Fields as Bodies: a unified presentation of spacetime and internal gauge symmetry

David Wallace*
Balliol College, University of Oxford
(Dated: February 23, 2015)

Using the parametrised representation of field theory (in which the location in spacetime of a part of a field is itself represented by a map from the base manifold to Minkowski spacetime) I demonstrate that in both local and global cases, internal (Yang-Mills-type) and spacetime (Poincaré) symmetries can be treated precisely on a par, so that gravitational theories may be regarded as gauge theories in a completely standard sense.

I. INTRODUCTION

In its simplest formulation, a field theory is a function from a base space representing spacetime to some space of field values. A field is not so much an entity as a way of talking about certain properties of the points of spacetime.

This is not simply a philosophical matter. Mathematically, field theories treat the spacetime and internal features of the theory — and their associated symmetries — in sharply distinct ways. The spacetime structure of the theory is represented by structure on the base space (a metric, an affine connection and so forth); the internal structure is represented by structure on the field-value space.

This division persists when we move from global to local internal symmetries (the gauge principle) and from global to local theories of spacetime (from special to general relativity). Analogies are found between the spacetime and internal structures, but at a fundamental mathematical level they differ. An extensive literature on the degree to which general relativity is a gauge theory (and if so, a gauge theory of what?) points to the significance of this mathematical difference. In particular, so-called “Poincaré gauge theory” differs from Yang-Mills gauge theories in a number of respects and these differences are generally attributed to the different nature of spacetime and internal symmetries.

In this paper I demonstrate that the difference represents a certain choice in how a field theory is formalised, and not a difference of kind. It is in fact possible to formulate field theories — local or global — so that the spacetime features of the theory are treated in precisely the same way as the internal features. Conceptually, the key move is to stop thinking of a field as a collection of properties of spacetime, and to consider it instead as an extended body, whose parts have both spatiotemporal properties and internal-space properties. Mathematically, this move is parametrisation: we remove the spacetime structure from the base space, and instead represent it as a further field — the “location field” on that space, giving the spacetime properties (in the first instance: the locations) of the various parts of the body. The move from a global to a local theory then proceeds in exactly the same way, mathematically, for spacetime as for internal symmetries; in this sense at least, gravity turns out to be fully a gauge theory.

Many aspects of this framework have been discussed previously. The standard mathematical treatments of relativistic particles and strings in a background spacetime proceed in this way: the base manifold is a (1- or 2-dimensional) bare manifold and its location in spacetime is a map from that base manifold to spacetime. Treating a four-dimensional field theory this way (again, with a fixed background spacetime) has been explored as a toy model for quantum gravity, notably in discussions of the problem of time and the nature of general covariance.

In the gauge principle is applied to the parametrised motion of a particle (cf also [13, 14]) while in discussions of Poincaré gauge theory it is known that additional fields can be added to the theory to ameliorate its differences from Yang-Mills theory, and that additional fields can be identified [15, 16] as the gauge versions of the spacetime-location field discussed above. So far as I am aware, however, there has not so far been an unified account of these various strands.

The structure of the paper is as follows. In sections II–III I briefly review the gauge principle as applied to internal symmetries and the reasons why theories with local spacetime symmetry in their normal formulations fail to be gauge theories in this strong sense. In section IV I give an account of the parametrised formalism applicable to global spacetime theories (in particular special relativity) and note some conceptual advantages of that process.

---

1 Some accounts of general relativity, e.g. [1, 2] embrace this distinction, treating spacetime as the (dynamically active) container against which physics plays out, and gravity as the curvature of that background. Others (e.g. [3, 4]) prefer to describe gravity as one more dynamical field, on a par with the fields of particle physics, and to downplay its distinctive aspects. I hope that the formalism of this paper provides a mathematical framework appropriate for the second conception.

2 The first proposal for a gauge theory of gravity appears to have been [5] and the theory was developed further by [6] and [7]; for technical details and further references see [8] and references therein.

3 The parametrised approach was first worked out in [9, 10]; for further development see, e.g., [11].
formalism. In section V I apply the gauge principle to the simplest of such theories: “pure location theory”, the trivial theory where the field has only locational properties; the resultant theory, kinematically speaking, is general relativity. In section VI I generalise this analysis to apply to matter fields, including vector, tensor and spinor fields, and in section VII I briefly consider some possible generalisations. Section VIII is the conclusion.

II. A REVIEW OF THE GAUGE PRINCIPLE

A (global, i.e. non-gauge) field theory can be understood as a function

$$\varphi : B \to V,$$  

(1)

where $B$ is the base space (usually physical spacetime) and $V$ is the target space whose points are the possible values of the field. For maximum generality I will assume $B$ is a differentiable manifold (possibly with additional structure, e.g. a metric) and that $V$ is a quite arbitrary mathematical space. In physics the most common choice is for $V$ to be a vector space of some description (perhaps equipped with an inner product or other such structure), or else the real or complex field; for illustrative purposes we will often use $V_N$, the $N$-dimensional real vector space equipped with a positive-definite inner product.

In most interesting cases, $V$ has a non-trivial group of automorphisms, i.e. one-to-one mappings that preserve its mathematical structure. The automorphisms of a vector space are the invertible linear maps; the automorphisms of a real vector space equipped with a positive-definite inner product.

Given a field value $\varphi(x) \in V$, we can distinguish between its absolute features, which are invariant under the automorphisms, and its relative features, which are not invariant. For $\varphi(x) \in V_N$, the absolute features are determined by the magnitude $|\varphi(x)|$; facts about the direction of $\varphi(x)$ in $V_N$ are relative, since they can be changed arbitrarily by the action of $\mathcal{G}_V$. Physically, the absolute features represent facts purely about the field at $x$, whereas the relative features represent facts about the field at $x$ compared to its value at other points: in the case of $\varphi(x) \in V_N$, it makes sense to speak of (and in principle measure) $|\varphi(x)|$ intrinsically, whereas the angle of $\varphi(x)$ can only be coherently discussed, or measured, relative to a coordinate system defined by the field elsewhere in space. (Note that in the physics literature global symmetry transformations are typically described as occurring relative to some external reference frame not itself acted on by the symmetry.)

