GENERAL COVARIANCE, ARTIFICIAL GAUGE FREEDOM AND EMPIRICAL EQUIVALENCE

A Dissertation

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

by

James Brian Pitts

Don Howard, Director

Graduate Program in History and Philosophy of Science
Notre Dame, Indiana
July 2008
GENERAL COVARIANCE, ARTIFICIAL GAUGE FREEDOM AND
EMPIRICAL EQUIVALENCE

Abstract

by

James Brian Pitts

This dissertation updates the debate over the nontriviality of general covariance for Einstein’s General Theory of Relativity (GTR) and considers particle physics in the debate over underdetermination and empirical equivalence. Both tasks are tied to the unexplored issue of artificial gauge freedom, a valuable form of descriptive redundancy.

Whereas Einstein took general covariance to characterize GTR, Kretschmann thought it merely a formal feature that any theory could have. Anderson and Friedman analyzed substantive general covariance as the lack of absolute objects, fields the same in all models. Some extant counterexamples and a new one involving the electron spinor field are resolved. However, Geroch and Giulini diagnose an absolute object in GTR itself in the metric’s volume element. One might instead analyze substantive general covariance as formal general covariance achieved without hiding preferred coordinates as scalar “clock fields,” recalling Einstein’s early views. Theories with no metric or multiple metrics make the age of the universe meaningless or ambiguous, respectively, so the ancient and medieval debate over the eternity of the world should be recast.

Particle physics provides case studies for empirical equivalence. Proca’s electromagnetism with some nonzero photon mass constitutes a family of rivals to
Maxwell’s theory. Whereas any Proca theory can be distinguished empirically from Maxwell’s, the Proca family approaches Maxwell’s for small masses, yielding permanent underdetermination with only approximate empirical equivalence. The weak nuclear force also displays a smooth massless limit classically, but not after quantization, recalling the instability of empirical equivalence under change of auxiliary hypotheses. The standard electroweak theory apparently permits a photon mass term and hence underdetermination, but possible further unification might not. The question of underdetermination regarding massive gravity is unresolved.

Physicists often reformulate a theory to install additional fields and new gauge symmetries preserving empirical content. Post-positivist philosophers might judge the result a distinct and inferior theory, but physicists consider it a perhaps superior formulation of the same theory. Evidently a step back from naïve realism towards mathematical-empirical equivalence is appropriate for physical ontology and theory individuation. Artificial gauge freedom licenses a generalized Kretschmann objection, but the clock field case suggests a resolution.
For Dilkushi, more precious than jewels.
## CONTENTS

ACKNOWLEDGMENTS ......................................................... v

CHAPTER 1: INTRODUCTION .................................................. 1
  1.1 Overview ............................................................... 1
  1.2 General Covariance .................................................. 4
  1.3 Empirical Equivalence ............................................... 10
  1.4 Artificial Gauge Freedom .......................................... 14

CHAPTER 2: GENERAL COVARIANCE, ANDERSON-FRIEDMAN ABSOLUTE OBJECTS, VARIATIONAL PRINCIPLES & CLOCK FIELDS 18
  2.1 Absolute Objects, Admissible Coordinates and Groupoids ........ 18
  2.2 Tensor Densities .................................................... 37
  2.3 Confined Objects and Global Space-time Topology .................. 49
  2.4 Jones-Geroch Vector Field Counterexample and Friedman’s Reply 50
  2.5 Redefinition with Variational Principle, Maidens’s Worry, and the Rosen-Sorkin Answer ........................................ 53
  2.6 Eliminating Local Irrelevance Excludes the Geroch-Jones Vector Field from Dust and Other Fluids ................................. 62
  2.7 Torretti’s and Norton’s Constant Curvature Space Counterexamples Do Have Absolute Objects ............................... 71
  2.8 Geroch-Giulini Scalar Density Counterexample $\sqrt{-g}$ in GTR ... 73
  2.9 Bach-Weyl Gravity Lacks $\sqrt{-g}$ .................................. 78
  2.10 Implications of Non-Metric or Multi-Metric Theories for Ancient-Medieval Debate Over the Eternity of the World .................. 80
  2.11 Which Fields to Test for Absoluteness? Equivalence, Irreducibility, Concomitants, etc. ............................................. 83
  2.12 Absoluteness and Nonvariationality? .............................. 97
  2.13 Tetrad-Spinor Counterexample and Resolution via Nonlinear Geometric Objects and the Ogievetsky-Polubarinov Spinor Formalism 103
  2.14 Unwanted Vanishing Absolute Objects: Concomitant Irrelevance . 115
  2.15 Conclusions about Anderson-Friedman Absolute Objects Program 120
ACKNOWLEDGMENTS

I am very grateful to my advisor Don Howard for his wise guidance and support at all stages of this work. Katherine Brading’s great interest has been most helpful. Ikaros Bigi has served as an oracle in particle physics whenever I have consulted him. In return I can only offer the conclusions that particle physics is important for the philosophy of space-time and the general philosophy of science. Anja Jauernig has helped greatly to extract the philosophical lessons out of technical issues. William McGlinn’s encouragement is appreciated.

Robert Geroch’s influence has been enormously beneficial on chapter 2. His counterexample to the Anderson-Friedman program radically transformed an ancestor of that chapter (Pitts, 2006a). The present work contains in section 2.11 a theorem and proof outline that he donated. Harvey Brown and Oliver Pooley have been helpful in many ways beyond what the text can show. Discussion with John Norton on underdetermination and empirical equivalence has been helpful. Ted Jacobson, Charles Misner and Aron Wall have contributed insightful probing on the material of chapter 2. I thank Roger Jones for discussion about his counterexample to the Anderson-Friedman project, David Malament for contributing helpful perspective on the Anderson-Friedman project, and many people with whom I have discussed the matter of spinors in curved space-time. Discussion with William Lane Craig regarding the debate over the eternity of the world is appreciated. All views expressed here are my own.
I thank the University of Notre Dame for a Presidential Fellowship that made the conditions of the production of this work pleasant. The Reilly Center for Science, Technology and Values has provided generous support for conference presentations. The Mustard Seed Foundation’s generous support for 2007-8 gave me considerable liberty for writing.

One of the great methodological lessons of this work is the importance of old eastern European particle physics and differential geometry literature, which fortunately is generally available in English or perhaps French or German. Where my needs exceeded the ample holdings on campus, the Notre Dame libraries generously tracked down materials by interlibrary loan without limit, allowing me to find many gems and ensuring that I did not lack a paper that I might need. The libraries of the University of Maryland kindly gave me a library card during my stay in the Washington, D.C. area. The Joint Institute for Nuclear Research in Dubna, Russia was able and willing to answer my request for preprint R-1890 containing the unpublished appendix of the spinor paper by Ogievetsky and Polubarinov (Ogievetski˘i and Polubarinov, 1965).

More personally, I thank my parents for their encouragement and support during my long years of education. My wife Dilkushi arrived late in the writing process, but has transformed my life.
CHAPTER 1

INTRODUCTION

1.1 Overview

This dissertation will address two longstanding issues in the philosophy of physics or the more general philosophy of science, namely, general covariance and empirical equivalence. Both of these issues are strongly tied to the largely untouched conceptual and technical question of artificial gauge freedom. It has been claimed, starting with Einstein in the 1910s and including workers in Loop Quantum Gravity today, that general covariance is the chief novel and distinctive feature of the General Theory of Relativity (GTR), the great lesson about space-time physics that all theorizing in the foreseeable future ought to respect. Whatever the details might be, general covariance is supposed to involve the absence of certain structures that are impervious to the contents and history of the world. Initially Einstein took general covariance to be manifested in the admissibility of arbitrary coordinate systems in GTR. However, almost immediately Erich Kretschmann cast doubt on Einstein’s analysis of general covariance and argued that it was a merely formal feature of GTR that any theory (of the sort that physicists might reasonably entertain), including GTR’s predecessors and competitors, could share if formulated with a bit of ingenuity (Kretschmann, 1917; Norton, 1995; Rynasiewicz, 1999).
The question whether there exist empirically equivalent but incompatible theories bears on how tightly empirical constraints from the progress of science might constrain our theorizing. While the issue of empirical equivalence has been widely discussed, philosophers’ discussions often have involved rather thin examples, perhaps generating new theory candidates by de-Ockhamizing (replacing one theoretical entity, perhaps “force,” by some combination of multiple entities such that only that combination plays a role in the theory, such as “gorce plus morce” (Glymour, 1977)), ad hoc deletion of some regions of space-time or objects therein while the remainder behaves just as in the mother theory, and the like. P. Kyle Stanford argues that resorting to these sorts of examples of underdetermination that philosophers employ, whether algorithmic or not, is a “devil’s bargain for defenders of underdetermination, for it succeeds only where it gives up any significant and distinctive general challenge to the truth of our best scientific theories” by collapsing scientific underdetermination into the familiar and perhaps insoluble problem of radical skepticism (Stanford, 2006, p. 12). While there are interesting issues in the vicinity of those examples, probably more interesting and certainly more novel examples are available in contemporary physics in the context of field theory. To be more specific, Maxwell’s electromagnetism, the Yang-Mills field theories such as describe the weak and strong nuclear forces, and Einstein’s GTR, along with their massive non-gauge and massive gauge relatives, with or without quantization, give examples of a number of interesting meta-theoretic phenomena, including the instability of empirical equivalence under change of auxiliary hypotheses (Leplin, 1997). Some of these phenomena have been anticipated in the philosophical literature but without such interesting instantiations; others might not have been contemplated previously by philosophers. Using physically
interesting examples also has the advantage that fairly detailed principles for theory identification are already suggested by standard techniques and attitudes in physics. Thus there is comparatively little risk that the question of theoretical equivalence will be resolved one way or another simply by *ad hoc* stipulation. Using examples from real physics avoids collapsing scientific underdetermination into radical skepticism, as Stanford has urged.

The themes of general covariance and empirical equivalence are tied together by a third theme, the presence or absence of gauge freedom, including the sort of "artificial" gauge freedom that particle physicists and general relativists are increasingly adept at installing by the addition of extra fields and gauge symmetries. Here gauge freedom is broadly construed as a theory formulation’s having arbitrary functions of space-time in the solution of the equations of motion. Given a suitable theory formulation that violates general covariance on some reading or that lacks a gauge symmetry, it is often possible to construct an empirically equivalent formulation that satisfies that analysis or has a gauge symmetry. The most famous and systematic procedure for installing artificial gauge freedom is the Batalin-Fradkin-Tyutin procedure (BFT) (Batalin and Fradkin, 1986, 1987; Batalin et al., 1989; Batalin and Tyutin, 1991), which takes constrained Hamiltonian theory formulations with broken gauge symmetries (expressed technically as “second class constraints”) as input and yields formulations with unbroken gauge symmetries (expressed technically as “first class constraints”). The philosophically interesting spirit can be generalized, as will appear below, not only to Lagrangian formulations of theories, but to theories that have no constraints. Given a field theory, one can add additional fields and additional symmetries (some examples of which will be more interesting than others) to get an empirically equivalent
theory formulation. Thus the BFT procedure and similar techniques provide an algorithm for constructing formulations exemplifying the Kretschmann-type worry that general covariance or gauge freedom is a merely formal rather than substantive feature of a theory formulation. Knowing of BFT-type procedures, one can show that certain examples of general covariance or gauge freedom are artificial and formal; this result might help to determine which, if any, examples are natural and substantive. Standard gauge-fixing techniques in effect reverse the BFT procedure and so suggest the theoretical equivalence of pre-BFT and post-BFT formulations. A more detailed look at general covariance, empirical equivalence and artificial gauge freedom will indicate what material will be discussed in the following chapters.

1.2 General Covariance

After conceding the truth of Kretschmann’s criticism that general covariance was a merely formal property reflective of the theorist’s ingenuity, Einstein then maintained that GTR was (substantively) generally covariant in the special sense that the field equations of GTR took their simplest form when expressed (formally) generally covariantly, thereby implying that other theories would look more complicated in that form. Thus Einstein in effect distinguished formal or weak general covariance from substantive or strong general covariance, if one may use terms that would arise later (Bergmann, 1957; Stachel, 1993). While something seems intuitively right about Einstein’s claim, the task of analyzing or replacing the “simplest form” criterion for substantive general covariance thus far has not received a universally satisfying resolution. Although it is widely believed that there is some unique nontrivial notion of general covariance that at least roughly
captures the main innovation of Einstein’s 1915-6 theory of gravity, perhaps matters are not so simple. Keeping in mind the possibility that the target of analysis might not exist, it seems prudent to continue looking for it, because one who does not seek almost surely will not find.

It is often suggested that physicist James L. Anderson’s absolute objects program (Anderson, 1964, 1967, 1971) gives the most successful analysis presently available of substantive general covariance, namely, that GTR lacks absolute objects. The concept of absolute objects is intended to capture flat metric tensors in Special Relativity, Newton’s absolute time (or its gradient), and other fields that are impervious to what goes on in the world. The project has received favorable attention and development from other physicists (Lee et al., 1974; Misner et al., 1973; Ohanian and Ruffini, 1994; Thorne et al., 1973; Trautman, 1966). Philosophers have also received the absolute objects project favorably (Friedman, 1973, 1983; Hiskes, 1984; Norton, 1993, 1995); Michael Friedman’s work was sufficiently thorough and influential that the project is sometimes called the “Anderson-Friedman” absolute objects program. However, the absolute objects project has been criticized, especially because of certain supposed counterexamples. Some of these have been resolved (Pitts, 2006a), but a new one due to Robert Geroch and (independently) Domenico Giulini seems insoluble (Giulini, 2007; Pitts, 2006a) without significantly modifying the Anderson-Friedman notion of absoluteness, weakening the connection between substantive general covariance and the lack of absolute objects, or giving up the claim that GTR is substantively generally covariant. This counterexample is the volume element $g = \text{def} \ det(g_{\mu\nu})$, a scalar density, which can be set to $-1$ in a neighborhood about any point, and so counts as absolute by Friedman’s criteria. This counterexample calls attention to the fact that some
types of geometric objects, including an irreducible part of the metric tensor, are inherently susceptible to satisfying the Anderson-Friedman definition of absoluteness, or, in brief, “susceptible.” Whether susceptibility to absoluteness renders the resulting absoluteness more significant, less so, or what, is presently unclear. It turns out that the collection of geometric objects that are susceptible to absoluteness was found (almost) in its entirety some time ago by Andrzej Zajtz (Zajtz, 1988b), though this fact was not applied to the question of absolute objects until now. The absolute objects program will be discussed in great detail in chapter 2. A further task of that chapter is to begin to fulfill a promissory note about how to avoid an absolute object in coupling spinors to GTR using the allegedly obligatory tetrad formalism (Pitts, 2006a) by using an alternative formalism with a symmetric square root of the metric (Ogievetskiǐ and Polubarinov, 1965).

While the Anderson-Friedman criterion unfortunately finds absoluteness in the susceptible geometric objects, much of Anderson’s analysis remains fruitful nonetheless. For example, Anderson and others (Thorne et al., 1973) believed that no absolute object is varied in an action principle (assuming that the formulation admits an action principle). They also believed that all dynamical (that is, non-absolute) fields are varied in an action principle to get the Euler-Lagrange equations of motion. Counterexamples to be discussed below show that absoluteness does not correlate with non-variation in an action principle in the expected way. GTR itself is a counterexample, in fact, because of $\sqrt{-g}$. Given that the Anderson-Friedman sort of absoluteness criterion gives wrong answers in a collection of interesting cases, it is noteworthy that the variational principle criterion might give the right answers if one prohibits irrelevant fields from the formulation as Anderson required, and also prohibits clock fields, which are preferred coor-
ordinates in disguise; these will be discussed shortly. Thus a variational principle might give an acceptable analysis of general covariance.

On the other hand, it is not entirely clear why the Euler-Lagrange equations obtained from varying a field susceptible to absoluteness are “about” the susceptible field, or even “about” a relation between the susceptible field and whatever other fields appear in that Euler-Lagrange equation. For a susceptible field, any small neighborhood of any model is just the same as any neighborhood in any other model. It is quite tempting to say that the susceptible field has no serious dynamics and constitutes a fixed background; its Euler-Lagrange equation merely tells how other fields relate to this fixed background.

Absolute objects appear to correspond to the sort of “background” that workers in Loop Quantum Gravity (LQG) are eager to avoid. Given the starring role that “background independence” plays in motivating that technically impressive project (Ashtekar and Lewandowski, 2004; Rovelli, 2004; Smolin, 2002, 2006; Thiemann, 2007), difficulties in defining what is a “background” might pose something of a conceptual de-motivation for some features of LQG. Here I have in mind more than the LQG polemic against string theory, which LQG theorists fault for using a background metric in various ways such as a perturbative splitting of the fields into a background part and a dynamical part for use in a series expansion. Rather, workers in the LQG project routinely appeal to background independence to constrain their quantization project to avoid any use at all of a background metric or even a background volume element. On the other hand, Giulini has wondered whether background independence is a completely clear notion (Giulini, 2007). Certainly the technical achievements of LQG would remain of considerable interest in achieving a gauge-invariant non-perturbative quantization of gravity, and per-
haps the dissolution of singularities (Ashtekar et al., 2006a,b; Bojowald, 2001a,b, 2002; Bojowald and Hinterleitner, 2002; Bojowald and Kagan, 2006), even if the drive to ‘take seriously the conceptual novelty of GTR’ were downplayed due to inability to say quite what that novelty was. Once the criteria of sameness in all models and not being varied in an action principle prove not to be coextensive, it becomes more difficult to identify uniquely the alleged conceptual novelty of GTR that all future theorizing should respect. Writers in the LQG project, while occasionally claiming to define background independence carefully, generally fail to achieve the level of detail of the Anderson-Friedman project. In Lee Smolin’s work one perhaps finds hints of two criteria that also appear importantly in Anderson’s discussion of absolute objects (Anderson, 1967), namely a sameness-in-all-models criterion and a criterion of being determined by equations of motion (if perhaps not a variational principle) (Smolin, 2006, p. 202); the equation-of-motion criterion is implemented by path integration in the quantum context. Perhaps LQG writers would take Anderson’s condition of no absolute objects as an analysis of background independence. However, the Geroch-Giulini counterexample suggests that GTR itself has a background volume element. Thus the motivation to avoid a ‘background’ volume element seems to be reduced. Avoiding a background volume element has played a discernible role in LQG work. Some authors have managed to find a Lie algebra for the constraints in canonical gravity by defining new constraints that are functions, perhaps non-polynomial, of the usual constraints from the Dirac-Arnowitt-Deser-Misner 3-metric canonical quantization project (Kouletsis, 1996; Kuchař and Romano, 1995; Markopoulou, 1996); Thomas Thiemann, however, denies the utility of these new constraints for his efforts at background-independent quantization due to the new constraints’ having density
weight 2 rather than 1 (Thiemann, 2006), and thus giving coordinate-dependent or background-dependent results. But if even GTR has a background volume element, as the Geroch-Giulini counterexample suggests, then perhaps Thiemann’s criteria are unnecessarily strict. Thus recognizing a bit of background dependence in GTR might help to resolve one problem in the canonical quantization of gravity, namely, achieving a Lie algebra for the constraints. In any case, those who are attracted to LQG by its alleged background independence have some stake in the outcome of the Anderson-Friedman project and the Geroch-Giulini counterexample.

Another suggestion sometimes made by physicists about GTR, that it is an “already parametrized theory,” suggests another fruitful way to analyze general covariance. Parametrization involves the use of preferred coordinates as scalar fields, known as “clock fields” (Arnowitt et al., 1959, 1962a,b; Bergmann and Brunings, 1949; Dirac, 1964; Kuchař, 1973, 1981; Lanczos, 1949; Teitelboim, 1973; Torre, 1992a; Westman and Sonego, 2007). Clock fields have recently come to the attention of some philosophers of physics, such as John Earman and John Norton (Earman, 2006; Norton, 2003). The idea behind the phrase “already parametrized” is that GTR does not need any explicit clock fields because they are already there, albeit obscured and mixed in with the true degrees of freedom. Hopes for digging out the clock fields from the metric, or even regarding GTR as literally parametrized without exhibiting the clock fields explicitly, have diminished somewhat over the years. However, whether clock fields are already in GTR somewhere does not matter for my purpose of analyzing general covariance. What matters is that no further clock fields are needed for GTR to achieve formal general covariance. At present it appears that the absence or presence of (non-redundant
(Bergmann and Brunings, 1949)) clock fields in a formally generally covariant theory correctly tracks expectations that a theory formulation is or is not substantively generally covariant, respectively, assuming that all non-variational fields have been replaced by functions of the clock fields and their derivatives. Clock fields are preferred coordinates in disguise, so the criterion that a formally generally covariant theory formulation is substantively generally covariant just in case it lacks (non-redundant) clock fields revives the spirit of Einstein’s mid-1910s claim that general covariance involves the lack of preferred coordinate systems. Most people nowadays agree that there is nothing very interesting about coordinate systems (Earman, 2006), so it is surprising that so old-fashioned a criterion turns out to work so well, and indeed better than the more sophisticated Anderson-Friedman project. The absence of clock fields and the absence of fields not varied in the action principle seem to be not merely coextensive, but very simply related conceptually. The functional form of a parametrized theory’s dependence on the clock fields determines the theory’s symmetry group in a simple way. A discussion of parametrization therefore concludes chapter 2 on general covariance.

