A Brief Summary of Kaluza-Klein Theory

Jonah Miller
jonah.maxwell.miller@gmail.com

Friday May 3^{rd}, 2013

Abstract

We briefly discuss the origins of Kaluza-Klein theory. First, we describe Kaluza's original "cylinder condition" and how this allows for the unification of Einstein gravity and electromagnetism in a five-dimensional framework. We then discuss possible justifications of the cylinder condition, with a special emphasis on Klein's compactification approach.

1 Introduction and Some History

As early as 1914, physicists dreamed of a geometric unification the fundamental forces of nature. Inspired by Minkowski's treatment of Maxwell's equations in special relativity [1], Gunnar Nordström unified electromagnetism and Newtonian gravity by embedding a symmetric energy-momentum tensor in five-dimensional Minkowski space [2, 3]. Since Nordström's idea relies on Newtonian gravity, it's not particularly relevant today. But given that Nordstöm published his idea before the discovery of general relativity, it does represent a remarkably prescient geometric approach to fundamental physics.

Only seven years later, now after the discovery of general relativity, Theodor Kaluza discovered that the metric tensor of five-dimensional curved spacetime can be viewed as the metric of four-dimensional spacetime coupled to the electromagnetic vector potential and a scalar field [2, 4]. He sent his idea to Einstein, who encouraged him to publish [2] and who later expanded on Kaluza's idea [2, 5].

Kaluza-Klein theory gets the other half of its name from Oscar Klein. As we will see in section 2, Kaluza's theory imposes a rather arbitrary condition—which he called the "cylinder condition"—on the geometry. In an innovative attempt to quantize electric charge, Klein suggested one explanation for this condition, where the extra fifth dimension is periodic and very small. See section 3 for details. The insights of Kaluza and Klein have inspired a plethora of higher-dimensional unification schemes, from supergravity and string theory [2, 6, 7] to Brane theory and higher-dimensional cosmologies [8, 9].

In this work, we discuss Kaluza's initial insight in section 2 and Klein's extension of the theory in 3. Klein's method is not the only physically compelling way to study the cylinder condition and we will briefly discuss the other approaches in that section as well. Finally, in section 4 we offer some concluding remarks.

2 Kaluza's Five-Dimensional Spacetime

We now present Kaluza's original idea with the arbitrary "cylinder condition." Although Kaluza originally used Einstein's tensor notation [4], Thiry demonstrated that the langauge of differential forms is better suited to the task at hand [10], so we will use that notation. We borrow the setup from Einstein and Bergmann [5] and then loosely follow Pope, but with more detail [6].

2.1 General Construction

Let $\mathcal{M}^{(5)}$ be a (4+1)-dimensional Lorentzian manifold and let

$$\hat{x} = x^A \partial_A \tag{1}$$

be a set of general coordinates on $\mathcal{M}^{(5)}$.[1] Without loss of generality, we choose ∂x_0 to be the basis element in a timelike direction, ∂_i to be three spatial coordinates, and ∂_4 to the be the fourth spatial coordinate, which will describe the extra fifth dimension. By convention we will choose timelike directions to have negative square norm and spacelike directions to have positive square norm. We will call the five-dimensional metric $\hat{g}_{AB}$.[2]

2.1.1 The Cylinder Condition

At this point, there is no geometric difference between any of the spacelike coordinates; ∂_i $i \in \{1, 2, 3\}$, and ∂_4 are indistinguishable. Kaluza's cylinder condition, however, will make them distinct, and we are now in a position to state it. Kaluza (rather apologetically) enforces that the five-dimensional metric be independent of x^4 [4, 2, 7]. In other words, the derivatives of the metric with respect to the fifth coordinate must vanish:

$$\partial_4 \hat{g}_{AB} = 0. \tag{2}$$

This is a statement about symmetry. The cylinder condition demands that the spacetime have a specific isometry,

$$x^4 \to x^4 + \varepsilon, \ \forall \ \varepsilon \in \mathbb{R}. \tag{3}$$

Furthermore, it enforces a specific parametrization where ∂_4 is the Killing vector associated with our isometry. Mathematicians call parametrizations of this type "Clairaut" [11], and the existence a Clairaut parametrization is not at all guaranteed. We are requiring that at least one spacelike Killing vector field exists in this spacetime and, since most spacetimes lack *any* isometries, this condition is quite restrictive.

This Clairaut parametrization *does* allow us to define tensors on $\mathcal{M}^{(4)}$, however. We can foliate our spacetime by spacelike hypersurfaces, each orthogonal to the Killing vector field ∂_5, such that these hypersurfaces are related by a global isometry. We call any such hypersurface $\mathcal{M}^{(4)}$. This foliation allows us to write down the topology of $\mathcal{M}^{(5)}$:

$$\mathcal{M}^{(5)} = \mathcal{M}^{(4)} \times \mathcal{M}^{(1)}, \tag{4}$$

where $\mathcal{M}^{(1)}$ is some one-dimensional Euclidean manifold [5]. The choice of hypersurface in a given foliation is arbitrary, since the five-metric will be identical on each one.

2.1.2 The Metric

We want to choose a geometrically natural parametrization and write down a metric. Before we do that, however, let's define a few quantities which we will use to build our line element. First, let the four-dimensional metric on a given hypersurface be $g_{\mu\nu}$. Let $\mathcal{A}_\mu$ be a four-covector field and let ϕ be a scalar field on $\mathcal{M}^{(4)}$.

Although they won't have tensorial properties on the general five-dimensional spacetime, we can certainly extend these quantities and define them on $\mathcal{M}^{(5)}$. However, we must impose that

$$\partial_4 \mathcal{A}_\mu = 0, \partial_4 g_{\mu\nu} = 0, \quad \text{and } \partial_4 \phi = 0. \tag{5}$$

Although these names are suggestive—$\mathcal{A}_\mu$ suggests a vector potential and ϕ suggests a scalar field—they currently contain no physical information. They will only take on physical meaning once we write down a metric and study the geometry of the space.