(One form of) the gauge argument begins with disquiet that the non-absolute features of $\psi(x)$ and $\psi(y)$ can be directly compared, no matter the spatiotemporal separation of $x$ and $y$; as [17] puts it, this seems “at odds with the ‘spirit’ of relativity”. Leaving aside the conceptual justification of this move, what makes it possible mathematically is that $\psi(x)$ and $\psi(y)$ lie in the same mathematical space $V$, and so we can remove this automatic comparability by equipping each point $x \in B$ with its own copy $V_x$ of $V$. A field is no longer a function of $B$ to $V$, but an assignment to each $x \in B$ of a point $\tilde{\varphi}(x) \in V_x$. Mathematically, this makes $\tilde{\varphi}$ a section of a $V$-fibre bundle over $B$ with structure group $\mathcal{G}_V$.

Formulating our field theory just as a section of a fibre bundle throws away too much structure: we need to restore the ability to at least compare field values at infinitesimally close points $x, x + \delta x$, and we do so by introducing a $\mathcal{G}_V$ connection on the bundle, which allows us to parallel-transport a field value in $V_x$ along a path to $V_y$ so as to compare it directly to another such field value in $V_y$. Its precise mathematical formulation in the current rather abstract framework will not be needed; cf., e.g., [18].) There is no conceptual requirement that the parallel transport rule be independent of the path taken from $x$ to $y$ (to impose such a requirement makes sense only if we have not abandoned the idea that direct comparisons of field values at distant points are permitted), so the connection in general introduces additional degrees of freedom into the theory. A theory of this kind, consisting of a field defined on a fibre bundle and a connection on that fibre bundle, is a local (or gauge) field theory.

The fibre-bundle formulation is elegant, but calculationally awkward compared to the formulation of global fields. At least locally, we can return to the global for-

---

4 This restriction to scalar-valued fields (that is, excluding vectors or tensors on $B$) may seem overly restrictive; I will show in section VI that it is no real restriction.

5 For instance, “complex fields” are often represented as maps to the complex plane, but since those theories normally have a phase symmetry, the real/imaginary distinction does no physical work, and the target space of the complex field is better taken as a one-dimensional complex vector space.

6 To be more precise: the bundle here is an associated bundle over $B$, with typical fibre $V$, defined by some principal $\mathcal{G}_V$-bundle over $B$; there may, of course, be many such principal bundles.

7 Whether it can be done globally depends on the topology of the
normal by a choice of gauge: a smooth choice for each \( V_x \) of an automorphism \( f_x : V_x \to V \) to (a fixed copy of) \( V \). The field \( \tilde{\varphi}(x) \in V_x \) can now be represented by \( \varphi(x) = f_x \cdot \tilde{\varphi}(x) \in V \). The choice of gauge also fixes a representation of the connection as a covariant derivative \( D_\mu \) on \( B \times V \):

\[
D_\mu \varphi = \partial_\mu \varphi + A_\mu \cdot \varphi,
\]

where \( A_\mu \) is a one-form taking values in the Lie algebra \( g_V \) of \( G_V \), and \( \cdot \) is the Lie algebra representation of \( g_V \) induced by the representation of \( G_V \) on \( V \). The field theory is then jointly represented, in this gauge, by \( \varphi \) and \( A_\mu \). In the case of \( V_N \), the field is now (kinematically) a Yang-Mills theory with structure group \( SO(N) \) (or \( O(2) \) for \( N = 2 \)).

The choice of gauge, however, is purely conventional. Any gauge transformation between two gauges can be represented by a map \( \Lambda : B \to G_V \), and transforms \( \varphi \) and \( A_\mu \) as follows:

\[
\varphi(x) \to \Lambda(x) \cdot \varphi(x)
\]

\[
A_\mu(x) \to \Lambda(x) A_\mu(x) - \partial_\mu \Lambda(x) \Lambda^{-1}(x)
\]

where \( \text{Ad} \) is the adjoint action and \( \Lambda^{-1} \) acts on \( \partial_\mu \Lambda \) by right translation (for matrix groups, \( \text{Ad}(X)Y = XYX^{-1} \) and the action of \( \Lambda^{-1} \) is ordinary multiplication). We can then calculate that

\[
D_\mu \varphi(x) \to \Lambda(x) \cdot D_\mu \varphi(x)
\]

Notice that these gauge symmetries are purely internal, acting (independently) on each fibre and leaving the points of the base manifold \( B \) invariant. The full symmetry group of the theory will be the gauge symmetries together with any symmetries of \( B \) itself.

In calculations, it can be useful to eliminate this gauge freedom by gauge-fixing: picking, for each equivalence class of \( (\varphi, A_\mu) \) pairs under gauge transformations, a single preferred member. Examples in \( V_N \) are the Lorenz gauge in which we require that \( \partial^\rho A_\rho = 0 \) and that \( A_\mu \) vanishes at spatial infinity, and the unitary gauge in which we pick a particular direction in \( V \) and require \( \varphi(x) \) to lie in that direction, so that the only residual degree of freedom of \( \varphi \) is \( |\varphi| \).

