual argument from gauge invariance, this vector field must be massless. This is exactly the conclusion we reached at the end of Section 1 by explicit computation: the off-diagonal components of g_{MN} contain a massless 4-dimensional vector. Now we have identified that vector as a gauge particle. Kaluza and Klein, in their original work, went a step further and identified this particle with the photon.

This method of generating gauge bosons from the gravitational field can be readily generalized to more involved situations. Let us, then, consider a general compact manifold described by the metric tensor $g_{\eta\varsigma}(y)$. Let $\phi^{\eta}(y)$ be a symmetry motion of the manifold, a set of local translations which leaves $g_{\eta\varsigma}$ unchanged. Using (4.4) and (4.15), we can write this condition as

$$\delta g_{\eta\varsigma} \;=\; \partial_{\eta}\phi^{\theta} g_{\theta\varsigma} \;+\; g_{\eta\theta}\partial_{\varsigma}\phi^{\theta} \;+\; \phi^{\theta}\partial_{\theta}g_{\eta\varsigma} \;=\; 0\,. \tag{5.8}$$

Such a motion is called a *Killing vector* or an *isometry*. Now consider making a coordinate transformation on the full space given by

$$\xi^{M} \;=\; \left(0, \vec{0}, \lambda(x)\cdot\phi^{\eta}(y)\right)\,. \tag{5.9}$$

Inserting (5.9) into the general transformation law for g, we find

$$\delta g_{\mu\eta} \;=\; \partial_{\mu}\lambda\cdot\phi_{\eta}(y)\,. \tag{5.10}$$

Thus, in general, the off-diagonal terms in the metric tensor contain 4-dimensional gauge fields for every symmetry of the compact dimensions:

$$g_{\mu\eta} \;=\; A_{\mu}(x)\cdot\phi_{\eta}(y)\,. \tag{5.11}$$

Let us illustrate this result with a concrete example. Consider a configuration ~~of~~ space-time in which the total dimensionality is $(N+3)$, with $N-1$ dimensions curled up to form the unit sphere in N dimensions. Two views of this space-time are shown in Fig. 15(a) and (b). The N-dimensional sphere is a space of high symmetry: it possesses $N-1$ symmetries of the form of Fig. 15(c), corresponding to $N-1$ orthogonal directions in which one can move the north pole of the sphere. In addition it possesses $(N-1)(N-2)/2$ symmetries of the form of Fig. 15(d), this being the number of planes in which one can rotate leaving the north pole invariant. In all, there are $N(N-1)/2$ Killing vectors, and so we must find this number of gauge symmetries and of 4-dimensional gauge bosons in the compactified theory. The counting is just that appropriate to a non-Abelian gauge theory with gauge group $O(N)$.

This mechanism for producing symmetry zero modes applies also to spin-$\frac{3}{2}$ fields. In this case, low-energy modes of the field are generated by *Killing spinors*, local supersymmetry motions $\epsilon_a(y)$ satisfying

$$\delta\psi_{\eta a}(y) = 0 . \tag{5.12}$$

If the space-time geometry contains a flat Minkowski space in addition to compactified dimensions, this condition is equivalent to $D_\eta\epsilon(y) = 0$. Given such a Killing spinor, one can consider modes of the d-dimensional gravitino field of the form

$$\Psi_{MA} = \psi_{\mu a}(x)\cdot\epsilon_b(y) . \tag{5.13}$$

Here A is the index of a d-dimensional spinor; such a spinor may be decomposed into the product of a 4-dimensional spinor and a $(d-4)$-dimensional spinor, and this decomposition has been used in writing (5.13). Then $\psi_{\mu a}(x)$ will appear as a massless gravitino field in 4 dimensions.

It is considerably more difficult to produce topology zero modes in compactified geometries. To set up the problem, consider the eigenvalue problem defined by the Dirac operator in a geometry of the form of (5.1). We may write

$$\gamma^M D_M\psi = \left[\gamma^\mu D_\mu + (\gamma^\eta D_\eta)\right]\psi . \tag{5.14}$$

The quantity in parentheses is the Dirac operator on the compactified space.