And now we'll do just that. We can choose a parametrization of the space so that the five-dimensional line element takes the form

$$d\hat{s}^2 = e^{2\alpha\phi} g_{\mu\nu} dx^\mu dx^\nu + e^{2\beta\phi}(dx^4 + \mathcal{A}_\mu dx^\mu)^2, \tag{6}$$

[1] In this work, Greek indices will sum from zero to three and upper-case Latin indices from zero to four.

[2] Later, we will define tensors on four-dimensional submanifolds of $\mathcal{M}^{(5)}$. We will denote five-dimensional tensors with hats on them and four-dimensional tensors without the hats.

where α and β are currently arbitrary constants [6]. This means that the components of the five-dimensional metric tensor $\hat{g}_{AB}$ take the form

$$\hat{g}_{\mu\nu} = e^{2\alpha\phi} g_{\mu\nu} + e^{2\beta\phi}\mathcal{A}_\mu\mathcal{A}_\nu,\ \hat{g}_{\mu 4} = e^{2\beta\phi}\mathcal{A}_\mu,\ \text{and}\ \hat{g}_{44} = e^{2\beta\phi}. \tag{7}$$

So long as $\beta \neq 0$, this metric corresponds to a valid Clairaut parametrization of the space. Although this metric looks a little funny, the reason for our choice will be clear soon.

To make the remaining calculations easier, let's transition to a coordinate-free description. Let be e^a be a tetrad basis for one of the $\mathcal{M}^{(4)}$ hypersurfaces such that

$$\eta_{ab} = e^\mu_a e^\nu_b g_{\mu\nu}\ \forall\ a, b \in \{0, 1, 2, 3\}.$$

We now want to write down a fünfbein basis for the five dimensional spacetime, satisfying

$$\hat{\eta}_{ab} = \hat{e}^A_a \hat{e}^B_b \hat{g}_{AB}\ \forall\ a, b \in \{0, 1, 2, 3, 4\}.$$

Thanks to our nice choice of coordinates we can see by inspection[3] that

$$\hat{e}^a = e^{\alpha\phi} e^a \text{ and } \hat{e}^z = e^{\beta\phi}(dx^4 + \mathcal{A}_\mu dx^\mu) \tag{8}$$

will work. Beware the abuse of notation! e^a is a tetrad basis element, but $e^{\alpha\phi}$ and $e^{\beta\phi}$ are Euler's constant raised to powers. In keeping with Pope, we will denote the coordinate-free basis elements by lower-case Latin indices. We set aside lower-case z for the basis element pointing in direction associated with the cylinder condition [6]. We will assume that that lower-case Latin indices range from zero to three. If they do not, the range will be indicated.

To calculate the curvature of the space, we turn to the Cartan structure equations [12],

$$\begin{aligned} \hat{T}^a &= d\hat{e}^a + \hat{\omega}^a{}_b \wedge \hat{e}^b,\ a, b \in \{0, 1, 2, 3, 4\} \\ \hat{R}^a{}_b &= d\hat{\omega}^a{}_b + \hat{\omega}^a{}_c \wedge \omega^c{}_b + \hat{\omega}^a{}_z \wedge \hat{e}^z,\ a, b, c \in \{0, 1, 2, 3, 4\} \end{aligned}$$

where $\hat{T}^a$, $\hat{R}^a{}_b$, and $\hat{\omega}^a{}_b$ are the five-dimensional torsion two-form, curvature two-form and spin connections respectively. Since our spacetime is Torsion-free, the structure equations reduce to

$$0 = d\hat{e}^a + \hat{\omega}^a{}_b \wedge \hat{e}^b + \hat{\omega}^a{}_z \wedge \hat{e}^z \tag{9}$$
$$0 = d\hat{e}^z + \hat{\omega}^z{}_b \wedge \hat{e}^b + \hat{\omega}^z{}_z \wedge \hat{e}^z \tag{10}$$
$$\hat{R}^a{}_b = d\hat{\omega}^a{}_b + \hat{\omega}^a{}_c \wedge \omega^c{}_b + \hat{\omega}^a{}_z \wedge \hat{e}^z \tag{11}$$
$$\hat{R}^a{}_z = d\hat{\omega}^a{}_z + \hat{\omega}^a{}_c \wedge \hat{\omega}^c{}_z + \hat{\omega}^a{}_z \wedge \hat{e}^z \tag{12}$$
$$\hat{R}^z{}_z = d\hat{\omega}^z{}_z + \hat{\omega}^z{}_c \wedge \hat{\omega}^c{}_z + \hat{\omega}^z{}_z \wedge \hat{e}^z \tag{13}$$

with our usual index conventions. We'll need the following one-form,

$$d\phi = \partial_A \phi dx^A = \partial_\mu \phi dx^\mu. \tag{14}$$

Let's use equations 9 and 10 to calculate the spin connection.

$$\begin{aligned} d\hat{e}^a &= d\left(e^{\alpha\phi} e^a\right) \\ &= de^{\alpha\phi} \wedge e^a + e^{\alpha\phi} de^b \\ &= \alpha e^{\alpha\phi} d\phi \wedge e^a - e^{\alpha\phi} \omega^a{}_b \wedge e^b \\ &= \alpha e^{a\phi} \partial_\mu \phi (dx^\mu \wedge e^a) - e^{a\phi} \omega^a{}_b \wedge e^b, \end{aligned}$$

but, from equation 9,

$$\begin{aligned} d\hat{e}^a &= -\hat{\omega}^a{}_b \wedge \hat{e}^b + \hat{\omega}^a{}_z \wedge \hat{e}^z \\ &= -e^{\alpha\phi} \hat{\omega}^a{}_b \wedge e^b - e^{\beta\phi} \hat{\omega}^a{}_z \wedge (dx^4 + \mathcal{A}) \end{aligned}$$

