

A note on the Kaluza–Klein theory

Li-Xin Li 

Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871,
People's Republic of China

E-mail: lxl@pku.edu.cn and prof.li.xin.li@gmail.com

Received 30 January 2023; revised 14 August 2023

Accepted for publication 22 August 2023

Published 5 September 2023



CrossMark

Abstract

We show that the Kaluza–Klein (KK) theory contains a fundamental problem: The four-dimensional metric tensor and the electromagnetic potential vector assumed in the KK theory belong to four-dimensional vector spaces that are not integrable in general, resulting that the four-dimensional physical variables and the corresponding field equations derived from the five-dimensional Einstein field equation (i.e. the four-dimensional Einstein field equation and the Maxwell equations) are not defined on a four-dimensional submanifold. That is, the four-dimensional spacetime assumed in the KK theory does not exist. No satisfactory solutions are found within the KK formalism. Perhaps the best approach to fix the problem is giving up the KK theory and looking for a new unified scheme for gravitational and electromagnetic interactions in the framework of a spacetime with extra dimensions, as having already been explored in some literature.

Keywords: Kaluza–Klein theory, extra dimension, string theory

(Some figures may appear in colour only in the online journal)

1. Introduction

The Kaluza–Klein (KK) theory represents the first attempt to unify the gravitational and electromagnetic interactions in the framework of general relativity extended to a spacetime with extra dimensions [1–3]¹. In the KK theory, the bulk spacetime is assumed to be five-dimensional and described by the five-dimensional Einstein field equation. By 4+1



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

¹ An even earlier attempt to unify electromagnetic and gravitational fields in a five-dimensional spacetime before the appearance of general relativity was given by Nordström in 1914 [4].

decomposition of the five-dimensional spacetime metric, a four-dimensional Einstein field equation and the Maxwell equations are derived from the five-dimensional Einstein field equation, which are assumed to describe the four-dimensional world where we live. How to make the extra dimension compact, small, and static has been a challenging problem in modern theoretical physics [5–9]. Nowadays, introducing (compact or noncompact but warped) extra dimensions in addition to the four dimensions of the spacetime where we live has been a popular strategy for unifying all fundamental interactions in nature, e.g. the theories of supergravity [10, 11], superstring [12–15], and brane gravity [16–18]. To test the existence of extra dimensions, the KK particles arising from the excitation of fields along the compact extra dimensions have been extensively searched by the Large Hadron Collider [19–21].

Despite the success of the KK theory in derivation of the Maxwell equations from the higher-dimensional Einstein field equation and its heavy influence on modern theories of unification, in this paper we show that the KK theory has a serious problem in its foundation: The four-dimensional metric tensor and the electromagnetic potential vector assumed in the KK theory are defined in vector spaces that are not integrable hence not tangent to any four-dimensional submanifold, unless the electromagnetic field antisymmetric tensor vanishes. That is, the four-dimensional spacetime assumed in the KK theory to support the effective four-dimensional theory of electromagnetism and gravity is defined only when the electromagnetic field vanishes, which conflicts the original aim of the KK theory in unifying the gravitational and electromagnetic interactions.

The influence of the problem just mentioned may not be limited to the KK theory. As is well known, one of the cornerstones of string theory—extra dimensions and compactification of extra dimensions—originated from the KK theory with an extension from one extra dimension to multiple extra dimensions. In string theory, a popular approach to derive gauge fields from higher-dimensional gravity is through the KK mechanism with an extension to spacetime of dimensions greater than five. For example, this is the case in the eleven-dimensional supergravity when it is connected to the low-energy limit of M-theory [13–15].

The paper is organized as follows. In section 2, we outline the KK theory and derive the representation of the four-dimensional metric tensor and the electromagnetic potential vector in the five-dimensional spacetime. The representation is uniquely determined by the self-consistency requirement of the theory. In section 3, we discuss the geometric interpretation of the above two KK variables and quantities derived from them (e.g. the electromagnetic field antisymmetric tensor). We show that the four-dimensional quantities are in vector spaces orthogonal to the direction of the extra dimension. In section 4, we prove that the vector spaces containing the four-dimensional variables are not integrable unless the electromagnetic field tensor vanishes. Thus, in general there does not exist a four-dimensional submanifold supporting the four-dimensional theory derived from the five-dimensional Einstein field equation.

Section 5 is devoted to discussion on the action principle and compactification of the extra dimension under the assumption of the cylinder condition. We show that after compactification, although the four-dimensional Einstein field equation and the Maxwell equations can be derived from the action principle, there still does not exist a four-dimensional submanifold supporting the four-dimensional field equations. Finally, in section 6, we summarize the results that we have obtained in this paper and discuss their implications.

Throughout the paper geometrized units with $G = c = 1$ are adopted unless otherwise stated, where G is the four-dimensional gravitational constant and c is the speed of light. In addition, we will take $(-, +, +, +, +)$ as the convention for the signature of the five-dimensional spacetime metric. The abstract index notation for vectors and tensors advocated

in [22] will be used. That is, vectors and tensors are denoted by letters followed by lower case Latin indices, e.g. v^a , g_{ab} , etc.

2. The KK formalism

The success of the KK theory relies on a specific decomposition scheme of the metric tensor of a five-dimensional bulk spacetime. Without loss of generality, in a five-dimensional spacetime $(\mathcal{M}, \tilde{g}_{ab})$ we take a coordinate system $\{x^0, x^1, x^2, x^3, x^4 = w\}$ and write the matrix representation of the five-dimensional metric tensor \tilde{g}_{ab} as²

$$\tilde{g}_{AB} = \begin{pmatrix} g_{\mu\nu} + \phi^2 A_\mu A_\nu & \phi^2 A_\mu \\ \phi^2 A_\nu & \phi^2 \end{pmatrix}, \quad (1)$$

where indices $A, B = 0, 1, 2, 3, 4$, and $\mu, \nu = 0, 1, 2, 3$. Capital Latin letters label coordinate components of five-dimensional vectors and tensors. Lower case Greek letters label coordinate components of four-dimensional vectors and tensors.

