Ref.TH.3259-CERN

KALUZA-KLEIN-KOUNTERTERMS *)

M.J. Duff **)

CERN - Geneva

and

D.J. Toms

Blackett Laboratory
Imperial College, London

A B S T R A C T

We review the reasons for treating eleven-dimensional supergravity as a Kaluza-Klein theory. In this case one should ask not whether $N = 8$ supergravity is finite in four dimensions but whether $N = 1$ supergravity is finite in eleven dimensions. With this motivation, we study the ultra-violet behaviour of Kaluza-Klein theories from both the higher and lower dimensional points of view. Generalized zeta function techniques allow one to express counterterms and anomalies for theories defined on a product manifold $M_1 \times M_2$ in terms of those corresponding to M_1 and M_2 separately. Certain special features of the effective four-dimensional theory, like vanishing anomalies, emerge by compactifying an odd number of dimensions. Some apparent paradoxes are resolved. We also comment on a "spin-one equivalence principle" inherent in Kaluza-Klein theories.

*) Talk given by M.J. Duff at the Second Euro-physics Study Conference on Unification of the Fundamental Interactions, Erice, Sicily, 6-14 October 1981.

**) Permanent address : Blackett Laboratory, Imperial College, London.

1. INTRODUCTION

The major obstacle in constructing a consistent quantum theory of gravitation is that Einstein's general theory of relativity seems, by conventional criteria, to be non-renormalizable and hence incapable of unambiguous prediction at the quantum level. Many ingenious suggestions have been made to circumvent this impasse but none with complete success¹. In this paper, we wish to explore another possibility: to examine the problem within the framework of higher dimensional theories of the Kaluza-Klein² type.

It may seem paradoxical to attempt to improve the ultra-violet behaviour of a theory by increasing the dimension of space-time, but not if we are looking for finiteness as opposed to power-counting renormalizability. Moreover, the chances of finiteness seem to be better if the theory is supersymmetric and/or the dimension of space-time is odd. Indeed, the work presented here was largely inspired by Witten's³ suggestion of obtaining an $SU(3) \times SU(2) \times U(1)$ gauge theory by applying Kaluza-Klein ideas to eleven-dimensional⁴ supergravity.

Kaluza-Klein theories

In its modern form, a Kaluza-Klein theory works as follows. One starts with gravity, described by the Einstein Lagrangian, coupled to matter fields of lower spin in d space-time dimensions, where $d > 4$. Ideally, the resulting field equations should then admit of a stable ground-state solution for which the extra dimensions are compactified. The ground-state metric would then describe a product* manifold $M_1 \times M_2$, where M_1 is four-dimensional space-time with the usual signature (possibly Minkowski space, but more generally allowing a non-vanishing cosmological constant) and M_2 is a compact "internal" space of Euclidean signature. One then performs a Fourier expansion of all the fields, including the gravitational fluctuations about the ground state, with a suitable set of harmonics on M_2 and expansion coefficients depending only on the four co-ordinates of M_1 . Integration over the extra co-ordinates then yields the field equations of an effective four-dimensional theory on M_1 describing a finite set of massless states, including the graviton, together with an infinite tower of massive states corresponding to the normal modes on M_2 . The masses will be quantized in units of m , where m^{-1} is the typical "size" of M_2 , and all states have spins ≤ 2 . Thus these extra dimensions need not conflict with one's everyday sensation of inhabiting a four-dimensional world (with its inverse square law of gravitational attraction) provided the massive states are sufficiently heavy, i.e., provided the compactified dimensions are sufficiently small.

On the microspic scale, we believe that the strong, weak and electromagnetic interactions are described by a Yang-Mills gauge theory. It is necessary therefore that M_2 be either a group space G or else a coset space G/H . In this case the spectrum of four-dimensional states will include, in addition to the massless graviton, massless Yang-Mills fields with gauge group G . Indeed, the whole beauty of Kaluza-Klein theories is that it is not necessary to postulate the separate existence of Yang-Mills fields; they are automatically begotten by gravity. Since G is presumed to contain $SU(3) \times SU(2) \times U(1)$, the number of extra dimensions must be enough to accommodate this**. Requiring that the corresponding gauge coupling constants be of order unity, or thereabouts, fixes

* We have in mind a global direct product, but it would be interesting to examine the consequences of a ground state which consists of a non-trivial fibre bundle. See Ref. 5. I am grateful to C.J. Isham for drawing my attention to this possibility.

** The most economical way to obtain the symmetry group G is to take $M_2 = G/H$ where H is a maximal subgroup of G . In any event, $d_2 = \dim M_2 = \dim G - \dim H$. Moreover, $\dim G$ cannot exceed that of a space of maximal symmetry, i.e., $\dim G \leq d_2(d_2+1)/2$. Hence $\dim H \leq d_2(d_2-1)/2$. I am grateful to D. Olive for discussions on this point.

the extra dimensions to be of the Planck size, i.e., $m^{-1} \sim 10^{-33}$ cm. This nicely meets our requirement that they be small. In consequence, the massive states have masses in excess of 10^{19} GeV and cannot be identified with any known particles. In this approximation, all observed particles are massless. In a realistic theory, one's initial choice of spinor fields in d dimensions would be such as to yield massless spin-1/2 fermions in the right representations of $SU(3) \times SU(2) \times U(1)$ to be identified with the known quarks and leptons. Their physical masses, and those of the W-bosons will then have to be provided by some other mechanism, perhaps by a Higgs-Kibble mechanism since in general scalar fields will also be fathered by gravity.

As far as low energy phenomenology is concerned, therefore, one can discard these massive Kaluza-Klein states and focus one's attention on the massless sector. However, the inclusion of an infinite tower of massive states will drastically affect the ultra-violet behaviour of the theory*. This is the subject of the present paper.

[To avoid confusion, one should contrast this picture of "spontaneous compactification"⁶ with that of "dimensional reduction"⁷: another device which also appears in modern treatments of higher dimensional theories (like the bosonic string in $d = 26$ or the fermionic string in $d = 10$). In its original form, dimensional reduction simply consists of taking all fields to be independent of the extra co-ordinates. In the above Kaluza-Klein picture, this corresponds to a flat ground-state $M_1 \times M_2$, where M_1 is Minkowski space, and M_2 is an n -torus, (i.e., $S^1 \times S^1 \times \dots S^1$) whose radii are then shrunk to zero to eliminate the massive states. An interesting modification, permissible when there is an internal symmetry, consists of taking the fields to be "twisted", i.e., not strictly periodic but only periodic up to a phase⁸. The special dependence on the extra co-ordinates then gives rise to symmetry breaking and mass generation. In neither form of dimensional reduction need one ascribe any physical significance to the extra dimensions, in sharp contrast with the spontaneous compactification described above.]

Supergravity

The Kaluza-Klein idea may, in principle, be applied to any theory of gravity plus matter in any dimension greater than four**.

* This infinite tower of massive states would presumably also have cosmological consequences.

** It has also been applied to theories with $s \leq 1$, but this seems the worst of all possible worlds. One acquires all the disadvantages of non-renormalizability without its only advantage: gravity.

Unfortunately, no-one yet knows whether there exists a choice of matter coupling and choice of dimension which fulfils all requirements of a realistic theory listed above*. However, it is particularly compelling to apply the idea to supergravity, for the following reasons:

1) Supergravity severely restricts both the choice of matter couplings and the choice of dimension. Such a restriction must be a feature of any successful attempt at Kaluza-Klein unification. Otherwise, the mystery of the correct matter fields in four dimensions would simply be swapped for the same mystery in d dimensions, with the extra headache of which d to pick. In supergravity the maximum dimension permitting a consistent** ($s \leq 2$) theory is eleven¹³. In this $d = 11$ theory, all fields are gauge fields: the gravitational elfbein $e_\mu^a(x)$, the rank-three antisymmetric tensor $A_{\mu\nu\rho}(x)$, and a single Rarita-Schwinger Majorana spinor $\psi_\mu^\alpha(x)$ (hence $N = 1$ supersymmetry).

2) In Kaluza-Klein, the mass matrix for the fermion fields on M_1 is given by the fermion operator (Dirac or Rarita-Schwinger) on M_2 . The number of massless fermions is therefore determined by the number of zero-eigenvalue modes of this fermion operator on the coset space G/H . With the exception of the n -torus (which featured in the dimensional reduction discussed earlier), all such spaces have positive scalar curvature (at least with their standard metric) and hence by a theorem due to Lichnerowicz¹⁴ there are no zero-modes of the Dirac operator. One may circumvent this theorem by including Yang-Mills gauge fields in the higher dimension*** but this would be contrary to the spirit of Kaluza-Klein whereby the Yang-Mills fields emerge automatically. Fortunately, Lichnerowicz's theorem does not apply to the Rarita-Schwinger operator****. Thus to

* In particular, the stability of the Kaluza-Klein vacuum is an especially knotty problem. See Ref. 9 for a treatment of this problem in the case of $M_1 = \text{Minkowski space}$ and $M_2 = S^1$.

** The problem of finding a consistent coupling of spin $\geq 5/2$ to gravity is qualitatively different from $3/2$, and probably has no solution. See Refs. 10, 11 and 12.