Interpretationally, the field theory represents (a) the absolute features of the field at spacetime points, and (b) the path-dependent relations between values of the field at distinct points. These features are represented jointly and redundantly by \( \varphi \) and \( A_\mu \), but there is no general division of labour: it is not the case, in particular, that \( \varphi \) represents (a) and \( A_\mu \) represents (b), though this may be true in a particular choice of gauge-fixing (it is true for \( V_N \) in the unitary gauge). In many cases (specifically, when \( g_V \) acts freely on \( V \), perhaps except for some subset of \( V \) of measure zero) specification of \( \rho \) and \( D_\mu \rho \) (that is, specification of the absolute features of the field and of the relations between infinitesimally close points) suffices in general to allow us to solve for \( \varphi \) and \( A_\mu \) up to gauge transformations, and thus to specify the entire theory.\(^9\)

Before continuing I should note that the gauge argument does not uniquely fix the dynamics of the newly local theory, even given the dynamics of the global theory; however, insofar as classical field theory is of interest mainly as a route to quantum theory this is of comparatively little significance given that renormalisation considerations generally fix the form of a theory’s Lagrangian almost uniquely (i.e., up to finitely many experimentally determined parameters) once the theory’s symmetries are known. Throughout this paper I will be content to recover the kinematic form of a theory, leaving its dynamics to be determined by symmetry and renormalisation considerations.

### III. WHY (NORMAL FORMULATIONS OF) GENERAL RELATIVITY ARE NOT GAUGE THEORIES (IN THIS SENSE)

General relativity (GR) is often introduced by starting with Minkowski spacetime, representing it in differential-geometric form via a metric and an affine connection, and then dropping the requirement that the metric and connection are curvature-free. There is a clear family resemblance to the gauge argument, and much ink has been spilled trying to understand GR in this way: it has been considered, variously, as a gauge theory of the diffeomorphism group, the general linear group \( GL(4, \mathbb{R}) \), the translation group \( R^4 \), the affine group of linear maps and translations (formally \( R^4 \times GL(4, \mathbb{R}) \)), the Lorentz group \( SO(3, 1) \), the Poincaré group of Lorentz boosts and translations (\( P = R^4 \times SO(3, 1) \)), and probably others besides. These various theories undoubtedly have much in common with ‘standard’ gauge theories, and whether they are ‘real’ gauge theories can quickly become an exercise in semantics.

It is, however, possible without semantic tangles to ask and answer the question of whether GR, in a given formulation (and with or without matter), is a gauge theory in the specific sense of section II: that is, whether it is a field theory specified by a section of some \( V \)-bundle over a base manifold \( B \) together with a connection on that bundle, or (equivalently) whether it is specified, for some \( V \), by a function \( \varphi : B \to V \) and a one-form \( A_\mu \) taking values in \( g_V \), with (3–4) as symmetries. In at least the

---

\(^8\) There is a background metric assumed here.

\(^9\) In Yang-Mills theory, this can be seen by adopting the unitary gauge ([19]; see, e.g., [20] for details), which is defined only for \( \varphi \neq 0 \).
standard formulations, the answer is ‘no’, and the reasons are instructive. (For simplicity, henceforth ‘gauge theory’ means exclusively ‘gauge theory in my sense.’)

The traditional formulations of GR regard it as the theory of a metric tensor field $g_{\mu\nu}$ and a covariant derivative $\nabla_V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\tau} V^\tau;\text{ GR proper requires that } \nabla_V g_{\mu\nu} = 0, \text{ and } \Gamma^\nu_{\mu\tau} = 0 \text{ (‘torsion freedom’)} \text{ which suffices to fix } \nabla.$ Einstein-Cartan theory (or Einstein-Cartan-Sciama-Kibble theory) is obtained by dropping the torsion-freedom requirement, and metric-affine gravity by additionally dropping the assumption of metric compatibility.$^{10}$

These fail to be gauge theories for several reasons. Most significantly (and even if $g$ is set aside entirely) $\nabla_\tau$, while a covariant derivative of some kind, is not a covariant derivative in the sense of section II, since it is defined on the tangent bundle $TB$, and this bundle has more structure than the fibre bundles previously considered.

To elaborate, we can think of $TB$ as a vector bundle, with $T_xB$ a 4-dimensional vector space. But there is a preferred identification between vectors in the fibre at $x$ and infinitesimal paths in $B$ through $x$ — indeed, in standard formulations of the tangent bundle the vectors in $T_xB$ are identified with those infinitesimal paths, or equivalently (if mathematically more conveniently) with derivative operators along those paths. It is this relation between the base manifold and the tangent bundle that makes it possible to regard the derivative $\gamma(\lambda)$ of a path $\gamma(\lambda)$ through $B$ as lying in $T_{\gamma(\lambda)}B$.

This shows up in the mathematical form of the connection coefficients $\Gamma^\nu_{\mu\tau}$. This can be thought of as a $GL(4,\mathbb{R})$-valued 1-form (with the $\mu,\tau$ coefficients interpreted as $GL(4,\mathbb{R})$ indices and $\nu$ as a 1-form index). But the indices really index the same space, so that in particular it makes sense to impose torsion-freedom, a requirement with no analogue for internal connections (since there the lower indices of the connection index coordinates in different spaces).

It shows up also in the automorphisms of the bundle, and hence in the symmetries of the theory. Whereas in section II there was a sharp separation between symmetries of the base manifold, and internal symmetries that leave base-manifold points unchanged, in this case there are no purely internal symmetries, since any non-trivial transformation of $T_xB$ will break the relation between vectors and infinitesimal curves. The only automorphisms of the tangent bundle are those induced by diffeomorphisms of the base manifold.

(The literature is inconstant as to whether the tangent bundle should be called a fibre bundle or not; for clarity, henceforth I call fibre bundles in the original sense internal fibre bundles.)

More flatfootedly, the metric field $g_{\mu\nu}$ is neither a section $\varphi$ of an internal fibre bundle, nor a connection of any kind. (It can be understood as a section of the $(0,2)$ tensor bundle over $B$, but this, like the tangent bundle, is not an internal fibre bundle).