The 4×4 matrix $g_{\mu\nu}$ is interpreted as the component of the metric on a four-dimensional spacetime (\mathcal{M}, g_{ab}) associated with the coordinate system $\{x^0, x^1, x^2, x^3\}$, the 4×1 matrix A_μ as the component of an electromagnetic potential dual vector, and the function ϕ as a scalar field in (\mathcal{M}, g_{ab}) . With the convention in equation (1), the five-dimensional spacetime metric tensor \tilde{g}_{ab} is represented in the coordinate system $\{x^\mu, w\}$ as

$$\tilde{g}_{ab} = \tilde{g}_{AB} dx_a^A dx_b^B = (g_{\mu\nu} + \phi^2 A_\mu A_\nu) dx_a^\mu dx_b^\nu + 2\phi^2 A_\mu dx_{(a}^\mu dw_{b)} + \phi^2 dw_a dw_b, \quad (2)$$

where the parentheses in the indices of a tensor denote symmetrization of the tensor about the indices inside the parentheses. The Einstein summation convention for tensor components is used, i.e. an index appearing in both subscripts and superscripts is summed over all dimensions represented by the index.

The inverse of the 5×5 matrix in equation (1), which is also the component matrix of the inverse five-dimensional metric tensor \tilde{g}^{ab} , is

$$\tilde{g}^{AB} = \begin{pmatrix} g^{\mu\nu} & -A^\mu \\ -A^\nu & \phi^{-2} + A_\rho A^\rho \end{pmatrix}, \quad (3)$$

where the 4×4 matrix $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$, i.e.

$$g_{\mu\nu} g^{\nu\rho} = \delta_\mu^\rho \quad (4)$$

where $\delta_\mu^\nu = 1$ if $\mu = \nu$ and 0 otherwise; and

$$A^\mu \equiv g^{\mu\nu} A_\nu. \quad (5)$$

Equations (4) and (5) automatically imply

$$A_\mu = g_{\mu\nu} A^\nu. \quad (6)$$

By equation (3), the inverse five-dimensional metric tensor is represented as

$$\tilde{g}^{ab} = g^{\mu\nu} \left(\frac{\partial}{\partial x^\mu} \right)^a \left(\frac{\partial}{\partial x^\nu} \right)^b - 2A^\mu \left(\frac{\partial}{\partial x^\mu} \right)^a \left(\frac{\partial}{\partial w} \right)^b + \left(\frac{1}{\phi^2} + A_\rho A^\rho \right) \left(\frac{\partial}{\partial w} \right)^a \left(\frac{\partial}{\partial w} \right)^b. \quad (7)$$

² The form of metric decomposition in equation (1) agrees with the general case for the KK theory generalized to a $(4+n)$ -dimensional spacetime to include non-Abelian gauge fields, where the ϕ^2 is replaced by a matrix g_{ij} with the indices i and j running from 1 to n for the extra dimensions [23–25].

It can be verified that the following reciprocal relation is satisfied

$$\tilde{g}_{ab}\tilde{g}^{bc} = \tilde{\delta}_a^c \equiv \delta_\mu^\nu dx_a^\mu \left(\frac{\partial}{\partial x^\nu} \right)^c + dw_a \left(\frac{\partial}{\partial w} \right)^c, \quad (8)$$

where $\tilde{\delta}_a^c$ is the identity map in the five-dimensional spacetime.

In fact, the 5×5 matrices \tilde{g}_{AB} and \tilde{g}^{AB} in equations (1) and (3) are inverse to each other if and only if (a) the 4×4 matrices $g^{\mu\nu}$ and $g_{\mu\nu}$ are inverse to each other (equation (4)), and (b) the A^μ and A_μ are related by equations (5) and (6). It should be noted that equations (4)–(6) are not independent, since any two of them can give rise to the other.

The KK 4-metric tensor g_{ab} and the electromagnetic potential dual 4-vector A_a are, respectively, a tensor and a vector in the five-dimensional spacetime $(\tilde{\mathcal{M}}, \tilde{g}_{ab})$. The question is how they are expressed in coordinate components in the five-dimensional coordinate system $\{x^\mu, w\}$. Since $g_{\mu\nu}$ are interpreted as the components of the four-dimensional metric in coordinates $\{x^\mu\}$, the μ - ν components of g_{ab} must be $g_{\mu\nu}$. Then, the general form of g_{ab} must be

$$g_{ab} = g_{\mu\nu} dx_a^\mu dx_b^\nu + 2g_{\mu 4} dx_a^\mu dw_b + g_{44} dw_a dw_b, \quad (9)$$

where $g_{\mu 4}$ and g_{44} are to be determined. Similarly, since A_μ is interpreted as the coordinate component of A_a in $\{x^\mu\}$, and $A^\mu = g^{\mu\nu} A_\nu$ as the coordinate component of A^a , we must have

$$A_a = A_\mu dx_a^\mu + A_4 dw_a \quad (10)$$

and

$$A^a = A^\mu \left(\frac{\partial}{\partial x^\mu} \right)^a + A^4 \left(\frac{\partial}{\partial w} \right)^a, \quad (11)$$

where A_4 and A^4 are to be determined.