[3]The power of working in a coordinate-free basis is that many quanities can be calculated "by inspection."

too. So,

$$\alpha e^{a\phi}\partial_\mu\phi(dx^\mu\wedge e^a) - e^{a\phi}\omega^a{}_b\wedge e^b = -e^{\alpha\phi}\hat{\omega}^a{}_b\wedge e^b - e^{\beta\phi}\hat{\omega}^a{}_z\wedge(dx^4+\mathcal{A}). \tag{15}$$

If we identify like terms,

$$\begin{aligned} \hat{\omega}^a{}_b\wedge e^b &= \omega^a{}_b\wedge e^b, && (16)\\ \text{and } \alpha e^{\alpha\phi}\partial_\mu(e^a\wedge dx^\mu) &= e^{\beta\phi}\hat{\omega}^a{}_z\wedge(dx^4+\mathcal{A}). && (17)\end{aligned}$$

Similarly,

$$\begin{aligned} d\hat{e}^z &= d\left[e^{\beta\phi}(dx^4+\mathcal{A})\right]\\ &= \beta e^{\beta\phi}d\phi\wedge(dx^4+\mathcal{A}) + e^{\beta\phi}d^2x^4 + e^{\beta\phi}d\mathcal{A}\\ &= \beta e^{\beta\phi}\partial_\mu\phi dx^\mu\wedge(dx^4+\mathcal{A}) + e^{\beta\phi}\mathcal{F}, && (18)\end{aligned}$$

where $\mathcal{F}=d\mathcal{A}$. The d^2x^4 term disappears because $d^2\Phi=0$ for any differential form Φ (x^4 is a zero-form). From equation 10,

$$\begin{aligned} d\hat{e}^z &= -\hat{\omega}^z{}_b\wedge\hat{e}^b + \hat{\omega}^z{}_z\hat{e}^z\\ &= -e^{\alpha\phi}\hat{\omega}^z{}_b\wedge e^b + \hat{\omega}^z{}_z\hat{e}^z.\end{aligned}$$

So,

$$-e^{\alpha\phi}\hat{\omega}^z{}_b\wedge e^b = \beta e^{\beta\phi}\partial_\mu\phi dx^\mu\wedge(dx^4+\mathcal{A}) + e^{\beta\phi}\mathcal{F}. \tag{19}$$

If we inspect equation 19 and re-absorb some of the scale factors into $\hat{e}^z$, we can infer that

$$\begin{aligned} \hat{\omega}^{az} = -\hat{\omega}^{za} &= -\beta e^{-\alpha\phi}\partial^a\phi\hat{e}^z - \frac{1}{2}\mathcal{F}^a{}_b e^{(\beta-2\alpha)\phi}\hat{e}^b, && (20)\\ \text{and } \hat{\omega}^{zz} &= 0. && (21)\end{aligned}$$

where $\partial_a = e^\mu{}_a\partial_\mu$ is the partial derivative translated into a coordinate-free basis with the help of the inverse fünfbeins on $\mathcal{M}^{(4)}$, $e^\mu{}_a$. Similarly, $\mathcal{F}_{ab}$ denotes the components of $\mathcal{F}$ in the coordinate-free basis [6]. If we plug this result into equation 16, we find that

$$\hat{\omega}^{ab} = \omega^{ab} + \alpha e^{-\alpha\phi}(\partial^b\phi\hat{e}^a - \partial^a\phi\hat{e}^b) - \frac{1}{2}\mathcal{F}^{ab}e^{(\beta-2\alpha)\phi}\hat{e}^z. \tag{22}$$

It's now relatively straightforward to compute the curvature two-form. For example, from equation 13,

$$\begin{aligned} \hat{R}^{zz} &= \hat{\omega}^z{}_c\wedge\hat{\omega}^{cz}\\ &= -\left[\beta e^{-\alpha\phi}\partial^a\phi\hat{e}^z + \frac{1}{2}\mathcal{F}^a{}_b e^{(\beta-2\alpha)\phi}\hat{e}^b\right]\wedge\left[\beta e^{-\alpha\phi}\partial_a\phi\hat{e}^z + \frac{1}{2}\mathcal{F}_{ac}e^{(\beta-2\alpha)\phi}\hat{e}^c\right]\\ &= -\beta^2e^{-2\alpha\phi}\left(\partial^a\phi\hat{e}^z\wedge\partial_a\phi\hat{e}^z\right) - \frac{1}{2}\mathcal{F}^a_b e^{(\beta-3\alpha)\phi}\hat{e}^b\wedge\partial_a\phi\hat{e}^z - \frac{1}{2}\mathcal{F}_{ab}e^{(\beta-3\alpha)\phi}\left(\partial^a\phi\hat{e}^z\right)\wedge\hat{e}^b\\ &\quad -\frac{1}{4}e^{2(\beta-2\alpha)\phi}\mathcal{F}^a{}_b\mathcal{F}_{ac}\hat{e}^b\wedge\hat{e}^c\\ &= -\beta^2e^{-2\alpha\phi}\left(\partial^a\phi\hat{e}^z\wedge\partial_a\phi\hat{e}^z\right) + \frac{1}{2}\mathcal{F}^a_b e^{(\beta-3\alpha)\phi}\hat{e}^b\wedge\partial_a\phi\hat{e}^z - \frac{1}{2}\mathcal{F}^a_b e^{(\beta-3\alpha)\phi}\hat{e}^b\wedge\partial_a\phi\hat{e}^z\\ &\quad -\frac{1}{4}e^{2(\beta-2\alpha)\phi}\mathcal{F}^a{}_b\mathcal{F}_{ac}\hat{e}^b\wedge\hat{e}^c\\ \Rightarrow \hat{R}_{zz} &= \beta^2e^{-2\alpha\phi}\Box\phi + \frac{1}{4}e^{2(\beta-2\alpha)\phi}\mathcal{F}^2, && (23)\end{aligned}$$