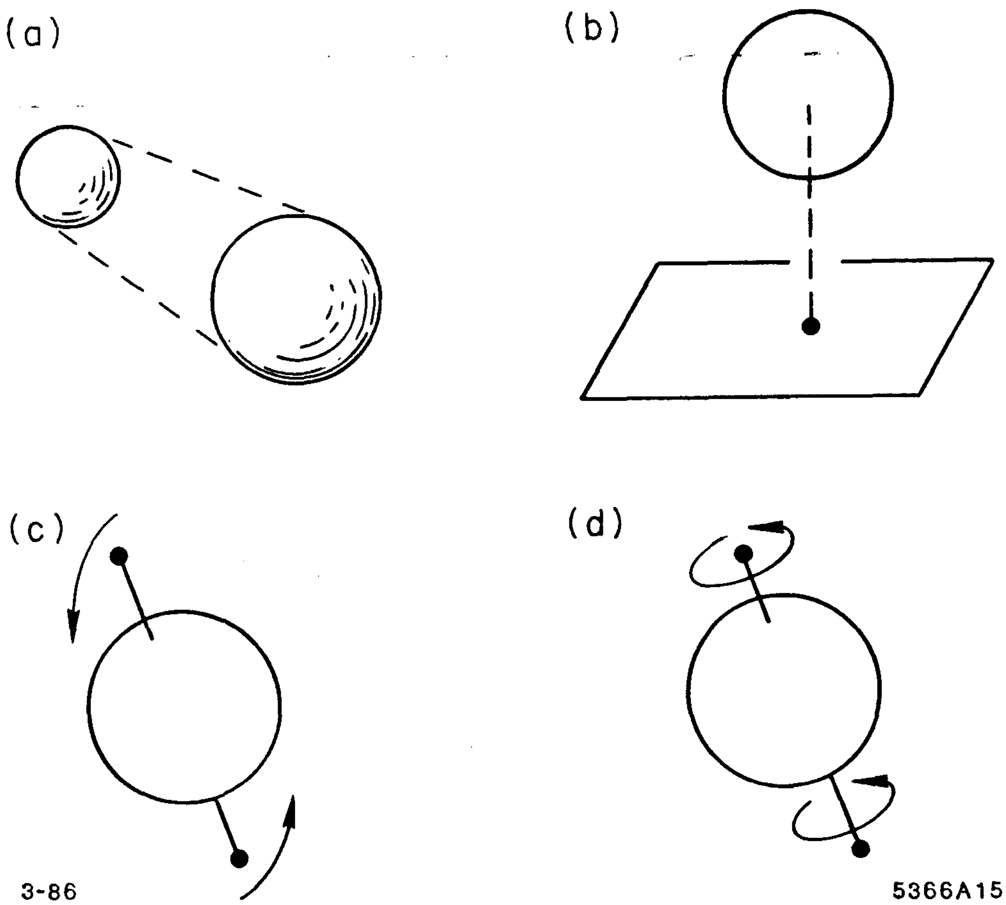

Fig. 15. A space-time with $N-1$ dimensions compactified into a sphere: (a) and (b) show two ways of visualizing this space; (c) and (d) show two classes of symmetry motions of the compactified subspace.

A zero mode of this operator yields a massless fermion in 4 dimensions. These zero modes are counted by the Index Theorem; to determine the number of zero modes, or at least to determine whether these zero modes have a chirality imbalance, we must discover what quantity stands on the right-hand side of the Index Theorem. If we were lucky, we might find there the Euler number

$$\chi \sim \int dx\, \epsilon^{\mu\nu\lambda\sigma\ldots} \epsilon_{\alpha\beta\gamma\delta\ldots} R_{\mu\nu}{}^{\alpha\beta} R_{\lambda\sigma}{}^{\gamma\delta} \cdots , \tag{5.15}$$

which counts the number of holes in the surface. But (5.15) is even under parity and so is an inappropriate candidate. To correct this difficulty, we should contract the indices $\alpha, \beta, \ldots$. In 4 dimensions, this produces an object called the Hirzebruch number, which does in fact appear in the appropriate index theorem:

$$N_+ - N_- = -\frac{1}{8\pi^2} \int d^4x\, \epsilon^{\mu\nu\lambda\sigma\ldots} R_{\mu\nu}{}^{\alpha\beta} R_{\lambda\sigma\alpha\beta} . \tag{5.16}$$

Unfortunately, a theorem of Linchnerowicz insists that the right-hand side of (5.16) can be nonzero only if the scalar curvature R of the compact manifold is somewhere less than 0. Actually, we really need a somewhat weaker requirement, that the right- and left-handed zero modes of the Dirac operator on the compact space, considered in terms of the gauge group defined by our earlier analysis of the symmetries of this manifold, should form inequivalent representations of this group. But this still cannot happen on a compact manifold of positive curvature.[3] Apparently, there is a conflict between the appearance of chiral fermions and the hypothesis that *all* gauge bosons arise as components of the gravitational field associated with Killing vectors of complex manifolds.