By equations (7) and (10), we have

$$A^a = \tilde{g}^{ab} A_b = (g^{\mu\nu} A_\nu - A^\mu A_4) \left(\frac{\partial}{\partial x^\mu} \right)^a - \left[A_\rho A^\rho - \left(\frac{1}{\phi^2} + A_\rho A^\rho \right) A_4 \right] \left(\frac{\partial}{\partial w} \right)^a. \quad (12)$$

By equation (5), comparison of equation (12) to equation (11) leads to

$$A_4 = 0, \quad A^4 = -A_\rho A^\rho. \quad (13)$$

Thus, we must have

$$A_a = A_\mu dx_a^\mu, \quad (14)$$

and

$$A^a = A^\mu \left(\frac{\partial}{\partial x^\mu} \right)^a - A_\rho A^\rho \left(\frac{\partial}{\partial w} \right)^a. \quad (15)$$

By equations (9) and (11), we have

$$g_{ab} A^b = (A_\mu + g_{\mu 4} A^4) dx_a^\mu + (g_{\mu 4} A^\mu + g_{44} A^4) dw_a \quad (16)$$

after submission of equation (6). Since g_{ab} and A^a are interpreted as, respectively, the metric tensor and the electromagnetic potential vector in a four-dimensional spacetime, we must have $g_{ab} A^b = \tilde{g}_{ab} A^b = A_a$. Then, comparison of equations (14) and (16) leads to

$$g_{\mu 4} = g_{44} = 0. \quad (17)$$

Thus, in coordinates $\{x^\mu, w\}$ the 4-metric tensor g_{ab} is represented as

$$g_{ab} = g_{\mu\nu} dx_a^\mu dx_b^\nu. \quad (18)$$

Then, by $g^{ab} = \tilde{g}^{ac} \tilde{g}^{bd} g_{cd}$ we get the inverse 4-metric tensor

$$g^{ab} = g^{\mu\nu} \left(\frac{\partial}{\partial x^\mu} \right)^a \left(\frac{\partial}{\partial x^\nu} \right)^b - 2A^\mu \left(\frac{\partial}{\partial x^\mu} \right)^a \left(\frac{\partial}{\partial w} \right)^b + A^\rho A_\rho \left(\frac{\partial}{\partial w} \right)^a \left(\frac{\partial}{\partial w} \right)^b. \quad (19)$$

From equations (18) and (19) we get

$$g_a^c = g_{ab} g^{bc} = \delta_\mu^\nu dx_a^\mu \left(\frac{\partial}{\partial x^\nu} \right)^c - A_\mu dx_a^\mu \left(\frac{\partial}{\partial w} \right)^c = g^c{}_a, \quad (20)$$

just as being expected.

Therefore, given the 4+1 decomposition of the five-dimensional metric in equation (1), the four-dimensional metric tensor g_{ab} and the electromagnetic potential vector A^a are uniquely determined by equations (18) and (15). They are determined by the assumed form of the five-dimensional metric and the self-consistency requirement of the theory, without additional assumptions.

3. Geometric interpretation of the KK variables

For the KK theory to be meaningful, the 4-metric tensor g_{ab} and the electromagnetic potential 4-vector A_a assumed in the KK theory, and the four-dimensional quantities derived from them (e.g. the four-dimensional Ricci tensor R_{ab} and the electromagnetic field antisymmetric tensor F_{ab}) must be defined on some four-dimensional manifold—or a four-dimensional submanifold embedded in the five-dimensional manifold $\tilde{\mathcal{M}}$. Such a submanifold \mathcal{M} should be a hypersurface in $\tilde{\mathcal{M}}$, since $\dim \mathcal{M} = 4 = \dim \tilde{\mathcal{M}} - 1$. Assuming that such a hypersurface has a unit normal n^a , which must be spacelike since (\mathcal{M}, g_{ab}) is supposed to be a four-dimensional spacetime. That is, all vectors in a vector space tangent to \mathcal{M} are orthogonal to n^a , and $\tilde{g}_{ab} n^a n^b = n^a n_a = 1$. Then, the 4-metric g_{ab} on \mathcal{M} must be related to the 5-metric \tilde{g}_{ab} on $\tilde{\mathcal{M}}$ by $g_{ab} = \tilde{g}_{ab} - n_a n_b$, or, equivalently,

$$g^{ab} = \tilde{g}^{ab} - n^a n^b. \quad (21)$$

The questions are: *does such a hypersurface exist?* If yes, *how is it defined?*

It appears that neither of the above two questions has been seriously considered in the literature, at least to the knowledge of the present author. In his original paper [1] (English translation in [26], page 61), Kaluza only wrote that ‘we are certainly free to consider our space-time to be a four-dimensional part of an R_5 ’. Kaluza used R_5 to denote a five-dimensional spacetime. In [2] (English translation in [26], page 76), Klein only stated that ‘four of the

coordinates, x^1, x^2, x^3, x^4 , say, are to characterize the usual space-time.’ (Klein’s x^1, x^2, x^3, x^4 are equivalent to our x^0, x^1, x^2, x^3 respectively, and his x^0 corresponds to our w coordinate.) How is ‘a four-dimensional part of an R_5 ’ defined? What is the exact meaning of ‘the usual space-time’ in mathematics? These questions have never been clearly answered.

In some references the authors have explicitly identified the four-dimensional spacetime with the hypersurface defined by $w = \text{const}$ without any proof or argument. For example, in [27] Einstein and Bergmann wrote that ‘We consider a four dimensional surface cutting each of the A -lines once and only once. We introduce on this surface 4 coördinates $x^a (a = 1 \dots 4)$ and assume x^0 equal zero on this surface.’ (section I of [27], subsection ‘The Special Coördinate System’). Note that their coordinate x^0 corresponds to our w , and their x^a correspond to our $x^\mu (\mu = 0, 1, 2, 3)$. Their ‘ A -lines’ correspond to our w -lines. To distinguish it from the electromagnetic potential vector A^a used in this paper, let us denote the 5-vector ‘ A ’ used in [27] by \tilde{A}^a . In our notations, $\tilde{A}^a = \phi^{-2}(\partial/\partial w)^a = \phi^{-1}n^a$, which is related to the 4-vector potential A^a by $\tilde{A}_a = A_a + dw_a$ (see equation (33) below, Einstein and Bergmann chose $\phi = 1$ so that $\tilde{g}_{ab}\tilde{A}^a\tilde{A}^b = 1$). Similarly, in [28] (English translation in [26], page 108), Thiry wrote that ‘Kaluza’s attempt at a unified theory consists of considering space-time as the $x^0 = \text{const.}$ subspace of a five-dimensional Riemann space, and of assuming this subspace cylindrical with respect to the fifth coordinate x^0 .’ (Thiry’s x^0 is equal to our w coordinate.)

