Fully covariant spontaneous compactification

Tom Lawrence

December 20, 2018

Abstract

We present a geometric field theory in which the Lagrangian has full general covariance in a higher-dimensional spacetime. Covariant constraints on a vector field cause the extra dimensions to compactify spontaneously. Changes of coordinates induce transformations under which the values of the covariant derivative of the vector form orbits. Constraints can be chosen which pick out orbits with a particular isotropy group. Tensors, including the metric, decompose into representations of this isotropy group. Consequently, the space can compactify into one isometric to a Cartesian product of spaces with geodesically complete factors, or into a generalisation of this.

We use this approach to construct the simplest possible generalisation of the Poisson equation for gravity consistent with general covariance and the equivalence principle. This can be cast into a form involving the Ricci tensor and metric, but unlike in General Relativity, these act as matrix operators on the vector field.

We show that for particular orbits, the field equations result in products of Einstein manifolds and that one solution is the product of Minkowski space and a two-sphere. There are natural coordinate bases to use in these product spaces. However, transforming to a generic basis introduces new terms in the Levi-Civita connection, including gauge potentials which represent 'fictitious forces'. In coordinate bases close to the natural ones, these potentials appear in the metric. It seems likely that with additional curvature caused by introducing matter fields, the fictitious forces would become real ones, representing gravity and other forces induced by the matter.

1 Introduction

1.1 Context

The current frontier of our understanding of how our universe works is based on two types of theory.

One is quantum field theory, which describes the working of the strong and electroweak interactions. In this framework, gauge potentials are coupled to matter fields by replacing partial derivatives with covariant derivatives.

The other is General Relativity (GR), describing gravity, which could be termed a 'geometric field theory'. In GR, there are also covariant derivatives, containing the Levi-Civita connection. This connection transforms in a similar way to gauge potentials. The depth of these parallels was first brought to light by Utiyama[1] in 1955.

In light of this, vast effort has gone into developing a quantum field theory of gravity. The problems with this are well known. Much of the motivation for this has been the success of the quantum field theory framework, both in its predictive power and in unifying interactions. For example, it has allowed us to understand that electromagnetism and the weak interaction are facets of a single electroweak interaction.

GR, however, is a very different type of theory. Its symmetries are not internal symmetries relating components of field multiplets, but the symmetries of four-dimensional spacetime. Moreover, its basic concepts are geometrical ones of connections and curvature, which are not manifest in the quantum field theories.

It seems far from obvious that a unified framework for handling both gravity and other interactions should be based on quantum field theory. It must contain quantum field theories and GR as limiting cases, but why should it not be a geometrical field theory or indeed some other kind of theory, as yet unknown?

Indeed, the first 'unification theories' were classical field theories. Maxwell's electromagnetism unifies magnetic and electrical effects in a single theory. While quantum mechanics was still being developed, Nordström[2, 3, 4], Kaluza[5] and Klein[6] proposed theories which sought to unify gravity and electromagnetism. These were based on adding an extra spatial dimension to the four-dimensional spacetime of relativity - further detail is given in the appendix.

Kaluza-Klein theory succeeded in reproducing the Einstein-Hilbert-Maxwell action, although its method for explaining charge quantisation was not consistent with observed masses[7]. Furthermore, after the Second World War, attention turned to the strong and weak nuclear interactions. The quantum field theory framework was found to be admirably suited to describing these. Theories of spontaneous symmetry breaking were developed within this framework. Initially, this work focused on global symmetries (see appendix), but this was extended to local symmetries[8, 9, 10, 58, 11], allowing electromagnetism to be unified with the weak interaction[12, 13] and then the strong interaction[14].

A further issue with Kaluza-Klein theory was that it provided no reason for one of the dimensions to be compact. This was addressed when the mechanism of spontaneous compactification was proposed in the late 1970s. The early papers utilised a scalar potential, based on the techniques of spontaneous symmetry breaking[15, 16]. By contrast, some papers in the early 1980s triggered compactification using a gauge field[17, 18, 19]. Those which are most relevant to this paper are described briefly in the appendix.

However, these papers sought to incorporate 'internal' interactions, which are described in terms of unitary symmetries. These are not symmetries of spacetime. The geometry to describe the gauge fields for these interactions is one based on fibre bundles. Therefore research focused on mechanisms which would result in these fibre bundle structures after compactification. Furthermore, O'Raifeartaigh's no-go theorem[20] limited the way in which these symmetries could be unified with those of the Poincaré group. O'Raifeartaigh looked at ways in which the Lie algebras of the Poincaré group and an internal symmetry could be embedded in a larger Lie algebra. He found that 'none of these (except the direct sum) seems to be particularly attractive from the physical point of view'. This conclusion was strengthened for quantum field theories by the Coleman-Mandula no-go theorem[21]. The only known exception to this was supersymmetry (which in addition provided a solution to the hierarchy problem), so most of the research either includes supersymmetry and supergravity explicitly or is constructed in such a way as to facilitate supersymmetric extensions.

1.2 Approach and structure of this paper

We take a different approach in this paper. Unitary groups can be defined in terms of the inner products of complex vectors. Unitary symmetries therefore arise naturally when considering spinor multiplets. They do *not* appear in the transformation properties of the tensor fields of GR. However, the outer product of a spinor and its conjugate or adjoint can always be decomposed into tensor fields. A unitary transformation of the spinor and its conjugate or adjoint induces an orthogonal or pseudo-orthogonal transformation of the vector component of the outer product. Such transformations are isometries on the tangent space, and are amongst the general linear transformations induced by changes of coordinates. We therefore take the approach that a geometric unified field theory should include vectors of orthogonal groups related to internal symmetries.

Unlike the theories based on fibre bundles, we start by asking what spontaneous compactification would look like if the extra dimensions are real spacetime dimensions on exactly the same footing as the four familiar ones.

For an N-dimensional theory, bases on a given tangent space are related by elements of a group isomorphic to $GL(N, \mathbb{R})$. The orthonormality of a frame basis is preserved by a subgroup isomorphic to its maximal pseudo-orthogonal subgroup. Considering the transformation relating a coordinate basis to a frame basis naturally leads one to the Weitzenböck connection of teleparallelism, but actually it turns out to be the Levi-Civita connection and the associated covariant derivative of a vector which are crucial to the spontaneous compactification described here. These connections and symmetries are studied in Section 2. It utilises a coset space decomposition of the general linear group with respect to its maximal pseudo-orthogonal subgroup. (An understanding of the Weitzenböck connection is not required in order to understand Sections 3 to 6. However, it may help researchers in this area avoid falling into potential traps and it may be essential for exploration of some of the issues raised in Section 7.)

Having established these foundations, we turn in Section 3 to finding a mechanism whereby the general linear symmetry is spontaneously broken to a direct product group, where one factor is the $GL(4, \mathbb{R})$ of four-dimensional spacetime.

This factor contains the familiar Lorentz group. To find this, we utilise the theoretical framework of non-linear realisations (see the appendix for references on this). In such models, symmetry breaking is triggered by constraints which are invariant under the full non-linearly realised group. In our case, we seek constraints which are invariant under the general linear group of the higher-dimensional spacetime. In contrast to other Kaluza-Klein theories, we demand that the constraints are built purely out of tensors of the higher-dimensional spacetime.

We show that an appropriate tensor for this is the covariant derivative of a vector. Changes of coordinates induce the general linear group to act on it by conjugation. Under this action, values of the covariant derivative form orbits which have the same eigenvalues. These can be grouped into strata, according to their isotropy groups. These groups are isomorphic to direct products of general linear groups. For appropriate choices of constraints, one factor of the isotropy group is $GL(4, \mathbb{R})$.

This has an impact on the topology of the spacetime. Denote coordinates on the four-dimensional spacetime y^μ and those on the additional dimensions y^X. In these coordinates, the metric factorises into $\mathrm{g}_{\mu\nu}$ and g_{XY}. If $\mathrm{g}_{\mu\nu}$ is independent of y^X and g_{XY} is independent of y^μ, then the pseudo-orthonormal part of the isotropy group represents the holonomy group of the Levi-Civita connection. In this case, the manifold becomes a direct product manifold with geodesically complete factors, analogous to the original Kaluza-Klein theory. If this is not the case, it appears that the manifold is still homeomorphic to such a direct product manifold.

To determine the local geometry of the manifold, we then need a field equation. This is the subject of Section 4. Our approach here is to remain faithful to the overall philosophy of GR, but to build up the field equation starting from Newtonian gravity. In Newtonian gravity, there is a simple relation between the gravitational potential and the resulting three-acceleration of a test particle. We start in a coordinate system in which this physics is valid and find the three-acceleration of a particle in free fall, according to the kinematics of GR. The components of this are components of the Levi-Civita connection. The simplest tensor containing the Levi-Civita connection is again the covariant derivative of a vector field. We use this to find a fully covariant equivalent of the Laplace equation and the Lagrangian that results in it. By giving the vector field a mass, we find the equivalent of the Poisson equation, which we take as the field equation. We show that this can be rewritten to include the Ricci tensor and metric.

This field equation is not that of GR but nonetheless for particular orbits results in the manifold compactifying such that internal space is an Einstein manifold while four-dimensional spacetime is flat. All tensor field multiplets, including the metric, then decompose into tensors of the submanifolds.

In Section 5 we show, by way of example, that when $N = 6$, the internal space can be a two-sphere. For this case, we also state the condition on the covariant derivative which causes compactification, both as an externally-imposed constraint, and as a symmetry-breaking potential. We find a solution to the constraint equations and show that this satisfies the field equation. The two-sphere has an $SO(3)$ symmetry, but only symmetries under its $SO(2)$ subgroup are manifest in the solution. We point out some issues to be considered if researchers seek to extend this to higher numbers of dimensions.

Most of this analysis is carried out in the y-coordinates. However, more general coordinate systems exist, in which the coordinates are all functions of *both* y^μ and y^X. In Section 6 we show that on transforming to such a coordinate system, additional terms appear in the Levi-Civita connection. These include a multiplet which transforms under the full $GL(N,\mathbb{R})$ symmetry as a Levi-Civita connection for Minkowski space and one which transforms as a gauge potential of the internal symmetry. These represent 'fictitious forces' resulting from our peculiar choice of coordinates. Besides these, there are multiplets which transform as charged vectors and tensors - it is not yet clear what these represent. In a coordinate basis close to that for the y-coordinates, these new multiplets appear in the metric, a feature this theory has in common with existing Kaluza-Klein theories[6, 22, 23].

Section 7 presents and discusses our conclusions. It highlights key insights provided by the theory and novel aspects of the model. It describes potential areas of future research and where we might expect this research to diverge further from the existing literature.

As referred to above, the appendix summarises some aspects of research into Kaluza-Klein theories, spontaneous compactification, non-linear realisations and spontaneous symmetry breaking which are particularly relevant to this paper.

In this paper, we do not attempt to include additional matter fields - it looks likely that could be a non-trivial issue. However, we would expect that in keeping with the precepts of GR, the inclusion of such fields would cause additional curvature in the spacetime. This would perturb the space away from that resulting from the field equations and constraints presented in this paper. Remote from such matter, we would expect the geometry to tend towards the solutions described below. But where the matter is denser, we would expect the Levi-Civita connection to have a non-zero field strength associated with all of its degrees of freedom. It seems reasonable to suggest that this would be observed as gravity and internal gauge fields associated with the additional matter.

We also do not attempt to quantise the model. Again, this could be a non-trivial issue, and as remarked above, it is not even clear that trying to force quantisation on this theory would be the appropriate thing to do.

Despite these limitations on the scope of this paper, this theory has several advantages over existing Kaluza-Klein theories:

- it is fully covariant under changes of coordinate on the higher dimensional spacetime, with no multiplets of lower-dimensional symmetries appearing in the Lagrangian;
- it therefore explains how the symmetries of the compactified spacetime arise;

- it clarifies the relations between the holonomy, the isometries and the gauge symmetries;
- it clarifies the geometric meaning of the gauge potentials;
- it offers an insight into how the geometry could vary in response to the inclusion of matter fields;
- it makes it clear that it is too simplistic to say that compactified spacetime is the low energy state of the system, and that further investigation is needed before making any definitive statement about the energy of this system.

1.3 Geometric notation

As mentioned above, all dimensions in this paper are considered on an equal footing (other than whether they are timelike or spacelike), so we have no need to use the conceptual structures or language of fibre bundles. Given that all group operations are induced by changes of coordinate, the language of general relativity is much more suitable, and this is how we proceed.

While we recognise that the properties of the tangent space and its dual allow us to identify bases on the tangent space with differential operators, we do not make use of such operators. Rather, we denote basis vectors throughout using the more abstract notation of a bold letter, such as $\mathbf{e}_K$, which physicists will usually be more familiar with. When the basis is orthonormal, we denote this with a hat, for example $\hat{\mathbf{n}}_K$.

We work with a variety of frame and coordinate bases which are all related by group operations. Consequently, we do not distinguish between frame and coordinate bases by using different indices, as this would obscure the group symmetries (and we would have difficulty finding enough sets of letters!).

Rather, we use different sets of indices to distinguish between different spacetimes dimensionalities. We will use capital Latin indices around the middle of the alphabet, I, J, K, L, M, N, to denote coordinates on the full N-dimensional spacetime. We will use Greek indices, μ, ν, ρ, σ, to denote coordinates on four-dimensional spacetime. (We avoid τ and λ which are reserved for proper time and a curve parameter.) Where we need to deal with just the spatial indices in this latter spacetime, we will use lower case Latin letters, i, j. Where $N > 4$ and we are considering the four-dimensional subspace and the extra dimensions separately, we will use capital Latin indices around the end of the alphabet, W, X, Y, Z to denote coordinates on the extra dimensions.

Where we need to specify which coordinate system a set of tensor components is in, we will do so by putting it in brackets in a superscript or subscript. For example, the components of a vector $\mathbf{V}$ in a coordinate system u'^M will be written $V^M_{(u')}$.

2 Tangent space symmetries: connections and coset decomposition

The theory presented in this paper started life as an attempt to apply the theory contained in Salam and Strathdee[11] to Utiyama's description of gravity[1], within the context of spontaneous compactification. However, Utiyama's description of gravity utilises a connection known as the Weitzenböck connection, and the relation between this and the Levi-Civita connection of GR has only recently been understood, as we explain below.