where

$$\mathcal{F}^2 = \mathcal{F}^{ab}F_{ab},$$

and

$$\Box\phi = \partial_a \partial^a \phi = \partial_\mu \partial^\mu \phi \tag{24}$$

is the Laplacian operator in four dimensions. The other calculations are similar. Pope gives the remaining terms:

$$\hat{R}_{ab} = e^{-2\alpha\phi}\left(R_{ab} + (\alpha\beta - 2\alpha^2)\partial_a\phi\partial_b\phi - \alpha\eta_{ab}\Box\phi\right) - \frac{1}{2}e^{2(\beta-2\alpha)\phi}\mathcal{F}_a{}^c\mathcal{F}_{bc} \tag{25}$$

$$\text{and } \hat{R}_{az} = \hat{R}_{za} = \frac{1}{2}e^{(\beta-\alpha)\phi}\nabla^b\left(e^{2(\alpha+\beta)\phi}\mathcal{F}_{ab}\right), \tag{26}$$

where R_{ab} is the curvature 2-form on $\mathcal{M}^{(4)}$, η_{ab} is the four-dimensional Minkowski metric, and where ∇^b is the covariant derivative on the coordinate-free basis [6].

Then the five-dimensional Ricci scalar comes from the trace of the curvature 2-form [6]:

$$\begin{aligned}\hat{R} &= \eta^{ab}\hat{R}_{ab} + \hat{R}_{zz} \\ &= e^{-2\alpha\phi}\left(\eta^{ab}R_{ab} + (\alpha\beta - 2\alpha^2)(\partial\phi)^2 - \alpha\eta^a{}_a\Box\phi\right) - \frac{1}{2}e^{2(\beta-2\alpha)\phi}\mathcal{F}^2 + \beta^2 e^{-2\alpha\phi}\Box\phi + \frac{1}{4}e^{2(\beta-2\alpha)\phi}\mathcal{F}^2 \\ &= e^{-2\alpha\phi}\left(R + (\alpha\beta - 2\alpha^2)(\partial\phi)^2 + (\beta^2 - 2\alpha)\Box\phi\right) - \frac{1}{4}e^{2(\beta-2\alpha)\phi}\mathcal{F}^2. \end{aligned} \tag{27}$$

2.1.3 Lagrangian Density

We would like to calculate the Lagrangian density for the theory. First,

$$\sqrt{-\hat{g}} = e^{(\beta+4\alpha)\phi}\sqrt{-g}. \tag{28}$$

So,[4]

$$\begin{aligned}\mathcal{L} &= \sqrt{-\hat{g}}\hat{R} \\ &= \sqrt{-g}e^{(\beta+2\alpha)\phi}\left[R + (\alpha\beta - 2\alpha^2)(\partial\phi)^2 + (\beta^2 - 2\alpha)\Box\phi\right] - \frac{1}{4}\sqrt{-g}e^{3\beta\phi}\mathcal{F}^2. \end{aligned} \tag{29}$$

We are now in a position to set α and β. We want our theory to include four-dimensional Einstein gravity, so the Lagrangian density better include a term that looks like $\sqrt{-g}R$. This tells us that we want [6]:

$$\beta = -2\alpha. \tag{30}$$

If we do this, our Lagrangian density becomes [6]:

$$\mathcal{L} = \sqrt{-g}\left[R - 6\alpha^2(\partial\phi)^2 + (4\alpha^2 - 2\alpha)\Box\phi - \frac{1}{4}e^{-6\alpha\phi}\mathcal{F}^2\right].$$

The term $-6\alpha^2(\partial\phi)^2$ is very reminiscent of the Klein-Gorden Lagrange density for a massless scalar field [6]. In the Klein-Gordon action, the term is $-\frac{1}{2}(\partial\phi)^2$ [6]. This informs our choice of α [6]:

$$\alpha = \frac{1}{\sqrt{12}}. \tag{31}$$

With α and β set, we can write down our final Lagrangian density. Since the Laplacian term $\Box\phi$, which gives a total derivative in $\mathcal{L}$ will have no effect on variational calculations, we drop it [6]:

$$\mathcal{L} = \sqrt{-g}\left[R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right]. \tag{32}$$

[4]For simplicity, we will work in units were $\frac{1}{16\pi G} = 1$.

From a five-dimensional theory *that contains only gravity*, we have constructed a four-dimensional Lagrangian density that includes Einstein gravity and a scalar field coupled to an object that looks an awful lot like the Maxwell tensor. Indeed, if we set $\phi = 0$, we would recover the Einstein-Maxwell Lagrangian density,

$$\mathcal{L} = \sqrt{-g}\left(R - \frac{1}{4}\mathcal{F}^2\right).$$

It would certainly be tempting to do so—we would get Einstein-Maxwell gravity and electromagnetism with no extra complications.[5] Unfortunately, this is forbidden [6]. To see why, let's work out the field equations.

2.1.4 Field Equations

For simplicity, we will assume that the five-dimensional spacetime has no boundary and thus all boundary terms will drop out. It is more convenient to use the variations $\delta g^{\alpha\beta}$ instead of $\delta g_{\alpha\beta}$ [14]. So,

$$\delta g_{\alpha\beta} = -g_{\alpha\mu}g_{\beta\nu}\delta g^{\mu\nu}. \tag{33}$$

We will also use the well-known variation of the determinant of the metric [14],

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\alpha\beta}\delta g^{\alpha\beta}. \tag{34}$$

Now, the variation of the action is

$$\begin{aligned}
\delta S_{kaluza} &= \int d^5x\delta\mathcal{L} \\
&= \int d^5x\delta\left[\sqrt{-g}\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)\right] \\
\Rightarrow \delta S_{kaluza} &= \int \sqrt{-g}\delta\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)d^5x \\
&\quad + \int\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)\delta\sqrt{-g}d^5x.
\end{aligned} \tag{35}$$