*** After all, we know that the Dirac operator on S^4 has zero modes in the presence of Yang-Mills instantons.

**** Zero modes of the Rarita-Schwinger operator on compact spaces of Euclidean signature have been treated before in the context of index theorems and axial anomalies. See Refs. 15, 16, 17 and 10.

explain spin-1/2 quarks and leptons in four dimensions via Kaluza-Klein, spin-3/2 fields seem almost* compulsory. But the only known consistent spin-3/2 theory is supergravity !

3) Most higher dimensional theories do not exhibit spontaneous compactification**. In $d = 11$ supergravity, not only does spontaneous compactification work but the number of space-time dimensions is naturally singled out as four²⁰! This is because a rank-three gauge field gives rise to a cosmological constant^{21,22} and because its field strength $F_{\mu\nu\rho\sigma}$ has four indices. The field equation for F

$$d^*F = F \wedge F \quad (1.1)$$

admits the solution ($c = \text{constant}$)

$$\begin{aligned} F_{\mu\nu\rho\sigma} &= c \, \epsilon_{\mu\nu\rho\sigma} \quad , \quad \mu, \nu, \rho, \sigma = 1, 2, 3, 4 \\ &= 0 \quad , \quad \text{otherwise} \end{aligned} \quad (1.2)$$

The resulting Einstein's equations then admit a ground state metric of the product form $M_1 \times M_2$, where M_1 is a four dimensional Einstein

* The absence of Dirac zero-modes on G/H has some fascinating exceptions involving non-standard metrics. It is not difficult to prove that on Einstein spaces ($R_{ij} = \Lambda \, \gamma_{ij}$) Killing vectors exist only when $R \geq 0$ and hence, by Lichnerowicz, there are no Dirac zero modes. However, if one is willing to relax the Einstein condition then spaces with negative R can admit both Killing vectors and Dirac zero modes. A classic example discussed by Hitchin in Ref. 18 is the three sphere S^3 . Relative to the standard $SU(2) \times SU(2)$ invariant metric, it has positive R and no Dirac zero modes. However, there exists a one-parameter (λ) non-standard metric with $SU(2) \times U(1)$ symmetry and negative R which not only admits Dirac zero modes but for which the number of such modes depends on the value of λ . Counting them turns out to be a non-trivial exercise in number theory. (I am grateful for discussions on these points with R. Coquereaux.) Within a Kaluza-Klein framework, this could (?) yield a unified electroweak theory for which the number of quarks and leptons depended, via the prime numbers, on the value of θ_w .

** The introduction of Yang-Mills gauge fields sometimes helps but, as mentioned before, this is contrary to the spirit of Kaluza-Klein. Another possibility is to invoke higher derivatives, see Ref. 19.

space ($R_{\mu\nu} = \Lambda g_{\mu\nu}$) with Minkowski signature and $\Lambda = -4c^2 < 0$ *, while M_2 is a seven-dimensional Einstein space with Euclidean signature and $\Lambda = 2c^2 > 0$ and hence compact.

4) Seven is not merely the maximum dimension of $M_2 = G/H$ permitted by supersymmetry, but also the minimum dimension that will accommodate $G \supset SU(3) \times SU(2) \times U(1)$. This remarkable numerical "coincidence" was first observed by Witten³, who showed that there are infinitely many seven-dimensional spaces with the right symmetry. A well-known example would be $CP^2 \times S^2 \times S^1$, though this example illustrates two potential problems. First, CP^2 does not admit of a spin structure, i.e., fermions cannot be globally defined. It does admit of a generalized (spin^C) structure²⁶, but it is unclear whether this is of any help in a Kaluza-Klein theory. Fortunately, there are other spaces with the right symmetry which do have a spin structure. Secondly, this way of obtaining the $U(1)$ factor requires that one of the extra dimensions, the S^1 circle, be flat. It is notoriously difficult to find solutions of the field equations with some dimensions curved and some flat**. For example, the seven-dimensional Einstein space of positive curvature discussed in (3) above would allow S^7 , $S^5 \times S^2$, $S^4 \times S^3$, $CP^2 \times S^3$ but not $CP^2 \times S^2 \times S^1$ or $S^6 \times S^1$ ***. (The product of two Einstein spaces is also an

* The appearance at the classical level of an enormous cosmological constant for space-time M_1 is a general feature of Kaluza-Klein theories whenever M_2 is non-flat, i.e., whenever G is non-Abelian. Supergravity does not appear to solve this problem unless one can argue that the ground state need not satisfy the classical field equations. A related problem which supergravity might cure concerns the enormous Casimir vacuum energies induced by compactified dimensions of the Planck scale. See Ref. 23. One might anticipate that the fermion and boson contributions to these vacuum energies will cancel. See Refs. 24 and 25.

** For the same reason that a ground state M_1 = Minkowski space is difficult to achieve when M_2 is not flat.

*** Another possibility would be to start not with $N = 1$ supergravity in $d = 11$, but with $N = 2$ supergravity in $d = 10$, which already has a $U(1)$ symmetry. (This theory has the distinction of being the zero-slope limit of the fermionic dual-string model.) One would then require compactification of the extra six dimensions, all of which may now be curved, to yield $SU(3) \times SU(2)$. This is possible since the $A_{\mu\nu\rho}$ field is also present in this theory. However, the fermions are neutral under this $U(1)$. Moreover, in contrast to $N = 1$ in $d = 11$, this theory is not finite even at one loop. See Section 5.

Einstein space, provided they have the same Λ .) $S^5 \times S^2$ has the properties of admitting a spin structure, solving the field equations and accommodating the right symmetries. It would yield an $SU(4) \times SU(2)$ Yang-Mills theory. (If we are forced by the field equations to give S^5 and S^2 the same Λ , this Yang-Mills theory would have only one independent coupling constant.) On the other hand, Witten³ has argued that the ground state need not necessarily be a solution of the field equations.

5) Finally, and this is the problem that concerns us most in this paper, any Kaluza-Klein theory must face up to the problem of non-renormalizability. The superficial degree of divergence of a Feynman graph in a d dimensional Einstein theory is given by¹

$$D = (d-2) L + 2 \quad (1.3)$$

where L is the number of loops. Now although D is greater for $d=11$ than $d=4$, this is not so significant since all $d > 2$ theories are anyway power-counting non-renormalizable. An entirely different possibility is that the theory is finite owing to a mutual cancellation of ultra-violet divergences. This is known to happen to at least two loop order in the $N=8$ supergravity in $d=4$ obtained by dimensional reduction from $N=1$ in $d=11$ (and also in the $N < 8$ theories obtained by truncation from $N=8$). However, if we are to take seriously the extra dimensions, the relevant question is not "Is $N=8$ supergravity finite in $d=4$?", but rather "Is $N=1$ supergravity finite in $d=11$?". Since the counterterm structure of simple and extended supergravity is still not well understood, finiteness in $d=11$ seems a priori just as likely as in $d=4$. The chances of a non-supersymmetric theory being finite, on the other hand, seem very remote whatever the dimension.

It is amusing to note from (1.3) that $d=11$ supergravity is automatically finite at odd-loop order because there are no local invariants with an odd number of derivatives which could serve as counterterms*. Consequently, one expects that the effective four-dimensional theory will also be finite at odd loops provided one retains the infinite tower of massive states: something which would not necessarily be true if these states are discarded as in dimensional reduction.

* We are assuming a regularization scheme with a dimensionless regularizing parameter like dimensional or ζ function regularization, as opposed to Pauli-Villars for example. It should be added, however, that even in this case, matter fields could spoil the odd-loop finiteness of pure gravity. Fortunately, this does not happen in supergravity. See Section 5.

Whether or not a Kaluza-Klein theory is supersymmetric, it seems bizarre that an infinite number of states can yield a finite result, when a finite number of states leads to infinities. Equally strange is the situation where a non-vanishing counterterm in the higher dimensional theory survives as a counterterm in the effective four-dimensional theory. The puzzle is that it will have more derivatives than four-dimensional power-counting would allow. Similar puzzles occur for the axial and trace anomalies.

In this paper we attempt to clarify these issues. Although motivated by the possibility of a physically realistic Kaluza-Klein theory, we shall not invoke the full apparatus of eleven dimensional supergravity but content ourselves with some simple toy models. These examples, albeit unrealistic, will serve to illustrate some of the apparent paradoxes which arise when Kaluza-Klein theories are viewed first from the higher, and then from the lower dimensional standpoints.