The basic variables in the gravitational sector of GR are the independent components of the metric. In an N-dimensional spacetime, there are $\frac{N(N+1)}{2}$ of these. Another key quantity is the connection - general relativity uses the Levi-Civita connection, which is uniquely defined on a given spacetime for a given coordinate system.

Einstein was keen to extend GR to incorporate electromagnetism in a geometric way. He tried a number of different approaches to this[24]. One approach was "Fernparallelismus", often called "distant parallelism" or "teleparallelism"[25]. He noted that an N-bein field has N^2 independent components. The components of the metric can be written as functions of these, but in addition there are $\frac{N(N-1)}{2}$ degrees of freedom contained in the N-bein field which describe invariances of the metric[26]. His idea was that these additional degrees of freedom could be used in describing electromagnetism. In defining an N-bein field across the spacetime manifold, he needed to use a new type of connection, which he discovered had already been investigated by Cartan, Weitzenböck and others[25, 27].

While this approach was unsuccessful in its aim, research into teleparallelism and its application to gravity has continued. It is now known that a theory of gravity can be based on the principles of teleparallelism which reproduces the field equations of general relativity. This is known as the Teleparallel Equivalent of General Relativity (TEGR). An excellent summary of the current state of knowledge can be found in a review by Pereira[28]; additional detail on aspects of this can be found in another review by Maluf[29].

The N-bein field components are the elements of a matrix transformation which maps a chosen frame basis into a coordinate basis, as described below. Such a matrix is an element of a general linear group. Furthermore, the invariances of the metric described by Einstein form a pseudo-orthogonal group, which is a subgroup of the general linear group. The general linear group can be partitioned into cosets of the pseudo-orthogonal group and this leads to a natural decomposition of the change of frame.

This section looks at the relationship between these connections and these tangent space symmetries, utilising the decomposition of the general linear group.

2.1 The tangent space at a point

General relativity treats spacetime as a curved pseudo-Riemannian manifold, that is, one which approximates to flat spactime at each point. This allows one to define a tangent space at each point, the elements of which are vectors. By taking outer products of the tangent spaces and their duals, one can define tensors of higher rank.

On a curved pseudo-Riemannian manifold, one needs to use curvilinear coordinates to parametrise a finite region of it. For a given region of any given manifold, there are an infinite number of possible curvilinear coordinate systems that could be used. General relativity is constructed to be generally covariant, making it easy to transform between different coordinate systems. The components of a vector can be transformed from one coordinate system to another using the Jacobian matrix for the transformation. Each such matrix is an element of a group isomorphic to $GL(4, \mathbb{R})$ (with matrix multiplication as the group operation). On an N-dimensional spacetime, Jacobian matrices are elements of a group isomorphic to $GL(N, \mathbb{R})$.

Consider an arbitrary N-dimensional Riemannian spacetime manifold $\mathcal{M}$. A region of it Ω may be parametrised using a system of curvilinear coordinates u^I. The vectors tangent to the curves of increasing u^1, $u^2, ...$ at a point A form a basis for the tangent space $T_A\mathcal{M}$, denoted $\mathbf{e}_M|_A$ - the "coordinate basis" for u^M. (The coordinate basis may be similarly defined at any other point in Ω. Where we are evaluating a quantity at a given point, we shall state explicitly which point it is evaluated at, as the greatest source of errors in carrying out this research has been due to confusing an expression giving the value of a quantity at a given point with an expression for the quantity as a function.) The value of a vector field at A may then be written as a linear sum of this coordinate basis, and indeed any other coordinate basis:

$$\mathbf{V}|_A \in T_A\mathcal{M} = V^M_{(u)}|_A \mathbf{e}_M|_A = V^N_{(u')}|_A {\mathbf{e}'}_N|_A \tag{1}$$

The infinitesimal displacement vector $du^M|_A \mathbf{e}_M|_A$ in these bases is used as a template to derive the fundamental transformation law

$$V^M_{(u)}|_A \mathbf{e}_M|_A = V^M_{(u)}|_A \left.\frac{\partial u'^N}{\partial u^M}\right|_A {\mathbf{e}'}_N|_A \tag{2}$$

We can see this as a transformation of either the basis or the components.

This being a Riemannian manifold, we can define a symmetric inner product for each tangent space:

$$(\mathbf{V}, \mathbf{W})_A = (\mathbf{W}, \mathbf{V})_A \in \mathbb{R} \tag{3}$$

The image of this map on the coordinate basis is the metric at A:

$$\mathrm{g}_{MN}|_A = (\mathbf{e}_M, \mathbf{e}_N)_A \tag{4}$$

and the inner product acts linearly over the tangent space. We can use this to find the transformation of the metric under a change of coordinates.

We can always define a set of coordinates x^I for which the basis is pseudo-orthonormal at our chosen point (with respect to the inner product). We will call this "frame basis" $\hat{\mathbf{n}}_I$:

$$(\hat{\mathbf{n}}_I, \hat{\mathbf{n}}_J)_A = \eta_{IJ} \tag{5}$$

The Jacobian matrix for transforming between this basis and the coordinate basis is an element of a group J_A which is isomorphic to $GL(N, \mathbb{R})$. We will denote the transformation between the chosen frame basis and the chosen (unprimed) coordinate basis $j_0|_A$:

$$(j_0)_M{}^I|_A = \left.\frac{\partial x^I}{\partial u^M}\right|_A \in J_A : \hat{\mathbf{n}}_M \mapsto \mathbf{e}_M = (j_0)_M{}^I \hat{\mathbf{n}}_I \tag{6}$$

while j will be used for a generic change of basis - for example,

$$j \in J_A : \mathbf{e}_M \mapsto \mathbf{e}'_M = j_M{}^N \mathbf{e}_N \tag{7}$$

Note that in this formalism, V^M consequently transforms according to:

$$j : V^M_{(u)}|_A \mapsto V^M_{(u')}|_A = V^N_{(u)}|_A (j^{-1})_N{}^M \tag{8}$$

Much of the theory in this paper will relate to two decompositions of j_0. We introduce one of these now. Let $\mathcal{M}$ have t timelike dimensions and s spacelike dimensions. Then the Minkowski metric (5) is invariant under spacetime rotations (including boosts) and spacetime inversions (such as reflections) and combinations of these, which make up a group I_A isomorphic to $O(t, s)$. J_A can be partitioned into cosets of the form $\Lambda_0 I_A$, so we can always write

$$j_0|_A = \Lambda_0|_A i_0|_A \tag{9}$$

where $i_0 \in I_A$. If we then define

$$\hat{\mathbf{k}}_K|_A = (i_0)_K{}^I|_A \; \hat{\mathbf{n}}_I|_A \tag{10}$$

we find that

$$(\hat{\mathbf{k}}_K, \hat{\mathbf{k}}_L)_A = \eta_{KL} \tag{11}$$

and

$$\mathbf{e}_M|_A = (\Lambda_0)_M{}^K|_A \; \hat{\mathbf{k}}_K|_A \tag{12}$$

and

$$\mathrm{g}_{MN}|_A = (\Lambda_0)_M{}^K|_A \; (\Lambda_0)_N{}^L|_A \; \eta_{KL} \tag{13}$$

2.2 Connections and covariant derivatives along a curve

Having examined the tangent space at a given point A, we now want to look at comparing the tangent spaces at different points. To do this, we need to use a connection.

General relativity uses a particular connection, the Levi-Civita connection, or Christoffel symbol. This has the advantages of being symmetric and being uniquely defined - on a given manifold in a given coordinate system, its components are single-valued at each point. However, when considering frame bases as we are here, it makes more sense to introduce the concepts by starting with connections on a curve, which can be generalised either to the Levi-Civita connection and its associated spin connection, or to those of teleparallelism.

Consider a curve $c(\lambda)$ through Ω parametrised by the single variable λ. We take λ to be invariant under changes of coordinate. Pick two points on it A and B. We define any map between the tangent spaces $T_A\mathcal{M}$ and $T_B\mathcal{M}$ which preserves linearity and the inner product as a "parallel map". There are an infinite number of these.

Now choose frame bases at both points, $\hat{\mathbf{n}}_I|_A$ and $\hat{\mathbf{n}}_I|_B$. Denote the parallel map $\tilde{}$ for which the image of $\hat{\mathbf{n}}_I|_A$ is $\hat{\mathbf{n}}_I|_B$:

$$\tilde{}: T_A\mathcal{M} \rightarrow T_B\mathcal{M} \tag{14}$$

$$\tilde{}: \hat{\mathbf{n}}_I|_A \mapsto \hat{\mathbf{n}}_I|_B \tag{15}$$

Then as $\tilde{}$ is a linear map,

$$\tilde{}: \mathbf{e}_M|_A \mapsto \tilde{\mathbf{e}}_M = (j_0|_A\, {j_0}^{-1}|_B)_M{}^N \mathbf{e}_N|_B \tag{16}$$

In the teleparallism formalism, this is valid regardless of how close or far apart A and B are. However, we are looking to define a connection. We therefore take A and B to be close to each other (in the language of relativity, the interval between these events is small). We then note that we can also define parallel maps to and from all the points on $c(\lambda)$ between these points - this set of parallel maps along this section of the curve constitutes a "parallelism". We choose this such that the transformation j_0 from the frame basis to the coordinate basis varies continuously with λ. (This means that not only must the coordinate basis and the frame basis be related by the same group J all along the curve, but j_0 must be in the same connected component of J at all points.) This allows us to carry out a Taylor expansion of ${j_0}^{-1}$ in λ, giving us

$$\tilde{\mathbf{e}}_M = \left(\mathbf{1} + \delta\lambda \left(j_0 \frac{\partial {j_0}^{-1}}{\partial \lambda}\right)_M{}^N \Bigg|_A \right) \mathbf{e}_N|_B + \mathcal{O}^2(\lambda) \tag{17}$$

From the linear nature of the parallel map, we then find the image of any vector $\mathbf{V}$:

$$\tilde{}: \mathbf{V}|_A \mapsto \tilde{\mathbf{V}} = V^N|_A \mathbf{e}_N|_B + \delta\lambda\, V^M|_A \left(j_0 \frac{\partial {j_0}^{-1}}{\partial \lambda}\right)_M{}^N \Bigg|_A \mathbf{e}_N|_B + \mathcal{O}^2(\lambda) \tag{18}$$

The quantity in brackets is our archetypal connection (up to a change in sign):

$$\left(\Gamma_\lambda^{(u)}\right)_M{}^N\Big|_A \equiv -\left(j_0 \frac{\partial j_0^{-1}}{\partial \lambda}\right)_M{}^N\Big|_A = \left(\frac{\partial j_0}{\partial \lambda} j_0^{-1}\right)_M{}^N\Big|_A \tag{19}$$

Under a change of curvilinear coordinates, from u^K to u'^K, we simply replace j_0 in these expressions by jj_0, where

$$j_M{}^N = \frac{\partial u^N}{\partial u'^M} \tag{20}$$

giving us

$$j : \left(\Gamma_\lambda^{(u)}\right)_M{}^N\Big|_A \mapsto \left(\Gamma_\lambda^{(u')}\right)_M{}^N\Big|_A = \left(j\Gamma_\lambda j^{-1}\right)_M{}^N\big|_A - \left(j\frac{\partial j^{-1}}{\partial \lambda}\right)_M{}^N\Big|_A \tag{21}$$

One possible change of coordinates is to the set x^I mentioned above, with pseudo-orthonormal basis at A. Then $j = j_0^{-1}$, so

$$\left(\Gamma_\lambda^{(x)}\right)_M{}^N\Big|_A = 0 \tag{22}$$

If $c(\lambda)$ is a geodesic, then x can have pseudo-orthonormal basis, and $\Gamma_\lambda^{(x)} = 0$, along the entire curve.

We can also look at changing parallelism. Consider a new parallelism $^-$, which again preserves orthonormality, so that

$$^- : \hat{\mathbf{n}}_I\big|_A \mapsto \bar{\mathbf{n}}_I\big|_B = i_I{}^J\, \hat{\mathbf{n}}_J\big|_B \tag{23}$$

If i is constant along $c(\lambda)$, Γ_λ is unaffected. But if i varies with λ (we take it to be in the same connected component of I at every point),

$$^- : (\Gamma_\lambda)_M{}^N\big|_A \mapsto (\Gamma'_\lambda)_M{}^N\big|_A = (\Gamma_\lambda)_M{}^N\big|_A - \left(j_0 i \frac{\partial i^{-1}}{\partial \lambda} j_0^{-1}\right)_M{}^N\Big|_A \tag{24}$$

We can use (18) to define a covariant derivative at A:

$$D_\lambda \mathbf{V}|_A = \lim_{\delta\lambda \to 0} \frac{\mathbf{V}|_B - \tilde{\mathbf{V}}}{\delta\lambda} \tag{25}$$

This can be extended to a field along the curve in the obvious way; the components of this field are

$$D_\lambda V^N = \partial_\lambda V^N + V^M (\Gamma_\lambda)_M{}^N \tag{26}$$

It is easy to show that this transforms covariantly:

$$j : D_\lambda^{(u)} V_{(u)}^M \mapsto D_\lambda^{(u')} V_{(u')}^M = D_\lambda^{(u)} V_{(u)}^N \left(j^{-1}\right)_N{}^M \tag{27}$$

In the x-coordinates, this simply becomes

$$D_\lambda^{(x)} V_{(x)}^M = \partial_\lambda V_{(x)}^M \tag{28}$$

2.3 Connections and covariant derivatives across Ω

It is possible to extend the way we defined Γ above to the whole of Ω. Rather than just defining a parallelism - a set of parallel maps - along a curve, we define a parallelism across the whole of Ω. This results in j_0 becoming a field over u^I. We can then define a connection field using the same approach as in (17), except we now Taylor expand in each of the curvilinear coordinates; this is known as the Weitzenböck connection:

$$\dot{\Gamma}_{LN}{}^{M}(u) \equiv -\left(j_0 \partial_L j_0{}^{-1}\right)_N{}^M \equiv \left(\partial_L (j_0) j_0{}^{-1}\right)_N{}^M \tag{29}$$

This is not the most general connection. Other rules for parallel transporting a vector exist, which do not take this form. More generally,

$$\tilde{}: \mathbf{V}|_A \mapsto \tilde{\mathbf{V}} = V^N|_A \mathbf{e}_N|_B - \delta u^L \, V^M|_A \Gamma_{LM}{}^N \mathbf{e}_N|_B + \mathcal{O}(\delta u)^2 \tag{30}$$

The transformation of $\Gamma_{LM}{}^N$ under a local change of basis is similar to the transformation for Γ_λ, except that we now need to act on the index L:

$$j(u) : \Gamma^{(u)\ N}_{LM} \mapsto \Gamma^{(u')\,N}_{LM} = j_L{}^K \left(j \Gamma^{(u)}_K j^{-1}\right)_M{}^N - j_L{}^K \left(j \partial_K j^{-1}\right)_M{}^N \tag{31}$$

where $(\Gamma_L)_M{}^N \equiv \Gamma_{LM}{}^N$.