6. Some Examples of Higher-Dimensional Models

Now that we have discussed the basic theoretical principles to be used in working out the low-energy spectrum of a theory with compactified dimensions, let us apply the understanding we have gained to analyze a series of models. We will first discuss, one last time, the original scheme of Kaluza and Klein. We will generalize this scheme in the grandest possible way, arriving finally at the 11-dimensional maximal supergravity theory. Then, motivated by the conflict discussed at the end of the previous section, we turn to a simple model with fundamental gauge fields in addition to gravitation. Finally, we discuss another grandly ambitious model which incorporates the best features of both approaches, the 10-dimensional superstring theory.

6.1 (Generalized) Kaluza-Klein Theory

Let us begin, then, by reconsidering and generalizing the model of Kaluza and Klein. The results of the previous section can be encapsulated in the following statement: if one considers the equations of gravity on a space of the form

$$(4 - \mathrm{d\ Minkowski\ space}) \cdot (\mathrm{compact\ } \mathbf{M}) \ , \tag{6.1}$$

the low-energy spectrum of this theory contains a graviton plus one 4-dimensional gauge boson for each isometry of $\mathbf{M}$. If the isometries of $\mathbf{M}$ form a non-Abelian group, as in the sphere example given above, we will find a non-Abelian gauge symmetry. As long as the gauge symmetry of the theory remains unbroken, these bosons remain exactly massless. In principle, such bosons can be given mass by the Higgs mechanism. It is a natural hypothesis that deformations of the geometry of $\mathbf{M}$ can act as Higgs bosons, and this mechanism for giving mass to Kaluza-Klein bosons has in fact appeared in some models.[18]

— Since the general logic of this program is quite clear, it seems not inappropriate to jump to its grandest known realization. To set up the theory in question, recall our result of Section 2 that a 5- or 6-dimensional spin-$\frac{3}{2}$ field yields, in a compactification to 4 dimensions, two 4-dimensional spin-$\frac{3}{2}$ particles This is a signal that the full theories obtained by compacitification contain 2 distinct 4-dimensional supersymmetries. Since the size of the spinor representation doubles at every even dimensionality, compactification of still higher-dimensional spaces will lead to theories with still more 4-dimensional supersymmetries. Fortunately, there is a limit to how much supersymmetry a theory can reasonably contain. Each 4-dimensional supersymmetry contains one helicity-lowering operator, and, in a theory with several supersymmetries, these operators act independently. A theory with a graviton (which has states of helicity ± 2) and particles of all lower spins can accommodate eight independent helicity-lowering operators:

$$-2 \leftarrow -\frac{3}{2} \leftarrow -1 \leftarrow -\frac{1}{2} \leftarrow 0 \leftarrow \frac{1}{2} \leftarrow 1 \leftarrow \frac{3}{2} \leftarrow 2\,, \qquad (6.2)$$

but more such operators can only be accommodated if the theory contains fields of spin > 2. If we are not to admit such high-spin fields (whose quantum theory is known to be extremely problematical[19]) the maximum number of supersymmetries we may allow is 8. A theory with this largest supersymmetry would arise from a theory with spinors of 16 components, or 32 Majorana components. Thus, this grandest and most symmetrical supergravity theory lives most naturally in 11 dimensions.

The 11-dimensional supergravity theory, constructed by Cremmer, Julia, and Scherk,[20] can be compactified in such a way as to maintain its maximal symmetry.[21] In principle, one would like to compactify this theory by dividing 11 dimensions into 4-dimensional Minkowski space plus a 7-dimensional sphere. Unfortunately, the energy-momentum generated by the compactification, set on the right-hand side of (4.13), requires that the 4-dimensional space have a curvature comparable to the curvature of the sphere. The resuting 4-dimensional geometry is a non-compact geometry of constant curvature—anti-de Sitter space.