Higher dimensional picture versus lower dimensional picture

In order to formulate some of these ultra-violet puzzles more precisely, consider the action for a single scalar field Φ on a five-dimensional manifold $M(5)$ with metric $g_{\hat{\mu}\hat{\nu}}$:

$$S = \frac{1}{2} \int d^5x \sqrt{\det g_{\hat{\mu}\hat{\nu}}} g^{\hat{\mu}\hat{\nu}} \partial_{\hat{\mu}} \Phi \partial_{\hat{\nu}} \Phi \quad (1.4)$$

Suppose that one of the dimensions is compactified so that

$$M(5) = M(4) \times S^1 \quad (1.5)$$

where $M(4)$ is unspecified. Let us write the co-ordinates

$$x^{\hat{\mu}} = (x^{\mu}, y) \quad (1.6)$$

where y is the S^1 co-ordinate

$$0 \leq y \leq \frac{2\pi}{m} \quad (1.7)$$

and m is the inverse radius of the circle. Since, in this simple example, the gravitational field acts as an external source with no dynamics of its own, we are free for illustrative purposes to choose the simple form

$$g_{\hat{\mu}\hat{\nu}}(x, y) = \begin{bmatrix} g_{\mu\nu}(x) & 0 \\ 0 & 1 \end{bmatrix} \quad (1.8)$$

Next, consider the Fourier expansion of the real scalar field

$$\Phi(x,y) = \sqrt{\frac{m}{2\pi}} \sum_{n=-\infty}^{\infty} \phi_n(x) \exp inmy \quad (1.9)$$

[Note that the expansion coefficients $\phi_n(x) = \phi_{-n}^*(x)$ have the correct canonical dimensions of a Bose field in four dimensions. In fact the canonical commutation relations for $\phi_n(x)$ follow from these of $\Phi(x,y)$ after integrating over y .] Substituting (1.8) and (1.9) into (1.4), and performing the y integration, one finds

$$S = \frac{1}{2} \sum_{n=-\infty}^{\infty} \int d^4x \sqrt{\det g_{\mu\nu}} [g^{\mu\nu} \partial_\mu \phi_n^* \partial_\nu \phi_n + n^2 m^2 \phi_n^* \phi_n] \quad (1.10)$$

Thus, instead of regarding the theory as that of a single real massless scalar in five dimensions, we are at liberty to interpret it as the four-dimensional theory of a single real massless scalar (the $n = 0$ mode) plus an infinite tower of complex massive scalars with quantized masses

$$m_n^2 = n^2 m^2 \quad (1.11)$$

Note that the masses of these complex scalars tend to infinity as we shrink the radius of the circle to zero.

The above example is a special case of a somewhat more general set-up where the scalar is defined on a product manifold

$$M = M_1 \times M_2 \quad (1.12)$$

with dimension

$$d = d_1 + d_2 \quad (1.13)$$

with co-ordinates

$$\begin{aligned} \hat{x}^{\hat{\mu}} &= (x^\mu, y^i) & \hat{\mu} &= 1, \dots, d \\ & & \mu &= 1, \dots, d_1 \\ & & i &= 1, \dots, d_2 \end{aligned} \quad (1.14)$$

and product metric

$$g_{\hat{\mu}\hat{\nu}}(x,y) = \begin{bmatrix} g_{\mu\nu}(x) & 0 \\ 0 & \gamma_{ij}(y) \end{bmatrix} \quad (1.15)$$

From the higher, d dimensional point of view the action is

$$S = \int d^d x \frac{1}{2} \sqrt{\det g_{\hat{\mu}\hat{\nu}}} \Phi \Delta \Phi \quad (\text{A.1})$$

where

$$\Delta = -\nabla^{\hat{\mu}} \nabla_{\hat{\mu}} \quad (\text{1.16})$$

is the Laplacian on M which splits as the sum

$$\Delta = \Delta_1 + \Delta_2 \quad (\text{1.17})$$

where Δ_1 and Δ_2 are the Laplacians on M_1 and M_2 respectively. Consequently, from the lower d_1 dimensional point of view the action is

$$S = \sum_n S_n = \sum_n \int d^{d_1} x \frac{1}{2} \sqrt{\det g_{\mu\nu}} [\phi_n \Delta_1 \phi_n + m_n^2 \phi_n \phi_n] \quad (\text{B.1})$$

where

$$m_n^2 = \lambda_n \quad (\text{1.18})$$

and λ_n are the eigenvalues of Δ_2 with eigenfunctions $Y_n(y)$

$$\Delta_2 Y_n(y) = \lambda_n Y_n(y) \quad (\text{1.19})$$

The fields ϕ_n given by

$$\Phi(x, y) = \sum_n \phi_n(x) Y_n(y) \quad (\text{1.20})$$

will belong to the appropriate representation of the symmetry group of M_2 . For example, if M_1 = space-time with $d_1 = 4$ and M_2 = the two-sphere S^2 with $d_1 = 2$ and radius m^{-1} , then Y_n are just the spherical harmonics

$$\Phi(x, \theta, \phi) = \sum_{\ell m} \phi_{\ell m}(x) Y_{\ell m}^{\ell}(\theta, \phi) \quad (\text{1.21})$$

and $\phi_{\ell m}(x)$ belong to the $(2\ell+1)$ dimensional representation of $SO(3)$ with masses

$$m_{\ell m}^2 = \ell(\ell+1)m^2 \quad (\text{1.22})$$

Let us now turn to the quantization of this scalar field still keeping classical the gravitational background. We can choose to view the quantization from standpoint A, starting with the action (A.1) or from standpoint B, starting with the action (B.1).

In case A, the effective action W is given by the functional integral

$$\exp - W[g] = \int d\Phi \exp - S[g, \Phi] \quad (A.2)$$

where, for convenience, we have Wick rotated to a Euclidean signature, and where S is given by (A.1). Performing the Gaussian integral yields

$$W = \frac{1}{2} \ln \det \Delta \quad (A.3)$$

W is the generating functional for all one-loop graphs with internal Φ lines and with the external $g_{\hat{\mu}\hat{\nu}}(x,y)$ field acting at the vertices.

In case B, the effective action is given by

$$\exp - W[g] = \int \prod_n d\phi_n \exp - \sum_n S_n[g, \phi_n] \quad (B.2)$$

where S_n are given by (B.1). Hence

$$W = \sum_n W_n = \frac{1}{2} \sum_n \ln \det (\Delta_1 + m_n^2) \quad (B.3)$$

W_n is the generating functional for all one-loop graphs with internal ϕ_n lines and with the external $g_{\mu\nu}(x)$ field acting at the vertices.

From the properties of determinants, it is clear that the quantum effective action on the right-hand side of (A.3) is identical to the quantum effective action on the right-hand side of (B.3). At this formal level, therefore, the quantum theories as well as the classical theories are the same whether one adopts viewpoint A or viewpoint B. What is far from clear, however, is whether the ultra-violet divergences depend on which picture one adopts, A or B. The following questions spring to mind:

Q1) In picture A, the power-counting rules are clearly those appropriate to d dimensions. But consider again the example mentioned previously. A massive scalar in a four-dimensional theory ($d_1 = 4$) would normally give rise to one-loop counterterms of the form $m^4\sqrt{g}$, $m^2\sqrt{g} R$, and $\sqrt{g}(\text{Riemann})^2$. Yet a counterterm like $\sqrt{g}(\text{Riemann})^3$ should appear in picture A if we chose $d = 6$. In picture B, therefore, do we count powers in d_1 or in d dimensions?

Q2) Similarly, if we are to use dimensional regularization, does the location of the poles giving rise to the counterterms change in going from one picture to the other?

Q3) At the classical level, shrinking the radii of the extra dimensions sends the masses of all the massive modes to infinity: the only surviving propagating modes are the massless ones. After renormalization, however, does one also recover the quantum theory of the massless modes in this limit?

Q4) The effective action $W = \sum_n W_n$ of Eq. (B.3) will be divergent and will require the addition of a counterterm δW to render it finite. Similarly the individual effective actions W_n will require counterterms δW_n . Is δW equal to $\sum_n \delta W_n$?

Q5) Is δW influenced by the choice of extra dimensions? Consider again our $d = 6$ example; would it make any difference to choose $M_2 = S^2$ or $M_2 = S^1 \times S^1$?

Finally, if d is odd the theory is finite in picture A, i.e., $\delta W = 0$. On the other hand, we know that a single massive scalar in four dimensions certainly has divergences, i.e., $\delta W_n \neq 0$. If picture B is to agree with picture A, how can a theory with an infinite number of massive fields be finite when with a finite number of massive fields it is infinite?

The most elegant techniques for resolving these puzzles are those of "generalized ζ functions" and the "heat kernel expansion" which we summarize in Section 2. In Section 3 we then apply these techniques to examine the problem of ultra-violet divergences in Kaluza-Klein theories from both the higher and lower dimensional points of view, establishing the equivalence of the two pictures and resolving the apparent paradoxes. The above puzzles also apply to the axial and trace anomalies. These are dealt with in Section 4, where a toy model with fermions is introduced and where we keep the spin 1 gauge fields in picture B in addition to the gravitational field. Of particular interest are the formulae describing the fermion zero modes on $M = M_1 \times M_2$ in terms of those on M_1 and M_2 . We also comment on the universality of charge to mass ratios in Kaluza-Klein theories: "The spin-one equivalence principle". Section 5 brings the reader up to date with new developments since the Erice conference.

2. ZETA FUNCTIONS AND THE HEAT KERNEL

In this section we summarize the properties of generalized zeta functions and the heat kernel expansion which we employ in Section 3. See Ref. 27 for a review, and Ref. 28 for further references.