For the Weitzenböck connection, just as for Γ_λ, we can apply a transformation j_0^{-1} to reduce it to zero - except that we can now do it over the whole of Ω. However, on a curved manifold, the frame bases defined by

$$\hat{\mathbf{n}}_I = \left(j_0^{-1}\right)_I{}^M \mathbf{e}_M \tag{32}$$

at each point do not represent the basis for any coordinate system.

This set of bases will not be of particular use in this work. However, it is worth noting what happens on a geodesic in more detail. If we consider a point particle moving along a geodesic, we can always base a set of coordinates x^I on its rest frame. The geodesic is parametrised by τ, the particle's proper time, which is proportional to x^0:

$$x^0 = \mathrm{c}\tau \tag{33}$$

These coordinates are "Riemann normal coordinates": they have pseudo-orthonormal basis along the entire geodesic, and indeed the first derivatives of the metric are zero. By comparison with (22), we therefore have

$$\left.\dot{\Gamma}^{(x)\ N}_{0M}\right|_{c(\lambda)} = 0 \tag{34}$$

For any connection $\Gamma^{(u)\ N}_{LM}$, we may define the covariant derivative of a vector, with components

$$D_L V^M = \partial_L V^M + V^N \Gamma_{LN}{}^M \tag{35}$$

The covariant derivative at a point A is an element of $T_A\mathcal{M} \otimes T^*_A\mathcal{M}$. Under a local change of basis, the inhomogeneous term in the transformation of Γ is cancelled by the inhomogeneous term in the transformation of $\partial_L V^M$. Consequently, $D_L V^M$ transforms covariantly:

$$j : D_L^{(u)} V_{(u)}^M \mapsto D_L^{(u')} V_{(u')}^M = j_L{}^K D_K^{(u)} V_{(u)}^N \left(j^{-1}\right)_N{}^M \tag{36}$$

This can be extended in the normal way to tensors of other ranks.

It is easy to show that any connection for which (30) preserves the inner product of vectors is metric compatible, that is

$$D_L \mathrm{g}^{MN} = 0 \tag{37}$$

However, it is not necessarily symmetric. For example, the Weitzenböck connection is metric compatible, but has a torsion:

$$\dot{T}_{LM}{}^N = \dot{\Gamma}_{LM}{}^N - \dot{\Gamma}_{ML}{}^N \neq 0 \tag{38}$$

The only symmetric, metric-compatible connection is the Levi-Civita connection:

$$\mathring{\Gamma}_{LM}{}^N = \mathring{\Gamma}_{ML}{}^N = \frac{1}{2}\mathrm{g}^{NK}\left(\partial_K \mathrm{g}_{LM} - \partial_L \mathrm{g}_{KM} - \partial_M \mathrm{g}_{KL}\right) \tag{39}$$

Now for any geodesic $c(\lambda)$, in the Riemann normal coordinates x^I, the derivatives of the metric are zero, so

$$\left.\mathring{\Gamma}^{(x)}_{LM}{}^N\right|_{c(\lambda)} = 0 \tag{40}$$

However, away from the geodesic the Levi-Civita connection is non-zero on a curved manifold, even in this coordinate system. Note that incorporating (34), we have

$$\left.\dot{\Gamma}^{(x)}_{0M}{}^N\right|_{c(\lambda)} = \left.\mathring{\Gamma}^{(x)}_{0M}{}^N\right|_{c(\lambda)} = 0 \tag{41}$$

We conclude this section by noting some further properties of the Weitzenböck and Levi-Civita connections. The Weitzenböck connection has zero field strength[28, 30]:

$$\partial_L \dot{\Gamma}_{NK}{}^M - \partial_N \dot{\Gamma}_{LK}{}^M + \dot{\Gamma}_{NK}{}^J \dot{\Gamma}_{LJ}{}^M - \dot{\Gamma}_{LK}{}^J \dot{\Gamma}_{NJ}{}^M = 0 \tag{42}$$

and (as noted above), it can be reduced to zero across Ω by a local change of basis. The scalar curvature (the Ricci scalar) may be constructed from its torsion tensor[28, 29]. For a given coordinate system on a given manifold, this connection is not unique - its definition depends on the parallelism chosen.

The field strength of the Levi-Civita connection is the Riemann curvature tensor:

$$R^M{}_{KLN} = \partial_L \mathring{\Gamma}_{NK}{}^M - \partial_N \mathring{\Gamma}_{LK}{}^M + \mathring{\Gamma}_{NK}{}^J \mathring{\Gamma}_{LJ}{}^M - \mathring{\Gamma}_{LK}{}^J \mathring{\Gamma}_{NJ}{}^M \tag{43}$$

and the connection cannot be reduced to zero across Ω by a local change of basis, except on a flat spacetime. For a given coordinate system on a given manifold, it is unique. The Riemann tensor can also be viewed in terms of the action of the covariant derivatives on a vector field:

$$\left[\mathring{D}_K, \mathring{D}_J\right] W^I = R^I{}_{LKJ} W^L \tag{44}$$

Finally, each connection $\Gamma_{LM}{}^N$ has an associated Lorentz connection or spin connection. Pereira[28] defines a Lorentz connection as a one-form assuming values in the Lie algebra of the Lorentz group. In N dimensions, this will be in the Lie algebra of $SO(1, N-1)$. This means that at least two of its indices must be frame indices. It therefore has two forms, one of which has all three indices as frame indices, while the other has two frame indices and one coordinate index. In the formalism of this paper, the Lorentz connection with three frame indices is considered to be the usual connection in the frame basis. The frame basis at a point A is the basis at that point for some set of Riemann normal coordinates x^I, so we can write this connection at this point as $\Gamma^{(x)}_{LK}{}^M\Big|_A$. The form with two frame indices and one coordinate index is considered to be in a mix of two different bases. We shall write this as follows:

$$\omega_M{}^{IJ} T_{IJ} \in SO(1, N-1) \tag{45}$$

where the first index is taken to be a coordinate index and the last two are frame indices.

This is easiest to see for the Weitzenböck connection, in situations where j_0 can everywhere be decomposed in the form (9) such that Λ_0 and i_0 are continuous functions of the coordinates. In this case,

$$\dot{\Gamma}^{(u)}_{LM}{}^N = (\Lambda_0)_L{}^K \eta_{KI} \dot{\omega}_M{}^{IJ} (\Lambda_0^{-1})_J{}^N + (\Lambda_0)_L{}^K \partial_M (\Lambda_0^{-1})_K{}^N \tag{46}$$

This equation can be inverted to give:

$$\dot{\omega}_M{}^{IJ} = \eta^{IK} [(\Lambda_0^{-1})_K{}^L \dot{\Gamma}^{(u)}_{ML}{}^N (\Lambda_0)_N{}^J + (\Lambda_0^{-1})_K{}^L \partial_M (\Lambda_0)_L{}^J] \tag{47}$$

Similar relations hold for other connections.

The Lorentz connection derived in this way is not unique: any local change of frame $i(u)$ (including $i_0(u)$) results in another Lorentz connection. $\omega_M{}^{IJ}$ transforms under a local change of frame according to:

$$i(u) : \omega_M{}^{IJ} \mapsto \omega'_M{}^{IJ} = \left(i \omega_M i^{-1}\right)^{IJ} - \left(i \partial_M i^{-1}\right)^{IJ} \tag{48}$$

where frame indices are raised and lowered using η^{IK} and η_{IK}. It transforms under a change of curvilinear coordinates according to:

$$j(u) : \omega_M{}^{IJ} \mapsto \omega'_M{}^{IJ} = j_M{}^N \omega_N{}^{IJ} \tag{49}$$

Note that the Weitzenböck spin connection can be reduced to zero everywhere by a local change of frame, whereas on a curved manifold the Levi-Civita spin connection cannot[28].

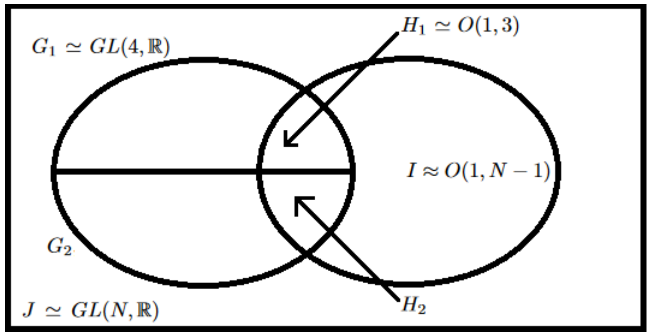

Figure 1: Decompositions of J

3 Breaking the general linear symmetry

The Lagragian density for any N-dimensional tensor theory should be invariant under changes of N-bein. It is built out of fields which transform as tensors of $J \simeq GL(N,\mathbb{R})$ (here $\simeq$ denotes isomorphism). We want our manifold to compactify to one for which all fields can be reduced to tensors of $G = G_1 \otimes G_2$, where $G_1 \simeq GL(4,\mathbb{R})$ and G_2 is a product of general linear groups. The metric will decompose into two factors, a four-dimensional metric and an $N-4$-dimensional metric (although the latter may itself be block diagonal). G_1 has a pseudo-orthogonal subgroup $H_1 \simeq O(1,3)$ under which the metric is invariant and similarly G_2 has an orthogonal subgroup H_2 under which the metric is invariant. This situation is represented in Figure 1.

A continuous symmetry of a physical system may be broken by applying algebraic constraints to fields in that system. We will say that the system is 'idealized' when the fields satisfy the constraints at every point in $\mathcal{M}$. The constraints can be regarded as basic parameters of the theory (as they are in non-linear sigma models). Alternatively, if one wishes them to be determined dynamically, this can be achieved by adding a potential to the Lagrangian which is minimised when they are satisfied (as in the Goldstone and Higgs mechanisms).

A crucial question, as identified by Isham[31], is which representation the fields involved in the constraints should belong to. The tensor must have enough components to include all of the coset space parameters associated with the 'broken generators'. In our case, there are $dim(J/G) = 8N - 32$ of these; for example, if $N = 6$, $dim(J/G) = dim(J) - dim(G_1) - dim(G_2) = 36 - 16 - 4 = 16$. A vector does not have enough components, but a rank-2 tensor always will.

In our case, we also want the symmetry breaking to induce curvature in the additional dimensions. Now we have seen that the Levi-Civita connection will reduce to zero on a geodesic in the appropriate coordinate systems. But a curved manifold may be distinguished by this connection varying from zero away from a geodesic. We therefore want the tensor appearing in the constraints to contain the Levi-Civita connection.

From the Section 2, it is clear that the covariant derivative of a vector satisfies both our requirements. We will denote this particular vector field which appears in the constraints $\mathbf{W}$. We can view the derivative of W^J at A as a matrix:

$$D_I{}^J \equiv \mathring{D}_I W^J|_A \tag{50}$$

This is an element of the space of all possible covariant derivatives at A, which we denote $\mathcal{D}_A$.

We now note two important properties of $D_I{}^J$.

3.1 Symmetric and anti-symmetric parts of $D_I{}^J$

$D_I{}^J$ is a real $N \times N$ matrix and therefore an element of $gl(N, \mathbb{R})$, the Lie algebra of $GL(N, \mathbb{R})$. We can exponentiate it to create an element of J_A:

$$(j_D)_I{}^J \equiv \exp(\mathrm{i}D_I{}^J) \equiv \delta_I^J + \mathrm{i}D_I{}^J + \frac{1}{2}D_I{}^K D_K{}^J + \dots \tag{51}$$

The action of j_D on the metric allows us to split $D_I{}^J$ into an element of the Lie algebra of I_A and an element in the part of $gl(N, \mathbb{R})$ orthogonal to this Lie subalgebra. This is done by identifying the part of $D_I{}^J$ which generates an isometry. Using 51 to first order, we have:

$$j_D : \mathrm{g}_{IJ}|_A \mapsto \mathrm{g}'_{IJ}|_A = (\delta_I^K + \mathrm{i}D_I{}^K)(\delta_J^L + \mathrm{i}D_J{}^L)\mathrm{g}_{KL}|_A + \mathcal{O}^2(D_I{}^J) \tag{52}$$

This is an isometry if and only if

$$\mathrm{i}D_I{}^K \mathrm{g}_{KJ}|_A + \mathrm{i}D_J{}^L \mathrm{g}_{IL}|_A = 0 \tag{53}$$

and hence

$$D_I{}^M = -\mathrm{g}^{MJ}|_A D_J{}^L \mathrm{g}_{IL}|_A = D^M{}_I \tag{54}$$

This means that the antisymmetric part of a general $D_I{}^J$

$$D_{[I}{}^{J]} \equiv \frac{1}{2}(D_I{}^J - D^J{}_I) \tag{55}$$

lies in the Lie algebra of I_A. Conversely, the symmetric part

$$D_{\{I}{}^{J\}} \equiv \frac{1}{2}(D_I{}^J + D^J{}_I) \tag{56}$$

lies in the subspace of the algebra of J_A orthogonal to this. (Consequently, if we were just breaking I, we could just use $D_{[I}{}^{J]}$ in our constraints.)