To make things simpler, we'll treat each integral separately. Let

$$\delta S_1 = \int \sqrt{-g}\delta\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)d^5x \tag{36}$$

$$\text{and } \delta S_2 = \int\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)\delta\sqrt{-g}d^5x. \tag{37}$$

The first integral becomes

$$\begin{aligned}
\delta S_1 &= \int \sqrt{-g}\delta\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2\right)d^5x \\
&= \int \sqrt{-g}\delta\left(g^{\mu\nu}R_{\mu\nu} - \frac{1}{2}g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - \frac{1}{4}e^{-\sqrt{3}\phi}g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma}\right)dx^5 \\
&= \int \sqrt{-g}\left[g^{\alpha\beta}\delta R_{\alpha\beta} + R_{\alpha\beta}\delta g^{\alpha\beta} - \frac{1}{2}(\partial_\mu\phi)(\partial_\nu\phi)\delta g^{\mu\nu} - \frac{1}{2}g^{\mu\nu}(\partial_\mu\phi)\partial_\nu\delta\phi\right]d^5x \\
&\quad -\frac{1}{4}\int\sqrt{-g}\left[-\sqrt{3}\mathcal{F}^2\delta\phi + 2se^{-\sqrt{3}\phi}F_{\mu\rho}F^{\nu\rho}\delta g^{\mu\nu} + 2e^{-\sqrt{3}\phi}F^{\mu\nu}\delta F_{\mu\nu}\right]d^5x.
\end{aligned} \tag{38}$$

[5]Historically, many people *did* set $\phi = 0$—likely because they distrusted scalar fields [2, 7]. However, neither Kaluza nor Klein made this simplification [2, 4, 13, 6].

Similarly, if we use equation 34, the second integral becomes

$$\begin{aligned} \delta S_2 &= \int \left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2 \right) \delta\sqrt{-g} d^5x \\ &= -\frac{1}{2}\int \sqrt{-g}\left(R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{4}e^{-\sqrt{3}\phi}\mathcal{F}^2 \right) g_{\alpha\beta}\delta g^{\alpha\beta} d^5x. \end{aligned} \tag{39}$$

If we combine δS_1 and δS_2, group variational variable and set the variation of the action to zero, we have the following equations:

$$0 = \int \sqrt{-g} g^{\mu\nu}\delta R_{\mu\nu} d^5x. \tag{40}$$

$$0 = -\frac{1}{2}\int \sqrt{-g} e^{-\sqrt{3}\phi} F^{\mu\nu}\delta F_{\mu\nu} d^5x \tag{41}$$

$$0 = \int \sqrt{-g}\left[\frac{1}{4}\sqrt{3}\mathcal{F}^2 - \frac{1}{2}(\partial^\mu\phi)\partial_\mu \right] \delta\phi d^5x \tag{42}$$

$$0 = \int \sqrt{-g}\left[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} - \frac{1}{2}\left((\partial_\mu\phi)(\partial_\nu\phi) - \frac{1}{2}(\partial\phi)^2 g_{\mu\nu} \right) - \frac{1}{2}\left(\mathcal{F}^2_{\mu\nu} - \frac{1}{4}\mathcal{F}^2 g_{\mu\nu} \right) \right] \delta g^{\mu\nu} d^5x, \tag{43}$$

where

$$\mathcal{F}^2_{\mu\nu} = \mathcal{F}_{\mu\rho}\mathcal{F}_\nu{}^\rho. \tag{44}$$

Equation 40 becomes an integral over the boundary and vanishes [14]. The remaining three integrals define our field equations [6, 7, 2, 4, 13, 15, 10, 5]:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{1}{2}\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}(\partial\phi)^2 g_{\mu\nu} \right) + \frac{1}{2}\left(\mathcal{F}^2_{\mu\nu} - \frac{1}{4}\mathcal{F}^2 g_{\mu\nu} \right), \tag{45}$$

$$0 = \nabla^\mu \left(e^{-\sqrt{3}\phi}\mathcal{F}_{\mu\nu} \right), \tag{46}$$

$$\text{and } \Box\phi = -\frac{3\sqrt{3}}{2}e^{-\sqrt{3}\phi}\mathcal{F}^2. \tag{47}$$

People sometimes eliminate the $-\frac{1}{2}Rg_{\mu\nu}$ term in equation 45 by subtracting the appropriate multiple of the trace to get [6]:

$$R_{\mu\nu} = \frac{1}{2}\partial_\mu\phi\partial_\nu\phi + \frac{1}{2}e^{-\sqrt{3}\phi}\left(\mathcal{F}^2_{\mu\nu} - \frac{1}{4}\mathcal{F}^2 g_{\mu\nu} \right). \tag{48}$$

This is looking pretty familiar! If we let the energy momentum tensor be

$$T_{\mu\nu} = \frac{1}{2}\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}(\partial\phi)^2 g_{\mu\nu} \right) + \frac{1}{2}\left(\mathcal{F}^2_{\mu\nu} - \frac{1}{4}\mathcal{F}^2 g_{\mu\nu} \right), \tag{49}$$

then equation 45 simply represents the field equations from Einstein gravity [12]. Similarly, although there's a re-scaling by ϕ, equation 46 looks very reminiscent of the source-free Maxwell equations in tensor notation [12]. (The other equation from electromagnetism, $\mathcal{F} = d\mathcal{A}$, is encoded in our definition of $\mathcal{F}$.) Equation 47 is new, though. This is the source term for the scalar field.

Now we can see why $\phi = 0$ is forbidden in general. Because the scalar and electromagnetic fields *interact*, $\phi = 0$ if and only $\mathcal{F}^2 = 0$. In other words, if $\phi = 0$, the entire system just reduces to four-dimensional Einstein gravity with no electromagnetic field at all. Since we can't get rid of ϕ, let's give it a name. We call it the *dilaton* [6].