The heat kernel

In a compact manifold, M , a second order Laplace-type operator of the form $\Delta = -\square + x$ will possess a discrete number of eigenfunctions ϕ_n and corresponding eigenvalues λ_n ,

*

Note incidentally that this Kaluza-Klein picture should be contrasted with the $N \rightarrow \infty$ limit of a $1/N$ expansion. There the infinitely many fields have the same mass and are in the same internal symmetry representations. Here, they have different masses and are in different representations.

$$\Delta \phi_n = \lambda_n \phi_n \quad (2.1)$$

Introduce a parameter t and define the traced "heat kernel"

$$\begin{aligned} Y(\Delta, t) &\equiv \text{Tr} \exp - \Delta t \\ &= \sum_n \exp - \lambda_n t \end{aligned} \quad (2.2)$$

It can be shown, and we omit the proof, that $Y(t)$ possesses an asymptotic expansion, valid as $t \rightarrow 0^+$, given by

$$Y(\Delta, t) \sim t^{-d/2} \sum_{k=0}^{\infty} A_k(\Delta) t^{k/2} \quad (2.3)$$

where d is the dimension of M , where

$$A_k(\Delta) = \frac{1}{(4\pi)^{d/2}} \int d^d x \sqrt{g} a_k(\Delta) \quad (2.4)$$

and where a_k are d independent invariants constructed from $g_{\mu\nu}$ and its derivatives of order k .

As an example, suppose Δ is the Laplacian on scalars of Section 1, then

$$\begin{aligned} a_0 &= 1 \\ a_2 &= \frac{1}{6} R \\ a_4 &= \frac{1}{180} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2) \\ a_6 &= \alpha R_{\mu\nu}^{\rho\sigma} R_{\rho\sigma}^{\alpha\beta} R_{\alpha\beta}^{\mu\nu} \\ &\quad + \beta R_{\mu\nu\rho\sigma} R^{\mu\alpha\rho\beta} R_{\alpha\beta}^{\nu\sigma} \\ &\quad + \gamma \nabla^\alpha R_{\mu\nu\rho\sigma} \nabla_\alpha R^{\mu\nu\rho\sigma} + \dots \end{aligned} \quad (2.5)$$

The precise form of a_6 is given in Ref. 29. The odd coefficients $a_1, a_3, a_5, \dots$ vanish when M is compact without boundary*.

* In Kaluza-Klein M_2 will always be compact without boundary, whereas space-time M_1 would in general be non-compact. The techniques described here are valid if M_1 is non-compact with negligible boundary. The method also generalizes to the case of boundaries but will not be discussed here.

Zeta functions

In analogy with the Riemann ζ function³⁰

$$\zeta_R(s) = \sum_{n=1}^{\infty} n^{-s} \quad (2.6)$$

define a generalized ζ function formed from the eigenvalues of Δ

$$\zeta(\Delta, s) = \sum_n \lambda_n^{-s} \quad (2.7)$$

This is related to $Y(\Delta, t)$ of (2.2) by a Mellin transform

$$\zeta(\Delta, s) = \frac{1}{\Gamma(s)} \int_0^{\infty} dt t^{s-1} Y(\Delta, t) \quad (2.8)$$

$$\begin{aligned} \zeta(\Delta, s) &= \frac{1}{\Gamma(s)} \int_0^1 dt t^{s-1} Y(\Delta, t) \\ &\quad + \frac{1}{\Gamma(s)} \int_1^{\infty} dt t^{s-1} Y(\Delta, t) \end{aligned} \quad (2.9)$$

The second integral converges for all s , since $Y \sim \exp - \lambda_{\min} t$ for large t , and the second term vanishes for $s = 0, -1, -2$. In the first term we may use the asymptotic expansion (2.3). This shows that $\zeta(\Delta, s)$ is analytic for $\text{Re } s > d/2$. When d is even, $\zeta(\Delta, s)$ has only $d/2$ simple poles at $s = (d-k)/2$ for $k = 0, 2, 4, \dots (d-2)$ with residue $A_{d-2s}/\Gamma(s)$. In particular, for $s = 0, -1, -2, \dots$ $\zeta(s, \Delta)$ is finite and given by

$$\zeta(\Delta, s) = (-1)^s (-s)! A_{d-2s}(\Delta) \quad (2.10)$$

When d is odd, $\zeta(\Delta, s)$ has an infinite number of simple poles at $s = (d-k)/2$ for $k = 0, 2, 4, \dots$ with residue $A_{d-2s}/\Gamma(s)$, and vanishes for $s = 0, -1, -2$.

We note in particular that

$$\zeta(\Delta, 0) = A_d(\Delta) \quad (2.11)$$

gives a meaning, via (2.7), to the number of eigenmodes of the operator Δ , a quantity which is formally infinite. We have assumed so far that Δ is positive. In general, of course Δ will have zero modes, which must first be projected out before forming the generalized ζ functions. In this case (2.11) generalizes to

$$A_d(\Delta) = n(\Delta) + \zeta(\Delta, 0) \quad (2.12)$$

where $n(\Delta) = \text{integer}$ is the number of zero modes. Thus $A_d(\Delta)$ counts all the modes and $\zeta(\Delta, 0)$, which is in general non-integer and sometimes even negative, "counts" the non-zero modes.

For example, suppose in our Kaluza-Klein theory of Section 1 Δ is the Laplacian on $M_2 = S^1$, i.e.,

$$\Delta = - \frac{\partial^2}{\partial y^2} \quad (2.13)$$

then

$$\lambda_n = m_n^2 = n^2 m^2, \quad n = 0, \pm 1, \pm 2, \dots \quad (2.14)$$

and

$$\begin{aligned} \zeta(\Delta, s) &= 2 \sum_{n=1}^{\infty} (n^2 m^2)^{-s} \\ &= 2 m^{-2s} \zeta_R(2s) \end{aligned} \quad (2.15)$$

where ζ_R is the Riemann ζ function (2.6). Thus the effective four-dimensional theory, given by the action (1.10) describes one massless state [$n(\Delta) = 1$], while the "number" of massive states is

$$\zeta(\Delta, 0) = 2 \zeta_R(0) = -1 \quad (2.16)$$

where we have used $\zeta_R(0) = -1/2$. See Ref. 30. Hence from (2.12)

$$A_1(\Delta) = 1 - 1 = 0 \quad (2.17)$$

in accordance with the vanishing of the $A_k(\Delta)$ coefficients when k is odd.

As another example, suppose Δ is the Laplacian on S^2 with radius m^{-1} and

$$\lambda_n = n(n+1)m^2 \quad (2.18)$$

with multiplicity $(2n+1)$. Then

$$\zeta(\Delta, s) = \frac{1}{m^{2s}} \sum_{n=1}^{\infty} \frac{(2n+1)}{[n(n+1)]^s} \quad (2.19)$$

Using the techniques described in Ref. 31, one can show that

$$\begin{aligned} \zeta(\Delta, 0) &= \left[-\frac{3}{4} - \zeta_R(-1) \right] \\ \zeta(\Delta, -1) &= \frac{m^2}{4} \left[-\frac{1}{8} + \zeta_R(-1) - 7 \zeta_R(-3) \right] \end{aligned} \quad (2.20)$$

$$\zeta(\Delta, -2) = \frac{m^4}{16} \left[\frac{1}{12} - \zeta_R(-1) + 14 \zeta_R(-3) - 31 \zeta_R(-5) \right]$$

Since $\zeta_R(-1) = 1/6$, $\zeta_R(-3) = 1/120$ and $\zeta_R(-5) = -1/252$, one finds

$$\begin{aligned}
 A_0(\Delta) &= m^{-2} \\
 A_2(\Delta) &= 1 + \zeta(\Delta, 0) = \frac{1}{3} \\
 A_4(\Delta) &= -\zeta(\Delta, -1) = \frac{m^2}{15} \\
 A_6(\Delta) &= \frac{1}{2} \zeta(\Delta, -2) = \frac{4m^4}{315}
 \end{aligned} \tag{2.21}$$

These results may be verified by inserting the curvature R_{ijkl} for S^2 with metric γ_{ij}

$$R_{ijkl} = m^2(\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}) \tag{2.22}$$

into the formula (2.4) and (2.5).

One-loop counterterms

Consider now a typical one-loop effective action given by

$$W = \frac{1}{2} \ln \det \Delta = \frac{1}{2} \sum_n \ln \lambda_n \tag{2.23}$$

W will be ultra-violet divergent. To regularize we may use dimensional regularization and replace the dimension d of M by $d + \epsilon$. Thus (2.3) becomes

$$Y(\Delta, t) \sim t^{-(d+\epsilon)/2} \sum_{k=0}^{\infty} A_k t^{k/2} \tag{2.24}$$

From (2.23) and (2.8),

$$\begin{aligned}
 W &= - \frac{d}{ds} \zeta(\Delta, s) \Big|_{s=0} \\
 &= - \frac{1}{2} \int_0^{\infty} dt t^{-1} Y(\Delta, t)
 \end{aligned} \tag{2.25}$$

Employing the same procedure as in (2.9) and using (2.24) we find, upon letting $\epsilon \rightarrow 0$, the one-loop counterterm

$$\delta W \equiv - W_{\text{DIVERGENT}} = - \frac{1}{\epsilon} A_d \tag{2.26}$$

Thus the calculation of the counterterm is trivial once we know the A_d coefficient. For example, massless scalars in $d = 4$ would give

$$\begin{aligned}
 \delta W &= - \frac{1}{\epsilon} A_4 \\
 &= - \frac{1}{\epsilon} \frac{1}{(4\pi)^2} \int d^4x \sqrt{g} \frac{1}{180} \left[R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2 \right]
 \end{aligned} \tag{2.27}$$

on using (2.5).