The tetrad formalism of GR$^{11}$ starts with the 4-dimensional Minkowski vector space $\mathcal{M}$, equipped with metric $\eta_{ab},$ and with the Lorentz group SO(3,1) as its automorphism group. (I use Roman letters here as indices in $\mathcal{M}$.) The theory is formulated on a fibre bundle over $B$ with typical fibre $\mathcal{M}$, and so looks rather more like a gauge theory; it is specified (again setting aside matter fields) by a tetrad $e$ (a one-form taking values in the fibres, i.e. $e(x) : T_xB \to \mathcal{M}, B$) and a covariant derivative on $\mathcal{MB}$. With respect to a particular gauge (and using Roman indices to index vectors in $\mathcal{M}$) the tetrad is represented by a 1-form $e^a$ taking values in $\mathcal{M}$, and the covariant derivative by a SO(3,1)-valued 1-form $\omega^a_b,$ as in

$$\nabla_\mu = \partial_\mu + \omega^a_b \partial_b.$$

From here there are several ways to connect connection and tetrad. Conventional GR in tetrad form is reproduced by imposing the torsion-freedom constraint

$$T[e, \omega] = de^a + \omega_a^b \wedge e^b = 0,$$

which determines $\omega$ uniquely. Teleparallel gravity$^{12}$ is reproduced by instead imposing the condition $\omega = 0$, so that the connection has no curvature in the usual sense ($\omega = 0$ is not invariant under local Lorentz transformations, so this is not an invariant specification.) In the more general case of Poincaré gauge theory$^{13}$ the tetrad and the connection are treated as independent fields.

$\omega$, by itself, is an SO(3,1) connection in the full sense of section II, with exactly the expected transformation properties: in particular, the theory has, as expected, local SO(3,1) transformations as symmetries. Furthermore, these remain symmetries even when $e^a$ is included: if $\Lambda : \mathcal{B} \to \text{SO}(3,1)$ is a gauge transformation, the overall symmetry is

$$e^a(x) \rightarrow \Lambda^a_b(x)e^b(x) \quad, \quad \omega^a_b(x) \rightarrow \Lambda^a_b(x)\omega^b_d(x)(\Lambda^{-1})^d_a(x) - \partial_d\Lambda^a_b(x).$$

But the $e^a$ field spoils tetrad gravity’s pretensions to be an SO(3,1) gauge theory in our sense. It is not a $\mathcal{M}$-valued function but an $\mathcal{M}$-valued one-form, and so cannot be understood as a section of the $\mathcal{M}$-bundle. In fact,

\[10\] For the traditional formalisms see, e.g., [1] or [21]; for the generalisations, see [8] and references therein.

\[11\] Reviewed in, e.g., chapter 2 of [3].


\[13\] x
it has a rather different natural geometric interpretation which belies the interpretation of that bundle as an internal bundle: for each $x$ it gives a 1-1 map from $T_xB$ to $M$ of $B$, allowing us to identify the Minkowski vector bundle with the tangent bundle. From this perspective the SO(3,1) freedom rather mundanely represents our freedom to coordinatise the tangent bundle as we like.

An alternative is to interpret $e^a$ as another connection: the translational part of the Poincaré group $P$. (This is the main conceptual idea behind teleparallel gravity.) From this perspective $(\omega,e)$ is a one-form taking values in the Lie algebra of $P$. Suggestively, if we formally calculate the curvature under this interpretation we get the $p$-valued 2-form $(R,T)$, where $R$ is the curvature of the $SO(3,1)$ connection and $T$ is the torsion. However, the theory still cannot be a gauge theory in our sense: the fibre $M$ is a vector space, and does not bear any representation of $P$. We can attempt to interpret the translations as acting on the base manifold $B$ instead, taking a point $x^\mu$ to $x^\mu + \delta x^\mu$ (this is the standard route used to arise at Poincaré gauge theory), but invariance under this transformation is just infinitesimal diffeomorphism invariance, and does not have the purely internal form we require for a gauge symmetry.

These various theories of gravity have in common that they represent gravitation and curvature on the base space of the theory, and hence make use of the tangent bundle. In the next section I will show how the concept of parametrisation lets us put gravitational physics on a par with other gauge theories, by moving the geometry entirely off the base space.

**IV. PARAMETRISED FIELD THEORY**

Field theories as standardly conceived are assignments of quantities to spacetime points, and as such differ sharply from particles or strings, which are extended bodies parts of which occupy spacetime points. A particle in special relativity is represented as a map from its one-dimensional worldline to Minkowski spacetime; a string, as a map from its two-dimensional worldsheet to spacetime. In each case, the underlying worldline or worldsheet is just a bare differentiable manifold, and the theory therefore has two distinct sets of symmetries: a diffeomorphism symmetry of the underlying manifold, and the global symmetries of the spacetime. In fact, particles and strings are global field theories in the sense of section II, and hence make use of the tangent bundle. In the next section I will show how the concept of parametrisation lets us reexpress this in the more familiar form

\[
\mathcal{L}(\bar{\phi}, \partial_\mu \phi, \psi, \partial_\mu \psi) = \frac{1}{2} \eta^{ab} \partial_a \psi \partial_b \psi - V(\psi),
\]

and writing $g_{\mu\nu} = \eta^{ab} e^a_\mu e^b_\nu$ lets us reexpress this in the more familiar form

\[
\mathcal{L}(g_{\mu\nu}, \psi, \partial_\mu \psi) = \left( \frac{1}{2} g^{\mu\nu} \partial_\mu \psi \partial_\nu \psi - V(\psi) \right) \sqrt{\text{det}g}.
\]

$\rho^\mu$ is the location field, which says where in spacetime any given part of the body manifold is.

Parametrisation has mostly been seen as a rather cheap way to make a theory generally covariant and discussed mostly as a toy theory for quantum gravity [9–11]. However, it has some virtues conceptually even aside from the gauge argument:

- It is most naturally thought of as representing a field as an extended body whose parts occupy various spacetime points and have various nonspatiotemporal properties. As such, it treats spacetime and internal degrees of freedom very much on

---

14 More formally, affine Minkowski space is a set together with a free transitive action on that set of $M$ regarded as an additive group.
a par: Minkowski spacetime and the ‘real’ internal space $\mathcal{V}$ both parametrise properties that given parts of the field might have.