1.3 Empirical Equivalence

Empirical equivalence is an issue that plays a key role in arguments about scientific realism. According to André Kukla,

[the main argument for antirealism is undoubtedly the argument from the underdetermination of theory by all possible data. Here is one way to represent it: (1) all theories have indefinitely many empirically equivalent rivals; (2) empirically equivalent hypotheses are equally believable; (3) therefore, belief in any theory must be arbitrary and unfounded. (Kukla, 1998, p. 58).]
On the question whether there exist incompatible but empirically equivalent theories, so that theories are underdetermined by data, a large literature exists (Bain, 2004; ben Menahem, 1990; Clendinnen, 1989; Cushing, 1994; Douven, 2000; Douven and Horsten, 1998; Earman, 1993; Ellis, 1985; Hoefer and Rosenberg, 1994; Jones, 1991; Kukla, 1993, 1996; Laudan and Leplin, 1991; Leplin, 1997; Mühlhölzer, 1991; Musgrave, 1992; Okasha, 1997; Quine, 1975; Sklar, 1985; Stanford, 2006; van Fraassen, 1980, 1989), but presently there seems to be considerable disagreement on various points. By “empirical equivalence,” I have in mind (unless otherwise qualified) precise equivalence for all models—not equivalence that might be broken with further experimental progress or the introduction of differing auxiliary hypotheses, or equivalence that holds in some but not all physically possible worlds. When I speak of approximate empirical equivalence, once again all models, not just some, are in view. During the heyday of logical empiricism, many influential people (including Hans Reichenbach, Hilary Putnam, Adolf Grünbaum, and Wesley Salmon) denied that distinct and incompatible but empirically equivalent theories existed (Glymour, 1970). But not only W. V. O. Quine’s work (Quine, 1975), but also the revival of scientific realism during the 1960s-70s, led to a revival of belief that distinct and incompatible but empirically equivalent theories exist. More recently the view that there do exist rival empirically equivalent theories has been somewhat widely held (Earman, 2006; Kukla, 1993; Musgrave, 1992), in contrast to the earlier positivist view that empirically equivalent theories say the same thing and so are merely linguistic variants. This dissertation aspires to address the question of empirically equivalent theories within the context of local classical and (to some degree) quantum field theories. In this context, physically interesting examples and sophisticated mathematical criteria for equivalence
are often available, in contrast to the more contrived examples often employed in the philosophical literature on underdetermination and theoretical equivalence. The latter sort of example also might have the disadvantage that it is too easy to stipulate that empirically equivalent formulations are the same theory.

Classical and quantum field theories provide philosophically interesting test cases for both approximate and exact empirical equivalence of physical theories. Chapter 3 considers approximate empirical equivalence using massless and massive versions of Maxwell’s electromagnetism, Yang-Mills field theories (describing the weak and strong nuclear forces), and Einstein’s GTR. By “massive theories,” I have in mind the standard particle physics-based sense, with some abuse of language, such that the theories’ equations of motion (using the standard potentials as variables) have a term that is algebraic and linear in the potential. A massless scalar field $\phi$ typically satisfies the relativistic wave equation

$$c^{-2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = 0. \quad (1.1)$$

A massive scalar field satisfies the Klein-Gordon equation

$$c^{-2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} + m^2 c^2 \hbar^{-2} \phi = 0. \quad (1.2)$$

The coefficient of the algebraic term has been written as $m^2 c^2 \hbar^{-2}$ in anticipation of the use of relativistic quantum field theory. In units with Planck’s constant (reduced by $2\pi$) $\hbar$ and the ‘speed of light’ $c$ set to 1, it becomes clear why the phrase “mass term” is used. For a classical theory, the ‘mass term’ involves not a mass, but an inverse length scale. In a quantum context, Planck’s constant can be used to convert the length-related scale to a mass scale. The length then corresponds
to the Compton wavelength. The Klein-Gordon equation and other massive field theories with Lagrangian densities have quadratic terms in the potential. Under quantization (if all goes well), the quanta, such as photons, have a nonzero rest mass; then light does not travel at the ‘speed of light.’

At the classical level, the massive theory families have representatives that approximate their massless relatives arbitrarily well empirically (though there is question about the health of the massive classical gravitational theories) for sufficiently small mass parameters, but the theoretical properties of massive theories differ greatly from those of their massless counterparts, except in the scalar case (and perhaps the spin $\frac{1}{2}$ case, for which the above discussion would require modification to discuss the first-order Dirac equation). For example, the massless electromagnetic, Yang-Mills and gravitational theories have mathematically indeterministic field equations, whereas the massive theories in the standard formulation have deterministic field equations. The fact that the mass of the photon (or vector boson or graviton, mutatis mutandis) is a free parameter shows that Proca’s massive electromagnetism is not one theory, but an infinite family of theories, one for each choice of value for the photon mass. This example therefore exemplifies permanent rather than transient underdetermination, and yet there is no possibility of identifying supposedly rival theories as really the same theory, because all theories involved (Maxwell’s and the various massive electromagnetic theories, for example) are empirically inequivalent. Under quantization, the massive electromagnetic theory remains healthy and continues to approximate its massless relative, thus giving an example of underdetermination in a significant quantum field theory.
The quantized massive Yang-Mills theories (excepting theories with the mass term pertaining to an Abelian sector) and quantized gravitational theories are theoretically problematic, and some argue that classical massive gravity is also defective. This phenomenon of a good classical theory that goes bad under quantization resembles or exemplifies Jarrett Leplin’s notion of instability of empirical equivalence under change of auxiliary hypotheses (Leplin, 1997). Here I have in mind taking quantization (or the lack thereof) as something like an auxiliary hypothesis combined somehow with one of the above mentioned field theories, construed in some skeletal form that is neither classical nor quantum.

Explicit and fairly realistic physical theories interrelated through advanced mathematics should yield insights not as readily available from the traditional discussions of abstract theory formulations $T_1$ and $T_2$ related by simple logical formulas. Earman has made a similar point in the context of determinism:

If philosophers had spent less time trying to achieve for determinism the superficial ‘precision’ afforded by formal symbolic notation and had spent more time studying the content of physical theories they might have confronted the truly fascinating substantive challenges that determinism must face in classical and relativistic physics. (Earman, 1986, p. 21)

Thus the traditional debates on underdetermination and empirical equivalence are not extremely relevant for the discussion that follows.

1.4 Artificial Gauge Freedom

Concerning exact empirical equivalence of rival (incompatible) theories, what is the philosophical significance of certain widely used techniques in particle physics and gravitation theory for introducing artificial gauge freedom, with the help of additional fields, into classical field theories? Can these techniques (or relatives
thereof) be used to generate physically interesting examples of empirically equivalent but incompatible theories? Or are the results merely alternative formulations of the same theory, as physicists would expect? Making the most natural application to electromagnetism, the question becomes: is Proca’s non-gauge massive electromagnetism (with some definite photon mass) really the same theory as Stueckelberg’s gauge massive electromagnetism (with the same photon mass)? The two theory formulations are empirically equivalent, but might seem ontologically distinct. Proca formulations have four fields and three degrees of freedom at every spatial point, with no gauge freedom, whereas Stueckelberg formulations have five fields and three degrees of freedom at each spatial point, with gauge freedom. Clearly $4 \neq 5$ and the absence of gauge freedom differs from the presence of gauge freedom, so there is at least some temptation to regard the theories as distinct. The presence of gauge freedom is often said by particle physicists to be of fundamental significance. Even if these claims are overstated (as Chris Martin argues with considerable plausibility (Martin, 2002)), it is noteworthy that gauge freedom by itself fails to distinguish Maxwell’s massless theory from the Stueckelberg formulation of massive electromagnetism. Empiricist-minded philosophers might share with most physicists the view that Proca’s and Stueckelberg’s families of theory formulations are really the same family of theories and that a Proca electromagnetism with a given mass and a Stueckelberg electromagnetism with that same mass are the same theory. The fact that one can gauge-fix Stueckelberg formulations to produce Proca formulations provides a technical implementation of the empirical equivalence of the two families. Alleged ontological differences that cannot have any empirical consequences and that disappear on gauge-fixing are not real, one might think. But then the presence or absence of gauge freedom
in massive electromagnetism is not a real feature of the theory, but merely a conventional choice of formulation. Thus massive electromagnetism is subject to a Kretschmann-style point: gauge freedom, if not presently initially, can be installed with a little effort. In chapter 3, more detailed application of electromagnetic theories to debates on underdetermination and empirical equivalence will be made. Given the deep similarities between electromagnetism on the one hand and Yang-Mills fields and GTR on the other, it will be interesting to see the extent to which these theories do and do not show the same pattern as electromagnetism. In installing artificial gauge freedom, I have in mind especially the classical theory formulated in Lagrangian terms. Quantization can make an important difference, even rendering a classically satisfactory theory unacceptable under quantization, due to failure of unitary or renormalizability, as will appear in the case of most massive Yang-Mills theories, excluding essentially Abelian theories.

Whereas the Stueckelberg trick (as it is often called) for installing gauge freedom into massive electromagnetism is basically an *ad hoc* or opportunistic move, it has become clear in recent decades that there is a general systematic procedure that achieves basically the same result. This is the Batalin-Fradkin-Tyutin *et al.* framework (Batalin and Fradkin, 1986, 1987; Batalin et al., 1989; Batalin and Tyutin, 1991). The BFT procedure in fact encompasses various tricks that are employed in standard examples. For massive Yang-Mills fields (not just the degenerate electromagnetic special case), Stueckelberg’s trick (Ruegg and Ruiz-Altaba, 2004) takes a massive theory with broken gauge symmetry due to the mass term, and then restores the symmetry with the introduction of one or more compensating fields. The BFT procedure, strictly construed, has some features that make application to massive gravity an unpleasant prospect, but some simple modifi-
cations suggested below yield a procedure that agrees with parametrized massive gravity with four clock fields.

Below I call attention to some disadvantages and limitations of the existing BFT technologies for installing artificial gauge freedom, and outline simpler procedures of both Lagrangian and Hamiltonian varieties. The procedures are illustrated using Proca’s massive electromagnetism, where they recover the usual Stueckelberg gauged massive electromagnetism rather more directly than does BFT. It is inconvenient that, due to the ‘boundary condition’ imposed on the new gauge compensation fields, the usual BFT procedure requires a canonical transformation to restore the typical velocity-momentum relationship and thus recover the Stueckelberg Lagrangian density for massive gravity. An alternative Lagrangian-friendly boundary condition is proposed below. I also outline a more general strategy of installing artificial gauge freedom. The BFT procedure adds extra fields only to convert theories with second-class constraints\(^1\) into gauge theories with only first class constraints. Furthermore, the BFT procedure converts all the second class constraints into first class constraints. But one might wish to convert only some second class constraints into first class constraints. More importantly, one might wish to add gauge freedom to theories with no constraints, such as by parametrizing a scalar field theory (Kuchař, 1973, 1981; Rovelli, 2004). Perhaps one could have both of these goals. In at least some cases, such goals can be realized with the addition of extra fields and gauge symmetries.

\(^1\)Constraints vanish for all dynamically possible trajectories and arise due to degenerate kinetic terms that obstruct the Legendre transformation from velocities to momenta. The term, somewhat confusingly, refers both to the expression that vanishes for dynamically possible trajectories (but not for all kinematically possible trajectories) and to the vanishing of that expression. Taking constraints as expressions of the phase space variables (and momentarily forgetting that they have the value of zero), one can take Poisson brackets of pairs of constraints (and then set the values to zero). Second class constraints have nonzero Poisson brackets with at least some other constraint(s). First class constraints have Poisson brackets that vanish, perhaps with terms proportional to the constraints themselves, with all the constraints (Sundermeyer, 1982).
2.1 Absolute Objects, Admissible Coordinates and Groupoids

James L. Anderson analyzed the novelty of Einstein’s so-called General Theory of Relativity (GTR) as its lacking “absolute objects” (Anderson, 1964, 1967, 1971; Anderson and Finkelstein, 1971). His absolute objects program has received favorable attention and development from both physicists (Lee et al., 1974; Misner et al., 1973; Ohanian and Ruffini, 1994; Thorne et al., 1973; Trautman, 1966) and philosophers (Friedman, 1973, 1983; Hiskes, 1984; Norton, 1993, 1995), especially Michael Friedman, so sometimes one speaks of the Anderson-Friedman absolute objects program. Often it has been said that absolute objects are what GTR was distinctive and novel in lacking. Metaphorically, absolute objects are often described as a fixed stage on which the dynamical actors play their parts.

A review of Anderson’s definitions will be useful. Absolute objects are to be contrasted with dynamical objects. The values of the absolute objects do not depend on the values of the dynamical objects, but the values of the dynamical objects do depend on the values of the absolute objects (Anderson, 1967, p. 83). Both absolute objects and dynamical objects are, mathematically speaking, geometrical objects or parts thereof. Geometric objects include things like tensors,
connections and tensor densities, but also more exotic entities. The theory of geometric objects, which fulfills a call by Ricci for objects more general than tensors in being perhaps nonlinear or involving higher derivatives of the coordinates (Nijenhuis, 1972), took recognizable form in the 1930s (Schouten and Haantjes, 1937; Veblen and Whitehead, 1932; Wundheiler, 1937). The history is sketched in various places (Gołab, 1974; Kucharzewski and Kuczma, 1964; Nijenhuis, 1972). The name reflects the original ambition that local differential geometry could be expressed entirely in terms of them. The importance of the fact that geometric objects are defined everywhere on the manifold will appear later in the context of the Jones-Geroch dust counterexample. (In more modern terms, the distinction between a global section and a local section of the relevant natural bundle is in view.) It will be shown that Anderson's definition is naturally amended to avoid this counterexample more or less as Friedman envisioned. The previously unnoticed (before (Pitts, 2006a)) fact that Anderson's analysis detects an absolute object in Torretti's constant curvature spaces example, as well as Norton's related conformally flat scenario, deprives them of force. A previously unnoticed counterexample (prior to (Pitts, 2006a)) involving spinor fields is proposed and resolved in outline using an alternative spinor formalism. Most significantly, as both Geroch and Giulini have pointed out (Giulini, 2007; Pitts, 2006a), GTR actually has an absolute object, using Friedman's local definition of absolute objects and an attractive choice of variables. Whether it is better to revise the notion of absolute object or to revise that claim that GTR lacks them is presently unclear.

Before absolute objects can be defined, the notion of a covariance group must be outlined. Here it will prove helpful to draw upon the unjustly neglected work of Kip Thorne, Alan Lightman, and David Lee (TLL) (Thorne et al., 1973); a
useful companion paper (LLN) was written by Lee, Lightman and W.-T. Ni (Lee et al., 1974). The TLL definition differs slightly from Anderson’s in its notion of faithfulness. According to TLL,

A group $G$ is a covariance group of a representation if (i) $G$ maps [kinematically possible trajectories] of that representation into [kinematically possible trajectories]; (ii) the [kinematically possible trajectories] constitute “the basis of a faithful representation of $G$” (i.e., no two elements of $G$ produce identical mappings of the [kinematically possible trajectories]); (iii) $G$ maps [dynamically possible trajectories] into [dynamically possible trajectories]. (Thorne et al., 1973, p. 3567)

One can now define absolute objects. They are, according to Anderson, objects with components $\phi_\alpha$ such that

(1) The $\phi_\alpha$ constitute the basis of a faithful realization of the covariance group of the theory. (2) Any $\phi_\alpha$ that satisfies the equations of motion of the theory appears, together with all its transforms under the covariance group, in every equivalence class of [dynamically possible trajectories]. (Anderson, 1967, p. 83)

Thus the components of the absolute objects are the same, up to equivalence under the covariance group, in every model of the theory. It is the dynamical objects that distinguish the different equivalence classes of the dynamically possible trajectories (Anderson, 1967, p. 84). One notices that the components of the absolute object need be the same, up to equivalence under the covariance group, for all dynamically possible trajectories, not all kinematically possible trajectories. Might this matter have gone otherwise? For most purposes this choice makes no difference, because typically those objects whose components are the same for all dynamically possible trajectories share the same feature for all kinematically possible trajectories. This condition fails, however, in the context of Rosen’s and Sorkin’s deriving the flatness of a metric using a variational principle with a Lagrange multiplier, as well as for the auxiliary fields of supergravity, as will appear below.
The appropriateness of the faithfulness clause in the definition of absolute objects would repay further consideration. Clearly the number 2 and the charge of the electron, which are invariant under the covariance group, should not count as absolute. However, the present faithfulness clause, depending on its interpretation, could have the unfortunate consequence that, for example, the flat metric tensor of special relativity with an electromagnetic field can fail to be absolute, because the flat metric is unaffected by electromagnetic gauge transformations.

There seems to be no compelling reason to require a covariance group instead of a mere covariance groupoid. A (Brandt) groupoid is a structure that would be a group if it were meaningful to multiply every pair of elements. Einstein’s equations on a background space-time, once one imposes a consistent notion of causality, have a covariance groupoid that is not a group (Pitts and Schieve, 2004). Local diffeomorphisms typically do not form a group, either, because their domains and ranges do not line up in the right way (Earman, 1974; Friedman, 1983; Kucharzewski and Kuczma, 1964; Nijenhuis, 1952). Groupoids sometimes appear in this context also, and perhaps other contexts as well, so it is worthwhile to discuss them briefly. According to Peter Hahn (Hahn, 1978) and Jean Renault (Renault, 1980), a (Brandt) groupoid is a set $G$ that is endowed with a product map $(x, y) \to xy : G^2 \to G$, where $G^2$ (a subset of $G \times G$) is the set of composable ordered pairs, and an inverse map $x \to x^{-1} : G \to G$ such that the following relations hold:

1. $(x^{-1})^{-1} = x$,

2. if $(x, y)$ and $(y, z)$ are both elements of $G^2$, then $(xy, z)$ and $(x, yz)$ are also elements of $G^2$ and $(xy)z = x(yz)$,

3. $(x^{-1}, x) \in G^2$, and if $(x, y) \in G^2$, then $x^{-1}(xy) = y$,
4. \((x, x^{-1}) \in G^2\), and if \((z, x) \in G^2\), then \((zx)x^{-1} = z\).

Thus groupoid multiplication is associative whenever it is defined, and each element has a 2-sided inverse. Instead of a single identity element, there are many little identity elements such as \(xx^{-1}\) that apply within the limits of compositability. One might have a groupoid but not a group if the collection of possible transformations depends on which field configuration one considers. To give a simple example, a man walking on an island has a groupoid of translations available, because depending on where he starts, some candidate translations will take him into the sea (c.f. (Vuillemin, 1973)). Hiskes also discusses how groupoids can arise in the context of absolute objects (Hiskes, 1984).

A more interesting example of a transformation groupoid comes from the inequalities for a causality-respecting “eigentliches” (proper) coordinate system once suggested by Hilbert (Brading and Ryckman, 2008; Hilbert, 1917; Pauli, 1921). Hilbert’s conditions are the following for the metric components:

\[
g_{11} > 0, \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, g_{44} < 0.
\]

Hilbert called any coordinate transformation from one proper coordinate system to another, an "eigentliche" (proper) coordinate transformation. Christian Møller (Møller, 1952, pp. 234, 235) presented similar inequalities for coordinates that could correspond to a real (subluminal) fluid. Thus coordinates can have qualitative significance even while lacking quantitative significance. Apart from the domain-range mismatch problem noted above, would the proper coordinate transformations (that is, such that the initial and final coordinates both satisfy Hilbert’s
inequalities) for some single metric tensor form a group? It is tempting to identify a coordinate transformation with a generating vector field, as in the infinitesimal relation \( x_2^\mu = x_1^\mu + \xi^\mu \), which does admit a finite generalization as the exponential map (Grishchuk et al., 1984) (though the finite form does not even cover all transformations near the identity (Freifeld, 1968)). It is evident that whether a given vector field \( \vec{\xi} \) generates a proper coordinate transformation depends on which proper coordinate system is used initially. If coordinate transformations are identified (at least in part) using their generating vector fields \( \vec{\xi} \) and \( \vec{\psi} \), then the composability of the \( \vec{\xi} \) and \( \vec{\psi} \) transformations is unknown until the initial coordinate system is specified. (Analogous considerations hold for Einstein’s equations on a real Minkowski background (Pitts and Schieve, 2004).) One need not pick out coordinate transformations by their generating vector fields (in cases where such exist), because one can simply specify the functional form of the new coordinates’ dependence on the old coordinates. But then a similar worry appears: a recipe that adds 3 to the value of the coordinate \( x^2 \) will have different effects depending on whether \( x^2 \) is an approximately Cartesian coordinate, a radial coordinate, or an angle, for example. Thus (even waiving the domain-range problem) the proper coordinate transformations for some single metric form a groupoid but not a group. Considering how often groupoids arise in space-time applications, perhaps one should speak generally of groupoids rather than groups, introducing groups only in special cases where their universal composability is relevant. For familiarity and brevity, however, I might talk of groups rather than groupoids.