3.2 Orbits and strata

The action of $j \in J$ on the covariant derivative of a vector is given by (36). The action on D is therefore a similarity transformation:

$$j : D \to D' = jDj^{-1} \tag{57}$$

Thus in every coordinate system, D has the same eigenvalues.

The action of $j|_A$ on $\mathcal{D}_A$ partitions it into orbits of derivatives with the same eigenvalues. The orbit structure of $gl(N,\mathbb{R})$ is much richer than that generally considered in the literature[32, 33, 34, 35]. For example, the algebra contains matrices which are not diagonalisable. Even if a matrix is diagonalisable, it may have complex eigenvalues. Given that the matrix is real and the group elements acting on it are also real, the diagonalised matrix does not lie in the orbit.

The action of j is not a free action. For example, pure reflections and inversions take the form

$$j_R = \begin{pmatrix} \pm 1 & & & \\ & \pm 1 & & \\ & & \dots & \\ & & & \pm 1 \end{pmatrix} \tag{58}$$

and are involutary. Consequently, it is easily seen that they stabilise any diagonal matrix.

The multiplicities of the eigenvalues determine the isotropy group of a matrix. Every matrix in the same orbit has the same isotropy group up to conjugation and orbits can be grouped into 'strata' according to their isotropy groups. These isotropy groups will be direct products of general linear groups.

As the eigenvalues are the roots of the characteristic equation for the covariant derivative, it is easy to use them to find the constraint equation resulting in a particular symmetry breaking pattern. We will see an example of this in Section 5. The eigenvalues are completely determined by the invariants appearing in the characteristic equation, which are constructed from the traces of the powers of the matrix. It is worth noting that the traces of the odd powers are zero for any matrix entirely in the $so(1, N-1)$ part of the algebra.

We are particularly interested in the stratum for which one factor of the isotropy group is $GL(4,\mathbb{R})$. We will want to show that for this stratum, the metric is reducible. But before this, it is worth briefly considering the stratum for which the general linear symmetry is not broken.

3.3 The unbroken symmetry stratum

If eigenvalues of D are all equal, then D is invariant under the whole of J_A. This means that there is only one matrix in each orbit and

$$D_I^{(u)J} = D_I^{(x)J} = \partial_I W_{(x)}^J|_A \tag{59}$$

for *every* coordinate system u^M.

Denote the single eigenvalue a. There is one diagonal matrix - and hence diagonal orbit - for each value of a:

$$D_I{}^J = a\delta_I^J \tag{60}$$

The traces of the powers of this matrix are entirely determined by a. In particular,

$$\mathrm{tr}(D^2) = (D_I W^J D_J W^I)|_A = Na^2 \tag{61}$$

As the components of the covariant derivative must be real, a is real, so this invariant cannot be negative; it is then minimised for

$$D_I W^J|_A = 0 \tag{62}$$

If we specify that $D_I W^J$ has a single distinct eigenvalue at every point in $\mathcal{M}$, then

$$D_I W^J = a(u^K)\delta_I^J \tag{63}$$

The covariant derivative of this is

$$D_K D_I W^J = (\partial_K a)\delta_I^J \tag{64}$$

This will be useful in Section 4.4.

3.4 Symmetry breaking strata

In other strata, the covariant derivative continues to transform under J_A according to (57). In particular,

$$j_0|_A : D^{(x)} \to D^{(u)} = j_0|_A D^{(x)} j_0^{-1}|_A \tag{65}$$

for any coordinate system u^M. In general, each basis vector in the coordinate basis is a linear sum of *all* of the basis vectors of the frame basis, as shown in (6). In this case, $D^{(x)}$ and $D^{(u)}$ are not equal. However, we can show that for the coordinate systems y^M mentioned in Section 1.2, $D^{(x)}$ and $D^{(y)}$ are equal. This is where we need to use the techniques of non-linear realisations.

As explained above, we are interested in strata for which the isotropy group is $G = G_1 \otimes G_2$. So far, we have decomposed $j_0|_A$ in the form (9), but G_A may also be used to partition J_A into cosets. We can therefore decompose j_0 as

$$j_0|_A = L_0|_A g_0|_A \tag{66}$$

where $g_0|_A \in G_A$ and $L_0|_A$ is an element of J_A which has no generators of G_A in its expansion. Denote a derivative in the frame basis which is invariant under $g_0|_A$ by $D_0^{(x)}$. Consequently any covariant derivative in the stratum can be written in the form

$$D_0^{(u)} = L_0|_A D_0^{(x)} L_0^{-1}|_A \tag{67}$$

for some $L_0|_A$ and some $D_0^{(x)}$. Note that all the N-bein degrees of freedom are contained in $L_0|_A$. This relation identifies each matrix $D_0^{(u)}$ with a matrix L_0. This means that if we denote this orbit

$$(\mathcal{D}_0)_A \subset \mathcal{D}_A \tag{68}$$

it is diffeomorphic to the space of all values of L_0 and hence the coset space J_A/G_A.

Any matrix in one of these strata has differing eigenvalues, with multiplicities determined by the isotropy group. The fact it has repeated eigenvalues simplifies its characteristic equation. For example, for $G_2 \simeq GL(N-4,\mathbb{R})$, there are four eigenvalues with the same value a and $N-4$ with the same value b. Thus a matrix in this stratum has characteristic equation

$$(D_0^{(u)} - a\mathbf{1})^4 (D_0^{(u)} - b\mathbf{1})^{N-4} = 0 \tag{69}$$

The simplified characteristic equation can be used to derive the constraint equation, as we will see in Section 5. This constraint equation will pick out from $\mathcal{D}_A$ the covariant derivatives with the required multiplicities of eigenvalues and hence determine the isotropy group. In more detail, it does this as follows. Any matrix $D \in \mathcal{D}_A$ can be written

$$D = j_0 D^{(x)} j_0^{-1} \tag{70}$$

for some $D^{(x)}$. Substituting in (66), we get

$$D = L_0 (g_0 D^{(x)} g_0^{-1}) L_0^{-1} \tag{71}$$

That is, we can obtain $\mathcal{D}_A$ by acting on the set of matrices

$$\mathcal{G}_A = \{g_0 D^{(x)} g_0^{-1} \;\; \forall g_0 \in G_A, D^{(x)}\} \tag{72}$$

with the set of all L_0. Conversely, acting with $(L_0)^{-1}$ on $\mathcal{D}_A$ reduces it to the set $\mathcal{G}_A$. We shall see below that the matrices in this set are tensors of G_A. Within this set, only the matrices $D_0^{(x)}$ satisfy the constraints.

In the stratum with $G_2 \simeq GL(N-4,\mathbb{R})$, there is one diagonal matrix for each pair of values a, b:

$$\mathfrak{D} = \begin{pmatrix} a & & & & & & & \\ & a & & & & & & \\ & & a & & & & & \\ & & & a & & & & \\ & & & & b & & & \\ & & & & & b & & \\ & & & & & & \dots & \\ & & & & & & & b \end{pmatrix} \tag{73}$$

If $D^{(x)}$ takes this form, then

$$D^{(u)} = L_0|_A \mathfrak{D} L_0^{-1}|_A \tag{74}$$

so

$$D_I^{(u)\,J} = a (L_0)_I{}^\mu|_A (L_0^{-1})_\mu{}^J|_A + b (L_0)_I{}^X|_A (L_0^{-1})_X{}^J|_A \tag{75}$$

It is worth noting that any diagonal matrix is symmetric and therefore lies entirely in the coset space part of the Lie algebra.

3.5 Reduction of the basis and tensor fields

It is clear that L_0^{-1} plays a key role in the symmetry breaking, just as it does in non-linear realisations. We now show how it is crucial to compactification. The decomposition (66) allows us to define a new basis for $T_A\mathcal{M}$:

$$\mathbf{l}_K|_A = (L_0^{-1})_K{}^M \mathbf{e}_M|_A = (g_0)_K{}^I|_A \,\hat{\mathbf{n}}_I \tag{76}$$

Note that as G_A is a direct product group, g_0 is block diagonal. Consequently,

$$\mathbf{l}_\mu|_A = (g_0)_\mu{}^\nu|_A \,\hat{\mathbf{n}}_\nu \tag{77}$$
$$\mathbf{l}_X|_A = (g_0)_X{}^Y|_A \,\hat{\mathbf{n}}_Y \tag{78}$$

The inner product for the new basis at A is also reducible:

$$(\mathbf{l}_\mu, \mathbf{l}_\nu)_A = (g_0)_\mu{}^\rho|_A (g_0)_\nu{}^\sigma|_A \eta_{\rho\sigma} \tag{79}$$
$$(\mathbf{l}_W, \mathbf{l}_X)_A = (g_0)_W{}^Y|_A (g_0)_X{}^Z|_A \delta_{YZ} \tag{80}$$
$$(\mathbf{l}_\mu, \mathbf{l}_X)_A = 0 \tag{81}$$

We can also write it in terms of the metric at A in u-coordinates:

$$(\mathbf{l}_K, \mathbf{l}_L)_A = (L_0^{-1})_K{}^M|_A \,(L_0^{-1})_L{}^N|_A \,\mathrm{g}^{(u)}_{MN}|_A \tag{82}$$

We shall call the coordinates associated with the new basis y^K.

Now if the constraints are satisfied over the whole of $\mathcal{M}$, we say that the system is 'idealized'. Whenever $D_I W^J$ meets the idealization conditions at a point, it can be reduced to $D_0^{(x)} = D_0^{(y)}$ using L_0^{-1}. Now the u-coordinates are consistently defined across a chart Ω. The sufficient and necessary condition for $D_I W^J$ to meet these conditions across Ω is for the y-coordinates - and hence L_0^{-1} - to be defined consistently across Ω. We want the y-coordinates to be continuous and $D_0^{(y)}$ to be a continuous function of them, so we require L_0^{-1} to be continuous over the chart.

Now the metric in u-coordinates is consistently defined over the chart Ω and is continuous over the chart. Then the inner product above can be extended to be a metric for the y-coordinates:

$$\mathrm{g}^{(y)}_{KL} = (L_0^{-1})_K{}^M (L_0^{-1})_L{}^N \mathrm{g}^{(u)}_{MN} \tag{83}$$

At any given point, however, this still satisfies (79)-(81). The metric thus takes a block diagonal form.

If $\mathrm{g}^{(y)}_{\mu\nu}$ is independent of y^X and $\mathrm{g}^{(y)}_{XY}$ is independent of y^μ, then the Levi-Civita connection is also reducible. By considering parallel transport around closed loops, it is easy to see that the holonomy group of the manifold is $H = H_1 \otimes H_2$, where $H_1 \simeq SO(1,3)$ and similarly H_2 is the orthogonal subgroup of G_2. Then by the de Rham decomposition theorem, we know that the manifold is isometric to a Cartesian product of manifolds, each of which is geodesically complete. (To aid understanding, we point out that a cylinder is an example of a manifold of this description. The original Kaluza-Klein theory was also on such a manifold; the extra dimension was 'circular' with constant radius in the four-space.) y^μ are the coordinates on the factor with holonomy group H_1 and y^X are the coordinates on the factor with holonomy group H_2.

If the independence condition is relaxed, then the resulting manifold can deform - it appears that it is still homeomorphic to the one with the reducible holonomy group, but there are new derivatives of the metric appearing in the Levi-Civita connection. This results in changes to the geodesics. (The analogy now is with a tube of varying radius - its geodesics diverge and converge along its length, reflecting the fact its intrinsic curvature differs from that of the cylinder. As remarked in the appendix, Jordan and Thiry constructed just such a generalisation of the original Kaluza-Klein theory.)

The u-coordinates are clearly not natural ones to use on the compactified manifold - in general, u^1 will be a function of both the y^μ *and* y^X coordinates, and similarly for the other u-coordinates. Using L_0^{-1}, we can associate any curvilinear basis $\mathbf{e}_I$ with a basis $\mathbf{l}_K$ in a coordinate system appropriate to the manifold after compactification. (In general, two different general curvilinear bases $\mathbf{e}_I$ and $\mathbf{e}'_I$ will be associated with two *different* bases $\mathbf{l}_K$ and $\mathbf{l}'_K$.)

We therefore want to transform any tensor fields in our system from a general curvilinear coordinate basis into one of the new bases appropriate to the manifold. In this new basis, these tensor fields will transform under J_A as representations of G_A. The method here follows the method of non-linear realisations exactly. For example, a vector $V|_A$ can be decomposed in the new basis:

$$\mathbf{V}|_A \in T_A\mathcal{M} = V^M_{(u)}|_A \mathbf{e}_M|_A = V^K_{(y)}|_A \mathbf{l}_K|_A \tag{84}$$

so that

$$V^K_{(y)} = V^M_{(u)} (L_0)_M{}^K \tag{85}$$

Now the action of $j \in J_A$ on a coset $L_0 G_A$ maps it to another coset $L' G_A$, so

$$j L_0 = L' g \tag{86}$$

for some $g \in G_A$. Using this, (8) and (87), it is easy to show that

$$j : V^K_{(y)} \mapsto V^K_{(y')} = V^L_{(y)} (g^{-1})_K{}^L \tag{87}$$

which may be written

$$j : V^\mu_{(y)} \quad \mapsto \quad V^\mu_{(y')} = V^\nu_{(y)} (g^{-1})_\nu{}^\mu \tag{88}$$

$$j : V^X_{(y)} \quad \mapsto \quad V^X_{(y')} = V^Y_{(y)} (g^{-1})_Y{}^X \tag{89}$$

Thus we have broken $\mathbf{V}$ into two multiplets. The induced action of J_A on these is to transform one of them under G_1 and the other under G_2.