- It provides for a rather clean separation between the substantive symmetries of the theory (which are represented by the automorphisms of spacetime $\mathcal{A}\mathcal{M}$ and of the internal space $\mathcal{V}$) and the diffeomorphism symmetry, which can be made a symmetry of any spacetime theory through appropriate reformulation. We have made the theory generally covariant, but without reducing the spacetime symmetries to special cases of the diffeomorphism symmetries.

- It makes for a cleaner understanding of Noether’s theorem, since the distinction between “dependent” and “independent” variables is eliminated: all variables are dependent variables, even the spacetime one.\(^{15}\) For instance, the infinitesimal symmetries of a parametrised scalar field theory $\varphi = (\rho^a, \psi)$ is just

$$\rho^a \rightarrow \rho^a + \xi^a; \quad \psi \rightarrow \psi$$

and the associated Noether current is

$$J^a = \frac{\partial L}{\partial \rho^a}.$$  \hspace{1cm} (15)

Assuming that $L = L(\rho^a)$ and that $L$ depends on $\rho$ only through $g_{\mu\nu}$, this yields the familiar-looking

$$J^\mu = e^\rho_\mu \left( \frac{\partial L}{\partial \rho^a} \right) \sqrt{\text{det} g}.$$  \hspace{1cm} (16)

It allows us to treat vector, tensor and spinor fields (in special relativity) straightforwardly as functions, as we will see in section VI.

V. GAUGING PURE LOCATION THEORY

The simplest parametrised field theory is the trivial ‘pure location theory’, where the only field is the location field $\rho : \mathcal{B} \rightarrow \mathcal{A}\mathcal{M}$. Under the additional requirement that $\rho$ is a smooth 1 : 1 mapping, this is just empty Minkowski spacetime in parametrised form; however, the gauge recipe of section II still applies directly to it and yields a highly non-trivial theory (at least kinematically; I continue to leave the dynamics unspecified). The details are as follows:

1. We localise the theory by replacing $\rho$ with a section $\tilde{\rho}$ of a bundle over $\mathcal{B}$. The resultant bundle $\mathcal{A}\mathcal{M}\mathcal{B}$ has typical fibre $\mathcal{A}\mathcal{M}$: it is an affine bundle, rather than the vector bundles more familiar from Yang-Mills theory. It is still, however, an internal bundle, with no particular connection to the tangent bundle $\mathcal{T}\mathcal{B}$.

2. We introduce a connection on $\mathcal{A}\mathcal{M}\mathcal{B}$ to allow us to compare the values of the location field $\tilde{\rho}$ at infinitesimally close points. With respect to a choice of gauge (that is, a smooth choice for every $x$ of a map $f_x : \mathcal{A}\mathcal{M}\mathcal{B} \rightarrow \mathcal{A}\mathcal{M}$, $\tilde{\rho}$ is represented by a function $\rho : \mathcal{B} \rightarrow \mathcal{A}\mathcal{M}$ and the connection by the covariant derivative

$$\nabla_\mu \rho^a = \partial_\mu \rho^a + (A_\mu \cdot \rho)^a,$$  \hspace{1cm} (18)

defined in terms of a one-form $A_\mu$ taking values in the Lie algebra of the Poincaré group.

Given some arbitrary choice 0 of origin of $\mathcal{A}\mathcal{M}$, we can decompose any Poincaré group element into a rotation around 0 and a translation, $A_\mu$ can then be broken into infinitesimal rotational and translational parts $(\omega^a_{\mu b}, \tau^a_\mu)$, and the partial derivative further decomposed as

$$\nabla_\mu \rho^a = \partial_\mu \rho^a + \omega^a_{\mu b} \rho^b + \theta^a.$$  \hspace{1cm} (19)

A gauge transformation is a map $\Lambda : \mathcal{B} \rightarrow \mathcal{P}$, which we can decompose as $\Lambda(x) = (R^a(x), \xi^a(x))$, and its action on the location field and the connection can be read off from (3-5), or else verified explicitly (for readability I use matrix notation and suppress the explicit Minkowski indices):

$$\rho(x) \rightarrow R(x) \rho(x) + \xi(x)$$  \hspace{1cm} (20)

$$\omega_\mu(x) \rightarrow \Lambda(R(x)) \omega_\mu(x) - \partial_\mu R(x) R^{-1}(x)$$  \hspace{1cm} (21)

$$\theta(x) \rightarrow R(x) \theta(x) - \Lambda(R(x)) \omega_\mu(x) \xi(x)$$  \hspace{1cm} (22)

$$\nabla_\mu \rho(x) \rightarrow R(x) \nabla_\mu \rho(x).$$  \hspace{1cm} (23)

(If you were expecting a translation term in (23), note that the derivative of a function to $\mathcal{A}\mathcal{M}$ takes values in the tangent space of $\mathcal{A}\mathcal{M}$, i.e. the Minkowski vector space $\mathcal{M}$, and the translational part of $\mathcal{P}$ acts trivially on $\mathcal{M}$. In the case of Yang-Mills theories where the target space is a vector space, this subtlety does not arise since a vector space is its own tangent space.)

This local pure location theory is a gauge theory in the full sense of section II, defined on an internal bundle and with “matter” field and connection transforming exactly as the gauge recipe stipulates. It has the full local Poincaré group (the group whose elements are maps

\(^{15}\) If desired we can treat the diffeomorphism symmetry as an independent-variable transformation — but the Noether currents associated to that symmetry vanish identically in any case.
from the base space to the Poincaré group) as a symmetry group, in addition to another symmetry group generated by the diffeomorphisms on the body manifold $B$, and the two have conceptually entirely different status: the former follows from the global Poincaré symmetry on the fibres and the imposition of locality, the second from the fact that points on the body manifold have no absolute properties of their own and so can be smoothly permuted without affecting the physics.