Given how deeply entrenched the admissibility of arbitrary coordinates presently is in space-time theory, it might be helpful to recall what sort of argument originally motivated such generality. In 1916 Einstein argued that
[t]he method hitherto employed for laying co-ordinates into the space-time continuum in a definite manner [yielding observable time or space intervals] thus breaks down, and there seems to be no other way which would allow us to adapt systems of co-ordinates to the four-dimensional universe so that we might expect from their application a particularly simple formulation of the laws of nature. So there is nothing for it but to regard all imaginable systems of co-ordinates, on principle, as equally suitable for the description of nature. (emphasis added) (Einstein, 1923, p. 117)

The defeasibility of this argument is obvious: one need only discover a way to adapt coordinates in such a way that an especially simple formulation of the laws of nature obtains, if possible. From time to time the harmonic coordinate condition \( \partial_{\mu}(\sqrt{-g}g^{\mu\nu}) = 0 \) of DeDonder and Fock (Fock, 1959) commends itself to some authors as a candidate, but without achieving widespread acceptance. While the resulting simplification of the Einstein equation to resemble the wave equation has its attraction, presumably the ultimate lack of appeal of this condition for more than pragmatic use is due to the fact that imposing this condition (or any other condition involving equations as opposed to inequalities) destroys one of the key motivations for adopting Einstein’s equations in the first place, namely, the invariance of the action under general coordinate transformations. On the other hand, imposing some inequalities as generalized coordinate conditions leaves untouched at least the infinitesimal transformations, which do most of the relevant work anyway. Thus the imposition of some inequalities to restrict the coordinates is less drastic a move than it might initially seem. The distinction is already made between large coordinate transformations and those connected to the identity (Alvarez-Gaumé and Witten, 1984; Visser, 1990); according to Matt Visser, the possibility of anomalous behavior under large diffeomorphisms is often but not justifiably ignored in canonical quantum gravity.
It has been asserted that the novel and nontrivial sense in which GTR is generally covariant is its lack of absolute objects (Anderson, 1967) or “prior geometry” (Misner et al., 1973, pp. 429-431). John Norton discusses this claim with some sympathy (Norton, 1992, 1993, 1995), though technical problems such as the Jones-Geroch dust and Torretti constant spatial curvature counterexamples are among his worries (Norton, 1993, 1995). Anderson and Ronald Gautreau encapsulate the definition of an absolute object as an object that “affects the behavior of other objects but is not affected by these objects in turn.” (Anderson and Gautreau, 1969, p. 1657) Depending on how one construes “affects,” this summary might be serviceable, but only if used very cautiously. On other occasions absolute objects are said to “influence” dynamical objects but not *vice versa* (Anderson, 1971, p. 169). Such terminology echoes Einstein and implies that absolute objects violate what Anderson calls a “generalized principle of action and reaction” (Anderson, 1967, p. 339) (Anderson, 1971, p. 169). Norton has argued, rightly I think, that such a generalized principle of action and reaction is hopelessly vague and arbitrary and that it should not be invoked to impart a spurious necessity to the contingent truth that our best current physical theory lacks absolute objects (Norton, 1993, pp. 848, 849). Harvey Brown has also expressed doubts about this generalized action-reaction principle (Brown, 2005b). It now turns out, however, that our best physical theory has an absolute object, thanks to the Geroch-Giulini counterexample, but that result in no way undermines Norton’s and Brown’s point. One might also doubt whether terms such as “affects,” “influence” and “act” adequately capture what absolute objects typically do. These terms suggest that the dynamical objects in question would have well-defined behavior if the absolute objects could somehow be ‘turned off,’ so to
speak (perhaps by replacing them with zero in the equations of motion), and that if the absolute objects were ‘turned on’ again, they would alter the well-defined behavior of the dynamical objects in much the way that an applied electric field alters the motion of a charged particle. But in important examples, such as Newtonian physics or special relativity, turning off many or all of the absolute objects destroys the theory: the equations of motion become degenerate or meaningless. The absolute objects do not so much alter an otherwise happy situation as provide conditions in which the dynamical objects can have well-defined behavior. Perhaps the stage metaphor for absolute objects is deeper than it seemed: presumably actors could put on a play on a stage consisting of a rubbery sheet or a giant pillow, or perhaps act in mid-air while falling freely, but it is easier to act on a firm wooden stage. Thus the claim that absolute objects have some defect knowable *a priori* easily may be taken too seriously, and often has been. The supposed fact that it is possible to do without them, as supposedly holds in Einstein’s theory, should be something of a surprise; the genuine fact that even GTR has an absolute object (given the current definition) in $\sqrt{-g}$ reopens the question whether it is even possible to do without absolute objects. The viability or otherwise of the Bach-Weyl theory below is quite relevant in that regard. Whether absolute objects “act” may be doubted, but it is illuminating to say that they are not acted upon.

In Anderson’s framework, an important subgroup of a theory’s covariance group is its symmetry group (Anderson, 1967, pp. 84-88). One first defines the symmetry group of a *geometrical object* as those transformations that leave the object unchanged. If the transformations are infinitesimal space-time mappings, then the Lie derivative of the geometrical object with respect to the relevant vec-
tor field vanishes for symmetries. The symmetry group of a physical system or theory—Anderson makes no distinction between them here—is

the largest subgroup of the covariance group of this theory, which is simultaneously the symmetry group of its absolute objects. In particular, if the theory has no absolute objects, then the symmetry group of the physical system under consideration is just the covariance group of this theory. (Anderson, 1967, p. 87)

Thus, roughly speaking, the fewer absolute objects a theory has, the more of its covariance transformations are symmetry transformations. For the example of a massive real scalar field obeying the Klein-Gordon equation in flat space-time in arbitrary coordinates, the covariance group is the group of diffeomorphisms, while the symmetry group is the 10-parameter Poincaré group corresponding to the ten Killing vector fields of Minkowski space-time. For a massive real scalar field coupled to gravity in GTR, the covariance group is again the diffeomorphisms. The symmetry group is also the diffeomorphisms, because any diffeomorphism leaves the set of absolute objects invariant, trivially, because there are no absolute objects (or so one thought until the Geroch-Giulini scalar density counterexample appeared). For a theory with no absolute objects, every member of the covariance group is also a member of the symmetry group, because the (empty) collection of absolute objects is left invariant by all of them. GTR has a rather large symmetry group (though not as large as one thought before the Geroch-Giulini counterexample) because of its paucity of absolute objects. The fact that the conformal metric density \( \hat{g}_{\mu\nu} \) and the metric \( g_{\mu\nu} \) might fail to have any symmetries in many models is true but irrelevant. One could make similar remarks about groupoids in relation to the symmetry group that were made above regarding the covariance group.

A word about active vs. passive diffeomorphisms is in order. Anderson and
TLL frequently speak of the “manifold mapping group.” At times I might use this terminology, or mix active and passive language freely. Strictly speaking, I intend all diffeomorphisms to be construed *passively* as coordinate transformations. While this position is contrary to widely held views nowadays, it follows from motivations that are quite generally recognized (as in (Earman, 1989)). Misner, Thorne and Wheeler take a passive view of all transformations up to chapter 41 of their exhaustive 44-chapter book (Misner et al., 1973, p. 1140), so one must be able to study GTR thoroughly without active transformations. In theories such as GTR, it is unclear that one can even *make sense* of active diffeomorphisms without introducing for points primitive mathematical individuation (mathematical haecceities) so robust that specific mathematical points can be moved relative to the fields defined on the manifold. The lesson of the hole argument, however, seems to be precisely that (meta)physical space-time points *cannot* be moved relative to the fields defined on the manifold. It seems awkward to take for granted a non-obvious mathematical framework of point individuation, only to deprive it of (meta)physical significance later. Why not anticipate and avoid the problem instead? Consideration of some rather baroque conceptions that have been employed to understand general covariance, illustrated in diagrams (Korté, 2006; Norton, 1989), illustrates this awkwardness quite effectively. Mathematicians have different goals from philosophers and physicists, so mathematicians’ implicit theory of mathematical individuation is hardly binding on philosophers and physicists, even at the mathematical level. The not-so-absolute position taken by Newton in *De Gravitatione et Aequipondio Fluidorum*, which certainly differs from Clarke’s view (Alexander, 1956), might strike a good balance in the spectrum of absolute-relational views:
...the immobility of space will be best exemplified by duration. For just as the parts of duration derive their individuality from their order, so that (for example) if yesterday could change places with today and become the later of the two, it would lose its individuality and would no longer be yesterday, but today; so the parts of space derive their character from their positions, so that if any two could change their positions, they would change their character at the same time and each would be converted numerically into the other. The parts of duration and space are only understood to be the same as they really are because of their mutual order and position; nor do they have any hint of individuality apart from that order and position which consequently cannot be altered. (Newton, 1962, p. 136)

Active transformations appear to make sense in STR, however, where one knows in advance what the coordinates mean and hence which space-time points are picked out by them (perhaps modulo rigid translations, rotations and boosts only). Thus the expectation that there is some important difference between STR and GTR regarding active vs. passive diffeomorphisms is vindicated on the view suggested here. Somewhat along these lines, Lusanna and Pauri (Lusanna and Pauri, 2006) replace active diffeomorphisms by passive and field-dependent ones. John Stachel and Mihaela Iftime have recently advocated a mathematically more elaborate procedure involving fiber bundles, in which (as in my suggestion) one cannot even formulate the hole argument. For Stachel and Iftime, one cannot formulate the hole argument because the mathematical points in the base manifold (which represents space-time) cannot be moved about independently of their fibers (which represent the fields defined thereon) (Iftime and Stachel, 2006; Stachel and Iftime, 2005). On the other hand, their procedure of defining space-time as the base space obtained by projection of a logically prior fiber bundle suggests greater ontological commitment to fiber bundles than one might have expected or wanted to make. One can preserve the priority of space-time while leaving the hole argument impossible to formulate simply by refusing to individuate mathematical points in-
dependently of the physical happenings; then active diffeomorphisms are (at least physically) impossible. It might be worthwhile to explore whether nontrivial active diffeomorphisms can exist for kinematically possible trajectories (“off-shell”) but become trivial for dynamically possible trajectories (“on-shell”) (c.f. (Fatibene et al., 2004)).

Finding Anderson’s definition obscure, Michael Friedman amended it in the interest of clarity (Friedman, 1973, 1983). Friedman takes his definition to express Anderson’s intuitions, so the target of analysis is shared between them. As it turns out, Friedman has made a number of changes to Anderson’s definitions, most of which seem to have received little comment by him or others, and some of which are not so helpful. As a result, some comparison between Anderson’s and Friedman’s works will be worthwhile.

First, though Friedman’s and Anderson’s equivalence relations are laid out somewhat differently, a key difference between them is that Friedman’s equivalence relation, which he calls \( d \)-equivalence, comprises only diffeomorphism freedom (Friedman, 1983, pp. 58-60), not other kinds of gauge freedom such as local Lorentz freedom or electromagnetic or Yang-Mills gauge freedom, in defining the covariance group. But local Lorentz freedom is a feature of the standard version of Einstein’s GTR + spinors, for example. Anderson, using a fairly standard terminology, calls such groups besides diffeomorphisms “internal groups” (Anderson, 1967, pp. 35, 36), though the term does not always fit perfectly for the examples available today.\(^1\) I find no argument for Friedman’s restricting the relevant

\(^1\)In cases such as electromagnetic or Yang-Mills gauge freedom or local Lorentz invariance of an orthonormal tetrad, the name “internal” fits well, because the transformations happen independently at each space-time point, due to lack of derivatives of the dynamical fields in the gauge transformation. However, some symmetries that are not diffeomorphisms resist being called internal. One example is a theory with Einstein’s equations formulated with a background metric tensor. Then there are two symmetries: diffeomorphisms and gauge transformations, both of which involve derivatives of the fields to arbitrarily high order (Grishchuk et al., 1984; Pitts and
equivalence relation to diffeomorphisms, apart perhaps from a suggestion that the interesting cases would be covered (Friedman, 1973). The goal is to distinguish physical sameness from conventional variation in descriptive fluff. Because these other symmetries involve descriptive fluff as much as diffeomorphisms do, it seems that Anderson was more successful than Friedman on this point. The role of internal groups in Anderson’s work also seems to have escaped Norton’s notice (Norton, 1993, pp. 847, 848). This point might be insignificant, except that an internal group is crucial in the tetrad-spinor counterexample.

Second, Friedman’s mathematical language is less general than Anderson’s and fails to accommodate some useful mathematical entities that Anderson’s older component language permits. Friedman follows most modern mathematicians in restricting his attention to that narrow collection of entities that all modern coordinate-free treatments of gravitation or (pseudo-)Riemannian geometry presently discuss, namely tensors and connections, but not, for example, tensor densities (especially of arbitrary real weight), which many such treatments neglect. This neglect is partly justified, as will appear in the next section, but only partly so, because sometimes mathematicians need to consider irreducible geometric objects (Zajtz, 1988b). Anne Hiskes also neglects densities; though she does not imply that they do not exist or are not geometric objects, she does imply that they are not among the most important kinds (Hiskes, 1981, p. 57). Considering how widespread the neglect of tensor densities is nowadays, and how important they are in treating the real or apparent counterexamples discussed

Scheive, 2004) and so are nonlocal in their finite forms. More famously, supersymmetry, which appears in supergravity and superstring theory, nontrivially combines internal and external symmetries (van Nieuwenhuizen, 1981). Both examples became known after Anderson’s work. I note in passing that the mixing of internal and external symmetries in supersymmetry is potentially more interesting and puzzling philosophically than the hole argument of GTR. The taxonomy of symmetries in TLL (Thorne et al., 1973) is more capacious, but still does not comfortably accommodate Einstein’s equations with a background metric.
in this work, I will discuss them in the next section. To accommodate the richness of geometric objects without using components, one needs to employ natural bundles (Fatibene and Francaviglia, 2003; Nijenhuis, 1972). However, the need for explicit computations and the emphasis on local equivalence (to be discussed shortly) make the component formalism especially convenient. It seems that no detailed treatment of the Ogievetsky-Polubarinov spinor formalism (Ogievetskiǐ and Polubarinov, 1965) has yet been given in any case.

Third, while Friedman considers variously rich and spare versions of what is intuitively one theory (Newtonian gravity) and states a methodological preference for spare theories, his treatment lacks the firm resolve of Anderson’s demand that “irrelevant” variables be eliminated. This requirement is also imposed by TLL (Thorne et al., 1973) and discussed by Norton (Norton, 1993). One can readily adopt the Andersonian prohibition of irrelevant variables to express Friedman’s intuitions about “natural” choices of variables (Friedman, 1983, p. 59) in relation to the Jones-Geroch dust counterexample.

A fourth difference pertains to the notion of standard formulations of a theory. Anderson argues (somewhat confusingly) that theories should be coordinate-covariant under arbitrary manifold mappings (Anderson, 1967, p. 82); this move seems to be offered as a substantive claim rather than a conventional choice. More understandably, TLL stipulate that the standard form of a theory be manifestly coordinate-covariant. Friedman, by contrast, takes as standard a form in which the absolute objects, if possible, have constant components (Friedman, 1983, p. 60) and so have limited coordinate freedom. Friedman implies that one can always choose coordinates such that the absolute objects (a) have constant components and (b) thus drop out of the theory’s differential equations, which then pertain
to the dynamical objects alone. However, claim (a) is falsified by the counterexample of (anti-) de Sitter space-time as a background (for example, (Logunov et al., 1991; Rosen, 1978)) for some specific curvature value. These space-times of constant curvature, at least for a fixed value of the curvature, satisfy Anderson’s and Friedman’s definitions of absolute objects for the space-time metric, but the components of the metric cannot be reduced to a set of constants. (See also various examples by Andrzej Zajtz (Zajtz, 1988b).) An analogous example with spatial curvature is also available. Anderson makes some effort to identify the ‘correct’ or best formulation of a theory, a task taken up in more detail by TLL (Thorne et al., 1973). The latter authors’ “fully reduced generally covariant representation” of a theory, unlike Friedman’s “standard formulation” (p. 60), retains the full coordinate freedom by leaving the absolute objects as world tensors (or tensor densities, connections, or whatever they are). Friedman’s expectation that absolute objects be expressible using constant components is too strong to apply in every example. Claim (b), that absolute objects with constant components drop out of the field equations, is falsified by the example of massive versions of Einstein’s theory (Babak and Grishchuk, 2003; Freund et al., 1969; Ogievetsky and Polubarinov, 1965; Pitts and Schieve, 2007; Visser, 1998). After a lull from the mid-1970s to the mid-1990s, massive variants of gravity have received considerable attention from physicists lately, especially particle physicists. In those theories such that the background space-time metric is flat, its components can be reduced to a set of constants globally by a choice of coordinates, but the background metric still does not disappear from the field equations because it appears in them algebraically, not merely differentially as Friedman apparently assumed tacitly. Especially because (a) is false, the Thorne-Lee-Lightman fully reduced
generally covariant formulation is therefore preferable to Friedman’s standard formulation, which fails to exist in some interesting examples. However, if one’s goal is more historical, so that Newtonian gravity and special relativity without gravity are the main theories of interest, then Friedman’s standard formulation suffices to illustrate the role of the Galilean and Poincaré groups, respectively.

Friedman’s expectation that the components of absolute objects could be reduced to constants in general, though incorrect, usefully calls attention to the role (or lack thereof) of Killing vector fields and the like in analyzing absolute objects. If the (anti-) de Sitter space-time examples show that constancy of components is too strict a criterion, the next best thing is to have still a maximal set of 10 Killing vector fields in four space-time dimensions, but with more non-commutation as in (anti-) de Sitter cases in comparison with Minkowski space-time. One could generalize requirements on Killing vector fields in various ways (Kramer et al., 1980). Because absolute objects need not be metric tensors, the general notion is not Killing vector fields, but generalized Killing vector fields, that is, vector fields such that the Lie derivative of the absolute objects vanishes. Certainly some notion of constancy is one of the core intuitions that one has about absolute objects, though it plays no role in Anderson’s definition of absolute objects, as John Earman has noticed (Earman, 1974). Newton’s claim that absolute space “remains similar and immovable” is suggestive of symmetry within a model (Earman, 1989), not merely similarity between models. Standard examples of absolute objects usually have a fair number of generalized Killing vector fields. In Anderson’s terminology, most typical theories will have fairly large symmetry groups. Usually at least a 7-parameter family of space- and time-translations and spatial rotations will be in the symmetry group, as in classical mechanics (Goldstein, 1980).² In GTR

²It is insufficiently often noticed that Galilean symmetry in classical physics follows from the
(including suitable matter fields), the lack (or scarcity, rather, given the Geroch-Giulini counterexample) of absolute objects implies a vast symmetry group. This large group of all diffeomorphisms (or all volume-preserving ones, rather, given the Geroch-Giulini counterexample) as symmetries of the absolute objects, in turn, leads to an embarrassment of riches concerning local conservation laws, albeit non-covariant and not unique, or else dependent on some auxiliary structure such as a vector field (Anderson, 1967; Bergmann, 1958; Komar, 1959; Lee et al., 1974). From this fact follows the so-called nonlocalizability of gravitational energy. If time translation invariance were required for absolute objects, then that criterion could exclude Norton’s counterexample involving Robertson-Walker metrics (Norton, 1993, p. 848). Absolute objects (or rather, nonvariational objects, given that the two will prove not to be coextensive) that lack translational Killing vector fields will generally fail to conserve energy-momentum (Trautman, 1966). The most typical and plausible examples of absolute objects do not apply forces that violate conservation laws; those that do, might well be called miraculous.

One change for the better made by Friedman (Friedman, 1973, 1983) and also by Hiskes (Hiskes, 1984) is the shift from Anderson’s criterion of global equivalence (modulo the covariance group) to local equivalence between arbitrary neighborhoods. In this way a theory does not avoid having absolute objects simply by admitting different spacetime topologies. With the question of absoluteness involving merely neighborhood-by-neighborhood equivalence, one can consider topology as a separate issue, or ignore it, as the case may be. One should note, however, that GTR likely fails to have $\sqrt{-g}$ as an absolute object by Anderson’s criterion of global equivalence. The vulnerability of GTR to the Geroch-Giulini $\sqrt{-g}$ quite contingent assumption of the absence of fourth or higher even powers of the velocity in the Lagrangian for a free particle. The restricted principle of relativity is not an *a priori* truth, perhaps unlike the homogeneity of time and the homogeneity and isotropy of space.
counterexample therefore depends to some degree on the shift from Anderson’s global equivalence criterion to the neighborhood-by-neighborhood equivalence criterion of Friedman and Hiskes. This is perhaps no small point, given that this counterexample makes GTR have an absolute object and thus falsifies the most famous application of the Anderson-Friedman absolute objects project. However, it seems on reflection that the local equivalence criterion really is preferable to Anderson’s global one, and one should not choose an inferior criterion merely to avoid counterexamples.

It is worth stressing (but was not in (Pitts, 2006a)) that the Friedman-Hiskes criterion is not simply logically weaker than Anderson’s criterion; in some respects their criterion is stronger than Anderson’s. While Anderson’s criterion requires sameness of a geometric object’s values (up to covariance transformations) between one whole space-time and another, nothing requires that every neighborhood be alike within a single space-time. By contrast, Hiskes’s Definition 17 requires equivalence between field values in a neighborhood about an arbitrary point \( p \) in manifold \( M \) and field values in a neighborhood about an arbitrary point \( p' \) in manifold \( M' \). It follows that the values are equivalent between any two points in \( M \) as well. Matters are perhaps less clear in Friedman’s book (Friedman, 1983) than in his earlier article (Friedman, 1973). In an appendix of the book he specifies that one attend to “global diffeomorphisms on a topologically ‘nice’ model of the theory” (Friedman, 1983, p. 360). In the main body Friedman’s definition DE (Friedman, 1983, p. 58) he employs two neighborhoods that both contain the point \( p \) in the manifold \( M \), thereby avoiding explicit trans-world comparisons, but requiring some implicit criterion nonetheless. Much depends on how one comes to identify the parts of the space-time. If space-time points get their identities
by virtue of what happens there, as relationalists tend to hold, then comparisons
around the same point \( p \) in \( M \) will tend to some degree to involve similar phe-
omena, because those phenomena help to constitute that point as \( p \). However,
trans-world identification of points might not be available, because Jeremy But-
terfield and Brent Mundy have both rejected trans-world identity in response to
the hole argument (Butterfield, 1989; Mundy, 1992). On the other hand, if space-
time points are individuated independently of what happens there, as absolutists
tend to hold, then dissimilar events are more apt to happen at the same point \( p \) in
different models. The difference between Anderson’s definition and Hiskes’s makes
a difference in, for example, the case of a modern field-theoretic analog of Aristo-
tle’s mechanics, if such a thing is possible. Both the bottom of the world and the
sphere of the moon are special locations in Aristotle’s mechanics. Supposing that
one could represent something along those lines using modern field theory, the
inhomogeneity of Aristotle’s space would violate Hiskes’s criterion of absoluteness
while perhaps satisfying Anderson’s. Hiskes’s definition implies that absolute ob-
jects are homogeneous throughout space-time and thus have (generalized) Killing
vector fields of some sort.