This result can naturally be extended to tensors. For example, a rank-2 tensor will break into a four-dimensional rank-2 tensor, a rank-2 tensor of the internal symmetry and charged vectors. As described above, this includes any matrix $D \in \mathcal{D}_A$. In particular, changing the basis allows us to define

$$D^{(y)\nu}_{\mu} = (L_0^{-1})_{\mu}{}^{M} D^{(u)N}_{M} (L_0)_{N}{}^{\nu} = (g_0)_{\mu}{}^{\rho} D^{(x)\sigma}_{\rho} (g_0^{-1})_{\sigma}{}^{\nu} \tag{90}$$

$$D^{(y)Y}_{X} = (L_0^{-1})_{X}{}^{M} D^{(u)N}_{M} (L_0)_{N}{}^{Y} = (g_0)_{X}{}^{W} D^{(x)Z}_{W} (g_0^{-1})_{Z}{}^{Y} \tag{91}$$

These can be extended to fields, which transform according to

$$j : D^{(y)\nu}_{\mu} \mapsto D^{(y')\nu}_{\mu} = g_{\mu}{}^{\rho} D^{(y)\sigma}_{\rho} (g^{-1})_{\sigma}{}^{\nu} \tag{92}$$

$$j : D^{(y)Y}_{X} \mapsto D^{(y')Y}_{X} = g_{X}{}^{W} D^{(y)Z}_{W} (g^{-1})_{Z}{}^{Y} \tag{93}$$

- multiplets which include the actual values the covariant derivative takes when the system is idealised.

In summary, whenever the system is idealized, the general linear symmetry is broken to $G_1 \otimes G_2$. This allows the y-coordinates to be defined consistently across Ω. The metric in these coordinates is block diagonal and all fields break into multiplets which transform as tensors of $G_1 \otimes G_2$. This defining of y-coordinates can be repeated for each chart used across $\mathcal{M}$. This determines the topology of $\mathcal{M}$ - either it is isometric to a Cartesian product manifold with geodesically complete factors or it is a generalisation of such a manifold in the way described above.

4 Field equations

The isotropy group of $D_I W^J$ determines the topology of the manifold, regardless of the form of the Lagrangian density. To find the geometry - more specifically, the curvature at each point - we need field equations. We then need to find solutions of these which are consistent with the idealization conditions.

We want field equations which are based on the fundamental principles and conceptual structure of GR. We also want them to be consistent with the field equations of Newtonian gravity in the appropriate non-relativistic limits. The approach we will take is to construct the simplest generalisation of the Poisson equation for gravity for a vector field which is consistent with general covariance and the equivalence principle. We will then see how the field equations simplify when the covariant derivative is constrained.

We start by generalising Laplace's equation. This is then extended to a generalisation of the Poisson equation.

4.1 Generalising Laplace's equation

We start by considering a test particle - one whose own gravitational field is negligible - moving in a background gravitational field. Newtonian gravitation has a scalar potential, ϕ. This means that the work done in moving the particle from one point to another is independent of path; around a closed loop it is zero. Denoting the particle's Newtonian velocity vector $\mathbf{v}$, the acceleration due to the field is

$$\frac{d\mathbf{v}}{dt} = -\nabla\phi \tag{94}$$

Laplace's equation for gravity simply equates the gradient of this to zero:

$$\nabla.(\nabla\phi) = 0 \tag{95}$$

In general relativity, if a test particle has no non-gravitational forces acting on it, it is considered to be in 'free fall'. For a particle with finite real mass, it moves on a timelike geodesic. Its relativistic velocity vector, whose components we shall denote C^M, is covariantly constant along the geodesic. Hence the free fall acceleration is entirely due to the variation in the transformation j_0 between the coordinate basis and the particle's rest frame along the path.

The result of parallel transporting the velocity vector from one point to another depends on the path taken, and parallel transporting it around a closed loop of intersecting geodesics induces a transformation. We therefore do not expect our generalisation of Laplace's equation to contain the derivative of a scalar potential.

From (35) and the covariant constancy of C^M,

$$\frac{\partial C^M_{(u)}}{\partial u^L} = -C^N_{(u)}\mathring{\Gamma}_{LN}{}^M \tag{96}$$

Note that in the rest frame coordinates,

$$C^0_{(x)} = 1, \;\; C^i_{(x)} = 0 \tag{97}$$

Now from (6), at any point A on the geodesic,

$$\mathbf{e}_0|_A = (j_0|_A)_0{}^0\hat{\mathbf{n}}_0|_A + (j_0|_A)_0{}^i\hat{\mathbf{n}}_i|_A \tag{98}$$

and

$$\mathbf{e}_i|_A = (j_0|_A)_i{}^0\hat{\mathbf{n}}_0|_A + (j_0|_A)_i{}^j\hat{\mathbf{n}}_j|_A \tag{99}$$

We now define a 'Newtonian coordinate system' v^M as one for which

$$(j_0)_0{}^0 = 1 + \mathcal{O}(\epsilon), \;\; (j_0)_0{}^i = \mathcal{O}(\epsilon), \;\; (j_0)_i{}^0 = \mathcal{O}(\epsilon) \tag{100}$$

at every point on the geodesic, where ϵ is a very small parameter. That is, the change of coordinates may mix up the spatial basis vectors to any extent, but the mixing of the spatial and timelike bases is very limited and $\mathbf{e}_0 \approx \hat{\mathbf{n}}_0$. Then using the usual transformation law for vectors together with (97), we have

$$C^0_{(v)} = 1 + \mathcal{O}(\epsilon), \;\; C^i_{(v)} = \mathcal{O}(\epsilon) \tag{101}$$

Substituting these into (96), we find

$$\frac{\partial C^M_{(v)}}{\partial v^L} = -\mathring{\Gamma}_{L0}{}^M + \mathcal{O}(\epsilon) \tag{102}$$

Thus the spatial components of the acceleration are given by

$$a^i \approx -c^2 \mathring{\Gamma}_{00}{}^i \tag{103}$$

where c is the speed of light. As remarked above, intrinsic curvature may be distinguished by the variation in the Levi-Civita connection with separation from the geodesic. It is therefore promising to see components of it appearing in this equation. However, this cannot function as a generalisation of the gradient of the gravitational potential, as it is not tensorial. It has been derived from the expression on the right hand side of (96), which at least has covariance in its indices. However, even that expression is still not tensorial and furthermore it contains a local vector which is only defined on the particle's path.

We therefore replace the local vector in this expression with the vector field which determines the topology, W^M. To ensure that our 'potential gradient' transforms as a tensor, we add on a term, to give us back the familiar covariant derivative:

$$D_L W^M_{(u)} = \partial_L W^M_{(u)} + W^N_{(u)} \mathring{\Gamma}_{LN}{}^M \tag{104}$$

This is the simplest possible generalisation of $\nabla\phi$ consistent with general covariance and the equivalence principle. To see the correspondence more clearly, observe that

$$D_0 W^i_{(u)} = \partial_0 W^i_{(u)} + W^0_{(u)} \mathring{\Gamma}_{00}{}^i + W^j_{(u)} \mathring{\Gamma}_{0j}{}^i \tag{105}$$

Then if v^M are now coordinates in which the equivalent of (101) holds for W^M, this contains the expression on the right hand side of (103), upto a constant:

$$D_0 W^i_{(v)} = \partial_0 W^i_{(v)} + \mathring{\Gamma}_{00}{}^i + \mathcal{O}(\epsilon) \tag{106}$$

Our generalisation of Laplace's equation then follows immediately, as a covariant divergence of (104). However, we have to be careful to take the divergence on the correct index. From (103) we see that the appropriate equation is

$$D_M D_L W^M = 0 \tag{107}$$

It can easily be shown that this can be obtained from a Lagrangian density

$$\mathcal{L} = k D_L W^M D_M W^L = k\,\mathrm{tr}(D^2) \tag{108}$$

where k is any constant of the correct dimensionality. The action integral uses the measure $|j_0|\, d^N u$, where $|j_0|$ has the property

$$D_I\, |j_0| = 0 \tag{109}$$

We then subject W^I to an active variation over Ω - one in which the coordinate basis is preserved - which vanishes at the boundary. Equation (107) then follows by using established procedures[36] (including making use of the fact that for a contravariant vector density of weight +1, the covariant derivative is equal to the partial derivative).

4.2 Generalising Poisson's equation

To arrive at a geometry with intrinsic curvature, we need to give the W^M field a mass. This will be described by a generalisation of Poisson's equation. We find this by adding a mass term to the Lagrangian density (108):

$$\mathcal{L} = kD_L W^M D_M W^K - \frac{1}{2}m^2 W^I W^J \mathrm{g}_{IJ} = k\mathrm{tr}(D^2) - \frac{1}{2}m^2(\mathbf{W},\mathbf{W}) \tag{110}$$

For full generality, we assume that m is a scalar function of the coordinates u^N. The same procedure then gives us the field equation

$$2kD_M D_L W^M = -m^2 W^M \mathrm{g}_{ML} \tag{111}$$

To find the ratio between k and m^2, we note that the timelike component of this may be rewritten

$$D_i D_0 W^i - D_0 D_0 W^0 = \frac{m^2}{2k} W^M \mathrm{g}_{M0} \tag{112}$$

In a coordinate system which becomes Newtonian at A, defining t by

$$t = v^0/c \tag{113}$$

we get

$$D_i D_0 W^i|_A - \frac{1}{c} D_t D_0 W^0|_A = \frac{m^2}{2k} W^0|_A \, \mathrm{g}_{00}|_A + \mathcal{O}(\epsilon) \tag{114}$$

The first term on the left is our generalisation of $\nabla^2\phi$ derived in the previous subsection - this can be expanded to give

$$D_i(\partial_0 W^i + \mathring{\Gamma}_{00}{}^i)|_A - \frac{1}{c} D_t D_0 W^0|_A = \frac{m^2}{2k} W^0|_A \, \mathrm{g}_{00}|_A + \mathcal{O}(\epsilon) \tag{115}$$

where $\mathring{\Gamma}_{00}{}^i$ is related by (103) to the three-acceleration caused by ϕ. The meaning of the next term is unclear, but it may be that it vanishes in the non-relativistic limit (unless there is a very rapid variation in $D_0 W^0$) due to the factor of $1/c$. Assuming this to be the case, comparing with Poisson's equation for gravity, we then expect the right hand side to reduce to $4\pi G\rho/c^2$ where ρ is the density of the **W** field. If we assume that in this coordinate system $W^0|_A = 1$, as we had for C^0 in the rest frame, we get

$$\frac{m^2}{2k} = \frac{4\pi G\rho}{c^2} \tag{116}$$

(Note that if m varies with u^N, so does ρ.) We therefore take as our generalisation of Poisson's equation

$$D_M D_L W^M = -\frac{4\pi G\rho}{c^2} W^M \mathrm{g}_{ML} \tag{117}$$

It can easily be verified that the two sides of this equation have the same dimensionality. This is the full field equation for a universe containing only the **W** field.

4.3 Ricci forms of the field equations

The field equations so far do not look much like those of GR. However, using (44) it is easy to show that

$$D_M D_L W^M = R_{KL} W^K + \partial_L (D_M W^M) \tag{118}$$

where R_{KL} is the Ricci tensor. Thus (107) and (117) can respectively be written as

$$R_{LK} W^K + \partial_L (D_M W^M) = 0 \tag{119}$$

and

$$R_{LK} W^K + \partial_L (D_M W^M) = -\frac{4\pi G\rho}{c^2}\, \mathrm{g}_{LM} W^M \tag{120}$$

Like the field equations of GR, these contain geometrical information in the Ricci and metric tensors. Unlike GR, these tensors do not appear on their own - they appear as matrix operators acting on $\mathbf{W}$.

4.4 Diagonal solutions

We saw in Sections 3.3 and 3.4 that there are orbits which contain diagonal matrices, leading to a simpler form for $D_I W^J$. In particular:

1. If $G \simeq GL(N, \mathbb{R})$, then in any coordinate system, $D_I W^J$ takes the form (63)

2. If $G \simeq GL(4, \mathbb{R}) \otimes GL(N-4, \mathbb{R})$, then in any coordinate system which respects the resulting compactification, $D_I W^J$ takes the form (73) everywhere

In both cases, this considerably simplifies the field equations.

Firstly, when $G \simeq GL(N, \mathbb{R})$, from (64) we have

$$D_J D_I W^J = \partial_I a \tag{121}$$

and

$$D_K D_I W^I = N \partial_K a \tag{122}$$

Inserting the first of these into (117) gives us

$$\partial_L a = -\frac{4\pi G\rho}{c^2}\, W^M \mathrm{g}_{ML} \tag{123}$$

Alternatively, the difference between these two gives us the Ricci tensor contracted with $\mathbf{W}$:

$$R_{JK} W^K = (1-N) \partial_J a \tag{124}$$

and this can be substituted into (120) to get (123). Note that combining (124) and (123), we get

$$R_{JK} W^K = \frac{4\pi G\rho}{c^2} (N-1) W^M \mathrm{g}_{MJ} \tag{125}$$

This is only possible for all values of the free index in all coordinate systems if

$$R_{JK} = \frac{4\pi G\rho}{c^2}(N-1)\mathrm{g}_{JK} \tag{126}$$

If ρ is constant, then so is the Ricci scalar and $\mathcal{M}$ is an Einstein manifold.

Note that as $\rho \to 0$, we find a tends to a constant value and $R_{JK} \to 0$; that is, the manifold becomes Ricci flat. Otherwise, a can only be constant if $W^M = 0$.

Secondly, when $G = G_1 \otimes G_2 \simeq GL(4, \mathbb{R}) \otimes GL(N-4, \mathbb{R})$, the covariant derivative matrix, the metric and the Ricci tensor all decompose into tensors of G_1 and tensors of G_2. We then find that

$$\partial_\mu a = -\frac{4\pi G\rho}{c^2}\, W^\nu \mathrm{g}_{\nu\mu} \tag{127}$$

and

$$R_{\mu\nu} W^\nu = -3\partial_\mu a \tag{128}$$

and similarly

$$\partial_X b = -\frac{4\pi G\rho}{c^2}\, W^Y \mathrm{g}_{XY} \tag{129}$$

and

$$R_{XY} W^Y = (5-N)\partial_X b \tag{130}$$

Combining these appropriately, we get

$$R_{\mu\nu} W^\nu = \frac{12\pi G\rho}{c^2}\, \mathrm{g}_{\mu\nu} W^\nu \tag{131}$$

and

$$R_{XY} W^Y = \frac{4\pi G\rho}{c^2}(N-5)\mathrm{g}_{XY} W^Y \tag{132}$$

which imply that

$$R_{\mu\nu} = \frac{4\pi G\rho}{c^2}\, \mathrm{g}_{\mu\nu} \tag{133}$$

and

$$R_{XY} = \frac{4\pi G\rho}{c^2}(N-5)\mathrm{g}_{XY} \tag{134}$$

Now if ρ is constant, both these manifolds are Einstein manifolds.