The curvature of the connection is as usual a Poincaré-valued 2-form, and can be decomposed into its rotational and translational parts $R_{a\mu\nu}, T_{a\mu\nu}$: in differential-form notation, it can be evaluated explicitly as

\[ R_b^a = d\omega_b^a + \omega_c^a \wedge \omega_d^b, \tag{24} \]
\[ T_a^{\mu\nu} = d\theta^a + \omega_b^a \wedge \theta^b. \tag{25} \]

Since the Poincaré group acts transitively on $\mathcal{AM}$, the location field $\rho(x)$ has no meaning by itself: any two locations are intrinsically identical. (This was also true before gauging the theory.) What has physical significance is the covariant derivative, which represents the displacement between two infinitesimally close points on the body manifold. Integrating this displacement over finite paths in the usual way for gauge theories (i.e., keeping track of noncommutativity) will give the displacements between finitely separated points, but now that the theory has been localised the displacement will in general be path dependent. The curvature then gives the infinitesimal anholonomy of the location field: the net displacement between a point and itself around an infinitesimal curved path:

\[ T_{a\mu\nu} = R_{b\mu\nu}^a \rho^b + T_{a\mu\nu}. \tag{26} \]

(It does so rather redundantly — ten components of curvature provide only a four-component displacement — reflecting the fact that the Poincaré group does not act freely on Minkowski space: given any two points, there are a large number of different Poincaré group actions that map one to the other. When we consider more complicated fields in section VI, this redundancy will be eliminated.) Anticipating later results, we call this infinitesimal anholonomy the torsion. Note that it cannot be identified simply with the translational part of the curvature: it contains both translational and rotational parts.

We can connect pure location theory to familiar physics by a choice of gauge. Specifically, let us adopt the stationary gauge, in which $\rho(x) = 0$ (that is, equals the arbitrarily chosen ‘origin’ of $\mathcal{AM}$) for all $x$. This is not a complete fixing of gauge: it breaks the local Poincaré symmetry down to the local Lorentz group, since any Lorentz transformation leaves the origin invariant. (There is an analogy with the breaking of $SU(N)$ down to $U(1)$ in internal symmetry breaking: $U(1)$ is the subgroup of $SU(N)$ under which the (arbitrarily-selected) lowest-energy state is invariant.)

In the stationary gauge, we have simply

\[ \nabla_{\mu} \rho^a = \theta_{\mu}^a \tag{27} \]
\[ T_{a\mu\nu} = (d\theta^a + \omega_b^a \wedge \theta^b)_{\mu\nu}: \tag{28} \]

in other words, in the stationary gauge the translational part of the connection represents the torsion of the location field. If we write

\[ e_{\mu}^a = \nabla_{\mu} \rho^a, \tag{29} \]

then since the rotational part of the connection is translationally invariant, we obtain a translationally invariant expression for the torsion:

\[ T^a = (de^a + \omega_b^a \wedge e^b). \tag{30} \]

The covariant derivative $e$ and the rotational connection $\omega$ provide a complete specification of the theory up to gauge invariance, and can of course be recognised (kinematically) as the tetrad and connection of tetrad-formalism gravitation. Or to put it another way around: the location field allows us to interpret the tetrad as the covariant derivative of the location field, and the connection as the rotational part of the Poincaré connection for that field. The tetrad can be identified directly with the translational part of the connection only in the stationary gauge (which is reflected in the well-known fact that it transforms homogeneously under translations).

Similarly, expression (30) can be recognised as the standard definition of the torsion. The parametrised formalism then illustrates in a very simple and direct way the interpretation of torsion as giving the infinitesimal translation around a closed loop.

The stationary gauge also allows us to deduce that pure location theory is specified entirely, up to gauge freedom, by the covariant derivative and the torsion. For (30) can be solved for $\omega$, yielding the full connection in the stationary gauge (and of course $\rho$ is trivially known in that gauge). In the case where the torsion is constrained to vanish, we (kinematically) recover general relativity. Conversely, by requiring the rotational part of the connection to vanish ($\omega = 0$) we can express the theory entirely in terms of $e$. Such a theory is a gauge theory of pure translation: since the translation group acts freely on $\mathcal{AM}$ we would expect the covariant derivative of $\rho$ to specify the complete theory, and indeed it does: torsion is given gauge-invariantly by

\[ T^a = d(\nabla \rho^a) \equiv de^a. \tag{31} \]

This theory can be recognised as teleparallel gravity (note that the $\omega = 0$ recipe is not invariant under local Lorentz invariance, reflecting the well-known Lorentz-dependence of the Weitzenbock connection of teleparallelism).
VI. VECTOR, TENSOR, AND SPINOR FIELDS

A vector field on Minkowski spacetime can be treated as a map $\varphi = (\rho, \psi)$ where $\rho$ takes values in Minkowski spacetime $\mathcal{AM}$ and $\psi$ in the Minkowski vector space $\mathcal{M}$. The map assigning an element $(x - y)$ of $\mathcal{M}$ to any $x, y \in \mathcal{AM}$ serves to ‘solder’ the two together and distinguishes the vector field from a purely internal vector field. Conceptually, each part of the body manifold has both a location in spacetime and a vector field value.

It will be convenient to regard vectors as living in their own copy of $\mathcal{M}$, equipped with a preferred identification with the copy of $\mathcal{M}$ used to define $\mathcal{AM}$, so that parametrised fields take values in $\mathcal{AM} \times \mathcal{M}$; similarly, tensor fields take values in the tensor products of Minkowski vector space equipped with such a preferred identification. I write $\mathcal{AM} \times \mathcal{V}$ for any vector or tensor space so linked to Minkowski space.

This framework is extremely similar to the way we represented scalar fields with internal degrees of freedom: in both cases the field is a map associating to each point of the base manifold both a location in $\mathcal{AM}$ and a field strength in some vector space. The difference is that the preferred identification means that the automorphism group of $\mathcal{AM} \times \mathcal{B}$ is not simply the product of the automorphism groups of $\mathcal{AM}$ and $\mathcal{B}$: rather, the automorphism group is the Poincaré group, with the translational part acting on $\mathcal{AM}$ alone and the rotational part acting jointly on $\mathcal{AM}$ and $\mathcal{B}$.