The spatial compactification

$$11-\mathrm{d} \rightarrow (4-\mathrm{d\ anti-de\ Sitter}) \times (7-\mathrm{d\ sphere}) \tag{6.3}$$

produces the decompositions

$$\begin{aligned} g_{MN} &\rightarrow g_{\mu\nu} + A_\mu{}^A + \ldots \\ \psi_{MA} &\rightarrow \psi_{\mu a}{}^I + \ldots \end{aligned} \tag{6.4}$$

The gauge bosons $A_\mu{}^A$ make up the bosons of an $O(8)$ gauge theory, in the matter that I described in the previous section. The 8 spin-$\frac{3}{2}$ particles $\psi_{\mu a}{}^I$ arise from 8 Killing spinors of the 7-sphere; these enter the 4-dimensional theory transforming as an **8** of $O(8)$ under the gauge symmetry. The 4-dimensional theory contains 56 spin-$\frac{1}{2}$ particles, which form an irreducible, but non-chiral, representation of $O(8)$. The maximal supergravity theory forms itself, then, into a beautiful and highly symmetrical structure, though one which still stands at some distance from the observed world.

6.2 Cremmer and Scherk's Monopole Model

As we discussed at the end of the previous section, it is extremely difficult to obtain a chiral multiplet of spin-$\frac{1}{2}$ particles from a compactified geometry unless the original, higher-dimensional theory contains fundamental gauge fields. It is worthwhile, then, to review the simplest example of a theory in which fundamental gauge fields are present and play an important, nontrivial role. This is a model due to Cremmer and Scherk,[22] which begins in a 6-dimensional space-time containing a fundamental $O(3)$ (isospin) gauge theory and a Higgs field ϕ^i transforming as a vector (**3**) of this $O(3)$.

Cremmer and Scherk found a solution to this model of the form indicated in Fig. 16. They split 6-dimensional space into 4-dimensional Minkowski space

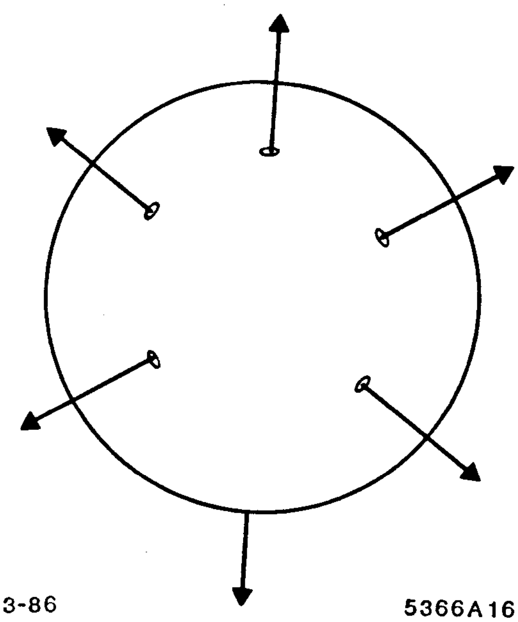

Fig. 16. Configuration of Higgs fields in the compactification model of Cremmer and Scherk.

and a 2-dimensional sphere. They then decorated this sphere with a Higgs fields whose direction in isospace is everywhere normal to the sphere

$$\phi^i \sim (\hat{r})^i \tag{6.5}$$

and with gauge fields whose field strength also has an isospin direction normal to the sphere

$$F_{\eta\varsigma}{}^{i} \sim \epsilon_{\eta\varsigma i} \cdot B \ . \tag{6.6}$$

This configuration of geometry and fields is completely invariant under combined space and isospin rotations, in just the way that the Skyrme model example discussed in Section 3 had an invariance under such combined rotations. Thus, the low-energy 4-dimensional spectrum of this model contains a full $O(3)$ gauge symmetry, whose gauge bosons are associated with the geometrical transformations shown in Fig. 17.