As $\rho \to 0$, *both* manifolds become Ricci flat. If, at the other extreme, both manifolds are curved, their curvature scalars are given by

$$R^{(4)} = \frac{48\pi G\rho}{c^2} \tag{135}$$

and

$$R^{(N-4)} = \frac{4\pi G\rho}{c^2}(N-5)(N-4) \tag{136}$$

- they are of the same order. Neither of these have the features we are looking for. However, if a is constant, the equations admit a solution in which the four-space is flat and the other submanifold is curved.

Naturally, a similar decomposition can be used with analogous results when G_2 is itself a direct product of general linear groups.

It is perhaps worth pointing out explicitly that the $\mathbf{W}$ is not the gravitational potential - gravity itself acts through the connection - but $D_M\mathbf{W}$ provides a medium through which it acts. There is no need to have a purely geometric term in the Lagrangian for the desired geometry to result. This justifies our choice of the covariant derivative as the tensor to which to apply the constraints.

5 Example: the two-sphere

In this section, we work through an example of applying the theory we have developed in the previous sections. In this example, $G_2 \simeq GL(2,\mathbb{R})$. This is of particular interest, as $H_2 \simeq SO(2)$, which is the transformation group of an electromagnetic charge doublet. The additional dimensions form a sphere. We calculate the constraints in spherical polar coordinates and find a field configuration on the sphere which satisfies them. We then show that this solution satisfies the field equations. We end the section with some remarks on the symmetries of this solution and some issues to consider when trying to generalise this method.

We start by identifying appropriate constraints to ensure a diagonal solution with $G_2 \simeq GL(2,\mathbb{R})$ and a flat four-space. This can be achieved if the diagonal covariant derivative is

$$D_K^{(y)L} = \begin{pmatrix} 0 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & b(y) & \\ & & & & & b(y) \end{pmatrix} \tag{137}$$

(A more general form would have the zeros replaced by a constant a, but we will keep things simple. This form takes $\mathbf{W}$ to be covariantly constant over time and the three macroscopic spatial dimensions.) This means that the covariant derivative matrix has characteristic equation

$$D^4(D - b\mathbf{1})^2 = D^6 - 2bD^5 + b^2D^4 = 0 \tag{138}$$

Taking traces, we have

$$\mathrm{tr}(D^6) - 2b\,\mathrm{tr}(D^5) + b^2\mathrm{tr}(D^4) = 0 \tag{139}$$

Now we do not want our constraints to specify a particular value of b - we need derivatives of b to be non-zero. Instead, we want to eliminate b from this to give a relationship between the invariants - we can do this using

$$\mathrm{tr}D = 2b \tag{140}$$

Our constraint equation is therefore

$$\mathrm{tr}(D^6) - \mathrm{tr}D\ \mathrm{tr}(D^5) + \frac{1}{4}(\mathrm{tr}D)^2\mathrm{tr}(D^4) = 0 \tag{141}$$

We can either view this as an external constraint applied to a system with the Lagrangian density (110) with (116), or we can start with the Lagrangian density

$$\mathcal{L} = k\mathrm{tr}(D^2) - 4\pi G\rho k(\mathbf{W},\mathbf{W}) - k'\left(\mathrm{tr}(D^6) - \mathrm{tr}D\ \mathrm{tr}(D^5) + \frac{1}{4}(\mathrm{tr}D)^2\mathrm{tr}(D^4)\right)^2 \tag{142}$$

Minimising the potential or satisfying the constraint then gives the required topology. The extra dimensions form a two-dimensional Einstein manifold. With the appropriate sign for ρ, this must be a two-sphere.

We can therefore adopt the coordinates $y^X = \theta, \phi$ for all of this manifold except for where $\theta = \pi$ and where $\phi = 0, 2\pi$. The basis for these coordinates has inner products

$$(\mathbf{l}_\theta, \mathbf{l}_\theta) = r_0^2 \tag{143}$$
$$(\mathbf{l}_\theta, \mathbf{l}_\phi) = 0 \tag{144}$$
$$(\mathbf{l}_\phi, \mathbf{l}_\phi) = r_0^2\mathrm{sin}^2\theta \tag{145}$$

where r_0 is the sphere's radius. This means that an orthonormal basis at A is given by

$$\hat{\mathbf{n}}_\theta|_A = \frac{1}{r_0}\mathbf{l}_\theta|_A \tag{146}$$
$$\hat{\mathbf{n}}_\phi|_A = \left.\frac{1}{r_0\mathrm{sin}^2\theta}\right|_A \mathbf{l}_\phi|_A \tag{147}$$

As noted in Section 2.3, the field of such basis vectors does not form a basis for any coordinate system. The corresponding g_2 is then simply

$$g_2 = \begin{pmatrix} r_0 & 0 \\ 0 & r_0\mathrm{sin}\theta \end{pmatrix} \tag{148}$$

This could be used to find the Weitzenböck connection for the parallelism defined by

$$^- : \hat{\mathbf{n}}_X|_A \mapsto \hat{\mathbf{n}}_X|_B \tag{149}$$

Instead, we calculate the Levi-Civita connection to find that its non-zero components are

$$\mathring{\Gamma}^\theta_{\phi\phi} = -\mathrm{sin}\theta\ \mathrm{cos}\theta \tag{150}$$
$$\mathring{\Gamma}^\phi_{\phi\theta} = \mathring{\Gamma}^\phi_{\theta\phi} = \mathrm{cot}\theta \tag{151}$$

The idealisation conditions in these coordinates are then

$$D_\theta W^\theta = \partial_\theta W^\theta = b(\theta,\phi) \tag{152}$$
$$D_\theta W^\phi = \partial_\theta W^\phi + W^\phi \cot\theta = 0 \tag{153}$$
$$D_\phi W^\theta = \partial_\phi W^\theta - W^\phi \sin\theta \cos\theta = 0 \tag{154}$$
$$D_\phi W^\phi = \partial_\phi W^\phi + W^\theta \cot\theta = b(\theta,\phi) \tag{155}$$

It is easy to see that a solution to these equations is

$$W^\theta = \xi \sin\theta \tag{156}$$
$$b = \xi \cos\theta \tag{157}$$
$$W^\phi = 0 \tag{158}$$

We remark on this solution at the end of this section.

From the Levi-Civita connection, we can calculate the Riemann tensor. Contracting this with the metric then reveals that the Ricci tensor has two non-zero components:

$$R_{\theta\theta} = 1 \tag{159}$$
$$R_{\phi\phi} = \sin^2\theta \tag{160}$$

Note that this means the Ricci tensor is proportional to the metric, as expected.

We then want to check whether the solution (156)-(158) satisfies the field equations. From these, (137), (143)-(145) and (159)-(160) we find that

$$R_{\theta Y} W^Y + \partial_\theta (D_X W^X) = -\xi \sin\theta = -\frac{1}{r_0^2} \mathrm{g}_{\theta X} W^X \tag{161}$$

and

$$R_{\phi Y} W^Y + \partial_\phi (D_X W^X) = 0 = \mathrm{g}_{\phi X} W^X \tag{162}$$

Therefore the two-space part of (120) is satisfied if

$$r_0^2 = \frac{c^2}{4\pi G\rho} \tag{163}$$

Note that to cause the extra dimensions to compactify so tightly, the **W** field has to have a staggeringly high density. For example, for the radius to be the Planck length, the density needs to be of the order of 10^{95} kg m^{-3}. Even if r_0 is only at nuclear scales, it would need to be of the order of 10^{56} kg m^{-3}. (With a rising number of extra dimensions, the necessary density falls away according to (136)). Note that r_0 will vary with y^μ if ρ does.

There are other aspects of this solution which are worthy of further discussion. The fact that we only have two additional dimensions means that the curvature of the compact manifold can be completely described using the Ricci scalar. We therefore know that it must be a sphere and that we can utilise polar coordinates. We have then stated a particular solution of the idealisation conditions and shown that it satisfies the field equations. An informative direction for future research might be to look into what we are able to say about the most general solution of the idealisation conditions which is consistent with the field equations - both for two dimensions and for higher dimensionality, where the Einstein manifold is not fully determined by the scalar curvature.

The two-sphere has an $SO(3)$ symmetry, which is not present in our initial Lagrangian, and it can be identified with the manifold $SO(3)/SO(2)$. However, both g_2 and our solution (156)-(158) are dependent only on one periodic component. All reference to the ϕ coordinate has dropped out and the ϕ-component of $\mathbf{W}$ has also dropped out of the solution. Consequently, only the $SO(2)$ symmetry is manifest in the solution. This is analogous to the finding of Volkov *et al*[37] that for Luciani-type compactification on G/H (where H is the holonomy group of the compact submanifold), the gauge potentials associated with G/H are non-dynamical and can be eliminated using a gauge transformation. Interestingly, in our case the solution depends on the θ coordinate on the sphere, and therefore has periodicity π.

If future research looks at general solutions in arbitrary numbers of dimensions, it could include this question of whether the additional symmetries of the compact space always drop out in this way. In searching for such solutions, researchers should be aware of points raised by Pons[38]. Firstly, one can ask which solutions are 'consistent truncations'. Secondly, in seeking a specific solution, it is possible to make simplifying assumptions (for example, to obtain particular field content) which are equivalent to imposing constraints. This may result in the specific solution having a lower degree of symmetry than a completely general consistent solution.

6 Fictitious forces

The most fundamental principle of general relativity is that matter causes spacetime to curve. We have seen how the theory presented here respects that principle. The $\mathbf{W}$ field triggers spontaneous compactification, thus determining the *background* geometry of spacetime. For a realistic model, though, we need to include familiar matter fields, particularly spinors. The mathematical framework for incorporating spinors is a current subject of my research. However, we can get a sense of what might happen if such fields are included, by changing from the y-coordinate system to another set of coordinates. In this section, we look first at connections in a generic coordinate system, then at the metric in a system that varies slightly from the y-coordinates.

On our idealised manifold $\mathcal{M}$, we can decompose j_0 at every point using (66) or its equivalent. This allows us to define the basis (77)-(78) at A and similarly define $\mathbf{l}_K|_B$.

As a first stage, we define a rather unconventional map between bases at A and B:

$$^{\times} : \mathbf{l}_K|_A \mapsto \overset{\times}{\mathbf{l}}_K = \mathbf{l}_K|_B \tag{164}$$

This map does not necessarily preserve orthonormality, as $\mathbf{l}_K|_A$ and $\mathbf{l}_K|_B$ are not in general orthonormal. However, we still take it to be linear. We can then use it to relate $\mathbf{e}_M|_B$ to $\overset{\times}{\mathbf{e}}_N$ with a Taylor expansion, in a similar way to (17):

$$\mathbf{e}_M|_B = \left(\delta_M^N + \delta u^L \left(\overset{\times}{\Gamma}_L\right)_M{}^N\bigg|_A\right)\overset{\times}{\mathbf{e}}_N + \mathcal{O}^2(u) \tag{165}$$

where

$$\overset{\times}{\Gamma}_L = -L_0\partial_L {L_0}^{-1} \tag{166}$$

Now L_0 has both pseudo-orthonormal and non-pseudo-orthonormal parts. Furthermore, $\overset{\times}{\Gamma}_L$ contains a term which lies in $\mathcal{G}$ and a term which lies in $\mathcal{J}/\mathcal{G}$ (where $\mathcal{J}$ and $\mathcal{G}$ are the Lie algebras of J and G). These subgroup and coset space parts are traditionally known as v_L and a_L respectively. From these definitions,

$$(v_L)_\mu{}^X = 0, \qquad (v_L)_X{}^\mu = 0 \tag{167}$$

$$(a_L)_\mu{}^\nu = 0, \qquad (v_L)_X{}^Y = 0 \tag{168}$$

Under the action of $j \in J$, the non-zero parts transform as follows. Those in v_L transform as gauge potentials of G_1 and G_2 (or more precisely, the embedding of these potentials in $\mathcal{G}$):

$$j : (v_\mu)_\nu{}^\rho \mapsto (v'_\mu)_\nu{}^\rho = (g_1)_\mu{}^\sigma \left((g_1)v_\sigma(g_1)^{-1} - (g_1)\partial_\sigma(g_1)^{-1}\right)_\nu{}^\rho \tag{169}$$

$$j : (v_X)_\nu{}^\rho \mapsto (v'_X)_\nu{}^\rho = (g_2)_X{}^Y \left((g_1)v_Y(g_1)^{-1} - (g_1)\partial_Y(g_1)^{-1}\right)_\nu{}^\rho \tag{170}$$

$$j : (v_\mu)_W{}^Z \mapsto (v'_\mu)_W{}^Z = (g_1)_\mu{}^\sigma \left((g_2)v_\sigma(g_2)^{-1} - (g_2)\partial_\sigma(g_2)^{-1}\right)_W{}^Z \tag{171}$$

$$j : (v_X)_W{}^Z \mapsto (v'_X)_W{}^Z = (g_2)_X{}^Y((g_2)v_Y(g_2)^{-1} - (g_2)\partial_Y(g_2)^{-1})_W{}^Z \tag{172}$$

where $g = g_1 \otimes g_2$ is defined by (86). Those in a_L, on the other hand, transform as spacetime vectors and tensors with internal symmetry charges:

$$j : (a_\mu)_\nu{}^X \mapsto (a'_\mu)_\nu{}^X = (g_1)_\mu{}^\sigma(g_1)_\nu{}^\rho(a_\sigma)_\rho{}^Y((g_2)^{-1})_Y{}^X \tag{173}$$

$$j : (a_Z)_\nu{}^X \mapsto (a'_Z)_\nu{}^X = (g_2)_Z{}^W(g_1)_\nu{}^\rho(a_W)_\rho{}^Y((g_2)^{-1})_Y{}^X \tag{174}$$

$$j : (a_\mu)_X{}^\nu \mapsto (a_\mu)_X{}^\nu = (g_1)_\mu{}^\sigma(g_2)_X{}^Y(a_\sigma)_Y{}^\rho((g_1)^{-1})_\rho{}^\nu \tag{175}$$

$$j : (a_Z)_X{}^\nu \mapsto (a_Z)_X{}^\nu = (g_2)_Z{}^W(g_2)_X{}^Y(a_W)_Y{}^\rho((g_1)^{-1})_\rho{}^\nu \tag{176}$$

These are parts of a connection which has zero field strength, so they represent 'fictitious forces', experienced due to a change of reference frame, from y-coordinates to u-coordinates. Both have pseudo-orthonormal and non-pseudo-orthonormal parts.