2.2 Symmetries

Out of a five-dimensional, purely gravitational system, we seem constructed the field equations for a four-dimensional system that includes gravity, electromagnetism, and a scalar field. If this is truly the case, however, the symmetries of the five-dimensional spacetime have to translate into the symmetries of a four-dimensional spacetime and the gauge-invariances of the vector potential and scalar field [6, 7]. Let's see if we can understand how these symmetries appear. As before, this derivation closely follows that of Pope [6].

The five-dimensional Einstein theory is invariant under five-dimensional general coordinate transformations, which can be written infinitesimal form as

$$\delta\hat{x}^M = -\hat{\xi}^M(\hat{x}), \tag{50}$$

where $\hat{\xi}^M$ is an arbitrary function of all five coordinates [6]. This induces a change in the five-dimensional metric [6]:

$$\delta\hat{g}_{MN} = \hat{\xi}^P\partial_P\hat{g}_{MN} + \hat{g}_{PN}\partial_M\hat{\xi}^P + \hat{g}_{MP}\partial_N\hat{\xi}^P. \tag{51}$$

However, Kaluza's cylinder condition is much more restrictive. It enforces that our coordinate system be Clairaut and that ∂_4 be a Killing vector (see section 2.1.1). Thus, in Kaluza's theory, our "arbitrary" function ξ^A takes on a particular form:

$$\hat{\xi}^\mu(\hat{x}) = \xi^\mu(x) \text{ and } \hat{\xi}^z(x^A) = cx^4 + \lambda(x), \tag{52}$$

where $\xi^\mu(x)$ is a general coordinate transformation on $\mathcal{M}^{(4)}$, $c \in \mathbb{R}$ is a constant, and $\lambda(x)$ is an arbitrary differentiable function on $\mathcal{M}^{(4)}$ [6]. These are the coordinate transformations allowed in the five-dimensional theory.

We can calculate the effects of an allowed coordinate on the four-dimensional metric, vector potential, and dilaton field by plugging the allowed transformations into equation 7. As it turns out, almost all of the four-dimensional symmetries we're looking for can be attained with $c = 0$. For simplicity, we will discuss this case first. Then we'll discuss the symmetry embodied by c.

First, let's look at $\hat{g}_{44}$:

$$\begin{aligned} \delta\hat{g}_{44} &= \hat{\xi}^P\partial_P\hat{g}_{44} + \hat{g}_{P4}\partial_4\hat{\xi}^P + \hat{g}_{4P}\partial_4\hat{xi}^P = \xi^\rho\partial_\rho\hat{g}_{44} + 2c\hat{g}_{44} \\ \Rightarrow \delta e^{2\beta\phi} &= \xi^\rho\partial_\rho e^{2\beta\phi} \\ \Rightarrow \delta\phi &= \xi^\rho\partial_\rho\phi. \end{aligned} \tag{53}$$

This implies that ϕ does indeed transform as a scalar under four-dimensional general coordinate transformations. We also see the first hints of gauge symmetry: ϕ is independent λ, which will be the gauge choice for our vector potential. These are the symmetries we would expect for a scalar field [6].

Similarly,

$$\begin{aligned} \delta\hat{g}_{\mu 4} &= \xi^\rho\partial_\rho\hat{g}_{\mu 4} + \hat{g}_{\rho 4}\partial_\mu\xi^\rho \\ \Rightarrow \delta e^{2\beta\phi}\mathcal{A}_\mu &= \xi^\rho\partial_\rho e^{2\beta\phi}\mathcal{A}_\mu + e^{2\beta\phi}\mathcal{A}_\rho\partial_\mu\xi^\rho \\ \Rightarrow \delta\mathcal{A}_\mu &= \xi^\rho\partial_\rho\mathcal{A}_\mu + \mathcal{A}_\rho\partial_\mu\xi^\rho + \partial_\mu\lambda(x). \end{aligned} \tag{54}$$

The $\xi^\rho\partial_\rho\mathcal{A}_\mu$ and $\mathcal{A}_\rho\partial_\mu\xi^\rho$ terms are appropriate for covector on $\mathcal{M}^{(4)}$, while the $\partial_\mu\lambda$ term reflects the appropriate symmetry for a $U(1)$ gauge field with gauge parameter λ. All this tells us that $\mathcal{A}_\mu$ behaves appropriately for the electromagnetic vector potential [6].

Finally,

$$\begin{aligned} \delta\hat{g}_{\mu\nu} &= \xi^\rho\partial_\rho\hat{g}_{\mu\nu} + \hat{g}_{\rho\nu}\partial_\mu\xi^\rho + \hat{g}_{\mu\rho}\partial_\nu\xi^\rho + \hat{g}_{4\nu}\partial_\mu\xi^4 + \hat{g}_{\mu 4}\partial_\nu\xi^4 \\ \Rightarrow \delta g_{\mu\nu} &= \xi^\rho\partial_\rho g_{\mu\nu} + g_{\mu\nu}\partial_\mu\xi^\rho + g_{\mu\rho}\partial_\nu\xi^\rho. \end{aligned} \tag{55}$$

This is, of course, the appropriate change in the metric under a general coordinate transformation. The metric is thus independent of the gauge parameter λ, as it ought to be if $\mathcal{A}_\mu$ is Maxwell's vector potential [6].

With $c = 0$ we've found all the symmetries of an Einstein-Maxwell system: covariance with a general coordinate transformation and gauge invariance under λ. However, we're still missing the symmetry associated with a scalar field; only the change in the dilaton should matter, not its overall value. This is the constant shift symmetry [6]. To see how this symmetry appears, we need to discuss one further symmetry of the five-dimensional theory.

In addition to general coordinate transformations, we can rescale the five-dimensional metric by a positive[6] constant factor [6],

$$\hat{g}_{MN} \to \kappa^2 \hat{g}_{MN},\ \kappa \in \mathbb{R}. \tag{56}$$

The Riemann tensor is inert under this transformation. However, when we contract its indices to form the Ricci scalar, we pick up one over the scale factor from the inverse metric [6]:

$$\hat{R} \to \kappa^{-2} \hat{R}. \tag{57}$$

This rescales the Lagrangian density by a constant factor. However, we can divide it away when we set the variation of the action to zero and it has no effect on the field equations [6].