There are also people taking different views. For example, Coquereaux & Esposito-Farese [29] interpreted the four-dimensional spacetime as a hypersurface orthogonal to the w -lines by stating that ‘Locally, the four-dimensional space orthogonal to this vector will be interpreted as the usual space-time’. Their ‘this vector’ corresponds to our $n^a = \phi^{-1}(\partial/\partial w)^a$, i.e. the vector n^a in equation (33) below. However, Coquereaux & Esposito-Farese did not provide evidence supporting their views. They did not even consider whether ‘the 4-dimensional space orthogonal to this vector’ exists or not. As will be shown later in this paper, such a hypersurface does not exist in general.

The view that the submanifold supporting the four-dimensional variables in the KK theory coincides with the hypersurface defined by $w = \text{const}$ was disproved in [30]. Let us use \mathcal{S} to denote the hypersurface defined by $w = \text{const}$, and write its unit normal as s^a . The s^a is defined by $\tilde{g}_{ab}s^a s^b = 1$ and $\tilde{g}_{ab}s^a(\partial/\partial x^\mu)^b = 0$ for $\mu = 0 \dots 3$. Thus we must have

$$s_a = (\tilde{g}^{44})^{-1/2} dw_a = \frac{1}{\sqrt{\phi^{-2} + A_\rho A^\rho}} dw_a. \quad (22)$$

By equations (15) and (19), we have

$$A^a s_a = -(\tilde{g}^{44})^{-1/2} A_\rho A^\rho = -(\tilde{g}^{44})^{-1/2} A_a A^a, \quad (23)$$

and

$$g^{ab}s_b = (\tilde{g}^{44})^{-1/2} \left[-A^\mu \left(\frac{\partial}{\partial x^\mu} \right)^a + A^\rho A_\rho \left(\frac{\partial}{\partial w} \right)^a \right] = -(\tilde{g}^{44})^{-1/2} A^a. \quad (24)$$

Hence, the four-dimensional variables g_{ab} and A^a are not orthogonal to s^a unless $A_a = 0$.

By equation (2) we have the metric tensor on $\mathcal{S}(w = 0)$

$$\hat{g}_{ab} = \tilde{g}_{ab} - s_a s_b = (g_{\mu\nu} + \phi^2 A_\mu A_\nu) dx_a^\mu dx_b^\nu + 2\phi^2 A_\mu dx_a^\mu dw_b + \frac{\phi^4 A_\rho A^\rho}{1 + \phi^2 A_\rho A^\rho} dw_a dw_b. \quad (25)$$

After restriction of the action of \hat{g}_{ab} on vectors tangent to \mathcal{S} , we get the metric tensor on \mathcal{S}

$$\hat{g}_{ab} = (g_{\mu\nu} + \phi^2 A_\mu A_\nu) dx_a^\mu dx_b^\nu. \quad (26)$$

The four-dimensional components of \hat{g}_{ab} are not $g_{\mu\nu}$, but $g_{\mu\nu} + \phi^2 A_\mu A_\nu$. Hence, the KK theory can be treated as *approximately* being defined on \mathcal{S} only if the electromagnetic field is sufficiently weak so that $|\phi^2 A_\mu A_\nu| \ll |g_{\mu\nu}| \sim 1$, i.e. only if

$$|\phi A_\mu| \ll A_{\text{crit}} \equiv \frac{c^2}{G^{1/2}} = 3.48 \times 10^{24} \text{ statvolt}. \quad (27)$$

If $|\phi A_\mu| \gtrsim A_{\text{crit}}$, the g_{ab} and A_a are not orthogonal to s^a and the KK theory cannot be treated as being defined on the hypersurface defined by $w = 0$.

To find a submanifold \mathcal{M} supporting the KK four-dimensional variables, we submit equations (7) and (19) into equation (21). We get

$$n^a n^b = \frac{1}{\phi^2} \left(\frac{\partial}{\partial w} \right)^a \left(\frac{\partial}{\partial w} \right)^b, \quad (28)$$

which immediately leads to a unique solution (up to a sign)

$$n^a = \phi^{-1} w^a, \quad w^a \equiv \left(\frac{\partial}{\partial w} \right)^a. \quad (29)$$

Thus, if the submanifold \mathcal{M} exists, its unit normal n^a must be a unit vector tangent to the coordinate lines of the extra dimension. It is easy to verify that both g_{ab} and A^a are orthogonal to n^a .

Let us denote the tangent space of the five-dimensional manifold \tilde{M} at a point $p \in \tilde{M}$ by \tilde{T}_p , $\dim \tilde{T}_p = \dim \tilde{M} = 5$. The disjoint union of \tilde{T}_p at all points of \tilde{M} is called the tangent bundle of \tilde{M} and denoted as $\tilde{T} = \tilde{T}(\tilde{M})$. Let \mathcal{T} be a rank-4 subbundle of \tilde{T} (called a rank-4 distribution or tangent distribution, or tangent subbundle [31]), described by a disjoint union of subspaces containing all vectors and tensors orthogonal to $w^a \propto n^a$ at all points of \tilde{M} , i.e. $\mathcal{T} = \coprod_{p \in \tilde{M}} \mathcal{T}_p$ with $\dim \mathcal{T}_p = 4$. The \mathcal{T} is a smooth distribution in the sense that for each $p \in \tilde{M}$ we can find an open neighborhood \tilde{O} of p such that in \tilde{O} , \mathcal{T} is spanned by smooth vector and tensor fields orthogonal to w^a . We have $g_{ab}, g_a^b, g^{ab} \in \mathcal{T}$. The g_{ab} is the metric tensor field in \mathcal{T} . The g^{ab} is the inverse metric tensor field, and g_a^b the identity map in \mathcal{T} ³. Since $A_a w^a = 0 = A^a w_a$, we have $A_a, A^a \in \mathcal{T}$, too.