While we have derived them using the unconventional connection (166), it is easy to see how these turn up in the Levi-Civita connection. If we take the Levi-Civita connection in y-coordinates then perform a transformation to u-coordinates, we find

$$L_0(u) : \overset{\circ}{\Gamma}{}^{(y)}_{LM}{}^N \mapsto \overset{\circ}{\Gamma}{}^{(u)}_{LM}{}^N = (L_0)_L{}^K\left((L_0)\overset{\circ}{\Gamma}{}^{(y)}_K(L_0)^{-1}\right)_M{}^N + (L_0)_L{}^K(v_K + a_K)_M{}^N \tag{177}$$

They also appear in the metric. Taking the inner products of (165), we find

$$(\mathbf{e}_M, \mathbf{e}_N)_B = (\overset{\times}{\mathbf{e}}_M, \overset{\times}{\mathbf{e}}_N) + \delta u^L \left((v_L + a_L)_M{}^K (\overset{\times}{\mathbf{e}}_K, \overset{\times}{\mathbf{e}}_N) + (v_L + a_L)_N{}^K (\overset{\times}{\mathbf{e}}_M, \overset{\times}{\mathbf{e}}_K) \right) + \mathcal{O}^2(u) \tag{178}$$

Now this coordinate basis is utterly general. However, we expect any familiar matter fields to cause far less curvature than that of the $\mathbf{W}$ field, due to having much lower density. We would therefore expect to be able to treat any additional curvature as a small perturbation to the background geometry. Given that we can use the y-coordinate basis across a chart of the idealised manifold, we would expect to be able to use a basis which is very close to it on the perturbed manifold. Furthermore, while we would no longer be able to decompose j_0 into L_0 and g_0 at every point on this chart with these both being continuous functions, we can still use this decomposition at a chosen point. This would allow the coordinate basis across the chart to reduce to $\mathbf{l}_K$ at any single point we choose.

While in this paper we are sticking to the idealised manifold, we can mirror this reasoning by specialising to a coordinate basis $\mathbf{l'}_K$ on it which differs only slightly from $\mathbf{l}_K$ and coincides with it exactly at point A. Using (164), this means that

$$\overset{\times}{\mathbf{l'}}_K = \mathbf{l}_K|_B \tag{179}$$

Replacing $\mathbf{e}_M$ with $\mathbf{l'}_K$ and $\overset{\times}{\mathbf{e}}_M$ with $\mathbf{l}_K|_B$ in (178) and using (81) and (167)-(168), we find

$$(\mathbf{l'}_\mu, \mathbf{l'}_\nu)_B = (\mathbf{l}_\mu, \mathbf{l}_\nu)_B + \delta u^L \left((v_L)_{\mu\nu} + (v_L)_{\nu\mu} \right)|_B + \mathcal{O}^2(u) \tag{180}$$

$$(\mathbf{l'}_\mu, \mathbf{l'}_X)_B = (\mathbf{l}_\mu, \mathbf{l}_X)_B + \delta u^L \left((a_L)_{\mu X} + (a_L)_{X\mu} \right)|_B + \mathcal{O}^2(u) \tag{181}$$

$$(\mathbf{l'}_X, \mathbf{l'}_Y)_B = (\mathbf{l}_X, \mathbf{l}_Y)_B + \delta u^L \left((v_L)_{XY} + (v_L)_{YX} \right)|_B + \mathcal{O}^2(u) \tag{182}$$

where we have lowered indices on v_L and a_L using $(\mathbf{l}_\rho, \mathbf{l}_\nu)$ and $(\mathbf{l}_X, \mathbf{l}_Y)$.

Note that in a coordinate system for which $L_0 \approx \mathbf{1}$, if we have chosen the constraints such that the four-space is flat, the Levi-Civita connection on it approximates to $(v_\mu)_\nu{}^\sigma$. $(v_\mu)_X{}^Y$, on the other hand, appears to be a gauge potential for an internal symmetry. In general, it will have both orthogonal and non-orthogonal parts - it is actually a gauge potential for G_2. However, as remarked in Section 2.3, every connection has an associated spin connection. The spin connection associated with this will be a one-form taking values in the Lie algebra of its maximal orthogonal subgroup ($SO(N-4)$ or a product of special orthogonal groups), that is, a gauge potential of that orthogonal symmetry. It seems reasonable to hope, as described in the final section below, that these pure gauge connections will gain a field strength.

7 Conclusions and discussion

We have shown that introducing a tensor field into an empty N-dimensional space can cause part of that space to compactify - even if it is not a multiplet of any other symmetry groups - but only if that field has the necessary properties. The covariant derivative of a vector field has the correct properties. It forms an orbit under the action (57) of the general linear group of the full spacetime. The symmetry breaking pattern - and hence the resulting topology - depends on which orbit the covariant derivative belongs to.

Furthermore, we have found a suitable Lagrangian (110) for such a system, such that the field equation (117) resulting from it is a simple generalisation of Poisson's equation. We have shown that such a system can admit a solution containing Minkowski space, where the remaining dimensions form a compact Einstein manifold. The constraint on the covariant derivative matrix which ensures this solution can be written in terms of the traces of its powers, and we have found this explicitly (141) in the case where the internal space is a two-sphere.

Finally, we have shown that when the system is idealised - the constraint is satisfied everywhere - all tensor fields decompose naturally into tensors of the subspaces when they are put in a coordinate system adapted to the manifold. But when we move to a different coordinate system, new terms appear in the full Levi-Civita connection; these include the Levi-Civita connection for the Minkowski space in the new coordinates and a gauge potential for the internal symmetry.

The key thing we have not done is to add fermionic matter to the system. This is likely to be a non-trivial undertaking. Now, the famous lift thought experiments in GR show that a non-inertial reference frame can be used to simulate the effects of a gravitational field locally. The difference between them is that in the non-inertial frame, the connection has no field strength, whereas it does in a gravitational field. We appear to have an analogous situation in Section 6. A change of coordinates has induced a connection with no field strength. Additional matter would cause a variation in the curvature, which would presumably give the connection a field strength. Then $(v_\mu)_\nu{}^\sigma$ (or some equivalent) would represent the gravitational field induced by the additional matter. Similarly, the spin connection associated with $(v_\mu)_X{}^Y$ would represent a gauge potential for the internal symmetry induced by the matter. However, while we would expect the Levi-Civita connection to differ only slightly from the values of (177), g_0, L_0 and $\mathbf{l}_K$ could no longer be defined continuously over a finite neighbourhood, so there could be issues with defining $(v_\mu)_\nu{}^\sigma$ and $(v_\mu)_X{}^Y$. It is therefore difficult to state this categorically without finding the correct coupling of spinors to the model and carrying out the analysis explicitly.

It is encouraging to see the internal gauge potential appear in the metric, as it does in the Klein metric and its non-Abelian generalisations[6, 22, 23]. However, way it appears is different - in particular, in the existing theories it appears quadratically in elements of the metric. This would seem to make it difficult for the N-dimensional Ricci scalar to contain the norm of the field strength for this gauge potential. However, it should be remembered that this is not an $SU(N)$ symmetry. If an $SU(N)$ gauge potential couples linearly to a spinor, then presumably it couples quadratically to the outer product of a spinor and its adjoint or conjugate. It therefore seems reasonable to hope that the covariant derivative of the vector in this product might contain this field quadratically. Finding the correct coupling of spinors to this model would allow us to clarify this relationship between the $SU(N)$ gauge potential and $(v_\mu)_X{}^Y$.

It is also worth remembering that $(v_\mu)_X{}^Y$ has a non-orthogonal part. In order for our model to reduce to anything like an Einstein-Yang-Mills form, we would presumably need the spinors to couple purely to the orthogonal part, allowing us to eliminate the non-orthogonal part with a gauge transformation. (That is, a change of coordinates.)

Overall, there are tantalising hints that what we have presented could be the tensor sector of a theory which fully unifies gravity and gauge symmetries in a geometric way.

While the correct coupling of spinors could conceivably eliminate the non-orthogonal part of $(v_\mu)_X{}^Y$, it is unclear whether a similar treatment is available for the charged vectors and tensors a_L. (Though we discuss a_X and v_X further below.) To appreciate the implications of this, consider the decompositions of an N-vector. It breaks into a neutral four-vector and a charged scalar. G_1 transforms the four-vector components and G_2 is a coordinate transformation on the charge space. The remaining components of J must relate to changes of both charge and spin - and a_L belongs to this part of the Lie algebra. Thus a field interacting with a_L would have its spin changed, by a single unit. This mixing is a novel feature of this model and further research needs to be done to determine its physical consequences.

These bosonic spin-changing symmetries are non-linearly realized, just as the fermionic spin-changing symmetries are non-linearly realized in the Volkov-Akulov model[39, 40]. (Gabrielli[41, 42] also looks at extending the Lorentz group to include symmetries which mix fields of different integer spin, but in four dimensions.) Whether fermionic spin changing symmetries may appear when fermions are coupled to the model - thus potentially giving a new geometric interpretation of supersymmetry - is an open question.

This raises the issues of O'Raifeartaigh's no-go theorem[20] and TEGR. Glibly, it might be assumed that the theorem rules out a model such as this, which combines Lorentz symmetry and an internal symmetry in a larger pseudo-orthogonal group. However, on closer inspection, things are not so clear cut. This is because the theorem is based on the *inhomogeneous* Lorentz group, or Poincaré group. This highlights something else which has not been covered in this paper - translation symmetries. In O'Raifeartaigh's analysis, he considers four cases, one of which includes the model presented here. For this case, it is the higher-dimensional translational symmetry that he views as a problem, as the quantum numbers which one would expect to be associated with it are not observed in particle physics.

However, his paper was written before non-linear realisations were fully understood. In the case where a symmetry is non-linearly realised, no quantum numbers will be observed relating to that symmetry. It is unclear at present whether translations are non-linearly realised in this theory or whether some other mechanism prevents the appearance of additional quantum numbers.

Understanding this would require a version of the model in which translation symmetries are manifest. This is where TEGR comes in. It was pointed out by Andrade *et al*[43] that gauge theories of the Poincaré and affine groups have more degrees of freedom than GR, as they contain both curvature and torsion which can vary independently. By contrast, TEGR uses torsion *in place* of curvature. It appears to be the equivalence principle - that inertial effects can (locally) be compensated by gravity - that provides the equivalence between teleparallelism and GR, at least for covariant derivatives. As the theory presented above is based on GR and uses the Levi-Civita connection, it may be possible to find a teleparallel equivalent for it. TEGR has been described as a gauge theory of translations[28, 29], so a teleparallel version of this theory may be an appropriate framework for handling translations.

Having discussed areas which are not covered in the body of this paper, we now point out several novel features of this theory and useful insights it provides, which do not seem to be discussed explicitly in the literature on spontaneous compactification.

Firstly, in this theory, spacetime which is isometric to a product of Minkowski space and a compact internal space can only occur when the fields that trigger compactification are the only ones present. In a realistic universe, this will be approximated only in deep space, where the densities of familiar particle fields are negligible. In the presence of normal gravitating and charged matter, the background geometry is perturbed. This may lead to geodesics which are not purely on one of the subspaces.

Consequently, Yang-Mills gauge fields appear in this theory in a different way from many Kaluza-Klein theories: rather than being associated with a compactified subspace, they represent deviations from the background geometry in the presence of charges. Again, further work is needed to establish the nature of these deviations. This puts a question mark over whether it is possible to use the usual harmonic expansions over the compact manifold. Until we have included charges and examined how to obtain a four-dimensional effective theory from the resulting model, it is not really possible to apply existing theory on the consistency of dimensional reduction. (By consistency, we mean that the there is a way of carrying out the reduction to a four-dimensional effective theory either from the Lagrangian or from the field equations, and the resulting field configurations are the same[38].) As this existing body of theory is largely based on either constraints being exactly satisfied or having an internal compact space, new theory may need to be developed to demonstrate such consistency.

If a method of applying dimensional reduction to this model were developed, it may eliminate the v_X and a_X fields, which relate to the parallel maps between tangent spaces as one moves in the extra dimensions. Also, it is worth recalling that Klein's model was developed with the idea that it might explain charge quantisation. As this model has many similarities to but also many differences from Klein's, it could be informative to investigate whether it also contains a mechanism for quantising charge, and if so, whether the defects of Klein's model are present in this one.

Finally, this theory makes it clear that in the absence of any matter, spacetime has its maximal symmetry. Consequently, describing the idealised manifold (with broken symmetry) as the 'ground state' or the 'low energy state' would be too simplistic. It would probably be accurate to state that for a particular field content, if the constraints are being applied using a potential, as in (142), the idealised manifold represents the minimum of the potential. However, the caveats here are crucial. Firstly, the constraints could equally be applied independently of the Lagrangian density. In this approach, the action is treated as the global invariant, supplemented by a set of local invariants given by the traces of the powers of the covariant derivative matrix. These local invariants could be viewed as fundamental universal constants. Secondly, the relativistic view of energy and energy density is as non-covariant quantities - components in tensors, namely mass-energy and energy-momentum density. In this view, mass-energy causes curvature and the greater the mass-energy, the greater the curvature. Thus when the densities of the matter fields are reduced to zero, the curvature is reduced to zero and the whole universe is flat in all N dimensions. However, for a full description of the energy of the system, we need to consider the gravitational energy. In GR, though, the equivalence principle makes it impossible to separate gravitational and inertial energy-momentum in a covariant way. To do this, once again the teleparallelism approach is required[44].