Now let's see what effect this rescaling has on the lower-dimensional theory. We can write this symmetry infinitesimally as

$$\delta \hat{g}_{MN} = 2a\hat{g}_{MN}, \tag{58}$$

where a is some constant parameter. Let's group this symmetry with the symmetry induced in the metric by a nonzero c:

$$\delta \hat{g}_{MN} = c\delta^4_M \hat{g}_{4N} + c\delta^4_N \hat{g}_{M4} + 2a\hat{g}_{MN}, \tag{59}$$

where δ^A_B is the Kronecker delta [6]. If we plug the five-dimensional line element (7) into this, we find that

$$\beta\delta\phi = a + c,\ \delta\mathcal{A}_\mu = -c\mathcal{A}_\mu,\ \text{and}\ \delta g_{\mu\nu} = 2ag_{\mu\nu} - 2\alpha g_{\mu\nu}\delta\phi. \tag{60}$$

This means that, if we rescale the five-dimensional metric, we can allow c to behave as the dilaton shift and keep the four-dimensional metric inert [6]. To do this, we enforce that our five-dimensional metric rescales as

$$a = -\frac{c}{3}. \tag{61}$$

If we rescale when we choose a new set of general coordinates, then the c transformation becomes [6]:

$$\delta\phi = -\frac{2\sqrt{3}}{3}c,\ \delta\mathcal{A}_\mu = -c\mathcal{A}_\mu,\ \text{and}\ \delta g_{\mu\nu} = 0. \tag{62}$$

This is exactly what we'd expect for a dilaton shift of $-2\sqrt{3}c/3$. Thus, with the additional constraint on metric rescaling, the system is symmetric under this transformation too [6].

So, if we impose Kaluza's cylinder condition on five-dimensional Einstein gravity, we recover field equations consistent with a four-dimensional system that contains Einstein gravity, Maxwell electromagnetism, and a scalar field. Furthermore, transformations in the higher-dimensional space that preserve the cylinder condition and the five-dimensional field equations result in the appropriate symmetries in the lower-dimensional spacetime.

3 Compactification and Other Tricks

Even with all the success of Kaluza's theory, the cylinder condition is a bit hard to swallow. With no justification, it seems rather contrived. Indeed, if we impose an isometry on the higher-dimensional theory, it is perhaps not surprising that this manifests as a gauge symmetry in the lower dimensional theory. Both

[6] In theory, the scale factor could be negative. However, this would reverse our sign convention for spacelike and timelike vectors.

the higher-dimensional isometry and the lower-dimensional gauge symmetry are manifestations of a one-dimensional Lie symmetry group, after all. It would therefore be nice if we could somehow justify Kaluza's mysterious condition.

The first successful attempt to justify the cylinder condition comes from Klein [2, 7, 13, 15], and this is where the name "Kaluza-Klein theory" originates. Klein discovered that, if the fifth dimension is a circle, the circumference can be made so small that any dependence of the metric on x^4 is unobservable [2, 7, 13, 15]. In modern notation, we say Klein's ansatz is that

$$\mathcal{M}^{(5)} = \mathcal{M}^{(4)} \times U(1), \tag{63}$$

where $U(1)$ is the circle group [6, 7]. Thus the fifth dimension is *compact* and we call Klein's scheme *compactification* [7]

With this topology for the five-dimensional spacetime, the metric must be periodic in x^4. This means that we can Fourier expand the five-dimensional metric as

$$\hat{g}_{MN}(x, x^5) = \sum_{n=-\infty}^{\infty} g_{MN}^{(n)}(x) e^{inx^4/L}, \tag{64}$$

where x are the coordinates of $\mathcal{M}^{(4)}$, n is the Fourier mode number, and L is the length scale of $U(1)$ [6]. As usual, higher-$|n|$ modes correspond to higher energy scales. In fact, the $n = 0$ modes correspond to massless fields, while the $n \neq 0$ modes correspond to massive fields [6]. This is easiest to see by studying a simpler model, elegantly explained by Pope [6].

Suppose we have a massless scalar field on $\mathcal{M}^{(5)}$, call it $\hat{\sigma}$.[7] Thanks to the Klein-Gordon equation, we have that

$$\hat{\Box}\hat{\sigma} = 0, \tag{65}$$

where $\hat{\Box} = \partial^M \partial_M$ [6]. If we Fourier expand $\hat{\sigma}$, we find that

$$\hat{\sigma}(x, x^5) = \sum_{n=-\infty}^{\infty} \hat{\sigma}_n(x) e^{inx^4/L}, \tag{66}$$

as with the metric [6]. If we apply the five-dimensional Lapplacian operator, we find that the each four-dimensional field satisfies

$$\Box \sigma_n = \frac{n^2}{L^2} \sigma_n, \tag{67}$$

which is the Klein-Gordon equation for a scalar field with mass $|n|/L$ [6]. Similar reasoning holds for the Fourier expansion of the metric tensor.

Klein saw the parallels between the periodicity of the compactified dimension and the azimuthal direction in an electron's orbit around a nucleus [2, 13, 15]. He rather ambitiously hoped to use the $n \neq 0$ modes to explain the quantization of electric charge in the same way that periodic boundary conditions explain the energy levels in an atom [2, 13, 15]. In Klein's scheme, the n^{th} harmonic of the metric holds an "electric" charge of

$$e_n = n \frac{2\sqrt{\pi G}}{\pi L},$$

where G is Newton's constant [2, 13, 15]. This would mean that the charge of an electron is

$$e = \frac{2\sqrt{\pi G}}{\pi L}.$$

We can choose L to match the observed electron charge, and it turns out to be about one hundred Planck lengths [2]. In this scheme, the cylinder condition holds only approximately, and it would break down at the

[7] We use $\hat{\sigma}$ to avoid confusing this field with the dilaton ϕ.

appropriate length scale [2, 7].[8] String theory and M-theory use a similar compactification approach, but with more compactified dimensions [7, 6].