If, on the other hand, the scalars were massive with mass m , then

$$\Delta \rightarrow \Delta + m^2 \quad (2.28)$$

$$Y(\Delta, t) \rightarrow e^{-m^2 t} \sum_n e^{-\lambda_n t} \quad (2.29)$$

$$A_d \rightarrow \sum_{k=0}^d \frac{(-m^2)^{(d-k)/2}}{[(d-k)/2]!} \quad (2.30)$$

where to obtain (2.30) we have expanded the exponential factor in (2.29) as

$$e^{-m^2 t} = 1 - m^2 t + \frac{m^4}{2} t^2 - \dots \quad (2.31)$$

In the above example, the counterterm would become

$$\delta W = -\frac{1}{\epsilon} \left[\frac{m^4}{2} A_0 - m^2 A_2 + A_4 \right] \quad (2.32)$$

$$= -\frac{1}{\epsilon} \frac{1}{(4\pi)^2} \int d^4 x \sqrt{g} \left[\frac{m^4}{2} - m^2 \frac{R}{6} + \frac{1}{180} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2) \right] \quad (2.33)$$

on using (2.5).

3. DIVERGENCES IN KALUZA-KLEIN THEORIES

Let us now apply the techniques of Section 2 to our model scalar field theory of Section 1, first of all from the point of view of the higher dimension.

Higher dimensional picture (A)

Starting from Eqs. (A.1), (A.2) and (A.3) of Section 1, the counterterm will be given by

$$\delta W = -\frac{1}{\epsilon} A_d(\Delta) \quad (3.1)$$

where Δ is the Laplacian on the d dimensional manifold M . When M is the product $M_1 \times M_2$, Δ splits into $\Delta_1 + \Delta_2$ as in (1.17). The beauty of this is that the heat kernel (2.2) factorizes

$$Y(\Delta, t) = Y(\Delta_1, t) Y(\Delta_2, t), \quad (3.2)$$

hence

$$t^{d/2} \sum_{k=0}^{\infty} A_k(\Delta) t^{k/2} = t^{d_1/2} \sum_{\ell=0}^{\infty} A_{\ell}(\Delta_1) t^{\ell/2} \cdot t^{d_2/2} \sum_{m=0}^{\infty} A_m(\Delta_2) t^{m/2} \quad (3.3)$$

Comparing powers of t ,

$$A_k(\Delta) = \sum_{\ell=0}^k A_{\ell}(\Delta_1) A_{k-\ell}(\Delta_2) \quad (3.4)$$

and so (3.1) yields

$$\delta W = -\frac{1}{\epsilon} \sum_{\ell=0}^d A_{\ell}(\Delta_1) A_{d-\ell}(\Delta_2) \quad (3.5)$$

(Note that the sum extends from 0 to d , the dimension of M). Equations (3.1) and (3.5) have a simple interpretation. The $A_d(\Delta)$ in (3.1) will just provide the d dimensional counterterms, i.e.,

$$\frac{1}{(4\pi)^{d/2}} \int d^d x \sqrt{g} (\text{Riemann})^{d/2} \quad (3.6)$$

Whereas in (3.5), the $A_{\ell}(\Delta_1)$ will be four-dimensional counterterms, i.e.,

$$\frac{1}{(4\pi)^2} \int d^4 x \sqrt{g} (\text{Riemann})^{\ell/2} \quad (3.7)$$

with numerical coefficients determined by the $A_{d-\ell}(\Delta_2)$. But this amounts to nothing more than computing the counterterm in d dimensions and then integrating out the extra y co-ordinates to obtain the counterterm for the effective four-dimensional theory - precisely the same procedure as adopted for the classical action.

Let us consider again the concrete example of $d = 6$:

$$\delta W = -\frac{1}{\epsilon} \left[A_0(\Delta_1) A_6(\Delta_2) + A_2(\Delta_1) A_4(\Delta_2) + A_4(\Delta_1) A_2(\Delta_2) + A_6(\Delta_1) A_0(\Delta_2) \right] \quad (3.8)$$

and suppose $M_2 = S^2$. Then from (2.5) and (2.21)

$$\begin{aligned} \delta W = & -\frac{1}{\epsilon} \cdot \frac{1}{(4\pi)^2} \int d^4 x \sqrt{g} \left[\frac{4m^4}{315} + \frac{m^2}{90} R \right. \\ & + \frac{1}{540} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2) \\ & \left. + \frac{1}{m^2} (\text{Riemann})^3 \right] \quad (3.9) \end{aligned}$$

where

$$(\text{Riemann})^3 = \alpha R_{\mu\nu}{}^{\rho\sigma} R_{\rho\sigma}{}^{\alpha\beta} R_{\alpha\beta}{}^{\mu\nu} + \dots \quad (3.10)$$

as in (2.5).

A few comments are now in order:

- 1) Comparison with (2.33) shows that the $\sqrt{g}$, $\sqrt{g}R$ and $\sqrt{g}R^2$ terms in (3.9) are only to be expected. The novel departure is the appearance of $\sqrt{g}(\text{Riemann})^3$ even though we are still at one loop.
- 2) Had we chosen $M_2 = S^1 \times S^1$ with inverse radii m and m' , then $R_{ijkl} = 0$ and all the $A_k(\Delta_2)$ in (3.8) would vanish save for $A_0(\Delta_2)$. In this case only the $\sqrt{g}(\text{Riemann})^3$ survives with coefficient $(m_1 m_2)^{-1}$.
- 3) Had we chosen d to be odd, then all the $A_k(\Delta_2)$ vanish irrespective of the choice of M_2 and $\delta W = 0$.
- 4) The above results indicate that we do not in general recover the corresponding four-dimensional massless theory as we shrink the extra dimensions. Compare (3.9) as $m \rightarrow \infty$ with (2.27).

Now let us repeat the above analysis from the point of view of the lower dimension.

Lower dimensional picture (B)

Starting from Eqs. (B.1), (B.2) and (B.3) of Section 1 the counterterms δW_n will be given by

$$\delta W_n = -\frac{1}{\epsilon} A_{d_1}(\Delta_n) \quad (3.11)$$

where

$$\Delta_n = -\Delta_1 + m_n^2 \quad (3.12)$$

and m_n^2 are the eigenvalues of Δ_2 . From (2.30), (3.11) and (3.12),

$$\delta W_n = -\frac{1}{\epsilon} \sum_{k=0}^{d_1} \frac{(-m_n^2)^{(d_1-k)/2}}{[(d_1-k)/2]!} A_k(\Delta_1) \quad (3.13)$$

Naively, one might suppose that, since $W = \sum_n W_n$, the total counterterms would be given by the sum of the individual counterterms, i.e.,

$$\delta W = \sum_n \delta W_n \quad (3.14)$$

Using (3.13) together with

$$\sum_n m_n^{-2S} = \zeta(\Delta_2, s) \quad (3.15)$$

and the relation (2.10) between ζ and A_k , one would then obtain

$$\delta W = -\frac{1}{\epsilon} \sum_{\ell=0}^{d_1} A_\ell(\Delta_1) A_{d-\ell}(\Delta_2) \quad (3.16)$$

This is almost the same result as that obtained in the high dimensional picture given by (3.5), but there is one vital difference: the sum in (3.16) extends only as far as d_1 .

In our $d = 6$ example (3.13) is

$$\begin{aligned} \delta W_n = & -\frac{1}{\epsilon} \int d^4x \sqrt{g} \left[\frac{1}{2} m_n^4 - m_n^2 \frac{R}{6} + \right. \\ & \left. + \frac{1}{180} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2) \right] \end{aligned} \quad (3.17)$$

and (3.16) would become

$$\begin{aligned} \delta W = & -\frac{1}{\epsilon} \frac{1}{(4\pi)^2} \int d^4x \sqrt{g} \left[\frac{4m^4}{315} + \frac{m^2}{90} R + \right. \\ & \left. + \frac{1}{540} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - R_{\mu\nu} R^{\mu\nu} + \frac{5}{2} R^2) \right] \end{aligned} \quad (3.18)$$

In other words, because the sum in (3.16) extends only as far as four and not six, the (Riemann)³ term in the high dimension result (3.9) is missing.

Naively, therefore, it seems that the high dimension picture yields a result expected by power-counting in d dimensions, whereas the low dimension picture yields a result expected by power-counting in d_1 dimensions. Yet formally, the two pictures should describe the same theory. What is wrong?