Appendix

This appendix summarises some aspects of research in two key areas, focusing on papers which are particularly relevant background for this one.

Kaluza-Klein theories and spontaneous compactification

While Einstein was working on General Relativity, Nordström[2, 3, 4] was working on a theory to unify gravity and electromagnetism. And just six years after GR was published, Kaluza[5] proposed what could be called the first geometric unification theory. His model incorporated both gravity and electromagnetism by adding a fifth coordinate to spacetime with a 'cylinder condition'. Klein[6] then assumed this to be a physical dimension and calculated the radius of this extra dimension to be of the order of 10^{-32}m.

The original Kaluza-Klein theory takes general relativity in five dimensions and puts it on a spacetime isometric to $M^4 \times S^1$ and identifies consistent solutions to the equations of motion. Jordan-Thiry theory[45] relaxes the isometry condition, generalising this to a spacetime homeomorphic to Klein's spacetime, by allowing the radius of the compact dimension to vary over four-dimensional spacetime.

The first paper on spontaneous compactification appears to be that by Cremmer and Scherk[15]. They proposed a model in six dimensions. In addition to a metric field with Einstein-Hilbert action, it included the gauge field of an $SO(3)$ internal symmetry and a scalar multiplet which transformed as the defining representation of $SO(3)$ as fundamental to the model. The scalar multiplet acted as a Higgs field which could cause two of the dimensions to spontaneously compactify to a two-sphere.

This model was generalised by Luciani[16], to one with an arbitrary number of extra dimensions, starting with the metric field, the gauge field of a group K and a Higgs multiplet. The compact space was now acted on directly by a group G. If $G = K$, and the symmetry breaking left $H \subset G$ unbroken, then the gauge fields after compactification could be associated with the Killing vectors of the one-parameter subgroups of G not in H. However, G could alternatively be a non-trivial subgroup of K. In this case, the maximal subgroup of K which commutes with G would also be left unbroken.

A series of papers in the early 1980s by Volkov, Sorokin and Tkach considered models with only metric fields and gauge fields - no scalar fields were required. The gauge field multiplet acts as a matter source for gravity which causes curvature. Symmetric internal spaces, with the form G/H, are found to be solutions to the field equations satisfying the ansatz that the gauge field strength is covariantly constant in all coordinate directions. H is the holonomy group of the internal space[19]. This may be either a simple group or a product of simple groups[46]. It is not necessary for the gauge field triggering the compactification to be one for the whole of G - the model can start with gauge fields for H only[18] or even an invariant subgroup of H[46]. The field equations have a solution in which the gauge fields are equal to the Lorentz connections associated with the Levi-Civita connection on the internal space[17, 37].

These models superficially look different from those of Cremmer and Scherk and Luciani. However, it was shown that when the Luciani model with $G = K$ is applied to a case where the internal space is symmetric, it reduces to the connection-based model[37]. This is because the components of the gauge fields of G which are associated with G/H are non-dynamical: they have zero intensity and can be eliminated using a gauge transformation.

These papers investigating mechanisms of spontaneous compactification also formed a basis for the further development of Kaluza-Klein theories. Various authors did not concern themselves with how compactification arose, but focused on the field content arising from it. Scherk and Schwarz[47] and Salam and Strathdee[23] addressed symmetry aspects of the resulting spacetime, starting with the geometrical fields - the vielbein, metric and connection - and then proceeding to other fields that might be in the system. Salam and Strathdee provided important theory on harmonic expansions on an internal quotient space.

Manton[48], on the other hand, considered symmetry breaking patterns for a gauge field for a simple, compact Lie group on $M^4 \times S^2$. He looked for solutions where the unbroken gauge group is $SU(2) \otimes U(1)$. The four-dimensional effective Lagrangian turned out to be just that for the bosonic part of electroweak theory.

Non-linear realisations and spontaneous symmetry breaking

Goldstone's first paper[49] on spontaneous symmetry breaking and Gell-Mann and Lévy's paper[50] which introduced the non-linear sigma model were both submitted for publication in Nuovo Cimento in 1960. However, these topics were studied in such different ways that it was not proved until after nearly a decade of research that they were two sides of the same coin.

Goldstone's paper and a follow-up with Salam and Weinberg[51] looked at potentials with degenerate minima constructed out of scalar fields. They found that whenever a Lagrangian has an invariance under a continuous global symmetry group which is not (fully) shared by its vacuum states, there will be spinless fields of zero mass present. These became known as Goldstone bosons.

Gell-Mann and Lévy considered three models relating to pion decays in a system of pions and nucleons. The third of these effectively took a multiplet of four scalar fields and constrained its length, allowing them to eliminate one of the fields from the Lagrangian. This was the first of many papers in the 1960s in which scalars were included non-linearly in the Lagrangian, so that the full symmetries of the system were not explicit. Much of the early work revolved around one particular realisation of a chiral group[52]. However, in 1969 Callan, Coleman, Wess and Zumino[53, 54] showed how a coset decomposition of a linear Lie group could be used to find the most general form of a Lagrangian in which a subgroup was linearly represented but the rest of the symmetries were realised non-linearly.

The geometry of these non-linear realisations was examined further by Isham [55], who introduced the concepts of Killing vectors and of a metric (prompted by Meetz[56]), and later by Boulware and Brown [57].

The extension of Goldstone's mechanism of spontaneous symmetry breaking to a gauge symmetry became known as the Higgs mechanism, following a series of papers in the mid 1960s[8, 9, 10], but the non-Abelian case was addressed by Kibble[58]. This again used a coset decomposition of the invariance group of the Lagrangian. It pointed out that the vacuum manifold could be identified with the coset space.

This led researchers to realise that non-linear realisations represented the low-energy effective theory where a global symmetry was spontaneously broken - this was shown by Honerkamp [59] in a specific case and by Salam and Strathdee[11] in the general case.

From this viewpoint, the non-linear realisation becomes a submanifold of the field space of the linear representation which is used to break the symmetry. It is therefore crucial to start with scalar fields in the appropriate representation to allow this. This issue was emphasised by Isham[31].

Acknowledgements

This paper is dedicated to the memory or Prof. K. J. Barnes. A wonderfully jolly and inspirational PhD supervisor, he introduced me to the fascinating world of non-linear realisations and set me looking at a symmetry breaking pattern which lies at the heart of the example in this paper. Ken, you never said it was going to be easy! Massive thanks to A. Unzicker and T. Case for their translations of Einstein's papers and for the former's website, which together gave me crucial insights, and to other translators. Huge thanks also to all the unsung heroes of physics: those who put helpful postings on sites such as StackExchange and ResearchGate (cheers, Uhtred, whoever you are!) and edit Wikipedia pages. Finally, thanks to Drs. L. Weston, J. Tighe and J. Hamilton-Charlton, for all the stimulating conversations, long walks and physics jokes over the years.

References

[1] Utiyama R 1955 *Phys. Rev.* **101** 1597–607

[2] G Nordström *Phys. Zeitschr.* **15** (1914) 504 (reprinted in arXiv:physics/0702221 [physics.gen-ph])

[3] G Nordström *Öfversigt af Finska Vetenskaps-Societetens Förhandlingar (Helsingfors)* **Bd. LVII** (1914-1915) Afd. A. N:o 4, p1–15 (reprinted in arXiv:physics/0702222 [physics.gen-ph])

[4] G Nordström *Öfversigt af Finska Vetenskaps-Societetens Förhandlingar (Helsingfors),* **Bd. LVII** (1914-1915) Afd. A. N:o 28, p1–21 (reprinted in arXiv:physics/0702223 [physics.gen-ph])

[5] T Kaluza *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* (1921) 966–972 (reprinted in arXiv:1803.08616 [physics.hist-ph])

[6] O Klein *Z.Phys.* **37** (1926) 895–906 (See also, for example, W O Straub (2014) http://vixra.org/abs/1406.0172; V H Satheesh Kumar and P K Suresh P K (2006) gr-qc/0605016)

[7] J.M. Overduin and P.S. Wesson *Phys. Rept.* **283** (1997) 303–380 (arXiv:gr-qc/9805018)

[8] Higgs P W 1964 *Phys. Lett* **12** 132–3

[9] Higgs P W 1964 *Phys. Rev. Lett.* **13** 508–9

[10] Higgs P W 1966 *Phys. Rev* **145** 1156–63

[11] Salam A and Strathdee J 1969 *Phys. Rev.* **184** 1750–68

[12] Salam A 1968 Weak and electromagnetic interactions *Elementary Particle Theory - Relativistic Groups and Analyticity* ed N Svartholm (John Wiley and Sons)

[13] Weinberg S 1967 *Phys. Rev. Lett.* **19** 1264–66

[14] H. Georgi and S.L. Glashow *Phys. Rev. Lett.* **32** (1974) 438–441

[15] E Cremmer and J Scherk 1976 *Nucl. Phys.* B **108** 409–16

[16] J F Luciani *Nucl.Phys.* **B135** (1978) 111

[17] D.V. Volkov, V.I. Tkach *JETP Lett.* **32** (1980) 668-670, *Pisma Zh.Eksp.Teor.Fiz.* **32** (1980) 681-684

[18] S. Randjbar-Daemi, R. Percacci *Phys.Lett.* **117B** (1982) 41

[19] D V Volkov, V I Tkach *Theor.Math.Phys.* **51** (1982) 427-431, *Teor.Mat.Fiz.* **51** (1982) 171-177

[20] O'Raifeartaigh L 1965 *Phys. Rev.* **139** B1052–62

[21] Coleman S and Mandula J 1967 *Phys. Rev.* **159** 1251–6

[22] E. Witten *Nucl. Phys.* **B186** 412–428

[23] Abdus Salam, J.A. Strathdee *Annals Phys.* **141** (1982) 316-352

[24] Tilman Sauer 2007 *Einstein's Unified Field Theory Program*

[25] Tilman Sauer 2006 *Historia Mathematica* **33** 399-439

[26] Albert Einstein 1928 *Sitzber. Preuss. Akad. Wiss.* **17** 217-22; translated into English in arXiv:physics/0503046 **[physics.hist-ph]** and [27] below

[27] *Selected Papers on Teleparallelism, translated and edited by* D. H. Delphenich

[28] J G Pereira 2012 arXiv:1210.0379 **[gr-qc]**

[29] José W Maluf 2013 *Ann. Phys. (Berlin)* **525** No.5, 339-57

[30] A.P. Balachandran, A. Stern, C.G. Trahern, *Phys. Rev. D***19**(1979) 2416-2446

[31] Isham C J 1969 *Nuov. Cim.* A **61** 729–36

[32] Louis Michel and Luigi A. Radicati, Ann. Inst. Henri Poincaré **Vol XVIII** (1973) 185

[33] L. O'Raifeartaigh, Rep. Pr. Phys. **42** (1979) 159

[34] Barnes K J, Hamilton-Charlton J and Lawrence T R 2001 *J.Phys. A* **34** (2001) 10881–900

[35] Hamilton-Charlton J D 2003 *Nonlinear Realizations and Effective Lagrangian Densities For Nonlinear σ-models* PhD Thesis (University of Southampton)

[36] D'Inverno R 1995 *Introducing Einstein's Relativity* (Oxford: Oxford University Press) p146–147

[37] D.V. Volkov, Dmitri P. Sorokin, V.I. Tkach *Theor.Math.Phys.* **56** (1984) 746–751, *Teor.Mat.Fiz.* **56** (1983) 171–179

[38] J. M. Pons *J.Phys.Conf.Ser.* **68** (2007) 012030 (*Preprint* hep-th/0610268)

[39] Volkov D V and Akulov V P 1973 *Phys.Lett* B **46** 109-110

[40] Tsuneo Uematsu, Cosmas K. Zachos 1982 *Nucl. Phys.* **B201** 250-268

[41] Gabrielli E 1991 *Phys. Lett* B **258** 151-5

[42] Gabrielli E 2001 *Ann. Phys.* **287** 229-59 (*Preprint* hep-th/9909117)

[43] V.C. de Andrade, H.I. Arcos and J.G. Pereira *PoS* **WC2004** (2004) 028 (*Preprint* gr-qc/0412034)

[44] R. Aldrovandi, Tiago Gribl Lucas, J.G. Pereira (*Preprint* arXiv:0812.0034 [gr-qc])

[45] R Coquereaux and G Esposito-Farese *Ann.Inst.H.Poincare Phys.Theor.* **52** (1990) 113–150

[46] D.V. Volkov, Dmitri P. Sorokin, V.I. Tkach *JETP Lett.* **40** (1984) 1162-1165, *Pisma Zh.Eksp.Teor.Fiz.* **40** (1984) 356-359

[47] Joel Scherk, John H. Schwarz *Nucl.Phys.* **B153** (1979) 61-88

[48] N.S. Manton *Nucl.Phys.* **B158** (1979) 141-153

[49] J. Goldstone, Nuovo Cimento **Vol XIX** (1961)154

[50] M. Gell-Mann and M. Lévy, Nuov. Cim. **Vol XVI** (1960) 705

[51] Jeffrey Goldstone, Abdus Salam and Steven Weinberg, Phys. Rev. **127** (1962) 965

[52] See, for example, J. Schwinger, Phys. Lett. **24B** (1967) 473; Jeremiah A. Cronin, Phys. Rev. **161** (1967) 1483; Benjamin W. Lee and H. T. Nieh, Phys. Rev. **166** (1968), 1507; Steven Weinberg Phys. Rev. **166** (1968) 1507

[53] S. Coleman, J. Wess and Bruno Zumino, Phys. Rev. **177** (1969) 2239

[54] C. G. Callan, S. Coleman, J. Wess and Bruno Zumino, Phys. Rev. **177** (1969) 2247

[55] C. J. Isham, Nuov. Cim. A **61** (1969) 188

[56] Kurt Meetz, J. Math. Phys. **10** (1969) 589

[57] David G. Boulware and Lowell S. Brown, Ann. Phys. **138** (1982) 392

[58] T. W. B. Kibble, Phys. Rev. **155** (1967) 1554–1561

[59] J. Honerkamp, Nucl. Phys. B **12** (1969) 227