Another consequence of this invariance is that it forces the Ricci tensor for this space to take the form

$$R_{AB} = \begin{array}{cc} \begin{pmatrix} -C & & & & & \\ & C & & & & \\ & & C & & & \\ & & & C & & \\ & & & & D & \\ & & & & & D \end{pmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \end{array} \ , \tag{6.7}$$

where C and D are constants. Equation (6.7) can be a solution to (4.13) if we put on the right-hand side a fixed background energy-momentum proportional to η_{AB} (a *cosmological constant*). The constant C depends on the curvature of 4-dimensional space as well as that of the sphere; to insure that our solution has an extended 4-dimension part, we must adjust the value of this cosmological constant so that (4.13) is satisfied for the C corresponding to 4-dimensional Minkowski

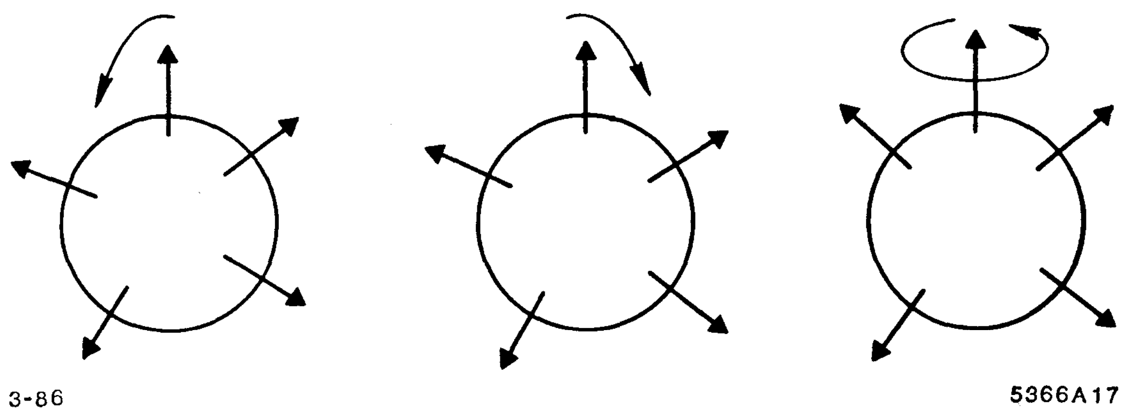

Fig. 17. Symmetry motions of the compactified geometry which give rise to the 4-dimensional gauge bosons of the Cremmer-Scherk model.

space. Thus, the Cremmer-Scherk configuration does provide a solution to the field equations of their model; however, it requires a fine tuning of one parameter to take a fully realistic form.

The Cremmer-Scherk model has one further very attractive feature. Since the gauge field strength and the Higgs field are always parallel in isospace, the flux of F^i, referred to the Higgs direction, has a nonzero integral over the sphere:

$$\int d^2x\, \epsilon^{\eta\varsigma} F_{\eta\varsigma}{}^i \phi^i \neq 0 \,. \tag{6.8}$$

The Atiyah-Singer index theorem implies that, if we add a Dirac field to this configuration of space-time, we will find chiral fermion zero modes. This model might, then, serve as a prototype of theories in which a realistic fermion content emerges from compactification.

6.3 COMPACTIFICATION FROM 10 DIMENSIONS

How can we combine the best virtues of this model with the grandeur of 11-dimensional supergravity? If such a grand fusion could be possible at all, it should take place in the dimensionality which contains the maximally symmetric system of supergravity coupled to matter. We might hope that this dimensionality is also the one which describes matter itself as symmetrically as possible. A multiplet of matter must necessarily, for our purposes, contain gauge bosons; it should also contain spin-$\frac{1}{2}$ fermions. In 4 dimensions, we know that we can form a complete supersymmetry multiplet from these two components. We should ask whether this is possible in any higher dimensions. In 4 dimensions, we can check that a gauge boson and a gaugino can form a complete supermultiplet by counting propagating fields: the gauge boson has two transverse polarization states, and a 2-component Weyl or 4-component Majorana fermion also has two physical states, half the number (counting spins for particle and antiparticle) of a Dirac fermion in 4 dimensions. This matching fails in 5 or more dimensions, except, magically, in 10 dimensions. There, as we noted in Section 2, we are allowed to place simultaneously a Weyl and Majorana condition on spinors. Thus, we may define 16-component Weyl-Majorana spinors which describe 8 propagating fermion states. This is exactly the number of transverse photon polarizations in 10 dimensions. Ten-dimensional supergravity with a Majorana-Weyl gravitino contains four 4-dimensional supersymmetries. By an argument analogous to that given in (6.2), this is the maximum number of supersymmetries allowed if we are to have supermultiplets containing only particles of spin 1 or lower. Thus, 10 dimensions provides the setting of maximal symmetry for theories of supergravity coupled to matter.