For spinors, we take $\mathcal{C}$ to be the 2-dimensional complex vector space, and fix a preferred bilinear map $e^{\alpha}_{\alpha'}$, from $\mathcal{C} \otimes \mathcal{C}^*$ to $\mathcal{M}$. Parametrised spinor fields now take values in the space $\mathcal{AM} \times \mathcal{C}$ equipped with this map; the automorphism group is now $\mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C})$, the double covering of the Poincaré group. I extend the notation $\mathcal{AM} \times \mathcal{V}$ to include the spinor case also.

Finally, for vector fields having additional internal degrees of freedom represented by some vector space $\mathcal{U}$, the field strengths take values in $\mathcal{M} \otimes \mathcal{U}$, with this copy of $\mathcal{M}$ again identified in a preferred way with the copy of $\mathcal{M}$ used to define $\mathcal{AM}$; this extends naturally to spinor and tensor fields with internal degrees of freedom, and again I extend the $\times$ notation to cover this case.

This somewhat abstract discussion is perhaps clearer in index form: a general tensor field on Minkowski spacetime is represented in the parametrised format by a location field $\rho^{\alpha}$ together with some object

$$\psi_{b_1 \ldots b_m \alpha_1 \ldots \alpha_k}^{a_1 \ldots a_n \beta_1 \ldots \beta_l}$$

where $a_i, b_i$ are spacetime indices and $\alpha_i, \beta_i$ are internal space indices. An infinitesimal internal symmetry acts on the $\alpha_i, \beta_i$ indices and leaves $\rho^\alpha$ invariant; an infinitesimal translation acts on $\rho^\alpha$ alone; an infinitesimal Lorentz transformation acts on the spacetime indices of $\psi$ and on $\rho^\alpha$. For spinor fields, the $a_i$ and $b_i$ become spinor indices.

Finally, the most general case of multiple such fields is represented by a tuple

$$(\rho, \psi_1, \ldots, \psi_n)$$

where the $\psi_i$ are vector or spinor fields of the above form.

A global theory formulated in this way has as its symmetry group, as we would expect, the product of (i) the Poincaré group $\mathcal{P}$; (ii) the internal symmetry group $\mathcal{G}_{\text{int}}$, which acts trivially on $\mathcal{AM}$; (iii) the diffeomorphism group.

Even a theory as complicated as this is still ultimately a map from $\mathcal{B}$ to some space $\mathcal{V}$ on which some symmetry group $\mathcal{G}$ (in this case $\mathcal{P} \times \mathcal{G}_{\text{int}}$) acts, and so can still be localised via the gauge argument. This yields as usual a theory on a fibre bundle with typical fibre $\mathcal{V}$, equipped with a connection $\mathcal{A}_\mu$ taking values in the direct sum of the Lie algebras of $\mathcal{P}$ and $\mathcal{G}_{\text{int}}$. This connection may be decomposed into a triple

$$A_\mu = (\theta_\mu, \omega_\mu, B_\mu)$$

where $B_\mu$ is a one-form taking values in $\mathfrak{g}_{\text{int}}$ and $\theta_\mu, \omega_\mu$ are the translational and rotational parts of the Poincaré connection introduced in section V. The covariant derivative defined by that connection acts in the standard way:

$$D(\rho, \psi_1, \ldots, \psi_n) = \partial(\rho, \psi_1, \ldots, \psi_n) + A_\mu \cdot (\rho, \psi_1, \ldots, \psi_n).$$

In the simple case of a single vector field $\psi^a$ with no internal degrees of freedom, for instance, we have

$$D_\rho(\rho^a, \psi^a) = (\partial_\rho \rho^a + \omega^a_{\mu \beta} \psi^\beta + \theta^a_{\mu \beta} \rho^\beta, \rho_\mu \psi^a + \omega^a_{\mu \beta} \psi^\beta) \equiv (e^a_\mu, D^\mu_\psi)$$

where following section V we continue to write

$$\epsilon^a_\mu \equiv \partial_\mu \rho^a + \omega^a_{\mu \beta} \psi^\beta + \theta^a_{\mu \beta}$$

and we define

$$D^\mu_\psi \psi^a \equiv \partial_\mu \psi^a + \omega^a_{\mu \beta} \psi^\beta$$

as the Lorentzian covariant derivative (it generalises in the obvious way to tensor and spinor fields). It follows from the general structure of gauge theory that this covariant derivative transforms via the tangent representation of $\mathcal{P}$ (that is, the Lorentzian part) under a gauge transformation $A(x) = (R^a_\mu(x), \xi^a(x))$: indeed,

$$(e^a_\mu, D^\mu_\psi) \rightarrow (e^a_\mu, D^\mu_\psi + (R^a_\mu e^a_\mu, R^a_\psi)).$$

The gauge and location fields, meanwhile, transform just as in section V, while $\psi$ transforms under the Lorentzian part of the gauge transformation. The theory, that is, is a gauge theory of the Poincaré group in a totally standard sense, with symmetry group being the semidirect product of the local Poincaré group and the diffeomorphism group.

If, however, we move to stationary gauge, we can eliminate $\rho$ entirely from the theory, and equate $e^a_\mu$ with the translational part of the gauge connection, the local Poincaré symmetry is broken down to a local Lorentz symmetry and the transformation laws become (suppressing indices)

$$\epsilon_\mu \rightarrow \epsilon_\mu + R \cdot \epsilon_\mu$$

$$\omega_\mu \rightarrow \omega_\mu + \text{ad}(\omega_\mu)R - \partial_\mu R$$

$$\psi \rightarrow \psi + R \cdot \psi$$

where $\epsilon_\mu$ is a 4-vector, $\omega_\mu$ is a 4-form, and $R$ is a 4x4 Lorentz matrix.
together with the usual action of the diffeomorphism group. This is the familiar symmetry group of tetrad gravity with dynamical connection: it is a hybrid theory, with some features of Lorentzian gauge theory but with the tetrad not identifiable straightforwardly either as a matter field or as a gauge connection.