2.2 Tensor Densities

Considering the neglect into which tensor densities have fallen in recent years,
not least in differential geometry texts, it is worthwhile to recall how useful they
are, if not essential in some applications. The conformal metric density \( \hat{g}_{\mu\nu} \)
with (anti)-unit determinant picks out the local null cones in the pseudo-Riemannian
case, without any extraneous information (Ehlers et al., 1972), such as pertaining
to volumes. Sometimes using tensor densities is merely a matter of convenience,
as in Joshua Goldberg’s Lagrangian density for GTR using $g^{\mu\nu} = \sqrt{-g}g^{\mu\nu}$ as the field variable (Goldberg, 1958; Halpern, 1963; Landau and Lifshitz, 1975; Scharf, 2001) in order to have fewer terms than the metric or its inverse yields. However, not infrequently more is at stake than mere convenience. Some time ago Bryce DeWitt discerned the significance of two of the three equations of the Dirac ‘algebra’ of the Hamiltonian and momentum constraint functions $\mathcal{H}_0$ and $\mathcal{H}_i$, respectively, in canonical GTR: the two Poisson brackets involving $\mathcal{H}_i$ have their precise form simply because $\mathcal{H}_0$ and $\mathcal{H}_i$ are a weight 1 scalar density and a weight 1 covector density, respectively (DeWitt, 1967a; Sundermeyer, 1982) (and there are no background geometric objects). In the literature on modern nonperturbative canonical quantization of gravity with Ashtekar’s new variables and the like, tensor densities are used routinely. Some authors write densities in a way that makes their weight manifest: a weight 2 density has two tildes over it, a weight $-1$ density has a tilde below it, etc. (Jacobson and Romano, 1992; Salisbury et al., 2000). Moreover, the use of a densitized lapse function has proven useful in 3+1-dimensional treatments of the initial value problem\textsuperscript{3} in GTR, the polynomialization of the constraints (Tate, 1992), and the dynamical preservation of the constraint equations (Anderson and York, 1998; Jantzen, 2004).

Perhaps these uses of densities are matters of convenience rather than necessity, because one can simulate tensor densities of integral weights using tensors with enough sets of antisymmetrized indices, as in the case of the volume form (though not all behaviors under transformations with negative determinant can be recovered). However, this procedure is not so obviously available for most densities

\textsuperscript{3}It is now customary in numerical general relativity to call the problem of inferring later or earlier states of a system from initial data the “Cauchy problem,” while the term “initial value problem” is reserved for the procedure of solving the constraint equations to get a set of initial data. This latter sort of problem exists only for constrained theories, such as GTR or Maxwell’s electromagnetism.
of non-integral weight; it is generally unclear, for example, what a quantity with a third of an index or \( \pi \) indices would mean. But tensor densities of fractional weight have been used in applications such as the conformal-traceless decomposition of André Lichnerowicz and James York in solving GTR’s initial value constraints in numerical general relativity (Brown, 2005a; York, 1972), unimodular variants of GTR (discussed in (Earman, 2003a; Unruh, 1989)), and quantum gravity (DeWitt, 1967b; Dirac, 1959; Katanaev, 2005; Leonovich and Mladenov, 1993; Peres, 1963). Spinors with density weight \( \frac{1}{2} \) have been used to avoid complex-valued canonical variables and hence to retain access to some important quantization techniques in modern canonical quantum gravity (Bojowald and Das, 2007; Bojowald et al., 2008; Nelson and Teitelboim, 1978; Thiemann, 1998a,b). Weight \( \frac{1}{2} \) fields including spinors have also been useful for path integration (Fujikawa, 1983; Fujikawa et al., 1985; Fujikawa and Yasuda, 1984) (and references therein). Some of these uses arise because breaking the metric tensor into irreducible pieces yields densities of fractional weights, which vary with manifold dimension. Densities with \textit{irrational} weights are, if not essential, at least very useful in deriving and expressing some massive variants of Einstein’s GTR (Ogievetsky and Polumbarinov, 1965; Pitts and Schieve, 2007) and some massive scalar variants (Pitts, 2009) of Nordström’s scalar gravity. Discovering or inventing these theories without densities of irrational weights would be difficult indeed, and for the scalar case, only one massive case has been discussed previously (Freund and Nambu, 1968). Perhaps this is merely a practical matter. For example, the tensor density \( G^{\mu\nu} \equiv_{def} g^{\mu\nu}(-g)^{\pi+\epsilon+\sqrt{2}} \) is equivalent as a geometric object to the inverse metric \( g^{\mu\nu} \) and also to the metric \( g_{\mu\nu} \). In terms of fiber bundles, \( G^{\mu\nu} \) is just another set of coordinates to cover the same fiber that the metric tensor and the inverse metric
cover—perhaps an eccentric choice of fiber coordinates for most purposes, excepting certain massive theories of gravity. One can see, then, that tensor densities do not need explicit consideration for at least some purposes of modern mathematicians. The component notation is doubtless a clumsier way to discuss the equivalence of geometric objects than is fiber bundle theory (Nijenhuis, 1972). It is the concept of irreducible geometric objects, as will appear below, that makes consideration of densities obligatory.

Hermann Weyl protested in 1920 against early clumsy efforts at component-free formalisms “which are threatening the peace of even the technical scientist” (Weyl, 1921, p. 54); fortunately a few modern authors have accommodated densities of arbitrary weight in a modern fashion\(^4\) (Calderbank and Pedersen, 1999; Cartier et al., 2002; Duval and Ovsienko, 1998; Fatibene et al., 1997; Fatibene and Francaviglia, 2003; Godina and Matteucci, 2003; Lang, 1995; Lee and Wald, 1990; Spivak, 1979; Trautman, 1970), though often not in a thorough and accessible way. One should also notice that tensor densities come in more than one kind; some can be of any real weight, while others are essentially of integral weight (Golab, 1974). By contrast to the present situation, tensor densities were once such standard parts of mathematical physics that they, in contrast to most types of geometric objects, shared the term “quantity” with tensors (Golab, 1974; Schouten, 1954). The Torretti and Norton (alleged) counterexamples and the scalar density counterexample (discussed below) that finds an absolute object in GTR are most readily discussed using tensor densities. Torretti himself elsewhere briefly discusses tensor densities, at least of weight 1 (Torretti, 1996, p. 317). Robert Dicke excluded densities for no apparent reason in his treatment of experimental gravitation (Dicke, 1965, p. 50), though he does admit doing so. If one has

\(^4\)I thank Robert Geroch for emphasizing this point.
a metric tensor already on hand (which Dicke did not), as in later frameworks for experimental gravitation (Earman, 1992; Thorne and Will, 1971), then neglecting densities does not exclude any theories because one can de-densitize the variables of any theory of interest with $\sqrt{-g}$ (assuming that one identifies as the “same theory” mathematically equivalent theory formulations in this way). One is still unable to recognize the absolute object contained in $\sqrt{-g}$ itself, however. As Thorne and Will note, Dicke implicitly excludes connections as well; one notices that if one already had a metric on hand, then excluding connections would also be innocuous, because an arbitrary connection could be expressed in terms of the metric, the torsion tensor and the nonmetricity tensor (Schouten, 1954). In the absence of a metric, however, Dicke’s restriction to scalars, vectors and tensors is judged by Thorne and Will “a dangerous policy” (Thorne and Will, 1971, p. 598).

Anderson did not neglect tensor densities (Anderson, 1967; Anderson and Finkelstein, 1971). He even noticed that the vanishing of the Weyl curvature tensor implied that the conformal metric density (which I write as $\hat{g}_{\mu\nu}$) is absolute, though the metric tensor is not (Anderson, 1967, pp. 77, 84). Anderson also distinguished irreducible and reducible geometric objects:

Whenever a geometric object can be broken up into parts that transform among themselves, we say that we have a reducible object. If no such decomposition is possible, we have an irreducible object. (Anderson, 1967, p. 20) (emphasis in the original)

John Stachel also mentions the notion of irreducibility of geometric objects (Stachel, 2002, p. 255). Rather than making a systematic move that rendered densities invisible, Anderson’s treatment has all the necessary resources, but simply was not quite consistently applied concerning the absoluteness of $\sqrt{-g}$. In any case, we have with densities examples of a problem noted by M. Ferraris, M. Francaviglia
and C. Reina:

In recent years, owing to their greater generality, geometric objects other than tensors began to enter physical applications, because in many cases using objects more general than tensors is essential [list of references omitted]. In fact, in spite of the widely known and systematic use of tensorial methods in mathematical physics, restricting ones [sic] attention to tensors may often turn out to be misleading. (Ferraris et al., 1983a, p. 120)

One should question the history implied here, as tensor densities have become less well known since the 1960s as component-based differential geometry has fallen somewhat out of favor and the more modern treatments often have not discussed densities. Given the tendency of Euler-Lagrange equations to come with nontrivial density weight \((1 - w)\) for a field of weight \(w\), it is difficult to avoid density weight (unless perhaps one makes the Lagrangian density a 4-form (Wald, 1984), hardly a move in the direction of simplicity). Mishandling or neglecting densities is dangerous but not uncommon.

Whatever the historical trends, the systematic lesson of Ferraris, Francaviglia, and Reina remains. Besides failing to accommodate densities, Friedman’s mathematical language is also inadequate to express the techniques used by V. I. Ogievetsky and I. V. Polubarinov (OP) in their atypical treatment of spinors coupled to gravity using a “square root of the metric” (Ogievetskiĭ and Polubarinov, 1965). Although they do not use the terminology, in effect they treat the spinor field and the metric together as a nonlinear geometric object (or nearly so, given the curious spinorial two-valuedness and perhaps some loose restrictions on coordinates). This OP spinor formalism should be useful in preventing the time-like leg (or a space-like leg) of the orthonormal tetrad, which is typically used with spinors, from counting as an unwanted absolute object. While the Geroch-
Giulini $\sqrt{-g}$ counterexample prevents GTR from lacking absolute objects, the Ogievetsky-Polubarinov formalism is useful in showing that coupling fermions to gravity does not introduce any absolute objects.

Sometimes one wants spinors, the “square root of the metric” and densities together. For example, the Dirac operator $\gamma^A e^\mu_A \nabla_\mu$ that acts on spinors is conformally covariant (Branson, 2005; Choquet-Bruhat and DeWitt-Morette, 1989). I notice that one can show using densities (including densitized spinors) that there is a conformally invariant Dirac operator lurking in the vicinity; the appropriate spinor turns out to have weight $\frac{3}{8}$ in four space-time dimensions or, more generally, $\frac{n-1}{2n}$ in $n$ space-time dimensions. That conformally invariant Dirac operator is $\gamma_\mu \hat{r}^{\mu\nu} \nabla_\nu \psi_w$, where $\gamma_\mu$ denotes a set of numerical Dirac matrices, $\hat{r}^{\mu\nu}$ is the symmetric square root of the inverse conformal metric density $\hat{g}^{\mu\nu}$, $\nabla_\nu$ is the Ogievetsky-Polubarinov covariant derivatives for spinors (Ogievetski˘i and Polubarinov, 1965) with the density weight term (with the weight altered to match the usual western rather than Russian conventions), and $\psi_w$ is a spinor with weight $w = \frac{n-1}{2n}$. No use is made of any scalar density in defining this operator, so it is a concomitant of just the weighted spinor and the conformal metric density. It turns out that the Lagrangian density in this formalism with the densitized variables is also manifestly conformally invariant in any dimension, because

$$L = \sqrt{-g} \bar{\psi} \gamma_\mu r^{\mu\nu} \nabla_\nu \psi = \bar{\psi}_w \gamma_\mu \hat{r}^{\mu\nu} \nabla_\nu \psi_w.$$  \hfill (2.1)

$\sqrt{-g}$ does not appear even algebraically in the latter expression. Analogous results hold for the conformally covariant coupling of scalar fields (c.f. (Choquet-Bruhat
and DeWitt-Morette, 1989; Wald, 1984)): using scalar (density) fields with weight

\[ w = \frac{n - 2}{2n}, \]  

(2.2)

\( \sqrt{-g} \) does not appear in \( L \). Much of the apparatus of conformal rescaling of fields (Choquet-Bruhat and DeWitt-Morette, 1989; Wald, 1984) can be avoided by the use of suitably weighted densities, which automatically depend on the metric in the appropriate way (or so one assumes initially, before taking the density as primitive) and do not change under conformal rescalings.

By using conformal invariance rather than conformal covariance, one avoids the introduction of surplus structure \( \sqrt{-g} \) introduced only for the purpose of noticing its irrelevance. Similar remarks could be made concerning conformal Killing vector fields. It is inefficient to introduce a metric, take its Lie derivative, notice that for some vector fields the result is proportional to that metric \( \mathcal{L}_\xi g_{\mu\nu} = fg_{\mu\nu} \) for some function \( f \), and then announce that such vector fields are conformal Killing vectors, though many books do this. Instead, recalling how to differentiate densities (Anderson, 1967; Schouten, 1954), one can simply note that the vectors in question give vanishing Lie derivative for the conformal metric density:

\[ \mathcal{L}_\xi \hat{g}_{\mu\nu} = 0. \]  

(2.3)

It is ironic that this tendency has resulted from the supposed elimination of surplus structure in modern-style differential geometry. Showing that one does not need any volume element at all, not simply that the choice of volume element is irrelevant, exemplifies a good policy (whether it is needed for scalar densities or not). Some kinds of geometric objects simply do not exist on certain manifolds,
such as a continuous nonvanishing vector field everywhere on the two-sphere (the “hairy ball theorem”) (Spivak, 1979, p. 94) (Dodson and Poston, 1991, pp. 183, 300) (Milnor, 1969, p. 30). Thus demonstrating the independence of the results from any specific choice of auxiliary object does not imply that the mere existence of the auxiliary object is an innocent assumption. By contrast, coordinates are not superfluous, because they are already introduced in the definition of a manifold (Wald, 1984).

As useful as it is to break metric tensors into their irreducible pieces (Dirac, 1959; Rovelli, 1989; York, 1972), which are densities, the (anti-)unit determinant condition on the components of conformal metric densities can cause complications in the context of variational principles, because the components are not all independent. For example, if one breaks the spatial metric $h_{ij}$ into its unimodular conformal part $\hat{h}_{ij}$ and (squared) volume element $h$ as $h_{ij} = \hat{h}_{ij} h^{\frac{1}{2}}$, the Poisson bracket between $\hat{h}_{ij}$ and its conjugate momentum is field-dependent, while the bracket between that momentum and itself fails to vanish, both of which behaviors are unusual. The correct answer has been given in the literature (Isenberg and Nester, 1980; Katanaev, 2006; Tate, 1992), but so has a tempting wrong answer.

This paean for tensor densities and lament over their neglect would not fully serve its purpose without a brief exposition of densities. Tensor densities, which are sometimes called relative tensors, come with as many collections of contravariant and covariant indices as do tensors, along with two further bits of information, namely, the density weight $w$ and the variously named character of either taking absolute values of the Jacobian in the transformation law (thus preserving the sign of the components) or not. All tensor densities can be obtained as products of tensors and scalar densities, so only scalar densities need much discussion at the
moment. A scalar density field $\phi$ of weight $w$ ($w$ being any real number) transforms, under a change of local coordinates from an unprimed set $x^\mu$ to a primed set $x'^\nu$, as

$$\phi' = \epsilon \left( \frac{\partial x}{\partial x'} \right) \phi,$$

where $\epsilon \left( \frac{\partial x}{\partial x'} \right) = 1$ for Weyl densities, $\epsilon \left( \frac{\partial x}{\partial x'} \right) = \text{sign} \left( \det \left( \frac{\partial x}{\partial x'} \right) \right)$ for the other important kind (Golab, 1974). The primes perhaps are opposite where one might have expected, but this is the usual convention (Anderson, 1967; Golab, 1974; Schouten, 1954) (though some authors, especially but not only Russian, define weight in the opposite fashion (Carroll, 2004; Katanaev, 2005; Ogievetsky and Polubarinov, 1965)). For $w = 0$ a Weyl density is a scalar, while the other kind of density is a pseudoscalar. Using these notions for scalar densities, one can define all sorts of tensor densities in the obvious ways. The covariant and Lie derivatives of densities have an extra term dependent on $w$ (but not on $\epsilon$). For a weight $w$ $(1,1)$ tensor density $\phi^\alpha_{\beta}$, the covariant derivative is (Anderson, 1967; Israel, 1979; Schouten, 1954)

$$\nabla_\mu \phi^\alpha_{\beta} = \phi^\alpha_{\beta,\mu} + \phi^\alpha_{\beta\gamma} \Gamma^\gamma_{\sigma\mu} - \phi^\alpha_{\sigma} \Gamma^\gamma_{\beta\mu} - w \phi^\alpha_{\beta\sigma} \Gamma^\gamma_{\sigma\mu},$$

(2.5)

while the Lie derivative with respect to a vector field $\xi^\mu$ is

$$\mathcal{L}_\xi \phi^\alpha_{\beta} = \xi^\mu \phi^\alpha_{\beta,\mu} - \phi^\mu_{\beta} \xi^\alpha_{,\mu} + \phi^\alpha_{\mu} \xi^\beta_{,\mu} + w \phi^\alpha_{\beta} \xi^\mu_{,\mu}.$$  

(2.6)

One can derive the Lie derivative from the coordinate transformation law by standard methods involving an infinitesimal transformation and Taylor expansion. One can show that the commutator of the covariant derivative of a scalar density is nonvanishing in general (Schouten, 1954, p. 140), but it vanishes for a con-
nection that is torsion-free and that has vanishing contraction on a certain pair of indices of its curvature tensor (which holds at least for connections compatible with some metric) (Edgar, 1992; Eisenhart and Veblen, 1922; Schouten, 1954; Wald, 1984).

If one wishes to extend the notion of a density to admit components with respect to an arbitrary ordered basis, not just a coordinate basis, one approach that suggests itself is that for a basis of vectors \( \vec{v}_B \) related to another basis \( \vec{v}_A \) by the relation \( \vec{v}_B = M^A_B \vec{v}_A \), the scalar density’s component with respect to each basis satisfies the relation

\[
\phi(2\vec{v}_B) = \epsilon(M) |\text{det} M|^w \phi(1\vec{v}_A),
\]

where \( \epsilon(M) = 1 \) or \( \epsilon(M) = \text{sign}(|\text{det} M|) \). The matrix \( M^A_B \) is any nonsingular \( n \times n \) matrix, so the density is expressed as a representation of \( \text{GL}(n) \), the general linear group in \( n \) dimensions. Just what the density does to the basis—what sort of machine it is that somehow eats vectors, to use the colorful metaphor of Misner, Thorne, and Wheeler (Misner et al., 1973)—is not terribly clear from this expression.

A more perspicuous answer to this question can be obtained by adapting an expression from Spivak (Spivak, 1979, p. 314). An \( n \)-form \( \omega \) (a totally antisymmetric rank \( n \) covariant tensor) on an \( n \)-dimensional manifold has \( n \) slots for contravariant vectors. The component of \( \omega \) with respect to an arbitrary ordered basis \( \vec{v}_A \) is \( \omega(\vec{v}_1, \ldots, \vec{v}_n) \). One can define a scalar density \( \phi \) of weight \( w \geq 0 \) (a restriction imposed to facilitate densities that can vanish, except where the meaninglessness
of 0\(^0\) intrudes) as

\[ \phi(\vec{v}_1, \ldots, \vec{v}_n) = \epsilon(\omega(\vec{v}_1, \ldots, \vec{v}_n))|\omega(\vec{v}_1, \ldots, \vec{v}_n)|^w, \]

(2.8)

where \(\epsilon(\omega(\vec{v}_1, \ldots, \vec{v}_n)) = 1\) or \(\epsilon(\omega(\vec{v}_1, \ldots, \vec{v}_n)) = sign[\omega(\vec{v}_1, \ldots, \vec{v}_n)]\). Thus if tensors are linear or multilinear maps, then densities are power-law maps (perhaps up to a sign). This construction fineses the problem of making sense of a fractional or irrational number of indices. To get densities \(\psi\) that have negative weight and that can vanish, one can perform the mirror-image construction with a totally antisymmetric rank \(n\) contravariant tensor \(\chi\) that is fed an ordered basis of covectors \(v^A\):

\[ \psi(v^1, \ldots, v^n) = \epsilon(\chi(v^1, \ldots, v^n))|\chi(v^1, \ldots, v^n)|^w, \]

(2.9)

where \(w \geq 0\) gives a weight \(-w\) density and \(\epsilon(\chi(v^1, \ldots, v^n)) = 1\) or \(\epsilon(\chi(v^1, \ldots, v^n)) = sign[\chi(v^1, \ldots, v^n)]\). By feeding the density a coordinate basis, one can derive the coordinate transformation formula. Whether these so-called ‘coordinate-free’ expressions are advantageous or not, they make contact with recent modes of expression in differential geometry. One possible source of inconvenience for regarding densities as machines that eat vector fields, as opposed to components with a certain transformation law, is that topological restrictions can inhibit the existence of the desired vector fields. A famous example in two dimensions with positive metric is the “hairy ball theorem” for the sphere \(S^2\) (Spivak, 1979, p. 94) (Dodson and Poston, 1991, pp. 183, 300) (Milnor, 1969, p. 30).
2.3 Confined Objects and Global Space-time Topology

While absolute objects and dynamical objects are mutually exclusive, TLL introduce a third category of “confined” objects (Thorne et al., 1973, pp. 3568, 3569). Examples given include universal constants. TLL clearly intend for confined objects not to overlap with absolute objects. Confined objects are there defined as “those which do not constitute the basis of a faithful representation of the [manifold mapping group]” (Thorne et al., 1973, p. 3568); this claim means (p. 3567) that there exist two distinct elements of the manifold mapping group that produce identical mappings of the confined objects. Above it was noted that Anderson requires that absolute objects “constitute the basis of a faithful realization of the covariance group of the theory.” (Anderson, 1967, p. 83) Nothing can transform faithfully and unfaithfully under the covariance group, which includes the coordinate transformations, so indeed nothing is both an absolute object and a confined object.