By equations (2) and (14), we have

$$\phi^2 A_a = \phi^2 A_\mu dx_a^\mu = \tilde{g}_{bc} \left(\frac{\partial}{\partial x^\mu} \right)^b \left(\frac{\partial}{\partial w} \right)^c dx_a^\mu. \quad (30)$$

By equation (8), we have

$$dx_a^\mu \left(\frac{\partial}{\partial x^\mu} \right)^b = \tilde{\delta}_a^b - dw_a \left(\frac{\partial}{\partial w} \right)^b. \quad (31)$$

Hence, we get

$$\phi^2 A_a = \tilde{g}_{bc} \left(\frac{\partial}{\partial w} \right)^c \left[\tilde{\delta}_a^b - dw_a \left(\frac{\partial}{\partial w} \right)^b \right] = \tilde{g}_{ac} \left(\frac{\partial}{\partial w} \right)^c - \phi^2 dw_a, \quad (32)$$

where we have used $\tilde{g}_{bc}(\partial/\partial w)^b(\partial/\partial w)^c = \tilde{g}_{44} = \phi^2$. Submitting the definition of n^a in equation (29) into equation (32), we get the relation

$$A_a = \phi^{-1} n_a - dw_a. \quad (33)$$

³ The $g_a^b = \tilde{g}_a^b - n_a n^b$ is also the projection operator mapping a vector (and a tensor) in \tilde{T} to a vector (and a tensor) in \mathcal{T} .

Equation (33) states that A^a is obtained from orthogonal decomposition of the vector field $\mu^a \equiv \tilde{g}^{ab} dw_b = (g^{44})^{1/2} s^a$: one component is along the direction of w^a (i.e. the direction of n^a by equation (29)), the other (which is $\propto -A^a$) is in the direction orthogonal to w^a .

By $\tilde{\nabla}_{[a} dw_{b]} = 0$, where the brackets in the indices of a tensor denote antisymmetrization of the tensor about the indices inside the brackets, equation (33) implies that $\tilde{\nabla}_{[a} A_{b]} = \tilde{\nabla}_{[a} (\phi^{-1} n_{b]})$, i.e.

$$\tilde{\nabla}_{[a} A_{b]} = n_{[b} \tilde{\nabla}_{a]} \phi^{-1} + \phi^{-1} \tilde{\nabla}_{[a} n_{b]} . \quad (34)$$

Although $\tilde{\nabla}_a$ is the derivative operator associated with the metric \tilde{g}_{ab} , equation (34) remains valid if $\tilde{\nabla}_a$ is replaced by any derivative operator. By equation (34), we have the antisymmetric tensor of the electromagnetic field

$$F_{ab} \equiv 2g_a^c g_b^d \tilde{\nabla}_{[c} A_{d]} = 2\phi^{-1} g_a^c g_b^d \tilde{\nabla}_{[c} n_{d]} , \quad (35)$$

where we have used the identity $g_a^b n_b = 0$. By definition, $F_{ab} \in \mathcal{T}$ since $F_{ab} w^b = 0 = F_{ab} w^a$.

If w^a were timelike and $n_a n^a = -1$, the quantity

$$\omega_{ab} \equiv g_a^c g_b^d \tilde{\nabla}_{[d} n_{c]} \quad (36)$$

would be called the vorticity tensor of the congruence of the integral curves of the vector field w^a [32]. Similarly, the expansion tensor would be defined by

$$\theta_{ab} \equiv g_a^c g_b^d \tilde{\nabla}_{(c} n_{d)} = \frac{1}{2} \tilde{\mathcal{L}}_n g_{ab} , \quad (37)$$

where $\tilde{\mathcal{L}}_n$ is the Lie derivative generated by n^a (as usual, the tilde above \mathcal{L}_n indicates that the operation is in the five-dimensional spacetime). Now w^a is spacelike and $n_a n^a = 1$. We can still define the ω_{ab} and θ_{ab} by equations (36) and (37) and call them the vorticity tensor and the expansion tensor, respectively, of the integral curves of w^a . Both ω_{ab} and θ_{ab} are $\in \mathcal{T}$, since they both are orthogonal to w^a . Then, equation (35) is equivalent to

$$F_{ab} = -2\phi^{-1} \omega_{ab} . \quad (38)$$

Therefore, the electromagnetic field antisymmetric tensor in the KK theory is proportional to the vorticity tensor of the congruence of the extra dimension coordinate curves.

4. Submanifold supporting the effective four-dimensional theory does not exist

As shown in section 3, the KK four-dimensional field variables g_{ab} and A_a are contained in a rank-4 distribution \mathcal{T} with members orthogonal to the extra dimension coordinate vector field w^a . In order that there exists an integral submanifold of \mathcal{T} , i.e. a submanifold $\mathcal{M} \subseteq \hat{\mathcal{M}}$ whose tangent space at any $p \in \mathcal{M}$ coincides with \mathcal{T}_p , the distribution \mathcal{T} must be involutive, i.e. for any $X^a, Y^a \in \mathcal{T}$ the commutator (Lie bracket) $[X, Y]^a \in \mathcal{T}$ (Frobenius's theorem, [31]). The condition that the tangent subbundle \mathcal{T} is integrable is equivalent to that the vector field w^a is hypersurface orthogonal, i.e. there exists a hypersurface orthogonal to w^a .