Having now settled on the dimensionality of space in our grand construction, we must ask, what is the gauge group? Ordinarily, this question has no definite answer; in 4 dimensions, we could choose any possible compact group. However, even in 4 dimensions, the specific representations to which we assign chiral fermions are constrained by the axial vector anomaly which we discussed at the end of Section 3. In general, not only the overall chirality current but also the gauge charge currents may have their conservation spoiled by an anomaly. Since the whole structure of a non-Abelian gauge theory depends on the conservation of the gauge charge currents, one must arrange that these anomalous terms cancel. This is the requirement that, in the $SU(5)$ grand unified theory, forces one to include equal numbers of $\mathbf{5}$ and $\overline{\mathbf{10}}$ multiplets of left-handed fermions.

In 10 dimensions, however, the cancellation of anomalies is much more difficult, for two reasons. First, the 10-dimensional anomalies arise from diagrams of the form of Fig. 18; 5 powers of the external field appear because in 10 dimensions there are 5 powers of $F_{\mu\nu}$ on the right-hand side of (3.31). The diagrams giving nonconservation of gauge currents thus contain a trace of 6 group matrices. Since this number is even, some invariants which appear in this trace will be necessarily positive for any group representation and so will not cancel when we sum over representations. Second, 10-dimensional Weyl-Majorana spinors are chiral with respect to their gravitational couplings and are their own antiparticles; this implies that we must consider gravitational as well as gauge contributions

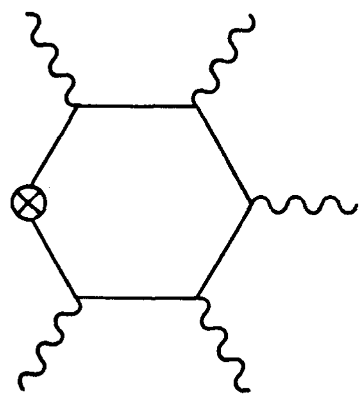

Fig. 18. A typical diagram contributing to the axial vector anomaly in 10 dimensions.

to the anomaly.[23] Until the summer of 1984, it was believed that *no* system of 10-dimensional supersymmetric matter multiplets allowed the cancellation of all gauge and gravitational anomalies. But then Green and Schwarz[24] discovered that a field in the 10-dimensional supergravity multiplet can also enter the expression for these anomalies, and that one can use this additional degree of freedom to arrange cancellations. Their mechanism works, however, only for two specific gauge groups: $O(32)$ and $E_8 \times E_8$. These two groups, then, provide the only possible choices for fundamental gauge groups of supersymmetric theories in 10 dimensions.

The group E_8 is the largest exceptional group in Cartan's classification. It contains in a natural way the groups which appear most often in discussions of grand unification; for example, one may decompose

$$E_8 \supset SU(3) \times E_6 \supset O(10) \supset SU(5) . \tag{6.9}$$

In the first step of this decomposition, $SU(3) \times E_6$, is a maximal subgroup of E_8. Though E_6 is somewhat less familiar that its subgroups $O(10)$ and $SU(5)$, it is a quite reasonable choice for a grand unifying group.[25] The fundamental 27-dimensional representation of E_6 contains one generation of quarks and leptons.

Let us, then, concentrate on the possible gauge group $E_8 \times E_8$. Candelas, Horowitz, Strominger, and Witten[7] have proposed the following compactification of this theory:

$$10 - \mathrm{d} \rightarrow (4 - \mathrm{d\ Minkowski}) \times (\mathcal{C}_6) . \tag{6.10}$$

Here $\mathcal{C}_6$ is a 6-dimensional space with the property that it possesses a spinor field configuration $\epsilon_a(y)$ satisfying $D_\eta \epsilon = 0$. Such a space is called a Calabi-Yau manifold.[26] Since this spinor field is a Killing spinor, the 4-dimensional theory resulting from the compactification will contain a gravitino and thus an unbroken local supersymmetry.