While the definition of confined objects allows them to transform in an unfaithful but sometimes nontrivial way under diffeomorphisms, many interesting examples are quantities that do not transform nontrivially at all. Some entities that seem intuitively absolute, but do not satisfy Anderson’s definition, fit into the category of confined objects. Besides TLL’s example of universal constants, some other things unaffected by coordinate transformations but useful in physical theories include the identity matrix, the Lorentz matrix $\text{diag}(-1,1,1,1)$ or any comparable signature matrix, fixed Dirac $\gamma^\mu$ matrices, the $2 \times 2$ Pauli spin matrices, Lie group structure constants, and Oswald Veblen’s “numerical tensors” (where “tensors,” in Veblen’s usage, included tensor densities). These are the

---

5To avoid confusion with philosophical terminology, as Roberto Torretti urged, I call these things “confined objects” rather than “confined variables” as in (Thorne et al., 1973).
Kronecker $\delta_{\mu}^{\nu}$ symbol, which is trivially a world tensor with components 1 and 0, and the Levi-Civita totally antisymmetric $\epsilon$ symbol with values 1, -1, and 0; these values are the components of both $e^{\alpha\beta\gamma\delta}$, a contravariant rank 4 tensor density of weight 1, and $e_{\alpha\beta\gamma\delta}$, a covariant rank 4 tensor density of weight -1 (Anderson, 1967; Misner et al., 1973; Spivak, 1979; Veblen, 1933). Thus at least some confined objects are tensor or tensor density fields and hence geometric object fields of the familiar kind. The various matrices listed above, including the numerical tensors, are coordinate scalars. Their being scalars makes them geometric object fields (though the field nature is trivial), but their being constant makes them transform unfaithfully under coordinate transformations; now even the transport term $\xi_{\mu}^{\alpha}\phi_{\beta\gamma\delta}$ in the Lie derivative vanishes. Thus the identity matrix, the Lorentz matrix, the Dirac and Pauli matrices, Lie group structure constants, and Oswald Veblen’s “numerical tensors” are all geometric objects which are also confined.

2.4 Jones-Geroch Vector Field Counterexample and Friedman’s Reply

With a clear grasp of absolute objects in hand, one can now consider the Jones-Geroch counterexample that claims that the 4-velocity of cosmic dust counts, absurdly, as an absolute object by Friedman’s or Anderson’s standards. Friedman concedes some force to this objection made by Geroch and amplified by Roger Jones, here related by Friedman:

...[A]s Robert Geroch has observed, since any two time-like, nowhere-vanishing vector fields defined on a relativistic space-time are $d$-equivalent, it follows that any such vector field counts as an absolute object according to [Friedman’s criterion]; and this is surely counter-intuitive. Fortunately, however, this problem does not arise in the context of any of the space-time theories I discuss. It could arise in the general relativistic theory of “dust” if we formulate the theory in terms of a quintuple $\langle M, D, g, \rho, U \rangle$, where $\rho$ is the density of the “dust” and $U$
is its velocity field. $U$ is nonvanishing and thus would count as an absolute object by my definition. But here it seems more natural to formulate the theory as a quadruple $\langle M, D, g, \rho U \rangle$ where $\rho U$ is the momentum field of the “dust.” Since $\rho U$ does vanish in some models, it will not be absolute. (Geroch’s observation was conveyed to me by Roger Jones, who also suggested the example of the general relativistic theory of “dust.” . . .) (Friedman, 1983, p. 59)

Here $D$ is the torsion-free covariant derivative compatible with $g$. Other sources, including what Roger Jones reported hearing from Geroch, indicate a qualification to local diffeomorphic equivalence of nonvanishing time-like vector fields (Jones, 1981b, pp. 167, 168) (Jones, 1981a) (Trautman, 1965, p. 84) (Wald, 1984, p. 18) (Dodson and Poston, 1991, pp. 198-200) (Olver, 1995, p. 19). In any case nothing in my argument will depend on global versus merely local equivalence between arbitrary neighborhoods. Jones also distinguishes the local diffeomorphic equivalence of nonvanishing time-like vector fields, which holds in general, from the (local) diffeomorphic equivalence of their covariant derivatives of various orders, which typically does not hold. The diffeomorphic equivalence of vector fields also has interesting consequences for energy-momentum conservation laws in GTR (Bergmann, 1958), because in a neighborhood one can take any non-vanishing vector field as a translation. Invariance under translations gives conservation laws, so the extraordinary richness of conservation laws in GTR readily follows. Anderson also discusses the matter of diffeomorphic equivalence of vector fields (Anderson, 1967, p. 426), but unfortunately does not draw any conclusions about the susceptibility of contravariant vector fields absoluteness.

Below I will argue that Friedman’s response is nearly satisfactory, though it has two weaknesses as he expressed it. First, the statement “$\rho U$ does vanish in some models” ought to have said “$\rho U$ does vanish in some neighborhoods in some models” to show that he is considering only genuine models of GTR + dust (in
which dust vanishes in some neighborhoods in some models), rather than some models with (omnipresent?) dust and some degenerate models which nominally have dust but actually have no dust anywhere. The latter would seem to be a cheat. As it stands, the reader is left to wonder whether such a cheat is doing important work for Friedman, though John Norton correctly read Friedman’s proposal as “relying . . . on the possibility that \( \rho \) vanishes somewhere” (Norton, 1993, p. 848). Clearly some models with dust have neighborhoods lacking dust, and it is these models which will prevent the dust 4-velocity from constituting an absolute object. Second, Friedman’s unfortunate notation \( \rho U^\mu \) suggests that the mass current density (which I will call \( J^\mu \)) is logically posterior to \( \rho \) and an everywhere nonvanishing time-like \( U^\mu \). If so, then one has not eliminated the absolute object after all. If a time-like nowhere vanishing \( U^\mu \) exists in the theory, then it is absolute even if \( \rho U^\mu \) vanishes somewhere and so is not absolute. Thus the significance of Friedman’s use of \( \rho U^\mu \) is left obscure. Instead one can take \( J^\mu \) to be the fundamental variable, while the timelike \( U^\mu \) is a derived quantity defined wherever \( \rho \neq 0 \); then \( U^\mu \) would not be a geometric object field. Alternatively, one can take \( U^\mu \) to be meaningful everywhere (and perhaps primitive), but vanishing where there is no dust; \( U^\mu \) would then change discontinuously at edges of the dust. If Friedman had said that \( J^\mu \) “does vanish in some neighborhoods in some models,” or that \( U^\mu \) “does vanish or fail to exist in some neighborhoods in some models,” then these two infelicities would have been avoided. Perhaps it was these expository imperfections that led Roberto Torretti to judge Friedman’s reply \textit{ad hoc} (Torretti, 1984) and John Norton to call it “a rather contrived escape” (Norton, 1993, p. 848). Once these problems are removed, the merit of Friedman’s
intuition shines brightly.\textsuperscript{6}

Below I shall review more discussion of this counterexample in the philosophical literature. Various neglected items from the physics literature will shed light on long-standing philosophical debates about absolute objects. Using the term “variational” for objects which are varied in an action principle (Gotay et al., 2004), and “nonvariational” in the obvious way, one can safely follow Anderson in making “absolute” and “dynamical” mutually exclusive, while leaving open the connection between absoluteness and nonvariationality. It will be shown that there exist theories with variational absolute objects, at least if one does not exclude the Rosen-Sorkin variational principle (Rosen, 1966, 1973; Sorkin, 2002) as somehow illegal. Such a theory can be obtained using Rosen’s trick to fulfill Maidens’s claim that the absolute special relativistic metric could be obtained variationally. However, these theories arguably violate Anderson’s demand to eliminate irrelevant variables. A natural extension of the proscription of irrelevant variables serves to eliminate the Jones-Geroch counterexample: the dust 4-velocity $U^\mu$ does not count as an absolute object for GTR + dust because $U^\mu$ does not exist where there is no dust.

2.5 Redefinition with Variational Principle, Maidens’s Worry, and the Rosen-Sorkin Answer

According to Anna Maidens (Maidens, 1998), Anne Hiskes (Hiskes, 1984) proposed amending the definition of absolute objects so that no field varied in a

\textsuperscript{6}It is perhaps worth noting that Sachs and Wu define dust and perfect fluids such that they have nonzero density everywhere (Sachs and Wu, 1977, pp. 103, 105). Clearly the modal properties of physically semi-realistic matter are not settled by such a definition, which presumably commends itself on grounds of technical or expository convenience.
theory's action principle would be regarded as absolute.\textsuperscript{7} The proposal that non-variationality be an \textit{additional} (nontrivial) necessary condition for absoluteness was predicated on the apparent failure of what had previously seemed to be a true generalization about absolute and dynamical objects, namely that absoluteness and nonvariationality were coextensive. This intuition of coextensivity was held by the master. Anderson wrote:

> In addition to the differences between absolute and dynamical objects discussed in Section 4-3 there is another important difference that appears to be characteristic of these two types of objects. The equations of motion for the dynamical objects can often be derived from a variational principle, especially if these objects are fields. On the other hand, it appears to be the case, although we can give no proof of the assertion, that the equations of motion for the absolute objects do not have this property... In the following discussion we will assume that the equations of motion for the dynamical objects of a theory follow from a variational principle and that those for the absolute elements do not. (Anderson, 1967, pp. 88, 89)

Thus Anderson suspected that most or all dynamical objects are variational, while no absolute object is variational. Similar intuitions are manifest in the TLL and LLN papers (Lee et al., 1974; Thorne et al., 1973); such a requirement also appears in their notion of being “Lagrangian-based” (Thorne et al., 1973, p. 3573). Recently John Earman has found it convenient to use “absolute” to mean non-variational (Earman, 2003a), as has Robert Wald (Wald, 1993) with Vivek Iyer (Iyer and Wald, 1994) (with some different terminology).\textsuperscript{8} Anderson was quite sensitive to the possibility of reformulating what intuitively seems like the same theory using various different sets, and indeed increasingly large sets, of variables in an action principle (Anderson, 1967, section 4.2). Unlike Hiskes, he strove to

\textsuperscript{7}It is not obvious where Maidens finds some of the most interesting ideas that she ascribes to Hiskes's paper.

\textsuperscript{8}I thank Ted Jacobson for calling ((Iyer and Wald, 1994)) to my attention.
define a unique correct formulation that gave the expected answers.

Maidens then provisionally rejected the idea that redefinition of absoluteness to include nonvariationality could be deployed to remove the Jones-Geroch counterexample (Maidens, 1998). If absolute objects must be nonvariational, while the dust 4-velocity is variational, then the dust 4-velocity is not absolute. Following Hawking and Ellis (Hawking and Ellis, 1973), Maidens indicates how the equations for the timelike vector field can be derived from a variational principle.\(^9\) However Maidens is also sensitive to the large variety of choices of variables and even the number of field components in an action principle for what intuitively counts as a single theory. Thus she expected such a use of this redefinition to fail, because it eliminates the Jones-Geroch counterexample at the cost of introducing a new one. More specifically, Maidens has suggested that there might be some way to reformulate special relativistic theories such that the flat metric, which surely ought to count as absolute, is varied in the action principle. If that could be done, then definition of absolute objects as including nonvariationality would prove to be too strict (the opposite problem from what the Jones-Geroch example suggests about Friedman’s), because it fails to count the metric tensor of special relativity as an absolute object. (Maidens presumably should envision a weakly generally covariant formulation of special relativity, though her notation is far from clear on that point.) “At this stage, however, we find a fly in the ointment, for its turns

\(^9\)One notices that Hawking and Ellis use a fluid variational principle with constrained variations, not the more familiar unconstrained variations. In some respects this is a disadvantage, though Schutz and Sorkin observe that it keeps one closer to the physical variables (Schutz and Sorkin, 1977). They also observe that in many cases, including this one, one can eliminate the constraints on the variation (not to be confused with constraints in the sense of gauge theories (Sundermeyer, 1982)) using Lagrange multipliers. It seems to me that John Ray’s variational principle might be preferable in the present context, because it involves varying \(U^\mu\) itself and uses unconstrained variations (Ray, 1972). There are some subtleties pertaining to the use of constrained variations; in this fluid example, one of two commonly used formalisms yields absurd conclusions (Bibbona et al., 2007).
out that given suitable starting assumptions we can derive the Lorentz metric from an action principle.” (Maidens, 1998, p. 262) Supporting such claims would involve actually displaying a suitable Lagrangian density whose Euler-Lagrange equations give the desired results or else citing a source where such work had been done. Surprisingly, she fails to do either one. Success would involve finding an action principle for which the flatness of the metric holds for all models (her case (c)), not just some (her case (a), p. 265). A bit later she finds that “it is an open question as to whether the metric of special relativity is derivable from an action principle.” (p. 266) Two pages later she once again claims that “some of the physically necessary fixed background, e.g. the Lorentz metric, can also be derived from an action principle.” (p. 268) It is not easy to harmonize these fluctuating statements.

Fortunately Maidens’s expectation that the flatness of a metric (for all models) can be derived from a variational principle is in fact correct. The question was resolved by Nathan Rosen in the 1960s (Rosen, 1966, 1973). He used an action principle involving a Lagrange multiplier field with 20 components, a trick recently used also by Rafael Sorkin and Domenico Giulini (Earman, 2006; Giulini, 2007; Sorkin, 2002). Thus requiring absolute objects to be nonvariational (while admitting the Rosen-Sorkin trick as part of an acceptable action principle) gives an excessively strict definition, so the Jones-Geroch counterexample is not adequately addressed thereby. Some objects that should count as absolute can be variational, as Maidens expected. Rosen includes the following term in an action principle (after a change in notation to $\eta_{\mu\nu}$ for the metric in question, which is a
priori arbitrary apart from the signature) to force \( \eta_{\mu\nu} \) to be flat:

\[
S = \int d^4x \sqrt{-\eta} R_{\rho\mu\nu\sigma}[\eta] P^{\rho\mu\nu\sigma}. \tag{2.10}
\]

This term is intended as a supplement to the action for a special relativistic theory, within which now \( \eta_{\mu\nu} \) would be subject to variation as well. \( P^{\rho\mu\nu\sigma} \), a tensor with the same symmetries as the Riemann tensor for \( \eta_{\mu\nu} \), serves as a Lagrange multiplier. Varying \( P^{\rho\mu\nu\sigma} \) immediately yields the flatness of \( \eta_{\mu\nu} \). Varying \( \eta_{\mu\nu} \) takes more work and gives an equation of motion especially involving the second derivatives of \( P^{\rho\mu\nu\sigma} \). That equation is not needed here. Rosen seems to make secret use of the Euler-Lagrange equations from \( P^{\rho\mu\nu\sigma} \) to discard terms involving \( R_{\rho\mu\nu\sigma}[\eta] \) in his equations 10, 11, and 12; if so, then his equation 12 is not an “identity” as he claims. Alternately, he might be taking the metric to be flat before the variation but curved after it, as Sorkin proposes (Sorkin, 2002), if that is an intelligible alternative.\(^{10}\) As was noted above, Anderson’s requiring component equality (up to equivalence under the covariance group) only for dynamically possible trajectories is relevant here. Using Rosen’s trick, one has a geometric object such that its components agree for dynamically possible trajectories (“on-shell,” as physicists say) but not for kinematically possible trajectories (“off-shell”), because the metric is not flat for all kinematically possible trajectories.

Anderson briefly states that one must remove irrelevant variables from the theory under analysis. He writes:

It is possible that a subset of the components of the [geometrical object characterizing the kinematically possible trajectories of the theory] do

\(^{10}\)It is worth noting that varying \( P^{\rho\mu\nu\sigma} \) gives a partial differential equation that describes the trajectory \( \eta_{\mu\nu} \), and varying \( \eta_{\mu\nu} \) gives an equation of motion primarily for \( P^{\rho\mu\nu\sigma} \). Thus expressions like “the equations of motion for \( \eta_{\mu\nu} \)” or “the equations of motion for \( P^{\rho\mu\nu\sigma} \)” could be ambiguous, if one assumes that varying a given field gives that field’s own evolution equations.
[sic] not appear in the equations of motion for the remaining components and furthermore can be eliminated from the theory without altering the structure of its equivalence classes. Such a subset is obviously irrelevant to the theory. We shall assume, therefore, that no subset of the components of [that geometrical object] is irrelevant in this sense.” (Anderson, 1967, p. 83)

Likewise TLL exclude the category of irrelevant variables (Thorne et al., 1973, p. 3569). Anderson observes that

one can always construct a hierarchy of theories all of which have the same equivalence-class structure in the sense that the equivalence classes of these theories can be put into one-to-one correspondence with each other. Two theories of such a hierarchy will differ both with regard to the mathematical quantities that describe their respective [kinematically possible trajectories] and their respective covariance groups. However, the set of mathematical quantities that describe the [kinematically possible trajectories] of a given theory in such a hierarchy will contain, as subsets, those of each theory that precedes it in the hierarchy. Likewise, its covariance group will contain, as a subgroup, the covariance group of each preceding theory…

The question then arises as to which theory of a hierarchy one should use to describe a given physical system. The answer rests, of course, in the final analysis, on the measurements that one can make on the system. It is necessary that each quantity used to describe the [kinematically possible trajectories] of a theory must, at least in principle, be measurable. (Anderson, 1967, p. 81)

Similar thoughts appear elsewhere in the text (pp. 306, 340). This requirement of observability, an unfortunate whiff of verificationism, presupposes that all the physics resides in the field equations.11 But typically, fields that do useful work are

---

11This last claim Anderson elsewhere implicitly appears to contradict when he considers boundary conditions (p. 75) and suggests (using “furthermore” on p. 83), surprisingly, that there could exist fields that do not appear in other fields’ equations of motion, but which help to determine the structure of the theory’s equivalence classes. As it happens, recent work on field formulations of Einstein’s equations provides an example: the flat metric does not appear essentially in the field equations, but it plays a role in the boundary conditions, topology, and gauge transformations (Pitts and Schieve, 2004). Boundary conditions are important in string theory as well (Braga et al., 2005). Thus Anderson is overly hasty in eliminating the background metric after deriving Einstein’s equations in flat space-time (Anderson, 1967, pp. 303-306) in
observable, and Anderson’s requirement of observability, if not entirely on target, at least emphasizes the importance of excluding idle fields, such as the Lagrange multiplier $P^{\rho\mu\nu\sigma}$ appears to be. Another important example of unobservable but relevant fields is given by the auxiliary fields in supersymmetric theories (Kaku, 1993) such as supergravity (the most basic version of which couples the spin 2 graviton with the spin $\frac{3}{2}$ gravitino in a unique way (Boulware et al., 1979; Deser, 1980)). Supersymmetry, whether global (constant over space-time) or local (varying from point to point), is a kind of transformation in which bosonic and fermionic fields are transformed into each other. One wishes to implement supersymmetry in a suitably deep (off-shell rather than merely on-shell (van Nieuwenhuizen, 1981)) way, so that kinematically possible trajectories, not just dynamically possible trajectories, are related by supersymmetric gauge transformations.\footnote{I thank Ikaros Bigi for advice along these lines.} The auxiliary fields in supersymmetry are needed to make the action invariant (up to a divergence) under supersymmetric transformations without use of the equations of motion, but the auxiliary fields (usually (van Nieuwenhuizen, 1981)) do not couple to anything else. They also vanish using their equations of motion (“on-shell”) (Kaku, 1993) and so violate a tempting necessary condition for physical significance. Fortunately Anderson’s criterion that irrelevant fields “can be eliminated from the theory without altering the structure of its equivalence classes” shows that the auxiliary fields in supersymmetry are relevant.

While Rosen’s trick vindicates Maidens’s assertion that building nonvariationality into the notion of absolute objects is unsuccessful, Andersonian resources

\footnote{the fashion of Kraichnan (Kraichnan, 1955). While Kraichnan’s use of a background metric in no way requires that quantization occur by covariant perturbation theory (Solov’ev, 1988), historically the two projects have been linked in the minds of many. Anderson critiqued perturbative approaches to Einstein’s equations in a verbal response to a paper by Richard Arnowitt (Arnowitt, 1963).}
might be invoked to exclude Rosen’s trick as a form of cheating. Anderson’s prohibition of irrelevant variables appears to exclude theories making use of Rosen’s trick, because the dynamical evolution of the Lagrange multiplier \( P^{\rho\mu\nu\sigma} \) has no effect on any other fields, whether gravitational or matter. \( P^{\rho\mu\nu\sigma} \) appears to do nothing useful by Anderson’s standards; one might have lower standards, however.

Making \( \eta_{\mu\nu} \) variational and yet absolute could perhaps be useful in that it lets one treat the theory readily using the existing constrained dynamics formalism (e.g., (Sundermeyer, 1982)), which has not made much room for nonvariational fields. Making \( \eta_{\mu\nu} \) variational, as Rosen noted, also allows one to define a conserved symmetric stress energy tensor without using the formal trick of the Rosenfeld approach, in which one replaces the flat metric by a curved one for taking a functional derivative and then restores flatness afterwards (Deser, 1970; Kraichnan, 1955; Pitts and Scheive, 2007; Rosenfeld, 1940). Whether Rosen’s trick or Rosenfeld’s is preferable is open to discussion in other contexts, but an Andersonian elimination of the Lagrange multiplier field as irrelevant would be at least a defensible view, and seems preferable in the context of discussing absolute objects. If the Lagrange multiplier field is rejected as irrelevant, then one cannot vary \( \eta_{\mu\nu} \) (at least not arbitrarily), lest flatness be lost. To avoid loss of dynamical information when a field previously varied in an action principle is rejected as irrelevant, it seems appropriate to impose the now-eliminated field’s Euler-Lagrange equation identically.\(^{13}\)

Excluding the Lagrange multiplier field \( P^{\rho\mu\nu\sigma} \) both eliminates the corresponding term from the action and gives up the Euler-Lagrange equation

\[
\frac{\delta S}{\delta P^{\rho\mu\nu\sigma}} = \sqrt{-\eta} R_{\rho\mu\nu\sigma} [\eta] = 0,
\]

\(^{13}\)There might be cases where the generalized Bianchi identities imply so much redundancy in the Euler-Lagrange equations that it is unnecessary to re-impose the whole of the now-lost Euler-Lagrange equations as identities.
so re-imposing as an identity the lost equation $\sqrt{-\eta} R_{\rho\mu\nu\sigma} [\eta] = 0$ makes the metric $\eta_{\mu\nu}$ flat even off-shell, or in other words flat for all kinematically possible trajectories. Variations of $\eta_{\mu\nu}$ that preserve flatness can be expressed in terms of “clock fields” $X^A$ using $\eta_{\mu\nu} = X^A_{\cdot \mu} \eta_{AB} X^B_{\cdot \nu}$. It makes no difference empirically whether one varies the clock fields, because the same equations will arise either way, either as Euler-Lagrange equations for the clock fields or consequences of the generalized Bianchi identities and the other fields’ Euler-Lagrange equations. If one bans clock fields from the absolute objects project, then it is best not to vary $\eta_{\mu\nu}$ at all. Thus Anderson’s exclusion of irrelevant fields provides a principled way to exclude the Rosen-Sorkin trick.