By Frobenius's theorem, w^a is hypersurface orthogonal if and only if $[X, Y]^a \in \mathcal{T}$ for all $X^a, Y^a \in \mathcal{T}$, which is mathematically equivalent to the condition that there is a dual vector v_a such that [22, 33]

$$\tilde{\nabla}_{[a} w_{b]} = w_{[a} v_{b]} . \quad (39)$$

Equation (39) is equivalent to the condition that the congruence of the spacelike curves tangent to w^a is vorticity-free, which is proved as follows. Since $w^a = \phi n^a$, equation (39) is equivalent to

$$\tilde{\nabla}_{[a} n_{b]} = n_{[a} v'_{b]} , \quad (40)$$

where $v'_a = v_a + \tilde{\nabla}_a \ln \phi$. Then we have

$$g_a{}^c g_b{}^d \tilde{\nabla}_{[c} n_{d]} = g_a{}^c g_b{}^d n_{[c} v'_{d]} = 0 , \quad (41)$$

since $g_a{}^c n_c = 0$. By equation (36) we get then

$$\omega_{ab} = 0 . \quad (42)$$

Equation (40) can also be derived from equation (42), since by $\tilde{\nabla}_a n_b = \theta_{ab} + \omega_{ba} + n_a a_b$ we have

$$\tilde{\nabla}_{[a} n_{b]} = \omega_{ba} + n_{[a} a_{b]} , \quad (43)$$

where $a_a \equiv n^b \tilde{\nabla}_b n_a$. Thus, $\omega_{ab} = 0$ implies $\tilde{\nabla}_{[a} n_{b]} = n_{[a} a_{b]}$, i.e. equation (40) if we take $v'_b = a_b$. Therefore, *a necessary and sufficient condition for a congruence of timelike or spacelike curves to be hypersurface orthogonal is that the congruence is vorticity-free.*

Only if w^a is hypersurface orthogonal the distribution \mathcal{T} in a neighborhood of any point is integrable and can be spanned by coordinate base vector fields [22]. By equations (38) and (42), w^a is hypersurface orthogonal if and only if

$$F_{ab} = 0 . \quad (44)$$

Thus, *the submanifold \mathcal{M} supporting the KK variables exists if and only if the electromagnetic field antisymmetric tensor vanishes.* This restriction on the KK theory is too strong, as there can be no unification of the gravitational and electromagnetic interactions if the electromagnetic field vanishes.

5. Action and compactification of the extra dimension

When the cylinder condition is satisfied, i.e. all components of the five-dimensional metric tensor are independent of the extra dimension coordinate [1], the Maxwell equations can easily be derived from the action principle⁴. The cylinder condition is equivalent to the requirement that w^a is a Killing vector of the five-dimensional spacetime, i.e.

$$\tilde{\mathcal{L}}_w \tilde{g}_{ab} = 0 . \quad (45)$$

When the above condition is satisfied, the Ricci scalar \tilde{R} associated with the five-dimensional bulk metric \tilde{g}_{ab} is related to the Ricci scalar R associated with the KK four-dimensional metric g_{ab} by [29, 34, 35]

$$\tilde{R} = R - \frac{\phi^2}{4} F_{ab} F^{ab} + \tilde{\nabla}_a v^a , \quad (46)$$

where v^a is a vector.

⁴ If the cylinder condition is dropped the derived field equations are much more complicated, see, e.g. [36].

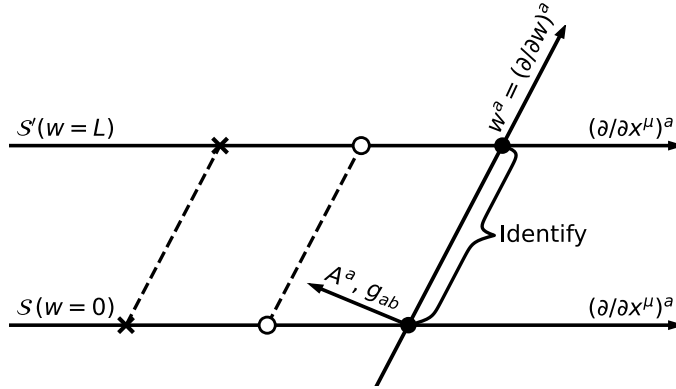


Figure 1. Under the cylinder condition the five-dimensional spacetime can be compactified along the direction of extra dimension [i.e. the direction of $w^a = (\partial/\partial w)^a$]. This way, the hypersurface $S(w=0)$ in $(\tilde{M}, \tilde{g}_{ab})$ is identified with the $S'(w=L)$ under the map generated by w^a . The black dot on S is identified with the black dot on S' , the circle on S identified with the circle on S' , and so on (as indicated by dashed lines). The KK variables A^a and g_{ab} are orthogonal to w^a hence not tangent to S , since in general w^a is not orthogonal to any hypersurface unless the electromagnetic field F_{ab} vanishes. Note that to make the extra dimension unaccessible to current experiments its circumference $C_w = \phi L$ has to be small.

The determinant of the five-dimensional metric, $\tilde{g} = \det \tilde{g}_{AB}$, is related to the determinant of the four-dimensional metric, $g = \det g_{\mu\nu}$, by $\tilde{g} = \phi^2 g$. Thus, we have $\sqrt{-\tilde{g}} = \phi \sqrt{-g}$ and the five-dimensional action of gravity

$$I_g = \frac{1}{\tilde{G}} \int \tilde{R} \sqrt{-\tilde{g}} d^4x dw = \frac{1}{\tilde{G}} \int dw \int \phi \left(R - \frac{\phi^2}{4} F_{ab} F^{ab} \right) \sqrt{-g} d^4x, \quad (47)$$

where \tilde{G} is the five-dimensional gravitational constant. The divergence term in equation (46), $\tilde{\nabla}_a v^a$, has been dropped off since it has no contribution to the action integral under appropriate boundary conditions.