(Given the transformations (40–42) it is possible, conversely, to ask how the theory can be modified to allow \( e_\mu \) to be interpreted as the translational part of a Poincaré connection. This can be done by introducing auxiliary fields\([15, 16]\), which from the present viewpoint may be identified as the components of the location field.)

The full framework of location field, ordinary field strength, and translational and rotational connection is, of course, more complicated than the normal tetrad form, and contains a redundancy that for calculational purposes is no doubt usually best eliminated by moving to the stationary gauge. It is gratifying, however, to see that when the full framework is kept in view, our theory (at least kinematically) is simply, straightforwardly, a gauge theory of the Poincaré group, and interesting to note how, as elsewhere in physics, choice of a specific gauge obscures the symmetry structure even as it makes the theory’s dynamical degrees of freedom more transparent.

Finally, note that the gauge theory we have constructed includes (or could include) two very different kinds of “vector fields”:

1. Matter fields where the field’s internal space includes a vector degree of freedom transforming under the Lorentz symmetry;

2. Gauge connections (Poincaré or internal) represented by one-forms.

Conceptually the former is a property of a spacetime point; the latter represents the infinitesimal relationship between points. For the former, the ability to represent it as a tangent-space object (via \( V^a = V^\mu e_\mu^a \)) is incidental and relies on an additional spatiotemporally extended object, the covariant derivative of the location field. For the latter, its (co)tangent nature is essential to its role in relating two different (infinitesimally separated) point.

It is therefore unsurprising (as well as following directly from our analysis) that the two sorts of vector field couple rather differently to the spacetime covariant derivative. In moving from flat to curved spacetime, ordinary derivatives of matter fields, including vector fields, get replaced by covariant derivatives, but (as is well known in the Poincaré gauge-theory literature; cf. the discussion in [26]) ordinary derivatives of gauge connections are not so replaced: indeed, the expression

\[
\nabla_\mu A_\nu - \nabla_\nu A_\mu = (\partial_\mu A_\nu - \partial_\nu A_\mu) - T^\sigma_{\mu\nu} A_\sigma
\]

is not gauge invariant unless the torsion vanishes. The gauge connection is itself part of a unified covariant derivative, representing that part of the derivative concerned with internal-space relations.

**VII. POSSIBLE GENERALISATIONS**

Richard Feynman once observed [27, p.168] that different formulations of the same theory, even if they are mathematically interchangeable, can point towards different generalisations to new physics. The parametrised approach to gravity suggests at least two possibilities for generalisation of extant physics.

Firstly, it provides an interesting way to understand the perennial idea that spacetime is discrete at a fundamental level. Typically, fundamental discreteness is understood as replacing Minkowski spacetime, or perhaps a Riemannian spacetime, with some discretised approximation; such a move inevitably breaks the translational, rotational and boost symmetries of spacetime. The parametrised approach offers an alternative: keep Minkowski spacetime continuous but discretise the base manifold \( B \). In this approach, that base manifold is replaced by a finite graph: a collection of points linked by nodes. If the symmetry group is \( G \) and the typical fibre is \( \mathcal{V} \), then

- Each point \( x \) gets its own copy \( \mathcal{V}_x \) of \( \mathcal{V} \);
- Matter fields assign to each \( x \) a point \( \varphi(x) \) in \( \mathcal{V}_x \);
- The gauge connection is an assignment, to each link \( x \rightarrow y \), of a structure-preserving map from \( \mathcal{V}_x \) to \( \mathcal{V}_y \);
- A choice of gauge is a map from each \( \mathcal{V}_x \) to (a shared copy of) \( \mathcal{V} \);
- Relative to a choice of gauge, matter fields assign a point of \( \mathcal{V} \) to every node; gauge connections assign an element of the automorphism group of \( \mathcal{V} \) to every link.

For internal gauge symmetries, this is in fact roughly the formalism already used in lattice gauge theory, though there the lattice is taken to be a lattice of points in Minkowski spacetime. There is also a certain similarity to the lattices used in loop-space quantum gravity, though there the connection takes values in the Lorentz rather than the Poincaré group.

Secondly, while in parametrised versions of extant field theories the translational part of the Poincaré group always acts on a pointlike “location field”, the general framework requires only a fibre \( \mathcal{V} \) on which that group acts in some way or other. We could, for instance, take \( \mathcal{V} \) to be the space of embeddings of a closed loop into Minkowski spacetime, on which the Poincaré group acts in the obvious way. A field theory with this fibre would assign a loop to every point on the base space, either lying in the same copy of Minkowski spacetime (for the global theory) or in a local copy. The Poincaré connection could be defined just as before, giving the relations between loops at infinitesimally separated spacetime points. The connections (if any) between this sort of “loop-valued field theory” and the superficially-similar string field theory lie beyond the scope of this paper.
VIII. CONCLUSION

The parametrised approach to field theories allows spacetime and internal symmetries to be treated very much on a par, and in particular allows general relativity and related theories of gravity to be seen as gauge theories of the Poincaré group in just the way that Yang-Mills theories are gauge theories of internal global symmetries.

This is not to say that there are no disanalogies between the two classes of theory. This paper has ignored entirely the global structure of general relativity, and in particular the various issues surrounding black holes and singularities. However, at the local level it is gratifying to see Poincaré and Yang-Mills gauge theories as simply applications of the same general formula to different symmetry groups.

This suggests in turn that other differences between the theories — again, at least at the local level — can be ascribed to differences in their respective dynamics rather than to differences at the kinematic level. In a companion paper [28] I provide a unified account of the dynamics of the two types of gauge theory from the Lagrangian view and obtain results in support of this suggestion.