Where does this dialectic leave us? Maidens proposed and rejected using a redefinition of absolute objects to exclude the Jones-Geroch counterexample to Friedman’s account of absolute objects. Maidens’s missing proof was supplied in advance by Rosen and more recently by Sorkin. But Rosen’s trick seems not to count against Anderson’s version of the intuition that absolute objects are non-variational, because Anderson wisely has criteria for eliminating irrelevant variables. Does it follow that Anderson’s intuition, in the larger context of his project that excludes irrelevant variables, is vindicated? That is, if we accept Anderson’s definitions and proscriptions, should we also accept his intuition that fields are variational if and only if they are dynamical? As it turns out, Anderson’s generalization survives this alleged counterexample but might be threatened by another in which all fields are variational but there is still an absolute object. I have in mind parametrized theories (Arkani-Hamed et al., 2003; Earman, 2003a; Kuchař, 1973, 1981; Norton, 2003; Schmelzer, 2000; Sundermeyer, 1982), in which preferred coordinates are rendered variational clock fields, as was just mentioned.
Perhaps some uses of clock fields could be excluded as irrelevant—not because
the fields themselves are irrelevant, but because their variationality is. On the
other hand, if clock fields are used to satisfy an appropriate notion of causality in
bimetric theories like massive variants of Einstein’s equations (Pitts and Schieve,
2004, 2007; Schmelzer, 2000), then their variationality is relevant. Parametrized
theories require more discussion than is appropriate here, however, so they will
be treated in a later section. The Geroch-Giulini $\sqrt{-g}$ scalar density example
below is another example of a variational yet absolute object. Whether or not one
wishes to address the Geroch-Giulini counterexample by invoking nonvariation-
ality as a necessary condition for absoluteness, this counterexample shows that
nonvariationality and sameness in all models come apart for GTR.

2.6 Eliminating Local Irrelevance Excludes the Geroch-Jones Vector Field from
Dust and Other Fluids

If Maidens’s proposed (and rejected) use of the redefinition to include non-
variationality in absoluteness is set aside for violation of Anderson’s prohibition
of irrelevant variables, then the Jones-Geroch counterexample still remains to be
addressed. Now it turns out that Anderson’s and TLL’s proscription of irrelevant
variables, if it does not quite remove the Jones-Geroch counterexample, at least
inspires a gentle amendment that does the job. This amendment seems especially
appropriate after one notices that TLL replace (Thorne et al., 1973, p. 3566) An-
derson’s notion of geometrical object (Anderson, 1967, pp. 14-16) with Andrzej
Trautman’s notion of a geometric object (Trautman, 1965). Presumably both
notions aim to capture the same intuition.

Given the relative inaccessibility of Trautman’s lectures, it will be worthwhile
to quote his definition of geometric objects in detail:

Let \( X \) be an \( n \)-dimensional differentiable manifold. Since tensors are not sufficient for all purposes in geometry and physics, for example scalar densities are not tensors, to avoid having to expand definitions and theorems whenever we need a new type of entity, it is convenient to define a more general entity, the geometric object, which includes nearly all the entities needed in geometry and physics, so that definitions and theorems can be given in terms of geometric objects so as to hold for all the more specialized cases that we may require.

Let \( p \in X \) be an arbitrary point of \( X \) and let \( \{x^a\}, \{x^a'\} \) be two systems of local coordinates around \( p \). A geometric object field \( y \) is a correspondence

\[
y : (p, \{x^a\}) \rightarrow (y_1, y_2, \cdots, y_N) \in R^N
\]

which associates with every point \( p \in X \) and every system of local coordinates \( \{x^a\} \) around \( p \), a set of \( N \) real numbers, together with a rule which determines \( (y_1', \cdots, y_N') \), given by

\[
y : (p, \{x^a'\}) \rightarrow (y_1', \cdots, y_N') \in R^N
\]

in terms of the \( (y_1, y_2, \cdots, y_N) \) and the values of \( p \) of the functions and their partial derivatives which relate the coordinate systems \( \{x^a\} \) and \( \{x^a'\} \). The \( N \) numbers \( (y_1, \cdots, y_N) \) are called the components of \( y \) at \( p \) with respect to the coordinates \( \{x^a\} \). (Trautman, 1965, pp. 84, 85)

Trautman then notes that spinors are not geometric objects. He also notes that some objects that are not themselves geometric objects are nonetheless parts of geometric objects; for example, the coordinate gradient of a contravariant vector \( V^\mu,\nu \) is not a geometric object, but \( \langle V^\mu, V^\mu,\nu \rangle \) is a geometric object. (Below I will deploy the OP spinor formalism (Ogievetski˘ı and Polubarinov, 1965) and show that spinors are very nearly parts of a nonlinear geometric object.) One can take a geometric object \( \phi \) (of whatever sort, perhaps with many components) and make new geometric objects using the derivatives of \( \phi \) (Nijenhuis, 1952); thus \( \langle \phi, \partial \phi/\partial x^\mu \rangle \) is a geometric object, and this procedure can be iterated to yield
geometric objects made of $\phi$ and its first $n$ derivatives. (These quantities are now called jets (Fatibene and Francaviglia, 2003).) Pace Friedman’s nonstandard usage (Friedman, 1983, p. 359), the class of geometric objects is not exhausted by tensors and connections. Trautman’s definition was fairly typical in its time, though a bit streamlined for physicists’ use. Geometric objects were considered with great thoroughness by Albert Nijenhuis (Nijenhuis, 1952) and by Kucharzewski and Kuczma (Kucharzewski and Kuczma, 1964). A more recent treatment of them using modern differential geometry has been given by Ferraris, Francaviglia, and Reina (Ferraris et al., 1983a).

The reader will notice that Trautman’s geometric objects are defined at every point in the space-time manifold. That fact is of special relevance for the dust example, because it implies that if a dust 4-velocity timelike unit vector field $U^\mu$ is used as a variable in the theory, then a dust 4-velocity timelike unit vector must be defined at every point in every model, even if no dust exists in some neighborhoods in some models. Here one recalls Anderson’s and TLL’s call for the elimination of irrelevant variables; Friedman also recognizes the value of eliminating surplus structure. It is not clear that existing notions of irrelevance apply to the present case. The dust 4-velocity is locally irrelevant, not globally irrelevant, one might say. Perhaps the authors had in mind fields that satisfy equations somewhat like the Klein-Gordon equation as their primary examples, as theoretical physicists often do; for such fields any irrelevance is likely to be global. But now that the question is raised, it does seem clear that wherever there is no dust, there ought not to be a dust 4-velocity timelike unit vector either—at least not if the task at hand is testing theories for absolute objects. In more modern terminology, Trautman considered only global sections, not local sections, of the natural bundles where
geometric objects live. One needs a global section (or a nowhere vanishing global section) of the tangent vector bundle for dust 4-velocities only if dust is present globally, however.

There seem to be three initially plausible alternatives concerning the dust 4-velocity where the dust has holes in some model, if one wants the primitive variables to be defined globally. First, one might retain a timelike 4-velocity vector even in holes in the dust, while expecting the 4-velocity values in the dust holes to be mere gauge fluff. It is noteworthy that at least some perfect fluid variational principles in the physics literature yield timelike unit vector 4-velocities even where there is no fluid (Ray, 1972). Perhaps mathematical convenience commends this option, though I find that Ray’s variational principle can be modified to lack a timelike 4-velocity in holes in the fluid by the use of the mass current density $J^\mu$ instead of the 4-velocity as the primitive variable. A densitized mass current density $\mathcal{J}^\mu$ (equal in value to $\sqrt{-g}J^\mu$) has been used in a fluid variational principle (Bibbona et al., 2007; Brown, 1993). Presumably one could show that the value of a timelike 4-velocity vector is in fact gauge fluff in dust holes by using the Dirac-Bergmann constrained dynamics technology (Sundermeyer, 1982), though one might perhaps run into technical challenges with changes of rank or with the noncanonical Poisson brackets that can appear in fluid mechanics (Morrison, 1998). In any case, the timelike dust 4-velocity in dust holes has no physical meaning, yet leads one to conclude that the theory has an absolute object. Clearly any absolute object whose existence is inferred only by using physically meaningless quantities is spurious. If one allowed physically meaningless entities into a theory while testing for absolute objects, then one could take any theory and construct an empirically equivalent theory with as many absolute objects as one
wants. One could concoct a version of GTR with Newton’s absolute space, for example. To permit such a procedure is just to give up Anderson’s program of analyzing the uniqueness of GTR, because analysis involves trying to get the intuitively known right answer as a consequence of some criteria. Anderson and TLL call for the elimination of irrelevant variables in order to address just this sort of problem. One might call the entities that they reject “globally irrelevant variables” because such entities play no role at any space-time point in any model. The Jones-Geroch example shows, I conclude, that one must also exclude “locally irrelevant variables,” entities that play no role in some neighborhoods in some models. One could consider whether mathematical entities that play no role at some space-time points or sets of measure zero should also be excluded as locally irrelevant, but there might be technical reasons for admitting them.

The two remaining options avoid this spurious absolute object in different ways. One option is to take the mass current density $J^\mu$ (or perhaps $\mathcal{J}^\mu$ or the like) to be the primitive variable and regard $U^\mu$ and the dust density $\rho$ as derived. Then $\rho$ is defined by $\rho = \sqrt{-J^\mu g_{\mu\nu} J^\nu}$. The 4-velocity $U^\mu$ is naturally defined by

$$ U^\mu = \frac{J^\mu}{\sqrt{-J^\nu g_{\nu\alpha} J^\alpha}} $$

or the like, so $U^\mu$ is only meaningful where the density is nonzero. That consequence is plausible on physical grounds and blocks the Jones-Geroch counterexample. The theory is thus formulated using a quadruple $\langle M, D, g, J \rangle$, not Friedman’s quadruple $\langle M, D, g, \rho U \rangle$ or the quintuple $\langle M, D, g, \rho, U \rangle$. In some models $J^\mu$ vanishes at some space-time points in some models of GTR + dust, so $U^\mu$ is undefined in such cases. Neither $J^\mu$ nor $U^\mu$ is a Gerochian nowhere-vanishing timelike vector field for all models. By contrast, the mass current density $J^\mu$, which is equal to
\( \rho U^\mu \) where \( \rho \neq 0 \), automatically vanishes where there is no dust and is continuous at the transition from dust to vacuum. Thus Friedman’s suggestion that it is more “natural” to use the mass current density, once freed from the two infelicities noted above, is seen to be very reasonable. If one is content for the primitive variable to be only a local section of a relevant bundle, then one can take \( U^\mu \) as primitive but vanishing outside the dust (or outside the closure of the support of \( \rho \) or the like).

The other option with globally defined primitive variables is to take \( U^\mu \) to be meaningful but vanishing in those places in certain models where the dust has holes. Although \( U^\mu \) exists everywhere, it vanishes in some places in some models, so not every neighborhood in every model has \( U^\mu \) that is gauge-equivalent to \((1, 0, 0, 0)\). Anderson’s definition of absolute object requires that, for any component \( \phi_\alpha \) of an absolute object in a theory, “[a]ny \( \phi_\alpha \) that satisfies the equations of motion of the theory appears, together with all its transforms under the covariance group, in every equivalence class of [dynamically possible trajectories].” (Anderson, 1967, p. 83) Even if we drop Anderson’s requirement of global equivalence in favor of Hiskes’s and Friedman’s local equivalence, \( U^\mu \) does not count as absolute. In dust-filled regions in a model, the dust 4-velocity \( U^\mu \) is diffeomorphic (at least in a neighborhood) to \((1, 0, 0, 0)\), but in dust holes \( U^\mu \) is diffeomorphic to \((0, 0, 0, 0)\) instead. Thus \( U^\mu \), like \( J^\mu \), is not an absolute object. One might tolerate as harmless the surplus structure embodied in the vanishing \( U^\mu \) vectors, though the mathematical discontinuity of the vector field makes it difficult to defend this option on grounds of mathematical convenience.

If one chooses to restrict one’s attention to models of GTR + dust that do have dust everywhere and always, such gerrymandering is simply changing the
subject to consider a different theory. If one takes a semantic view of theories, then restricting attention to such a set of models is just to discuss some new theory besides GTR + dust, namely GTR + omnipresent dust. Manifestly GTR + omnipresent dust is a proper subset of GTR + dust. GTR + omnipresent dust has the peculiar feature of describing “dust” with such attributes as necessary existence, omnipresence and eternity, attributes more suited to a Deity than to dust. Moreover, GTR + omnipresent dust is not the set of cosmological models of GTR. For example, one can write down cosmological models in which dust is present but not omnipresent (Feynman et al., 1995, p. 166) (Klein, 1971; Smoller and Temple, 2003). More realistic models include eras of radiation domination and perhaps dark energy, so dust is not even a good description of matter in every non-empty region of space-time in cosmological models in GTR. In short, GTR + omnipresent dust has no essential physical relevance to cosmology. Having suitably deflated expectations regarding the theory’s physical import, one can proceed to test it for absolute objects. The new theory GTR + omnipresent dust has an absolute object. But why shouldn’t it? Surely no one has well-founded intuitions to the contrary. Any matter with the attributes of necessary existence, omnipresence and eternity just isn’t much like dust, but rather has the vaguely theological flavor that some friends and foes of absolute objects (such as Newton and Einstein in his Machian aspect, respectively—if the reader will pardon the anachronism of ascribing to Einstein and Newton attitudes towards a concept that they never encountered precisely and explicitly) have sensed.\textsuperscript{14} Anderson

\textsuperscript{14}This remark should be qualified by recalling that Newton’s and Gassendi’s absolute space owes something to the ancient atomists, who were hardly the most theistic of the ancient Greek philosophers, while Aristotle, Descartes, Leibniz and Berkeley count as theistic while preferring relational notions of space. If there is any pattern to be found here concerning the influence of theological sympathies on attitudes towards space, perhaps the question turns on whether a given author thinks of absolute space as somehow a reflection of the divine (and thus attractive for the theist like Newton and repulsive for the non-theist like Mach), or rather as a sort of
anticipated the fact that one could consider a proper subset of models for which some field would count as absolute without counting as absolute for the full set of models. He wrote:

We should perhaps emphasize that we are discussing here universal absolute objects, which must appear in the description of every [dynamically possible trajectory] of our space-time description. It is quite possible that, for a subclass of [dynamically possible trajectories], one or more dynamical objects satisfy the criteria of Section 4-3 and so play the role of absolute objects for those [dynamically possible trajectories].... The existence of such special subclasses of [dynamically possible trajectories] as those discussed above does not, of course, constitute a violation of the principle of general invariance as we have formulated it. Only the existence of universal absolute objects would do so. (Anderson, 1967, pp. 339, 340)

Thus Anderson reminds us that absolute objects are universal, not (so to speak) provincial like the dust 4-velocity. While the dust 4-velocity constitutes an absolute object for the theory GTR + omnipresent dust, it does not constitute an absolute object for GTR + dust due to the failure of universality. Thus Friedman’s intuition, as modified above, is vindicated. The alleged Jones-Geroch counterexample fails to count as an absolute object for GTR + dust and thus fails to undermine Friedman’s analysis after a slight amendment using Andersonian resources.

One might summarize Friedman’s reply, as amended above, as follows: Geroch’s merely mathematical vector field is irrelevant and eliminable because it does no physical work, while Jones’s dust application of the vector field does physical work but fails to be both physically meaningful and everywhere nonvanishing in all models and so violates the diffeomorphic equivalence needed for absoluteness. At

rival to the Judeo-Christian God (and thus perhaps repulsive to Christians like Leibniz and Berkeley and attractive to some who do not believe in the God of the Abrahamic traditions). The Leibniz-Clarke correspondence (Alexander, 1956) might reflect theistic versions of these two attitudes toward the relationship between space and deity. But these comments are speculative.
this stage a summary might be useful. Physics literature previously unappreciated by philosophers of physics has been shown to shed light on the Jones-Geroch counterexample to Friedman’s (and likely Anderson’s or TLL’s) definition of absolute objects. An old result from Rosen vindicates Maidens’s claim that a redefinition of absolute objects to include nonvariationality could not be used to eliminate the Jones-Geroch counterexample without generating a new counterexample. The neglected but valuable paper by TLL and some infrequently noticed parts of Anderson’s book proscribe irrelevant variables, a fact with important consequences. This proscription perhaps can be used to exclude Rosen’s trick for deriving flat space-time from a variational principle. Then Anderson’s generalization that absolute objects are variational and *vice versa* would seem to be rehabilitated, at least provisionally, though the clock fields of parametrized theories pose further questions, as does the Geroch-Giulini $\sqrt{-g}$ scalar density example below. If variationality cannot be invoked to remove the Jones-Geroch counterexample, then some new move is required. Again the Anderson-TLL proscription of irrelevant variables is helpful, in spirit if not in letter. Excluding locally irrelevant values of the field $U^\mu$, which purports to be the 4-velocity field of dust, would imply that $U^\mu$ is undefined wherever the dust vanishes, while the mass current $J^\mu$ vanishes there. Alternatively, $U^\mu$ and $J^\mu$ both vanish there. Either way, GTR + dust fails to have an everywhere nonvanishing timelike vector field that exists in all models. Thus a slight amendment of the Anderson-Friedman tradition using the Andersonian opposition to irrelevant variables eliminates the Jones-Geroch counterexample.
2.7 Torretti’s and Norton’s Constant Curvature Space Counterexamples Do Have Absolute Objects

A second long-standing worry concerning the Anderson-Friedman absolute objects project was suggested by Roberto Torretti (Torretti, 1984). He considered a theory of modified Newtonian kinematics in which each model’s space has constant curvature, but different models have different values of that curvature. Because every model’s space has constant curvature, such a theory surely has something rather like an absolute object in it, Torretti’s intuition suggests. Though contrived, this example is relevantly like the cases of de Sitter or anti-de Sitter background metrics of constant curvature that are sometimes discussed in the physics literature (e.g., (Logunov et al., 1991; Rosen, 1978)), where one often lumps together space-times with different values of constant curvature. The failure of the metrics to be locally diffeomorphically equivalent for distinct curvature values entails that the metric tensor does not satisfy Anderson’s or Friedman’s definition of an absolute object (or TLL’s, for that matter). Thus Torretti concludes that Anderson’s project is not adequate for achieving the goals that Friedman has or ought to have.

Though neither Torretti nor later writers seem to have noticed, Anderson’s analysis, when applied to Torretti’s example, does yield a very specific and reasonable conclusion involving an absolute object. Though the spatial metric is not absolute, the conformal spatial metric density, a symmetric \((0, 2)\) tensor density of weight \(-\frac{2}{3}\) (or its \((2, 0)\) weight \(\frac{2}{3}\) inverse) is an absolute object. This entity, when its components are expressed as a matrix, has unit determinant. It appears routinely in the conformal-traceless decomposition used in finding initial data in numerical studies of GTR. It defines angles and relative lengths of vectors at a
point, but permits no comparison of lengths of vectors at different points. In three dimensions, conformal flatness of a metric is expressed by the vanishing of the Cotton tensor (Aldersley, 1979; Garcia et al., 2004; Hall and Capocci, 1999), not the Weyl tensor, which vanishes identically. That the conformal metric density is an absolute object is shown in the following way. Every space with constant curvature is conformally flat (Misner et al., 1973; Robertson and Noonan, 1968; Wolf, 1967). For conformally flat spatial metrics, manifestly the conformal parts are equal in a neighborhood up to diffeomorphisms. The conformal part just is the conformal metric density, so the conformal metric density is the same (within a diffeomorphism) locally for every model in Torretti’s theory. Perhaps Anderson’s analysis captures as absolute everything that it ought to capture. Concerning Norton’s modification of Torretti’s example to Robertson-Walker metrics (Norton, 1993, p. 848), analogous comments could be made: these space-times are conformally flat (Infeld and Schild, 1945; Kuchowicz, 1973; Padmanabhan, 1993; Tauber, 1967) and so have as an absolute object the space-time conformal metric density. The density weight of the conformal metric density is $-\frac{2}{n}$ in $n$ manifold dimensions.

The conformal metric density is not the only absolute object in the Torretti example. Another absolute object in Torretti’s theory, reflecting the constancy of the curvature (not just conformal flatness), is $\nabla_\mu R$, which vanishes, however. How to regard vanishing absolute objects is a question that has been addressed insufficiently. This particular vanishing absolute object seems less trivial than some that one might think up. As will appear shortly, the metric’s determinant is also absolute in both of these theories and any others in which such an entity appears, a fact with deep consequences for the Anderson-Friedman program.
Unimodular GTR was invented by Einstein and was discussed by Anderson along with David Finkelstein (Anderson and Finkelstein, 1971). While it is rather well known today (Earman, 2003a; Henneaux and Teitelboim, 1989), still a consideration of unimodular GTR helps one to reach the startling conclusion that not only it, but GTR itself, has an absolute object on Friedman’s definition. This fact was pointed out recently by Geroch (Pitts, 2006a) and Giulini (Giulini, 2007).