Reconciliation of pictures A and B

The answer lies in our seemingly innocuous assumption (3.14). This naive result would be perfectly legitimate if the summation extended over a finite number of fields. Because in our Kaluza-Klein theory the number of fields is infinite, however, then

$$\delta W \neq \sum_n \delta W_n \quad (3.19)$$

To understand the source of the error in the naive assumption (3.14), recall that in deriving (3.16) we employed (2.30), which in turn involved (2.31). In other words, we expanded

$$e^{-\lambda_n t} = 1 - \lambda_n t + \frac{\lambda_n^2}{2} t^2 - \dots \quad (3.20)$$

So far, so good. It is then false, however, to claim

$$\begin{aligned} \sum_n e^{-\lambda_n t} &= \sum_n 1 - \sum_n \lambda_n t + \sum_n \frac{\lambda_n^2}{2} t^2 - \dots \\ &= \zeta(\Delta_2, 0) - \zeta(\Delta_2, -1)t + \zeta(\Delta_2, -2)\frac{t^2}{2} - \dots \\ &= \sum_{k=0}^{\infty} A_{k+d_2}(\Delta_2) t^{k/2} \end{aligned} \quad (3.21)$$

on using (2.10), because a comparison with the asymptotic expansion of the heat-kernel (2.3) shows that (3.21) has omitted all the negative powers of t .

The physical interpretation of this is that the individual effective actions W_n may have finite terms involving inverse powers of m_n which, when summed over n , yield infinities which therefore contribute to the infinite part of W . In our $d = 6$ example, W_n involves finite terms of the form $m_n^{-2} R^3$ which when summed over n yield $\zeta(\Delta_2, 1) R^3$ which is divergent (see Section 2). To handle this divergence we note that in $d_2 + \epsilon$ dimensions $\zeta(\Delta_2, s)$ has a pole at $\epsilon = 0$ for $s = 1, 2, \dots, d_2/2$ with residue $A_{d_2-2s}(\Delta_2)$. In particular for $d_2 = 2$, $\zeta(\Delta_2, 1) R^3$ is then replaced by $\epsilon^{-1} A_0 R^3$ which was just the missing term in (3.18).

With sufficient care, therefore, picture B gives the same result as picture A. We are now in a position to answer the questions posed at the end of Section 1.

Questions answered

A1) In both pictures, the power counting rules are those appropriate to the higher dimension, (as may also be understood by treating Σ_n like a momentum integration). Thus even the effective four-dimensional theory may have one-loop counterterms like R^3 , etc.

A2) Similarly, with regard to the question of dimensional regularization, by the replacement

$$\int d^{d+\epsilon} x \rightarrow \sum_n \int d^{d_1+\epsilon} x \quad (3.21)$$

we see that the ϵ in the A counterterm (3.1) is the same ϵ as in the B counterterm (3.11). Dimensional regularization in $d = d_1 + d_2$ dimensions is seen to be equivalent to dimensional regularization in d_1 dimensions together with zeta-function regularization in d_2 dimensions.

A3) One does not recover the massless quantum theory by first performing the regularization and then shrinking the size of the extra dimensions. As the examples (2.27) and (3.9) illustrate, the divergences of the two theories are very different. Even the R^2 term, which is independent of m , has a different coefficient. Similar remarks apply to the finite contributions to the effective action, in particular the vacuum energies mentioned in Section 1. A single massless field in four dimensions induces no vacuum energy, whereas including all the massive modes and then sending $m \rightarrow \infty$ would yield an infinite vacuum energy (this is after the removal of ultra-violet divergences).

A4) Nor is it true that $\delta W = \sum_n \delta W_n$. As we have seen, it sometimes happens that

$$\sum_n (\text{finite})_n = \text{infinite}.$$

A5) The counterterms are influenced by the choice of extra dimensions e.g., $M_2 = S^2$ or $M_2 = S^1 \times S^1$. However, this is really an artefact of the toy models we have employed, in that the gravitational field was taken to be external with no dynamics of its own. In a realistic theory with all fields, including the gravitational fluctuations about the ground state, dynamical and depending on both x^μ and y^i this would not be the case. Different ground states would give different-looking counterterms in picture B, but all should be equivalent, in much the same way that the spontaneous breaking of a conventional gauge theory does not change its renormalizability properties. One would not expect finiteness in one phase and non-finiteness in another.

A6) Finally, although as mentioned in (A4) above, it sometimes happens that

$$\sum_n (\text{finite})_n = \text{infinite}$$

so it sometimes also happens that

$$\sum_n (\text{infinite})_n = \text{finite}$$

This explains how in picture B, the theory of an infinite number of massive states can be finite when the same theory of a single state is divergent. For example, taking the extra dimension to be a circle can yield poles in ϵ with residue

$$\sum_{n=-\infty}^{\infty} 1 \text{ and } \sum_{n=-\infty}^{\infty} n^2$$

But using ζ functions

$$\sum_{n=-\infty}^{\infty} 1 = 1 + 2\zeta_R(0) = 0 \quad (3.22)$$

and

$$\sum_{n=-\infty}^{\infty} n^2 = \zeta_R(-2) = 0 \quad (3.23)$$

These results have an interesting application to axial anomalies when fermions are included in the Kaluza-Klein scheme.

4. FERMIONS AND THE AXIAL ANOMALY

In this section we describe a simple model with fermions, again unrealistic but nevertheless useful for illustration.

The fermion model

Consider the action for a single massless Dirac spinor in interaction with an external gravitational field in five dimensions

$$S = \int d^5x \det \hat{e} \bar{\Psi} i \gamma^{\hat{\mu}} \hat{D}_{\hat{\mu}} \Psi \quad (4.1)$$

where $\hat{e}_{\hat{\mu}}^{\hat{a}}(x)$ is the fünfbein field ($\mu, a = 1, 2, 3, 4, 5$), $\gamma^{\hat{a}}$ are the five-dimensional Dirac matrices

$$\{\gamma^{\hat{a}}, \gamma^{\hat{b}}\} = 2\eta^{\hat{a}\hat{b}} \quad (4.2)$$

and $\hat{D}_{\hat{\mu}}$ is the five-dimensional covariant derivative formed from the spin connection.

As in Section 1 we again consider the fifth dimension to be compactified to a circle with radius m^{-1} , hence

$$\Psi(x, y) = \sqrt{\frac{m}{2\pi}} \sum_{n=-\infty}^{\infty} \psi_n(x) \exp inmy \quad (4.3)$$

This time, however, we generalize the ansatz for the five-dimensional gravity field to allow a non-vanishing gauge field in the effective four-dimensional theory:

$$\hat{e}_{\hat{\mu}}^{\hat{a}}(x, y) = \begin{bmatrix} e_{\mu}^a(x) & \kappa A_{\mu}(x) \\ 0 & 1 \end{bmatrix} \quad (4.4)$$

where κ^2 is chosen to be Newton's constant times $16\pi^2$ in order that the five-dimensional Einstein action yield the correct Einstein-Maxwell four-dimensional action. Substituting (4.3) and (4.4) into (4.1) and performing the y integration yields

$$S = \sum_{n=-\infty}^{\infty} \int d^4x \det e \bar{\psi}_n \left[i \gamma^{\mu} (D_{\mu} + iq_n A_{\mu}) + \right. \\ \left. + m_n \gamma^5 + \frac{i\kappa_F}{8} \gamma_{\mu\nu}^{\gamma^{\mu}\gamma^{\nu}\gamma^5} \right] \psi_n \quad (4.5)$$

where all quantities are now four-dimensional. Thus the effective theory describes one neutral massless fermion (the $n=0$ mode) together with an infinite tower of charged massive fermions with quantized charges

$$q_n = n \kappa m \quad (4.6)$$

and quantized masses

$$m_n = nm \quad (4.7)$$

All states also couple non-minimally to electromagnetism via the Pauli term ($F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$).

Spin-one equivalence principle

The remarkable feature of the charged massive states is that from (4.6) and (4.7), the charge to mass ratio

$$\frac{q_n}{m_n} = \kappa \quad (4.8)$$

is a universal constant for all particles*. Thus even when considering purely electromagnetic effects, one can not invoke any decoupling theorem to argue that a particle with mass a billion times the Planck mass, say, is any less important than a particle whose mass is "only" 10^5 grams, because this particle will couple a billion times more strongly. For example, the scattering of light by light is describable by a low-frequency Euler-Heisenberg effective Lagrangian of the form $q_n^4 F^4 / m_n^4$ and each massive mode would contribute equally to this process. Of course the net effect is still small and so presumably negligible for phenomenology at current energies.

There is a simple physical interpretation of this universality effect. In the five-dimensional theory, the equivalence principle is at work ensuring that gravity couples to everything with equal strength, and in the effective four-dimensional theory the photon is "remembering" its five-dimensional ancestry. Similar effects would also occur for higher dimensions and non-Abelian gauge theories.

This "spin-one equivalence principle" will always be a feature of Kaluza-Klein theories.

* A similar phenomenon, though in the context of twisted fields and dimensional reduction rather than Kaluza-Klein, is responsible for Scherk's "antigravity". See Refs. 32 and 33.