Since w^a is a Killing vector field, the five-dimensional spacetime can be compactified along the direction of extra dimension, i.e. the direction of w -lines (figure 1). That is, a spacetime point $\{x^\mu, w\}$ in $(\tilde{M}, \tilde{g}_{ab})$ is identified with the spacetime point $\{x^\mu, w+L\}$ in $(\tilde{M}, \tilde{g}_{ab})$, where $L > 0$ is a constant. Then, the value of w is restricted to the region of $[0, L)$, which leads to $\int dw = L$. To have the circumference of the extra dimension—which is $C_w \equiv \phi \int dw = \phi L$ —to be constant, ϕ must be constant [34]⁵. For the extra dimension to be unaccessible to current experiments the length $C_w = \phi L$ must be sufficiently small [2]. Then, equation (47) becomes

⁵ It should be noted that ϕ cannot be constant if the vacuum Einstein field equation $\tilde{R}_{ab} = 0$ is imposed. This follows because the equation $\tilde{R}_{ww} = 0$ entails that

$$\tilde{\nabla}^a \tilde{\nabla}_a \phi = \frac{\phi^3}{4} F^{ab} F_{ab},$$

as first noted by Jordan [37] and Thiry [28].

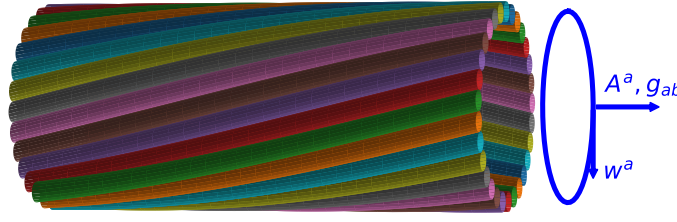


Figure 2. The Kaluza–Klein tube obtained by compactification of the extra dimension in the direction of w^a . The five-dimensional spacetime tube is made of twisted wires, each wire representing a hypersurface defined by $w = \text{const}$. The blue circle on the right represents the w -coordinate line with a length of $C_w = \phi L$. The KK variables A^a and g_{ab} are contained in a four-dimensional tangent subbundle \mathcal{T} orthogonal to w^a .

$$I_g = \frac{C_w}{\tilde{G}} \int \left(R - \frac{\phi^2}{4} F_{ab} F^{ab} \right) \sqrt{-g} d^4x. \quad (48)$$

The appearance of the term $-(\phi^2/4)F_{ab}F^{ab}$ in the Lagrangian density in equation (48) guarantees that the Maxwell equations can be derived from the five-dimensional Einstein field equation by variation of the action I_g with respect to the potential vector A^a . In fact, if we identify \tilde{G}/C_w as the four-dimensional gravitational constant $G = 1$ (i.e. $C_w = \tilde{G}$) and $\phi = 2$, the action in equation (48) reduces to the total action of gravity and electromagnetic fields in a four-dimensional spacetime (see [22], appendix E). Note that, however, all the quantities (F_{ab} , R , and g) appearing in the integral of equation (48) are defined in the tangent subbundle \mathcal{T} orthogonal to w^a , since they all are derived from A_a , g_{ab} , and the derivative operator ∇_a associated with g_{ab} .

According to the results in section 4, the rank-4 distribution \mathcal{T} is not integrable and hence g_{ab} , A_a , and quantities derived from them are not tangent to any four-dimensional submanifold embedded in \mathcal{M} unless the electromagnetic field F_{ab} vanishes. Although under the cylinder condition the four-dimensional Einstein field equation and the Maxwell equations are successfully derived from the five-dimensional Einstein field equation through the action principle, these equations are not supported by a four-dimensional submanifold hence do not define a four-dimensional spacetime.

The manifold structure of the KK theory, after the extra dimension is compactified, is depicted in figure 2. The five-dimensional spacetime ‘tube’ is made of twisted four-dimensional ‘wires’, with each wire representing a hypersurface $w = \text{const}$. The transverse cross-section of the spacetime tube corresponds to the w -coordinate lines, i.e. curves whose tangent vectors are $w^a = (\partial/\partial w)^a$, as indicated by the blue circle in the figure. The KK variables A^a , g_{ab} , and the associated distribution \mathcal{T} , are in the longitudinal direction along the tube (i.e. the direction perpendicular to w^a). They are not tangent to any four-dimensional submanifold. Thus, the action in equation (48) is defined in the rank-4 tangent subbundle or distribution \mathcal{T} , but not defined on a four-dimensional submanifold.

6. Summary and discussion

All existing physical theories are defined on a smooth manifold with or without a well-defined spacetime metric. In the KK theory, the five-dimensional theory is defined on a five-dimensional manifold with a Lorentz metric determined by the five-dimensional Einstein field equation. The four-dimensional metric tensor and the electromagnetic potential vector

assumed in the KK theory must be defined on a four-dimensional submanifold (i.e. a hypersurface) embedded in the five-dimensional manifold, in order for the derived four-dimensional theory (including the four-dimensional Einstein field equation and the Maxwell equations) to be able to describe the four-dimensional world where we live and do physical experiments. But this is not the case, as has been shown in this paper.

In general, the four-dimensional KK variables g_{ab} , A_a , and other geometric quantities derived from them (e.g. the four-dimensional Ricci tensor R_{ab} and the electromagnetic field antisymmetric tensor F_{ab}) are in a four-dimensional subbundle that is not tangent to any four-dimensional submanifold, since by the KK construction g_{ab} and A_a are orthogonal to the vector field w^a generating the extra dimension but w^a is not hypersurface orthogonal unless the electromagnetic field vanishes. Thus, the results presented in the paper lead us to such a paradox: the KK theory is valid mathematically only if the electromagnetic field derived from the KK theory vanishes. This is a general conclusion, independent of the cylinder condition adopted for derivation of the four-dimensional field equations.