Unimodular GTR comes in (at least) two flavors: the coordinate-restricted version in which only coordinates that fix the determinant of the metric components matrix to \(-1\), and the weakly generally covariant version that admits any coordinates with the help of a nonvariational scalar density (usually of weight 1 or 2, but any nonzero real weight suffices) and a dynamical conformal metric density, which is a \((0,2)\) tensor density of weight \(-\frac{2}{n}\) or a \((2,0)\) tensor density of weight \(\frac{2}{n}\) in \(n\) space-time dimensions. As Anderson and Finkelstein observe, a metric tensor as a geometric object is reducible into a conformal metric density and a scalar density. They have in mind an equation along these lines:

\[
g_{\mu\nu} = \hat{g}_{\mu\nu} \sqrt{-g}^\frac{2}{n}
\]  

(2.12)

As usual, \(g\) is the determinant of the matrix of components \(g_{\mu\nu}\) of the metric tensor in a coordinate basis; \(g\) is a scalar density of weight 2 and (at least for even \(n\)) takes negative values because of the Lorentzian signature of the metric tensor. \(\hat{g}_{\mu\nu}\) is the conformal metric density. The new variables \(\hat{g}_{\mu\nu}\) (or its inverse) and \(\sqrt{-g}\) (or any nonzero power thereof) are equivalent to those of Anderson and Finkelstein. They further observe that this scalar density is an absolute object in unimodular
GTR. This observation seems unremarkable because that scalar density is not variational. For comparison, one recalls that Asher Peres rewrote the Lagrangian density for GTR in terms of the conformal metric density and a scalar density (Peres, 1963); recently this idea was reinvented by M. O. Katanaev (Katanaev, 2005). Surely the result is still GTR and not some other theory. To my knowledge, no one (prior to Geroch, in effect) has ever considered whether the scalar density, even if varied in an action principle for GTR, might still count as an absolute object.\textsuperscript{15} Once the question is raised about GTR with the Peres-type variables, a positive answer seems obvious: GTR has an absolute object! This absolute object is a scalar density of nonzero weight, because every neighborhood in every model space-time admits coordinates (at least locally) in which the component of the scalar density has a value of $-1$ (Golab, 1974).

Interesting conclusions follow. First, either the famous Anderson-Friedman claim that GTR’s novelty lay in its lack of absolute objects, or that program’s analysis of absolute objects, is flawed. Second, the scalar density is absolute despite being variational. Perhaps some people tacitly have assumed that any field varied in an action principle is dynamical (that is, not absolute), even while officially employing Anderson’s definition of absoluteness. Third, it would be useful to combine hints from Anderson and Finkelstein about the (ir)reducibility of geometric objects with the notion of equivalent geometric objects to accommodate changes of basic variables or, as a field theorist might say, field redefinitions. (An uncommonly geometric treatment of field redefinitions in particle physics was

\textsuperscript{15}As Don Howard points out, it might be useful to look at Einstein’s various attitudes towards setting $\sqrt{-g} = 1$ during the 1910s to see if Einstein might have sensed this sort of problem. The matter is discussed by various authors (Brading and Brown, 2002; Brown, 2005b; Mehra, 1973; Pais, 1982) as well as some of Einstein’s most available work (Einstein, 1923). There is a sort of irony that this condition reappears as privileged in some sense even in GTR after so many years, though on more sophisticated mathematical grounds. I thank Ted Jacobson for bibliographic assistance.
given some time ago by John M. Cornwall, David N. Levin, and George Tiktopoulos (Cornwall et al., 1974a). Two geometric objects are called equivalent if each is a function of the other and the functions do not depend on a coordinate system (Golab, 1974; Kucharzewski and Kuczma, 1964; Nijenhuis, 1952; Zajtz, 1968). Some authors have worked toward a more modern or ‘intrinsic’ formulation, from which standpoint the components of a geometric object field are just coordinates used to describe the fibers of a bundle (Ferraris et al., 1983a,b; Haantjes, 1954; Haantjes and Laman, 1953; Kolár et al., 1993; Kucharzewski and Kuczma, 1963; Kuiper and Yano, 1955; Matteucci, 2003; Nijenhuis, 1960, 1972; Salvioli, 1972; Trautman, 1972; Zajtz, 1988a,b), which Nijenhuis called a “natural bundle.” In that context the equivalent geometric objects of the component approach amount to alternative coordinatizations of the bundle. Once one follows Hiskes and Friedman in seeking only local equivalence between arbitrary neighborhoods for absoluteness, rather than global equivalence, the test for absoluteness becomes a local procedure; thus some of the advantages of the fiber bundle language become less important. However, the use of gauge-natural bundles (Ferraris et al., 2003) is helpful to accommodate the internal gauge transformations which, as I noted above, were included in Anderson’s covariance group but tacitly omitted by Friedman.

Finally, though some philosophers of physics profess to know absolute objects when they see them, even without an analysis, the case of GTR formulated using a conformal metric density and a scalar density suggests otherwise. (The volume element in Torretti’s and Norton’s theories will count as absolute for the same reason.) Evidently no one has spotted the absolute scalar density in GTR simply by inspection until recently. It follows that either one sometimes does not know an
absolute object when one sees it, or that the Andersonian analysis of absolute objects gives the wrong answer for this example. If the latter horn is accepted, then Peres’s version of GTR in terms of a conformal metric density and a scalar density (both varied in the action principle) has no absolute object, whereas unimodular GTR in terms of a conformal metric density and a nonvariational scalar density has an absolute object, although the theories have the same geometric objects and nearly the same field equations (supplemented with Noether identities). Such a claim is perhaps a bit surprising. Do those who claim to spot absolute objects by inspection merely detect nonvariational objects in this instance? Whether a theory has nonvariational objects is, at least in some important examples, merely a question of its formulation, because tricks such as Rosen’s Lagrange multiplier or the parametrization of preferred coordinates into clock fields can be employed to turn nonvariational fields into variational ones.\footnote{One hesitates to generalize too broadly on this matter. In GTR in terms of the conformal metric density (or its inverse) and a scalar density, the latter counts as absolute, so one might be tempted not to vary it in the action principle. But then the field equations are changed: a cosmological constant enters as a constant of integration, as is well known. The reason pertains to the mathematical form of the Lie derivative of a scalar density (Israel, 1979; Schouten, 1954): for weight $w$, $\mathcal{L}_\xi \phi = \xi^\mu \phi_{,\mu} + w \phi \xi_{,\mu}$, and the $w$ term opens the door to the constant of integration. For scalars, it makes no difference whether one varies them as clock fields or not, because the form of the generalized Bianchi identities (Sundermeyer, 1982) and the linear independence of the gradients of the clock fields ensures that the same equations hold either way. (For future reference, I note that it doesn’t matter if one varies the Stueckelberg scalar field in gauged massive electromagnetism, either, because the resulting equation would hold anyway on account of the generalized Bianchi identities and the equation for the electromagnetic vector potential $A_\mu$.)} If the having or lacking of absolute objects is merely a formal feature that depends on a theory’s formulation, then some new way of escaping the Kretschmann objection to the physical vacuity of general covariance (Norton, 2003) is needed. Absent much healthy
competition, the Andersonian project is worthy of attention even if its widely advertised diagnosis of the novelty of GTR is incorrect.

If the novelty of GTR does not consist in its lacking absolute objects (given Anderson’s definition of them), still Anderson’s project remains useful, while it might be reparable for diagnosing the special novelty of GTR, if there is any. There are indeed interesting novel features of GTR that Anderson’s framework uncovers or suggests. For example, GTR apparently was novel in having an external symmetry group involving arbitrary functions of space and time, in having a group as large as the volume-preserving diffeomorphisms, and in having simultaneity described by a dynamical rather than absolute object. (Whether this last feature might be a liability for quantization, where one might want equal-time commutation relations for canonical quantization (van Nieuwenhuizen, 1977), is worth considering.) While the earlier proposal to invoke variational principles was too crude, some more sophisticated effort that bans irrelevant fields might succeed. It is not presently clear whether it is best to admit that GTR has an absolute object or to redefine absolute objects to keep GTR from having any, if possible, but it seems worthwhile to consider the question.

It might also be worth noting that less common choices of variables for Einstein’s equations leave open the possibility of degenerate metrics, for which $\sqrt{-g}$ can vanish in some places, and hence fail to be absolute.\(^1\) The Ashtekar variables allow for a degenerate space-time metric (Jacobson and Romano, 1992), as do the variables used by Deser (Deser, 1970) consisting of an independent connection and the weight 1 inverse metric density $g^{\mu\nu}$. If one takes as fundamental a set of variables along these lines, then the resulting theory lacks an absolute object. In both cases the polynomiality of the action and the use of a connection independent of

\(^{17}\)This remark was inspired by a question from Oliver Pooley.
the metric \textit{a priori} are crucial. At present it seems unwise to appeal to such a difficult and poorly understood phenomenon as a degenerate space-time metric to rescue the Anderson-Friedman project from the Geroch-Giulini counterexample, however.

2.9 Bach-Weyl Gravity Lacks $\sqrt{-g}$

Mathematicians have considered the question of which sorts of geometric objects are concomitants of some set of primitive geometric objects. For the purpose of setting up an action principle that admits arbitrary spacetime coordinates, one wants a concomitant of the fields of interest that is a weight 1 scalar density, perhaps up to a divergence. Is there a theorem on concomitants such that one could not build a formally generally covariant action principle without $\sqrt{-g}$ or some similar absolute quantity? If there isn’t yet such a theorem, might one be proved? In fact there can be no such theorem, because there is a counterexample. This is the Bach-Weyl conformal gravity (Bach, 1921; Dzhunushaliev and Schmidt, 2000; Edery et al., 2006; Elizondo and Yepes, 1994; Fiedler and Schimming, 1980; Kazanas and Mannheim, 1991; Schimming and Schmidt, 1990; Schmidt, 2007). This theory’s field equations can be derived from a Lagrangian density that is (up to a constant factor)

$$\hat{g}^{\rho\sigma} \hat{g}^{\alpha\beta} C_{\rho\sigma\mu} \nu C_{\alpha\beta\mu} \nu,$$  \hspace{1cm} (2.13)

where $C_{\alpha\beta\mu} \nu$ is the Weyl curvature tensor in its primordial form and the absence of $\sqrt{-g}$ has been made manifest by the use of the inverse conformal metric tensor density $\hat{g}^{\mu\nu}$ of weight $\frac{1}{2}$. Much like the Maxwell electromagnetic kinetic term $\hat{g}^{\mu\alpha} \hat{g}^{\nu\beta} F_{\mu\nu} F_{\alpha\beta}$, this Bach Lagrangian density is a weight 1 scalar density (as is required) without $\sqrt{-g}$. Of course one must also consider matter coupling for
a practical theory, but at least some kinds of matter certainly are admissible. The source-free Maxwell equations are manifestly conformally \textit{invariant} because the Lagrangian density can be written with only a space-time conformal metric density rather than the whole metric; see (Fulton et al., 1962) for the situation with sources. A massless scalar field can be coupled conformally covariantly if one wishes (Wald, 1984, pp. 447, 448).

At least for vacuum gravity, the Bach-Weyl theory avoids GTR’s Anderson-Friedman absolute object, so the Anderson-Friedman definition of absolute objects is not so loose that any local field theory, or any theory of gravity admitting sources, must have one. If one thinks that absolute objects are objectionable—perhaps agreeing with Einstein that entities that ‘act without being acted’ upon are “contrary to the mode of thinking in science” (Einstein, 1956, pp. 55, 56) or the like—and if one is satisfied with the Anderson-Friedman analysis (which shows that GTR has an absolute object), then one has a largely \textit{a priori} argument that Bach-Weyl conformal gravity is correct, and that GTR is false. I fail to endorse this argument for two reasons. First, I am not convinced that absolute objects are bad. Second, even if there were reasons to be suspicious of them, the empirical and explanatory virtues of GTR are so compelling that a choice between GTR and avoiding absolute objects ought to be resolved in favor of GTR. On the other hand, the empirical status of the Bach-Weyl theory has been criticized by Elizondo (Elizondo and Yepes, 1994) and Flanagan (Flanagan, 2006). The latter takes the theory to be ruled out empirically by observations of the solar system, but Mannheim claims to have answered this criticism (Mannheim, 2007). While I do not take sides in this ongoing controversy, it appears that foes of absolute objects (as defined by Anderson) need for Mannheim’s side to be right. Philosophers and
physicists who accept GTR, at any rate, need to learn to embrace absolute objects or redefine them so as to keep GTR from having any.

2.10 Implications of Non-Metric or Multi-Metric Theories for Ancient-Medieval Debate Over the Eternity of the World

The possibilities disclosed by contemporary physics call attention to a tacit false premise in the ancient and medieval debates over the eternity of the world. Whereas Aristotle and his followers argued that necessarily the world is eternal, John Philoponus and others perhaps influenced by him (Davidson, 1969; Sorabji, 1987) (including Saadya, Ghazâlî, and Bonaventure) argued that necessarily the world is finitely old. While both streams of argument already failed to convince some authors (such as Maimonides and Aquinas) long ago, and post-Kantian philosophy does not see great interest in this debate, it is noteworthy that 20th century geometry and physics pose a novel difficulty for both positions. Both positions require that there exists with metaphysical necessity a unique temporal metric capable at least of distinguishing a finite past from an infinite one. Clearly the temporal metrics of Newtonian physics and of STR differ, for example, but both give unambiguous answers regarding finite vs. infinite age.

However, the existence of a temporal metric fails for the Bach-Weyl theory, which has only the conformal metric density $\hat{g}_{\mu\nu}$, and thus does not define distances. At best, it defines ages and lengths only up to an arbitrary position-dependent factor. There seems to be no reason that this arbitrary factor should be bounded so as to ensure that even finitude vs. infinitude is factual rather than conventional. Thus the Bach-Weyl theory, which certainly appears to be metaphysically possible and is regarded by a few physicists as a candidate for
the physics of the actual world, lacks any temporal metrical structure sufficient even for distinguishing a finite past from an infinite one. Thus the existence claim presupposed by both sides as a necessary truth, fails.

On the other hand, the uniqueness claim can fail in some physical theories, especially theories with more than one metric. Scalar-tensor theories are perhaps the best known locus for the question “Which metric is the physical metric?” (Kaloper and Olive, 1998; Magnano and Sokolowski, 1994; Santiago and Silbergleit, 2000; Weinstein, 1996). While this question seems not to need an answer for most purposes (such as those involving only the field equations), the question is more urgent in discussing singularities, boundary conditions, positive energy and quantization (Catena et al., 2007; Faraoni and Nadeau, 2007; Kaloper and Olive, 1998; Santiago and Silbergleit, 2000; Sotiriou et al., 2008). Theories with multiple metrics might have different types of matter coupling to gravity in different ways, or one metric might be the best choice in the gravitational sector and another in the matter sector, or both. One metric might yield finite age, but another infinite age, in which case there seems to be no answer to the question of the age of the universe, even if the options at hand are merely ‘finite’ and ‘infinite.’ It turns out that the actual universe probably does not behave in accord with a scalar-tensor theory, given the empirical confirmation of the various principles of equivalence in gravity (Will, 1993), which make it difficult for theories empirically distinguishable from GTR in weak or moderate gravitational fields to be empirically viable. Radical ambiguity regarding the age of the universe might occur even in GTR construed as a field in flat space-time (Pitts and Schieve, 2003). In any case, the arguments at hand presuppose that existence and uniqueness hold necessarily, so their premise is refuted by scalar-tensor and other bimetric or multi-metric theo-
ries and by the Bach-Weyl theory, whether or not any such theory describes the actual world.

A few contemporary philosophers defend the tradition of Philoponus et al. regarding the impossibility of an infinite past and hence the necessity of creation. The persuasiveness of this Kalām cosmological argument for theism (Copan and Craig, 2004; Craig, 1979) is affected adversely by the possible non-existence and possible radical non-uniqueness of the age the universe (Pitts, 2008). This argument takes as premises that whatever begins to exist has a cause and that the universe began to exist. The premise that whatever begins to exist has a cause, if it is an item of knowledge, is plausibly construed as an a priori insight into causation and hence a necessary truth. Then the argument needs a notion of beginning that necessarily is meaningful. As has just appeared, both the existence and the uniqueness of the age of the universe, concerning even the mere distinction between finitude and infinitude, are contingent. By contrast Copan and Craig simply take for granted the existence, and perhaps the uniqueness, of the metric for timelike curves in setting up the Kalām argument (Copan and Craig, 2004, p. 199):

To assess the truth of the premise [that the temporal series of past physical events is not beginningless], it will be helpful to define some terms... In order that all the events comprising the temporal series of past events be of equal duration, we arbitrarily stipulate some event as our standard. . . . By a ‘beginning,’ one means a first standard event. It is therefore not relevant whether the temporal series had a beginning point (a first temporal instant).

If one wishes to recast the debate over the eternity of the world in light of the possibilities disclosed by modern physics, three possibilities come to mind for both sides. First, one might construe “beginning” or replace “finite age” in terms of a
first moment, a topological rather than a metrical notion. Singularities such as the Big Bang are unlikely to help with this argument, because they are generally taken as evidence of breakdown of the theory that predicts them (Pitts, 2008; Wald, 1984) (and references therein) rather than as features of reality. A second replacement for finite vs. infinite age might be inextensibility of the space-time. It is not clear how the arguments for the impossibility of an infinite past could be generalized plausibly to argue that necessarily space-time is past-extensible, however, so this option might not be useful for Philoponus, Copan and Craig. A third possibility is conditionalizing the arguments to apply to theories in which the existence and uniqueness of an adequate temporal metric apply. But then the conclusion might fail for the actual world. These options merit further exploration, though they seem unlikely to yield obviously sound arguments with important conclusions.

2.11 Which Fields to Test for Absoluteness? Equivalence, Irreducibility, Con-comitants, etc.

The theory of geometric objects has some more facets that merit exposition. While the terminology is fairly intuitive, defining terms will not only bring greater clarity, but will also bring to mind mathematical possibilities that otherwise could go unnoticed. A number of further distinctions have been made, but which are rarely needed. Most of what one needs to know about geometric objects (apart from modern ‘coordinate-free’ reformulations) is available in either of two older lengthy treatments (Kucharzewski and Kuczma, 1964; Nijenhuis, 1952). The theory includes various results which are difficult to summarize and sometimes curious, especially in one- or two-dimensional spaces (Aczél and Gołab, 1960). There
are also interesting general concepts, some of which I will now summarize. One curious fact is that, apart from Anderson’s use of the definition of geometric objects (and awareness of the idea of nonlinear geometric objects (Anderson, 1967, p. 15)), previously there has been essentially no interaction between the philosophers and physicists interested in absolute objects and the mathematicians interested in geometric objects, apart from the occasional but important involvement of Geroch.

A first distinction is between linear and nonlinear geometric objects. All the most common geometric objects, including tensor densities of all sorts (scalars, vectors, tensors, and nontrivial densities thereof) and connections in metric, metric-affine, or affine geometries, are linear, meaning that the (finite) rule for changing the components under coordinate transformations contains a linear term (and perhaps a zeroth order term) in the components, but no higher order terms or terms that are not integral powers of the components. The vast majority of the detailed work on geometric objects pertains to linear geometric objects. This situation is due partly due to ease of calculation and partly due to the apparent fact that, with a little effort, one can do everything that one needs with linear geometric objects (DeWitt, 1965). However, one might not be able to do everything economically with only linear geometric objects, because linearity might require extra fields, a fact of considerable importance to philosophers interested in either the Anderson-Friedman absolute objects program or in ontology. Thus I observe that, though evidently the symmetric “square root of the metric” \( r_{\mu \nu} \) (Ogievetskiï and Polubarinov, 1965) (on which much more below) apparently has never been discussed in the context of geometric objects, it is a nonlinear geometric object (apart, perhaps, from coordinate restrictions), because the coordinate transforma-
tion formula for $r_{\mu\nu}$ is nonlinear. (This symmetric square root $r_{\mu\nu}$ of the metric $g_{\mu\nu}$, being equivalent to $g_{\mu\nu}$, is therefore "quasi-linear" (Aczél and Golab, 1960, p. 79).) In fact that transformation law for $r_{\mu\nu}$ is an infinite series, a result which is not at all surprising when one recalls that the group composition law requires the composition of two coordinate transformations to be another coordinate transformation; clearly a nonlinear monomial or polynomial law would, upon iteration, yield ever higher powers, in violation of the group property. The Lie derivative of a linear geometric object is also a geometric object of either the same type or a ‘nicer’ (homogenized) type (Nijenhuis, 1952, p. 93) (Schouten, 1954, pp. 68, 106). For example, the Lie derivative of the connection coefficients $\Gamma^\alpha_{\mu\nu}$ is a tensor (Schouten, 1954, p. 152) (Nijenhuis, 1952, p. 102). The Lie derivative $L_\xi \phi$ of a nonlinear geometric object $\phi$ is not a geometric object (Yano, 1957), but the pair $\langle \phi, L_\xi \phi \rangle$ is a geometric object (Nijenhuis, 1952, p. 93) (Schouten, 1954, p. 107) (Szybiak, 1966; Yano, 1957). It turns out that the analogous result also holds typically for covariant differentiation (Szybiak, 1963), although the metric-spinor case (Ogievetskiï and Polubarinov, 1965) is cleaner due to a conjunction of unusual circumstances, especially the vanishing (Levi-Civita) covariant derivative of the metric.

Among the linear geometric objects, a second distinction is between homogeneous and inhomogeneous geometric objects. Homogeneous geometric objects have no term that is zeroth order in the components in their coordinate transformation law, whereas inhomogeneous geometric objects have such a term. The obvious inhomogeneous geometric objects are connection coefficients $\Gamma^\alpha_{\mu\nu}$ and their contractions $\Gamma^\alpha_{\mu\alpha}$ and $\Gamma^\alpha_{\alpha\mu}$ (which might be unequal due to torsion). A less obvious linear inhomogeneous geometric object is the gravitational potential $g^{\mu\nu} - \eta^{\mu\nu}$
(where $\eta^{\mu\nu}$ is the matrix $\text{diag}(-1, 1, 1, 1)$) (Ogievetskiǐ and Polubarinov, 1965); this object has only first derivatives of coordinates with respect to other coordinates in its transformation law, and is equivalent to the metric. It is sometimes worth pointing out whether a nonlinear geometric object has no zeroth order term in its component transformation law, so I generalize the term “homogeneous” to mean merely the lack of a zeroth order term. It turns out that the square root of the metric $r_{\mu\nu}$ is nonlinear but homogeneous, a result that seems not to follow from its series expansion (due to lack of known convergence of the series in this extreme context), but follows from its equivalence to $g_{\mu\nu}$.