The axial anomaly

Let us now consider the anomaly in the divergence of the axial vector current, following the approach of Refs. 34 and 35. There it was noted that in the functional integral approach to quantization, the origin of the anomaly may be understood by noting that the fermion measure is not invariant under chiral transformations. Let $d\mu[\bar{\Psi}_n, \psi_n]$ represent the measure for the n th mode in the theory described by the action (4.5). Then under the transformation

$$\psi_n \rightarrow (\exp \gamma_5 \alpha) \psi_n, \quad (4.9)$$

The measure transforms

$$d\mu[\bar{\Psi}_n, \psi_n] \rightarrow d\mu[\bar{\Psi}_n, \psi_n] e^{2i\alpha v_n} \quad (4.10)$$

where

$$v_n = \frac{q_n^2}{16\pi^2} \int d^4x \sqrt{g} \star F_{\mu\nu} F^{\mu\nu} - \frac{1}{384\pi^2} \int d^4x \sqrt{g} \star R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} \quad (4.11)$$

is a topological invariant, and the stars denote the dual tensors. The complete functional measure in this theory is

$$\prod_{n=-\infty}^{\infty} d\mu[\bar{\Psi}_n, \psi_n].$$

From (4.10) this transforms as

$$\prod_{n=-\infty}^{\infty} d\mu[\bar{\Psi}_n, \psi_n] \rightarrow \prod_{n=-\infty}^{\infty} d\mu[\bar{\Psi}_n, \psi_n] \cdot e^{2i\alpha \sum_{n=-\infty}^{\infty} v_n} \quad (4.12)$$

To interpret

$$\sum_{n=-\infty}^{\infty} v_n$$

we use the charge quantization (4.6) and ζ function regularization. Then from (4.11),

$$\begin{aligned} \sum_{n=-\infty}^{\infty} v_n &= 2\zeta_R(-2) \cdot \frac{\kappa^2 m^2}{16\pi^2} \int d^4x \sqrt{g} \star F_{\mu\nu} F^{\mu\nu} \\ &\quad - [1 + 2\zeta_R(0)] \frac{1}{384\pi^2} \int d^4x \sqrt{g} \star R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} = 0 \end{aligned} \quad (4.13)$$

on using (3.22) and (3.23). Thus the functional measure for this infinite tower of fermions is invariant under chiral transformations, even though that for a single fermion is not. Therefore the effective four-dimensional theory is anomaly-free which is

entirely consistent with the five-dimensional viewpoint: there are no chiral anomalies in odd dimensions*.

The relation between the anomaly for a single mode (the lower, d_1 dimensional theory) and that for all the modes (the higher, d dimensional theory) for an arbitrary theory in arbitrary dimensions may be simply read off from the formulae for the tensor product of elliptic complexes. See Ref. 36. First we recall that in even dimensions the Atiyah-Singer theorem yields

$$N^+ - N^- = \nu \quad (4.14)$$

where N^+ and N^- are respectively the number of right and left handed zero modes of the Dirac operator and ν is topological invariant, of which (4.11) is a special case. Although there will in general be zero modes in odd dimensions also, the split into right and left is not defined. If the number of zero modes on $M = M_1 \times M_2$ is denoted by N and those on M_1 and M_2 by N_1 and N_2 respectively, then the relevant formulae are**

a) d_1 even, d_2 even

$$\begin{aligned} N^+ &= N_1^+ N_2^+ + N_1^- N_2^- \\ N^- &= N_1^+ N_2^- + N_1^- N_2^+ \end{aligned} \quad (4.15)$$

b) d_1 even, d_2 odd

$$N = (N_1^+ + N_1^-) N_2 \quad (4.16)$$

c) d_1 odd, d_2 odd

$$N^+ = N^- = N_1 N_2 \quad (4.17)$$

In case (a) we have from (4.14) and (4.15)

$$\nu = \nu_1 \nu_2 \quad (4.18)$$

* Here, as in Section 3, one can even turn the argument around: consistency of the higher and lower dimensional points of view demands the ζ function interpretation of divergent quantities.

** The formulae are in fact more complicated when M is not globally the direct product $M_1 \times M_2$, i.e., when the spin-one gauge fields in picture B have non-trivial topology. Incidentally, in conventional gauge theories some authors are happy to sum over different gauge field topologies, but balk at the idea of summing over different space-time topologies. It is worth mentioning that, in Kaluza-Klein, the latter is necessary for the former.

In case (b), v is not defined, and as illustrated in our five-dimensional example above there is no anomaly even when v_1 is non-vanishing. In case (c), v always vanishes.

The invariance of the measure in the effective four-dimensional theory when d is odd means that the θ parameter is zero. This has implications for the CP problem.

Equations (3.22) and (3.23), used to prove the vanishing of the axial anomaly, may also be used to prove the vanishing of the one-loop trace anomaly and β function, a result in any case expected since the theory is odd-loop finite*.

5. RECENT DEVELOPMENTS

Here we bring the reader up to date with some recent developments in Kaluza-Klein theories:

Mode expansions on G/H

The expansion of the fields in a Kaluza-Klein theory in terms of normal modes on the compact space M_2 is quite a non-trivial matter in the case that M_2 is a quotient space G/H , especially for fields with spin. This problem has recently been analyzed in some detail by Salam and Strathdee³⁸, who pay particular attention to the problem of embedding H in the tangent space group $SO(d_2)$. Of particular interest is their result that the massive particles belong to infinite dimensional representations of non-compact groups. For example, in the original $4 + 1$ pure gravity theory the massive states are charged purely spin-two particles belonging to infinite dimensional representations of $O(1,2)$. It would be interesting to see what kind of non-compact groups would emerge from eleven-dimensional supergravity

Ultra-violet divergences

Further discussion of the kinds of problems arising with divergences and anomalies in higher dimensional theories, including the effects of taking the fields to be "twisted" in the higher dimensions, may be found in a recent paper by Duff and Toms²⁸.

Another development, due to Green, Schwarz and Brink³⁹ concerns the emergence of $N = 1$ supersymmetric Yang-Mills in $d = 10$ and $N = 2$ supergravity in $d = 10$ as the zero-slope limits of the fermionic string models. By considering the extra dimensions to

*The β function for a theory with an infinite number of fermions has also been studied in a different context by Nicolai³⁷.

be compactified as circles and then letting the radii of the circles tend to zero, they are able to analyze the ultra-violet divergences of the theories for different values of d . Interestingly, they find that at one-loop both theories are ultra-violet finite for $d < 8$ and infra-red finite for $d > 4$ *, the infra-red divergence being milder in the gravitational case.

Note however that this analysis does not include $N = 1$ in $d = 11$ which differs from $N = 2$ in $d = 10$ by an infinite tower of massive states. Indeed the eleven-dimensional theory is finite at all odd loops, in particular at one-loop. [Actually, it is not quite so obvious as it is for pure gravity that there are no invariants with an odd number of derivatives. For example, one could imagine a counterterm containing the $F_{\mu\nu\rho\sigma}$ of Eq. (1.1) raised to an odd power. Fortunately there is a discrete symmetry, (time reversal together with $A_{\mu\nu\rho} \rightarrow -A_{\mu\nu\rho}$) which rules these out.] Possible on-shell counterterms do exist at even loop order, however. They are of the symbolic form

$$(\kappa^2)^{L-1} \int d^{11}x d^{32}\theta [D^2]^\beta [W]^\gamma \quad (5.1)$$

where $W_{[\mu\nu\rho\sigma]}(x, \theta)$ is the super field strength⁴⁰, D stands for its supercovariant derivative, and the θ 's are the fermionic co-ordinates of superspace. The requirement that (5.1) be dimensionless and compatible with the above discrete symmetry yields

$$\beta + \gamma = 9L - 14 = \text{positive even integer} \quad (5.2)$$

Thus $\beta + \gamma = 4 \bmod 18$ and candidate counterterms exist at $L = 2, 4, 6$ etc. [I am very grateful to P. Howe, R. Kallosh and M. Sohnius for discussions on these points. An analysis of supergravity counterterms in both $d = 4$ and $d > 4$ supergravity has also been given recently by Grisaru and Siegel⁴¹. These authors make plausible the assumption that only integrals over all superspace, as in (5.1), ever appear as counterterms.] One should still not be too pessimistic about the finiteness of $d = 11$ supergravity, however, since time and again in supersymmetry and supergravity one finds that a priori possible counterterms nevertheless appear with zero coefficient¹.

More about Kaluza-Klein supergravity

In Section 1 we described some of the things supergravity can do for Kaluza-Klein. We finish this paper by considering what Kaluza-Klein can do for supergravity. After all, whether or not a realistic $SU(3) \times SU(2) \times U(1)$ theory emerges in the way described,

* Should one invoke the avoidance of infra-red divergences as another reason for Kaluza-Klein theories ?

the Kaluza-Klein idea still provides a deeper understanding of the structure of supergravity in four dimensions.