If we would prefer the cylinder condition to be absolute, we can set L to be even smaller, on the order of one Planck length, so that the massive modes have such high energies that they can never be observed. In this case the $n = 0$ mode is the only observable part of the metric, and it satisfies cylinder condition [6]. Klein's conserved charge ensures we can do this [13, 15, 2, 6]. Zero charge can never spontaneously become nonzero charge [6].

An easy way to understand this is that the Fourier basis functions $e^{inx^4/L}$ are representations of $U(1)$ [6]. The $n \neq 0$ mode functions are paired off by complex conjugates. The function corresponding to mode number n is the complex conjugate of the function corresponding to mode number $-n$. The $n = 0$ mode function is thus a representation of the trivial subgroup of $U(1)$. Since it's a subgroup, it's closed and no amount of multiplying $n = 0$ modes together will cause us to jump to an $n \neq 0$ mode [6].

Although Klein's compactification scheme is by far the most popular way to justify the cylinder condition, it is not the only option. $(3+1)$-dimensional projective space can be thought of as a hypersurface in $(4+1)$-dimensional Minkowski space. In 1930, Veblen and Hoffmann used this idea to justify Kaluza's fifth dimension and the cylinder condition, where the fifth dimension actually reflects the projective nature of spacetime [16]. More than thirty years later, Joseph discussed this possibility as well [17].

People have also argued that the cylinder condition is only an approximation, even on accessible energy scales. The biggest advocate of this approach is Wesson, who has explored some of the cosmological implications of such a scheme [18, 19, 8, 9].

4 Concluding Remarks

From a five-dimensional theory of pure gravity, we've attained a four-dimensional system containing gravity, electromagnetism, and a scalar field. Although Kaluza's cylinder condition appears contrived at first, Klein and others have demonstrated that it can be justified. This is a remarkable accomplishment, and one that has garnered a great deal of interest over the years.

Of course, Kaluza-Klein theory has some clear weaknesses. Although the theory predicts the massless scalar dilaton, we haven't observed any such particle. Since it's coupled to the electromagnetic field, we would have observed it if it existed. Additionally, the theory gets much more complicated when we add the strong and electroweak forces. Each force adds at least one dimension subject to its own "cylinder condition," and the whole theory becomes highly nontrivial.

Despite these difficulties, Kaluza-Klein theory remains seductive. Kaluza's geometric approach to unification appeals to our aesthetic sensibilities and Klein's contribution, the compactification, offers the tantalizing possibility of explaining quantum mechanics geometrically—the whole thing hints at a theory of quantum gravity. This is certainly why string theory and M-theory use compactified dimensions. It is highly likely that this beautiful theory will continue to inspire new ideas in and new approaches to gravitational physics.

Acknowledgments

I'd like to thank Michael Schmidt for the proof-reading he performed. I'd also like to thank Professor Oliver DeWolfe for suggesting I start my foray into Kaluza-Klein Theory with Overduin and Wesson [7] and Pope [6].

[8] Although this energy scale is very much inaccessible at the moment... and probably forever.

References

[1] H. Minkowski. Die grundgleighungen gür die elektromagnetischen vorgänge in bewegten körpern. *Nachrichten Von Der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 21:53–111, 1908.

[2] T. Appelquist, A. Chodos, and P. G.O. Freund, editors. *Modern Kaluza-Klein Theories*. Frontiers in Physics. Addison-Wesley, 1987.

[3] G. Nordström. On the possibility of a unification of the electromagnetic and gravitational fields. *Helsingfors*, page 504, March 1914. [Translation in [2]].

[4] T.H. Kaluza. On the unity problem in physics. *Sitzungsberichte der K. Preussischen Akademie der Wissenschaften zu Berlin*, page 966, December 1921. [Translation in [2]].

[5] A. Einstein and P. Bergmann. On a generalization of kaluza's theory of electricity. *Annals of Mathematics*, 39(3):683, April 1938.

[6] C. Pope. Kaluza-klein theory. `http://people.physics.tamu.edu/pope/ihplec.pdf`. [Online; accessed 16-April-2013].

[7] J. P. Overduin and P.S. Wesson. Kaluza-klein gravity. *Physics Reports*, 283(5–6):303–378, April 1997.

[8] P.S. Wesson. *Five-Dimensional Physics*. World Scientific, 2006.

[9] P.S. Wesson. *Space-Time-Matter*. World Scientific, 2007.

[10] Y. Thiry. The equations of kaluza's unified theory. *Comptes Rendus (Paris)*, 226:216, 1948. [Translation in [2]].

[11] J. Oprea. *Differential Geometry and Its Applications*. The Mathematical Association of America, 2007.

[12] R.M. Wald. *General Relativity*. University of Chicago Press, 1984.

[13] O. Klein. Quantum theory and five dimensional theory of relativity. *Z.F. Physik*, 37:895, April 1926.

[14] E. Poisson. *A Relativist's Toolkit: The Mathematics of Black Hole Mechanics*. Cambridge University Press, 2004.

[15] O. Klein. The atomicity of electricity as a quantum theory law. *Nature*, 118:516, 1926.

[16] O. Veblen and B. Hoffmann. Projective relativity. *Physical Review*, 36(5):810–822, 1930.

[17] J.T. Joseph. Coordinate covariance and the particle spectrum. *Physical Review*, 126(1):319–323, 1962.

[18] P.S. Wesson. A new approach to scale-invariant gravity. *Astronomy and Astrophysics*, 119:145–152, 1983.

[19] P.S. Wesson. An embedding for general relativity with variable rest mass. *General Relativity and Gravitation*, 16(2):193–203, 1984.