The compactification of Candelas, Horowitz, Strominger, and Witten turns out to have a number of mathematical features which produce remarkable effects in the physics. First, the existence of a covariantly constant spinor implies that the Ricci tensor vanishes, both on the 6-dimensional space and on the full 10-dimensional space. This allows the space (6.10) to solve Einstein's equations without invoking an adjustable cosmological constant term. Second, the rotation group in 6 dimensions, $O(6)$, is locally isomorphic to $SU(4)$, with the $\mathbf{4}$ of $SU(4)$ corresponding to the chiral spinor representation of $O(6)$. At each point, the covariantly constant spinor picks out a preferred direction in this representation, leaving over an orthogonal $SU(3)$. All of the rotation done by the curvature tensor must be within this subspace. This allows one to construct a solution of the Cremmer-Scherk type, with an $SU(3)$ gauge field proportional to the spin connection on the Calabi-Yau space

$$\omega_\eta{}^{\alpha\beta} \;=\; A_\eta{}^i\,(T^i)^{\alpha\beta} \;, \tag{6.11}$$

and a corresponding field strength proportional to the curvature tensor

$$R_{\eta\varsigma}{}^{\alpha\beta} \;=\; F_{\eta\varsigma}{}^i\,(T^i)^{\alpha\beta} \;, \tag{6.12}$$

where (T^i) is a representation matrix of $O(6)$ belonging to the $SU(3)$ subgroup we have picked out. We have already noted that $SU(3)$ is naturally a subgroup of E_8. Fermions in the adjoint representation of E_8 decompose under $SU(3)\times E_6$ in the following way:

$$(\mathrm{adj}E_8) \;\rightarrow\; (\mathrm{adj}E_6, \mathbf{1}) \;+\; (\mathbf{27}, \overline{\mathbf{3}}) \;+\; (\overline{\mathbf{27}}, \mathbf{3}) + (\mathbf{1}, \mathbf{8}) \;. \tag{6.13}$$

Third, Calabi-Yau spaces have no isometries. This means that the compactification (6.10) will produce no Kaluza-Klein gauge bosons. This is no problem, since we began with a fundamental $E_8 \times E_8$ gauge theory, we have more than enough already. Fourth, the right-hand side of the Atiyah-Singer Index Theorem in 6 dimensions contains the integral of a product of three powers of the field strength. Using (6.12), this integal may be converted to the structure

$$\int d^6y\, \epsilon^{\eta\varsigma\theta\pi\rho\sigma} \epsilon_{\alpha\beta\gamma\delta\epsilon\iota} R_{\eta\varsigma}{}^{\alpha\beta} R_{\theta\pi}{}^{\gamma\delta} R_{\rho\sigma}{}^{\epsilon\iota} \ . \tag{6.14}$$

More explicitly, the Index Theorem appropriate this field configuration is:

$$N_{\mathbf{3}} - N_{\bar{\mathbf{3}}} = \frac{1}{2}\chi \ , \tag{6.15}$$

where χ is the Euler number of the Calabi-Yau space. In a generic situation, we would have only zero modes for fermions in the **3** of $SU(3)$. According to (6.13), fermions belonging to the **3** of $SU(3)$ also belong to the **27** of $E(6)$. Thus, the fermion zero modes should yield 4-dimensional fermions corresponding to $\frac{1}{2}\chi$ **27**'s of E_6, that is, $\frac{1}{2}\chi$ generations of quarks and leptons. In this model, then, the number of quark and lepton generations reflects directly the topology of the compactified component of the higher-dimensional space-time.

Some difficulties remain in making this 10-dimensional compactification into a complete unified theory of Nature. We should, for example, worry about whether the theory underlying this computation is renormalizable, so that systematic corrections can be calculated. We might also ask whether the fundamental gauge group $E_8 \times E_8$ has a geometrical origin, so that the model really does unify matter and gravity. These problems seem extremely difficult to address, but, in fact, they are answered and perhaps solved if the true underlying theory is taken to be a theory of superstrings.[6,27] One might ask whether such a compactification can correctly predict the spectrum of quark and lepton masses. The answer to question is not known, but the 10-dimensional compactification has suggested some concrete ideas for attacking this long-standing problem.[28] One might also worry about another fundamental question, why one particular compactification which yields 4 extended dimensions and 3 generations is chosen over all others. I think it is fair to say that no one has any idea of how to solve this problem.

I hope, though, that this stunning example has succeeded in persuading you that Nature could well have more than 4 dimensions, and that the study of such higher-dimensional theories is not simply speculation but also physics. Even if the size of the extra, compact, dimensions is extremely small, their basic geometry can determine crucial features of the physics at much larger distances. Through specific examples, I have shown that the isometries of the compact dimensions can determine the gauge symmetries of Nature, that the spinor isometries of these dimensions can determine the amount of supersymmetry in Nature, and that the topology of these dimensions can determine the number of quark and lepton generations. The idea that there exist more than 4 space-time dimensions is a fundamental departure from previous theories of the basic forces, but perhaps it is just the turn required to resolve some of the fundamental mysteries that these theories still present us.