When the electromagnetic field is weak and has a negligible effect on the spacetime metric, i.e. when condition (27) is satisfied, the four-dimensional metric tensor and the electromagnetic potential vector can be regarded as *approximately* being defined on the hypersurface of $w = \text{const}$. But then the KK theory becomes an approximate and weak-field limit theory, conflicting the original spirit of unification of gravitational and electromagnetic interactions. In addition, without a precisely defined four-dimensional submanifold supporting the four-dimensional variables, it is hard to accept the approximate theory since it is not well defined in mathematics. An ultimate solution to the problem raised in this paper may be given by a different 4+1 decomposition of a five-dimensional spacetime metric as having been proposed in [30], where the four-dimensional spacetime is defined on a hypersurface that is not orthogonal to the extra dimension, but then the theory is different from the KK theory since an electromagnetic field equation with a curvature-coupled term is derived.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The author thanks the anonymous adjudicator for a very enlightening report, which has stimulated the author to think more deeply and more widely about the problem discussed in the paper. The report has also helped to improve the presentation of the paper. This work was supported by the NSFC Grants program (No. 11973014).

ORCID iD

Li-Xin Li  <https://orcid.org/0000-0002-8466-321X>

References

- [1] Kaluza T 1921 Zum unitätsproblem der physik *Itzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921** 966–72
- [2] Klein O 1926 Quantentheorie und fünfdimensionale relativitätstheorie *Z. Phys.* **37** 895
- [3] Klein O 1926 The atomicity of electricity as a quantum theory law *Nature* **118** 516

- [4] Nordström G 1914 Über die möglichkeit, das electromagnetische feld und das gravitationsfeld zu vereinigen *Phys. Z.* **15** 504–6
- [5] Cremmer E and Scherk J 1976 Spontaneous compactification of space in an Einstein–Yang–Mills–Higgs model *Nucl. Phys. B* **108** 409
- [6] Cremmer E and Scherk J 1977 Spontaneous compactification of extra space dimensions *Nucl. Phys. B* **118** 61
- [7] Luciani J F 1978 Space-time geometry and symmetry breaking *Nucl. Phys. B* **135** 111
- [8] Englert F 1982 Spontaneous compactification of eleven-dimensional supergravity *Phys. Lett. B* **119** 339
- [9] Li L-X and Gott J R 1998 Inflation in Kaluza–Klein theory: relation between the fine-structure constant and the cosmological constant *Phys. Rev. D* **58** 103513
- [10] van Nieuwenhuizen P 1981 Supergravity *Phys. Rep.* **68** 189–398
- [11] Duff M J, Nilsson B E W and Pope C N 1986 Kaluza–Klein supergravity *Phys. Rep.* **130** 1–142
- [12] Green M B, Schwarz J H and Witten E 2012 *Superstring Theory* 25th anniversary edn vol I and II (Cambridge University Press)
- [13] Polchinski J 2003 *String Theory* vol I and II (Cambridge University Press)
- [14] Becker K, Becker M and Schwarz J H 2007 *String Theory and M-Theory: A Modern Introduction* (Cambridge University Press)
- [15] Kiritsis E 2019 *String Theory in a Nutshell* 2nd edn (Princeton University Press)
- [16] Randall L and Sundrum R 1999 Large mass hierarchy from a small extra dimension *Phys. Rev. Lett.* **83** 3370
- [17] Randall L and Sundrum R 1999 An alternative to compactification *Phys. Rev. Lett.* **83** 4690
- [18] Maartens R and Koyama K 2010 Brane-world gravity *Living Rev. Relativ.* **13** 5
- [19] Nath P and Yamaguchi M 1999 Probing the nature of compactification with Kaluza–Klein excitations at the large hadron collider *Phys. Lett. B* **466** 100
- [20] Bhattacharyya G, Datta A, Majee S K and Raychaudhuri A 2009 Exploring the universal extra dimension at the LHC *Nucl. Phys. B* **821** 48
- [21] Deutschmann N, Flacke T and Kim J S 2017 Current LHC constraints on minimal universal extra dimensions *Phys. Lett. B* **771** 515
- [22] Wald R M 1984 *General Relativity* (The University of Chicago Press)
- [23] de Witt B S 1964 *Dynamical Theory of Groups and Fields (Relativity Groups and Topology, Les Houches 1963)* (Gordon and Breach Science Publishers) p 585
- [24] Kerner R 1968 Generalization of the Kaluza–Klein theory for an arbitrary non-abelian gauge group *Ann. Inst. Henri Poincaré* **9** 143
- [25] Cho Y M and Freund P G O 1975 Non-Abelian gauge fields as nambu-goldstone fields *Phys. Rev. D* **12** 1711
- [26] Appelquist T, Chodos A and Freund P G O (eds) 1987 *Modern Kaluza–Klein Theories* (Addison-Wesley)
- [27] Einstein A and Bergmann P 1938 On a generalization of Kaluza’s theory of electricity *Ann. Math.* **39** 683
- [28] Thiry M Y 1948 Les équations de la théorie unitaire de Kaluza *Compt. Rend. Acad. Sci. Paris.* **226** 216
- [29] Coquereaux R and Esposito-Farese G 1990 The theory of Kaluza–Klein–Jordan–Thiry revisited *Ann. Inst. Henri Poincaré* **52** 113
- [30] Li L-X 2016 A new unified theory of electromagnetic and gravitational interactions *Front. Phys.* **11** 110402
- [31] Lee J M 2012 *Introduction to Smooth Manifolds* 2nd edn (Springer)
- [32] Hawking S W and Ellis G F R 1973 *The Large Scale Structure of Space-Time* (Cambridge University Press)
- [33] Straumann N 2013 *General Relativity* 2nd edn (Springer)
- [34] Bailin D and Love A 1987 Kaluza–Klein theories *Rep. Prog. Phys.* **50** 1087
- [35] Williams L L 2015 Field equations and Lagrangian for the Kaluza metric evaluated with tensor algebra software *J. Grav.* **2015** 901870
- [36] Overduin J M and Wesson P S 1997 Kaluza–Klein gravity *Phys. Rep.* **283** 303
- [37] Jordan P 1947 Erweiterung der projektiven relativitätstheorie *Ann. Phys., Lpz.* **436** 219