The Cremmer-Julia⁴ version of $N = 8$ supergravity in $d = 4$ was obtained as follows. First, one adopts the ground state solution of Minkowski space $\times$ 7-torus. The gauge group is then simply 7 $U(1)$'s. Secondly, one discards the massive modes. The remaining massless states belong to representations of a rigid $SO(7)$. In particular, some of the spin-0 fields are realized as antisymmetric tensors. Finally, only after making (topologically non-trivial) duality transformations, which transform antisymmetric tensors to scalars, does one obtain the version with the E_7 (rigid) $\times$ $SU(8)$ (local) invariance. The geometrical origin of these hidden symmetries is thus far from clear.

Although in this pioneering work Cremmer and Julia were aware of the possibility of choosing the extra seven dimensions to be something other than a 7-torus, they confined their explicit calculations to this case. So leaving aside the question of a realistic gauge theory for the moment, it is still of interest for a supergravity enthusiast to ask what kind of $d = 4$ theories would emerge with different ground states. After all, there are potentially infinitely many such theories and it would be nice to have a concrete example of just one.

The task of carrying out the complete Kaluza-Klein reduction is very difficult but the most important piece of information is the number of massless states, their spins, and internal quantum numbers. Rules for counting these massless states will be given in a forthcoming publication by Duff and Pope⁴². Though straightforward in principle these rules can be complicated to apply in all but the simplest cases, and next to the 7-torus the simplest ground state solution, as discussed in Section 1, is the 7-sphere with its standard metric*, $S^7 = SO(8)/SO(7)$. This yields an effective $d = 4$ theory with local $SO(8)$ invariance and negative cosmological constant. Applying the counting rules for the massless modes, which fall naturally into $SO(8)$ representations, we find 1 graviton, 8 gravitini, 28 vectors (from the Killing vectors on S^7), and 56 spin 1/2 fermions. A direct count of the massless spin-0 fields is rather more difficult, but indirectly one can argue that it must be 70 owing to the 8-fold supersymmetry, which in turn is ensured by the presence of 8 massless spin-3/2 particles. [The difference in Betti numbers of the 7-sphere from the 7-torus turns out to guarantee that these 70 states will be genuine scalars all coming from the elfbein $e_\mu^a(x)$, rather than antisymmetric tensors coming from the 3-index field $A_{\mu\nu\rho}(x)$.]

* Seven-spheres hold great fascination for the mathematicians. See, for example, "On manifolds homeomorphic to the 7-sphere" by Milnor in Ref. 43. Will all this differential geometry eventually be important for supergravity?

But an $N = 8$ supergravity theory in $d = 4$ with all the above attributes already exists ! It is the recently constructed gauged theory of De Wit and Nicolai⁴⁴ which has, in addition to the obvious local $SO(8)$ invariance, a hidden local $SU(8)$. One recovers the E_7 (rigid) $\times$ $SU(8)$ (local) version as one switches off the gauge coupling. Thus Kaluza-Klein puts into perspective the different $N = 8$ supergravities known to exist in $d = 4$ (and which have a different ultra-violet behaviour¹.) Taking a 7-sphere and shrinking its radius to zero leads directly to the $E_7 \times SU(8)$ theory, which was obtained only indirectly by starting from the 7-torus. Indeed, one might now hope that S^7 will yield a more transparent geometrical origin of these hidden symmetries*.

The 7-torus and the 7-sphere are, however, very special cases. A generic ground state would have fewer than 8 gravitini, (i.e., fewer than 8 supersymmetries) and perhaps none at all. Indeed, this may be the most welcome feature of Kaluza-Klein ideas applied to eleven-dimensional supergravity: it would give a supersymmetric explanation for the absence of any observed supersymmetry in particle physics at current energies.

ACKNOWLEDGEMENTS

We are very grateful for conversations with A. Chockalingham, R. Coquereaux, S. Ferrara, G. Gibbons, P. Howe, C. Isham, R. Kallosh, W. Mecklenburg, H. Nicolai, D. Olive, D. Pollard, C. Pope, A. Salam, M. Sohnius, K. Stelle, P. Stephenson, J. Strathdee, P. Townsend, A. Trautman, N. Warner, P. West, C. Wetterich and E. Witten.

REFERENCES

1. For a recent review, see M.J. Duff, Ultraviolet Divergences in Extended supergravity, CERN preprint TH.3232 (1982), to appear in "Supergravity 81", Eds. S. Ferrara and J.G. Taylor, Cambridge University Press.
2. Th. Kaluza, Sitzungsber. Preus. Akad. Wiss. Berlin, Math. Phys. K1:966 (1921);
O. Klein, Z. Phys. 37:895 (1926).
3. E. Witten, Nucl. Phys. B186:412 (1981); this paper gives a comprehensive list of references to earlier work. See also Ref. 28.
4. E. Cremmer and B. Julia, Nucl. Phys. B159:141 (1979).
5. A. Chockalingham, Ph.D. Thesis, Imperial College (1981).
6. J. Scherk and J.H. Schwarz, Phys. Lett. B57:463 (1975).
7. For a review see E. Cremmer, preprint LPTENS 81/18 (1981).
8. J. Scherk and J.H. Schwarz, Phys. Lett. B82:60 (1979).
9. E. Witten, Instability of the Kaluza-Klein vacuum, Princeton preprint (1981).

* I am very grateful to Sergio Ferrara and Hermann Nicolai for discussions on this point.

10. S.M. Christensen and M.J. Duff, Nucl. Phys. B154:301 (1979).
11. F.A. Berends, J.W. van Holten, B. de Wit and P. van Nieuwenhuizen
Nucl. Phys. B154:261 (1979).
12. C. Aragone and S. Deser, Phys. Lett. 85B:161 (1979).
13. W. Nahm, Nucl. Phys. B135:149 (1978).
14. A. Lichnerowicz, C.R. Acad. Sci. Paris Sér. A-B 257:7 (1963).
15. S.M. Christensen and M.J. Duff, Phys. Lett. 76B:571 (1978)
16. S.W. Hawking and C.N. Pope, Nucl. Phys. B146:381 (1978).
17. N.K. Nielsen, M. Grisaru, H. Römer and P. van Nieuwenhuizen,
Nucl. Phys. B140:477 (1978).
18. N. Hitchin, Adv. Math. 14:1 (1974).
19. C. Wetterich, Spontaneous compactification in higher dimensional
gravity, CERN preprint TH.3239 (1982).
20. P.G.O. Freund and M.A. Rubin, Phys. Lett. B97:233 (1980).
21. M.J. Duff and P. van Nieuwenhuizen, Phys. Lett. B94:179 (1980).
22. A. Aurelia, H. Nicolai and P.K. Townsend, Nucl. Phys. B176:509
(1980).
23. S. Unwin, Phys. Lett. B103:18 (1981).
24. B. Zumino, Nucl. Phys. B89:535 (1975).
25. P.C. West, Nucl. Phys. B106 (1976) 219.
26. S.W. Hawking and C.N. Pope, Phys. Lett. 73B:42 (1978).
27. S.W. Hawking, Commun. Math. Phys. 55:133 (1977).
28. M.J. Duff and D.J. Toms, Divergences and anomalies in
Kaluza-Klein theories, CERN preprint TH.3248 (1982). Talk
given by M.J. Duff at the Second Quantum Gravity Seminar,
Moscow, October 1981.
29. P.B. Gilkey, J. Diff. Geom. 10:601 (1975).
30. E.T. Whittaker and G.N. Watson, A course of modern analysis,
Cambridge University Press, (1927).
31. S.M. Christensen and M.J. Duff, Nucl. Phys. B170 [FS1]:480
(1980).
32. J. Scherk, From supergravity to antigravity, in Supergravity,
Eds. D.Z. Freedman and P. van Nieuwenhuizen, North Holland
(1980).
33. D. Pollard, Antigravity and classical solutions of five-
dimensional Kaluza-Klein theory, Imperial College preprint
ICTP/81/82-16 (1982).
34. K. Fujikawa, Phys. Rev. Lett. 42:1195 (1979)
35. S. Deser, M.J. Duff and C.J. Isham, Phys. Lett. 93B:419 (1980).
36. M.F. Atiyah and R. Bott, Ann. of Math. 86:374 (1967).
37. H. Nicolai, Phys. Lett. B84:219 (1979).
38. A. Salam and J. Strathdee, On Kaluza-Klein theories, ICTP
Trieste preprint IC/81/211 (1981).
39. M.B. Green, J.H. Schwarz and L. Brink, N = 4 Yang-Mills and
N = 8 supergravity as limits of string theories, Caltech
preprint CALT-68-880 (1981).

- 40. E. Cremmer and S. Ferrara, Phys. Lett. 91B:61 (1980);
L. Brink and P. Howe, Phys. Lett. 91B:384 (1980);
S. Ferrara, Eleven-dimensional supergravity in superspace,
in Unification of the Fundamental Particle Interactions,
Eds. S. Ferrara, J. Ellis and P. van Nieuwenhuizen, Plenum
Press (1980).
- 41. M.T. Grisaru and W. Siegel, Supergravity II, Caltech preprint
CALT-68-882 (1982).
- 42. M.J. Duff and C.N. Pope, to be published.
- 43. J. Milnor, Ann. of Math. 65:399 (1956).
- 44. B. de Wit and H. Nicolai, Extended supergravity with local
 $SO(8) \times SU(8)$ invariance, CERN preprint TH.3183 (1981).