REFERENCES

1. Th. Kaluza, *Sitz. Preuss. Akad. Wiss. Berlin, Math. Phys.* **K1,** 966 (1921).
2. P. van Nieuwenhuizen, in *Relativity, Groups, and Topology II (Les Houches, 1983)*, B. S. DeWitt and R. Stora, eds. (North-Holland, 1984).
3. E. Witten, in *Shelter Island II,* R. Jackiw, N. N. Khuri, S. Weinberg, and E. Witten, eds. (MIT Press, 1985).
4. M. J. Duff, B. E. W. Nilsson, and C. N. Pope, *Phys. Repts.* **130,** 1 (1986).
5. O. Klein, *Z. Phys.* **37,** 895 (1926).
6. J. H. Schwarz, *Phys. Repts.* **89,** 223 (1982); M. B. Green, in *Proceedings of the 1985 International Symposium on Lepton and Photon Interactions at High Energy,* M. Konuma and K. Takahashi, eds. (Kyoto University, 1986).
7. P. Candelas, G. Horowitz, A. Strominger, and E. Witten, *Nucl. Phys.* **B258,** 46 (1985).
8. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics.* (McGraw-Hill, 1964).
9. F. Gliozzi, J. Scherk, and D. Olive *Nucl. Phys.* **B122,** 253 (1977).
10. T. H. R. Skyrme, *Proc. Roy. Soc.* **A260,** 127 (1961), *Nucl. Phys.* **31,** 556 (1962).
11. A. P. Balachandran, V. P. Nair, S. G. Rajeev, and A. Stern, *Phys. Rev. Lett.* **49,** 1124 (1982), *Phys. Rev.* **D27,** 1153 (1983).
12. E. Witten, *Nucl. Phys.* **B223,** 422, 433 (1983).
13. D. Finkelstein, *J. Math. Phys.* **7,** 1218 (1966).
14. M. F. Atiyah and I. M. Singer, *Ann. Math.* **87,** 484 (1968).
15. G. 't Hooft, *Phys. Rev. Lett.* **37,** 8 (1976).

16. J. Schwinger, *Phys. Rev.* **128,** 2425 (1982).

17. J. S. Bell and R. Jackiw, *Nuov. Cim.* **60A,** 47 (1969); S. L. Adler, *Phys. Rev.* **177,** 2426 (1969).

18. M. J. Duff, *Nucl. Phys.* **B219,** 389 (1983).

19. F. A. Berends, J. W. van Holten, B. de Wit, and P. van Nieuwenhuizen, *J. Phys.* **A13,** 1643 (1980).

20. E. Cremmer, B. Julia, and J. Scherk, *Phys. Lett.* **B76,** 469 (1978).

21. M. J. Duff and C. N. Pope, in *Supersymmetry and Supergravity 82,* S. Ferrara, J. G. Taylor, and P. van Nieuwenhuizen, eds. (World Scientific, 1983).

22. E. Cremmer and J. Scherk, *Nucl. Phys.* **B108,** 409 (1976), **B118,** 61 (1977).

23. L. Alvarez-Gaumé and E. Witten, *Nucl. Phys.* **B234,** 269 (1983).

24. M. B. Green and J. H. Schwarz, *Phys. Lett.* **149B,** 117 (1984).

25. F. Gürsey, P. Ramond, and P. Sikivie, *Phys. Lett.* **60B,** 177 (1976).

26. E. Calabi, in *Algebraic Geometry and Topology: A Symposium in Honor of S. Lefschetz.* (Princeton Univ. Press, 1957); S.-T. Yau, *Proc. Nat. Acad. Sci.* **74,** 1798 (1977).

27. D. Gross, J. Harvey, E. Martinec, and R. Rohm, *Phys. Rev. Lett.* **54,** 502 (1985), *Nucl. Phys.* **B256,** 253 (1985).

28. A. Strominger and E. Witten, *Comm. Math. Phys.* **101,** 341 (1985); A. Strominger, *Phys. Rev. Lett.* **55,** 2547 (1985).