A useful terminology due to Goloab is the “type” $(m, n, r)$ of a geometric object, listing in order the number of components $m$, the dimensionality $n$ of the manifold, and the differential class $r$ indicating how many derivatives of one coordinate system with respect to the other appear in the component transformation law. Most interesting geometric objects are of class $r = 1$, but connection coefficients have an inhomogeneous transformation law with second derivatives of one coordinate system with respect to the other, and so are of class $2$, as are their contractions. Thus the Lie derivative of a connection involves second derivatives of the generating vector field (Israel, 1979; Szybiak, 1966). It turns out that there are no linear inhomogeneous geometric objects with $m = n$ of differential class $r = 2$ inequivalent to contracted connections (Luchter, 1973). Some authors introduce a further number (Nijenhuis, 1972), the dimension of the fiber of the object, which in some cases can fail to equal the number of components (Kucharzewski and Kuczma, 1963). Presumably the conformal metric density $\hat{g}_{\mu\nu}$ has fiber dimension of 9 (or $\frac{n(n+1)}{2} - 1$ more generally), despite having 10 (or $\frac{n(n+1)}{2}$) components, given that $\text{det}(\hat{g}_{\mu\nu}) = \pm 1$. 

86
A geometric object \( \psi \) is called a concomitant of another geometric object \( \phi \) if the components of \( \phi \) determine those of \( \psi \) and do so in the same way in every coordinate system. If the dependence is purely algebraic, then \( \psi \) is an algebraic concomitant of \( \phi \); if there is also dependence on derivatives of \( \phi \), then \( \psi \) is a differential concomitant of \( \phi \). Among algebraic concomitants, it sometimes happens that \( \psi \) is a concomitant (algebraic) of \( \phi \) and \textit{vice versa}. Then one speaks of the “equivalence” (or occasionally “similarity”) of the two geometric objects \( \psi \) and \( \phi \) (Aczél, 1960; Golab, 1950; Kucharzewski and Kuczma, 1963, 1964; Nijenhuis, 1952; Zajtz, 1968). Thus the metric tensor \( g_{\mu\nu} \) is equivalent to infinitely many densitized relatives \( g_{\mu\nu}\sqrt{-g^{\omega w}} \ (w \neq \frac{2}{n}) \), the inverse metric \( g^{\mu\nu} \), and infinitely many densitized inverse metrics \( g^{\mu\nu}/\sqrt{-g^{\omega w}} \ (w \neq \frac{2}{n}) \). The metric \( g_{\mu\nu} \), its symmetric square root \( r_{\mu\nu} \) (linear with respect to the Lorentz group \( O(3,1) \)), almost all of the densitized inverse square roots, \textit{etc.} The conformal metric density \( \hat{g}_{\mu\nu} = g_{\mu\nu}\sqrt{-g^{\omega \omega}} \) and its inverse are also concomitants of the metric, but not \textit{vice versa}. Thus they are not equivalent to the metric. Whereas the metric has the nontrivial algebraic one-component concomitant \( g = \det(g_{\mu\nu}) \), a scalar density, the conformal metric densities have only the trivial scalar concomitant \(-1\). Thus the conformal metric densities are irreducible. (Irreducible geometric objects are sometimes called “primitive” (Zajtz, 1966), but I reserve the term “primitive” for philosophers’ use: primitive fields are those that are not defined in terms of something else. Perhaps one could choose reducible geometric objects such as the metric to be primitive, as has been done commonly, albeit without much premeditation, in the use of the metric \( g_{\mu\nu} \).) The only irreducible geometric objects of differential class \( r = 2 \) are equivalent to the contracted connection \( \Gamma^\mu_{\mu\nu} \) (where one can split off the torsion so that only one contracted connection exists) or the projective connection (Zajtz,
1966), which preserve geodesics but not their affine parametrization.

One lesson from the Geroch-Giulini $\sqrt{-g}$ counterexample and the Torretti
and Norton alleged counterexamples is the need to consider irreducible geometric
objects such as $\sqrt{-g}$ and $\hat{g}^{\mu\nu}$, not (just?) the metric $g_{\mu\nu}$, for example. If one
fails to consider irreducible geometric objects, then one regards as genuine what
are in fact spurious counterexamples to the Anderson-Friedman program, such as
the Torretti and Norton cases, while overlooking a genuine counterexample due
to Geroch and Giulini. It is tempting to conclude that only irreducible geometric
objects should be tested for absoluteness, especially if one makes some tacit
analogy with the idea of a basis of vectors. Such a conclusion, however, would
be wrong. Consider two theories, Nordström’s (second) scalar theory of gravity
with conformally flat metrics which typically are not flat (Beck and Howard, 1996;
Bergmann, 1956; Deser and Halpern, 1970; Einstein and Fokker, 1914; Hawking
and Ellis, 1973; Ni, 1972), and Proca’s massive electromagnetism in the context
of STR (Jackson, 1975). In both theories the conformal metric density $\hat{g}_{\mu\nu}$ is
absolute, because the Weyl tensor vanishes. In both theories $\sqrt{-g}$ is absolute,
not because of some special flatness property of $\sqrt{-g}$, but simply because, as
the Geroch-Giulini counterexample vividly showed, $\sqrt{-g}$, being a nonvanishing
scalar density, has a constant component value in some coordinate systems in
some neighborhood about an arbitrary point. Satisfying the condition of local
sameness in all models is something that nonvanishing scalar densities as such
do: they are susceptible to absoluteness. But in STR there is a further absolute
object, namely, the metric $g_{\mu\nu}$, whereas the metric is not absolute in Nordström’s
theory. If one only tested the irreducible geometric objects for absoluteness, then
one would fail to capture the difference between flat and merely conformally flat

88
metric tensors. Thus checking the irreducible geometrical objects, though needful, does not suffice. The absoluteness of the metric in STR comes not from any flatness property of $\sqrt{-g}$, but from a relation between the nonvanishing scalar density $\sqrt{-g}$ (which is automatically absolute by itself) and the (already assumed to be conformally flat) conformal metric density $\hat{g}_{\mu\nu}$. This relation is expressed in Ricci-flatness $R_{\mu\nu} = 0$, which is 10 equations at each space-time point, though not all are independent given the contracted Bianchi identity $\nabla_{\mu}(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R) = 0$.

In identifying with GTR the theory involving Einstein’s equations and the variables $\sqrt{-g}$ and $\hat{g}_{\mu\nu}$, despite the more usual use of the metric $g_{\mu\nu}$ or its inverse as the primitive variable in GTR, one implicitly rejects the policy of attaching deep significance to a particular choice of primitive variables. Given the now-established need to test both irreducible geometric objects like $\sqrt{-g}$ and reducible geometric objects like $g_{\mu\nu}$, clearly no choice of primitive variables suffices as the only collection of geometric objects to be tested or absoluteness. It appears satisfactory to test for absoluteness all combinations of inequivalent irreducible geometric objects. Thus for vacuum GTR, one would test $\sqrt{-g}$ (or some power thereof), $\hat{g}_{\mu\nu}$ (or its inverse), and, for example, $\langle \sqrt{-g}, \hat{g}_{\mu\nu} \rangle$. This last geometric object is equivalent to $\langle \hat{g}_{\mu\nu}, \sqrt{-g} \rangle$, $\langle \hat{g}^{\mu\nu}, \sqrt{-g} \rangle$, $g_{\mu\nu}$, $g^{\mu\nu} = \sqrt{-g}g^{\mu\nu}$, etc. For the Einstein-Aether theory of Ted Jacobson and collaborators (Eling et al., 2005; Jacobson and Mattingly, 2001), with a nonvanishing unit timelike vector field $U^\mu$, the geometric objects $\sqrt{-g}$ and $U^\mu$ each satisfy local equivalence between neighborhoods for all models, but the geometric object $\langle \sqrt{-g}, U^\mu \rangle$ is not absolute; the scalar density concomitant $\mathcal{L}_U \sqrt{-g} = (U^\mu \sqrt{-g})_{;\mu}$ vanishes in some models but not others.\(^{18}\)

A more serious question pertains to the status of those geometrical objects that by nature satisfy the definition of absoluteness as equality of components, up

\(^{18}\)I thank Ted Jacobson for suggesting this point.
to gauge equivalence, between arbitrary neighborhoods. I say that these geometric objects are susceptible to absoluteness, or more briefly, “susceptible.” Geroch has contributed a great deal to identifying which geometric objects are susceptible to absoluteness. Some time ago he noted that nonvanishing timelike vector fields were susceptible, thus contributing to the Jones-Geroch counterexample (Friedman, 1983). More recently he (Pitts, 2006a), and then Giulini (Giulini, 2007), called attention to the fact that \( \sqrt{-g} \) can be set to 1 in neighborhood about any point—a fact quite well known but not previously applied in this context—to conclude that even GTR has an absolute object. Giulini, knowing that nonvanishing vector fields are susceptible to absoluteness, contemplated formulating theories in terms of covector fields instead. Clearly not all covector fields are diffeomorphically equivalent, because some of them have vanishing curl, but others have nonzero curl. (I observe that curl-free or “closed” nonvanishing covectors, being the gradients of functions locally, that is, “exact,” are locally diffeomorphically equivalent \((1, 0, 0, \ldots)\). One need simply choose \( x^0 \) as the function of which the covector is the gradient. Thus curl-free nonvanishing covectors are susceptible to absoluteness.) This matter is discussed in some detail by Peter Olver (Olver, 1995). However, pace Giulini’s suggested choice of variables, if one can avoid having a set of vector fields that contain absolute objects simply by making a different choice of variables involving only covectors, then one can also avoid having \( \sqrt{-g} \) count as absolute by using \( g_{\mu\nu} \) as the variables and excluding \( \sqrt{-g} \). Such a reliance on the significance of a choice of primitive variables, a strategy of concealment rather than revelation, seems inadvisable.

Given Geroch’s and Giulini’s remarks about vectors and covectors, respectively, one can ask about (co)vector densities. Much as covectors permit a ten-
sorial curl without extra structure, contravariant (or “tangent”) weight 1 vector densities admit a divergence that is a scalar density, without any extra structure:

$$\nabla_\mu \tilde{v}^\mu = \partial_\mu \tilde{v}^\mu. \quad (2.14)$$

Thus weight one contravariant vector densities are not all diffeomorphically equivalent, because their divergence can invariantly vanish or fail to vanish. Geroch has shown that

[t]he class of nonvanishing tangent vector densities of weight $w$ is locally indistinguishable if and only if $w$ is not 1. The class of nonvanishing covector densities of weight $w$ is locally indistinguishable for no $w$. (Robert Geroch, personal correspondence, June-July 2006)

The proofs to come follow outlines provided by Geroch (except for divergence-less weight 1 tangent vector densities). As will appear below, one can say more about the nonvanishing tangent vector densities of weight 1 if one divides them into the collection with nonvanishing divergence and the collection with vanishing divergence.

One shows the distinguishability of covector densities $v_\mu$ of weight $w$ by noting the absence of the symmetric connection used in the covariant derivative $\nabla$ as follows:

$$v_{[\mu} \nabla_{\nu} v_{\alpha]} = v_{[\mu} (\partial_{\nu} v_{\alpha]} - \Gamma_{\nu \alpha] \beta} v_\beta - w v_\alpha \Gamma_{\nu] \beta} = v_{[\mu} \partial_{\nu} v_{\alpha]}. \quad (2.15)$$

This quantity is a concomitant of $v_\mu$ without any need of a connection; the result is a 3-form density of weight $2w$. (While antisymmetrizing on two indices suffices for the weight 0 case, antisymmetrizing again with another power of the covector density is necessary for $w \neq 0$ to get rid of the connection.) Thus no covector type of any density weight (including 0) is susceptible to absoluteness, because
the concomitant above invariantly vanishes for some covectors of a given type, while invariantly failing to vanish for others. This fact shows that not merely the number of components (\( n \) in \( n \)-dimensional spacetime), but also the form of the transformation law plays an important role. For tangent vector densities of weight \( w \neq 1 \), evidently there is no comparable concomitant to be had.

Using the coordinate transformation laws for vector densities and making a succession of coordinate transformations, one can demonstrate the local diffeomorphic equivalence of vector densities of all weights \( w \neq 1 \) and all kinds with respect to changing signs with the Jacobian. The transformation law is (Golab, 1974, p. 84)

\[
v^{\mu'} = v^\nu \frac{\partial x^{\mu'}}{\partial x^\nu} \left| \frac{\partial x}{\partial x'} \right|^w \epsilon. \quad (2.16)
\]

Here \( \epsilon = 1 \) for (Weyl) W-densities of any real weight, \( \epsilon = \text{sign} \left( \frac{\partial x'}{\partial x} \right) \) for G-densities of any real weight, and \( \epsilon = \text{sign} \left( \frac{\partial x'}{\partial x} \right)^w \) for J-densities of any integral weight. (One notices that J-densities of even weight are a special case of W-densities, whereas the J-densities of odd weight are a special case of G-densities. The variety of kinds of densities bears a loose relation to the distinction between polar and axial vectors used in electromagnetic theory (Jackson, 1975) and particle physics, because the setting of the coordinate transformation matrix determinant to \( \pm 1 \) makes some of these distinctions collapse; for scalar densities, particle physicists’ distinction between scalars and pseudoscalars is nearly accommodated. However, without the use of a metric to distinguish changes of time coordinate from changes in spatial coordinates, one cannot distinguish spatial parity inversion from time reversal, for example.) I use \( |x| \) for the absolute value of a number or the determinant of a matrix, while \( ||X|| \) takes the absolute value of a matrix’s determinant. For an arbitrary nonvanishing vector density of any weight \( w \neq 1 \) (and any kind with
respect to Jacobian signs), the goal is to show local diffeomorphic equivalence to \( \delta^\mu_0 = (1,0,0,\ldots) \). First, one can choose coordinates \( x^i \) (and \( x^i' \)) such that \( \tilde{\nu} \cdot \partial x^i = 0 \) for \( i = 1, \ldots, n-1 \), in \( n \)-dimensional spacetime, just as in the more familiar \( w = 0 \) case, by taking the coordinates \( x^\mu \) (and \( x^\mu' \), etc.) to have \( \tilde{\partial}/\partial x^0 \) parallel to \( \tilde{\nu} \). Thus \( v^i = 0 \) (and \( v^i' = 0 \)) without loss of generality, giving a simplified formula

\[
\nu^0' = v_0^0 \frac{\partial x^0'}{\partial x^0} \left\| \frac{\partial x}{\partial x'} \right\|^1_v \epsilon. \tag{2.17}
\]

For any sort of density, flipping the sign of a suitably chosen coordinate (\( x^0 \) or \( x^1 \), say) achieves \( \nu^0 > 0 \) (and \( \nu^0' > 0 \)). Now the desired \( x^0' \) is such that \( \partial x^0'/\partial x^0 > 0 \). For \( w \neq 1 \) one sets \( x^0' = x^i \), so the result is

\[
0 < \frac{\nu^0'}{\nu^0} = \left( \frac{\partial x^0'}{\partial x^0} \right)^{1-w}. \tag{2.18}
\]

For \( w \neq 1 \), one can easily take the positive \((1-w)\)th root and integrate with respect to \( x^0 \) for \( x^0' \). Setting \( \nu^0' = 1 \) yields the desired time coordinate (up to choice of origin on each curve of constant \( x^i \)), so indeed every nonvanishing vector density of weight \( w \neq 1 \) is locally diffeomorphically equivalent to \( \delta^\mu_0 \). Thus all tangent vector densities with \( w \neq 1 \) are susceptible to absoluteness, just as tangent vectors are. One could therefore imagine Gerochian vector density counterexamples, perhaps.

For divergenceless tangent vector densities of weight 1, I note that one can say more. The freedom to change ‘time’ coordinates \( x^0 \) serves no further purpose. Using \( \partial_\mu \tilde{v}^\mu = 0 \) and \( \tilde{v}^i = 0 \) (for ‘spatial’ coordinates \( x^i \), \( i = 1 \ldots n-1 \)), it follows that the nonvanishing component depends only on the spatial coordinates. Then one can use the ‘spatial’ coordinate freedom remaining in the reduced transformation.
to good effect. This equation is just the transformation law for a weight 1 scalar density on an \( n - 1 \)-dimensional manifold, so the susceptibility to absoluteness of scalar densities completes the proof of susceptibility for nonvanishing divergenceless tangent vector densities of weight 1. Thus these quantities also can always be expressed as \((1, 0, \ldots, 0)\).

Do the nonvanishing scalar densities \((w \neq 0)\) and nonvanishing contravariant vector densities (including \(w = 0\) but not \(w = 1\)), and divergenceless nonvanishing contravariant vector densities of weight 1 exhaust the collection of susceptible geometric objects? To answer that question might require knowing all the possible geometric objects with no more than \(n\) components in an \(n\)-dimensional space-time. In fact there are known some rather unfamiliar geometric objects including, for example, two vector-like entities discovered by Jakubowicz (Golab, 1974, pp. 98, 99), one with a transformation law that makes sign changes for alternating components, another with a transformation law involving the matrix minor of the usual transformation matrix for vectors. Fortunately at least the former sort of object is equivalent to a vector (Golab, 1974; Kucharzewski and Zajt, 1966). Another susceptible geometric object is the contracted Christoffel symbols \(\Gamma^{\nu}_{\mu
u}\), which are the contracted connection coefficients for a torsion-free metric-compatible connection. Of course one can generate further more geometric objects by adjoining the partial derivatives of susceptible geometric objects already in hand, thereby obtaining jets. In 2 dimensions, all metrics are conformally flat (Wald, 1984), so the conformal metric density \(\hat{g}_{\mu\nu}\), with only two independent components, is susceptible to absoluteness. Given some susceptible geometric object, one might
get another one using one of the totally antisymmetric $\epsilon$ tensor densities. Thus the dual of a contravariant weight 1 vector density is a 3-form, which admits an invariant curl.

It turns out that nearly all the above results (apparently excepting that concerning the divergenceless tangent vector density) and some others were in fact obtained some time ago by Andrzej Zajtz (Zajtz, 1988b), who has solved the problem of susceptible geometric objects almost entirely. Not surprisingly, only geometric objects with no more independent components $m$ than there are manifold dimensions $n$ are susceptible to absoluteness; if there were more independent components $m$ than spacetime dimensions $n$, then were would be too few arbitrary functions worth of coordinate freedom for any chance at susceptibility to absoluteness. In Zajtz’s terminology, which makes greater contact with modern-style differential geometry, certain geometric objects are “canonically integrable” and have a “homogeneous fibre”: these have constant components locally for some coordinate system. Others are locally diffeomorphically equivalent, but not consistent, having canonical forms that depend on one or more coordinates. Above I noted, pace Friedman, that this condition is sufficient but not necessary for absoluteness. Zajtz gives canonical forms for a number of geometric objects with no more components than there are manifold dimensions ($m \leq n$). Whereas (contravariant) vectors and vector densities (for $w \neq 1$ or restricted to the class with vanishing divergence for $w = 1$) have constant components in some coordinate system, Zajtz’s canonical form (locally, in a neighborhood of the coordinate origin) for weight 1 tangent vector densities with nonzero divergence, for example, depends in a simple way on one coordinate:

$$\tilde{v}^\mu = (1 + x^0, 0, \ldots, 0).$$
Their having this canonical form (or indeed any canonical form) in a neighborhood shows that the contravariant weight 1 vector densities with nonzero divergence are also susceptible to absoluteness. This result has the happy consequence, for my purposes, that the absoluteness of the nonvanishing scalar density $\partial_\mu \tilde{v}^\mu$ follows from the absoluteness of the (nonvanishing) weight 1 tangent vector densities with nonzero divergence, rather than being a novelty. If there were an example of an absolute differential concomitant from a non-absolute geometric object, then it would seem necessary to investigate differential concomitants for absoluteness quite generally.

If one wants to attribute philosophical significance to the presence or absence of absolute objects, as proponents of the Anderson-Friedman program do, then there is a difficult question to face, namely, what is the significance of the absoluteness of susceptible objects? If a nonvanishing vector field cannot avoid satisfying the absoluteness condition, can it be blamed for absoluteness? Perhaps not; perhaps susceptible absolute objects should be exempted from testing for absoluteness. Or should its absoluteness count even more, as the inevitable and appropriate expression of its character? Presently I have no answer to such a question. (Alas, the amusing parallel to questions in free will and determinism appears to shed no light on the answer.) One should, however, at least call attention to the susceptibility of any susceptible objects in one’s theory. Susceptibility is relevant to the status of the nonvanishing timelike vector field in the Einstein-Aether theory of Ted Jacobson and collaborators (Eling et al., 2005; Jacobson and Mattingly, 2001). They explicitly claim that the theory is generally covariant, without saying what general covariance is in much detail. Analogously, Brendan Foster claims that the aether’s being dynamical makes the theory preserve diffeomorphism in-
variance (Foster, 2006); presumably this is supposed to be a claim of substantive general covariance. However, clearly this theory has a Gerochian vector field that is both absolute and susceptible, so the theory is not generally covariant in the Anderson-Friedman sense. (Neither is it generally covariant in Wald’s admittedly vague sense that no preferred vector fields “pertaining only to the structure of space . . . appear in any law of physics” (Wald, 1984, p. 57).) If one doubts that susceptible geometric objects’ absoluteness should really count as absoluteness, then perhaps one need not worry about the general covariance of the Einstein-Aether theory; the same move would resolve the Geroch-Giulini counterexample for GTR. On the other hand, one might accept the Anderson-Friedman analysis and deny that the Einstein-Aether theory is generally covariant in some deep sense. If so, then no escape has yet been found for the general covariance of GTR, analyzed in the sense of lacking Anderson-Friedman absolute objects, either.

2.12 Absoluteness and Nonvariationality?

It has often been accepted that absolute objects are nonvariational and *vice versa* (Anderson, 1967; Thorne et al., 1973). At least for ‘reasonable’ field theories that claim to describe the whole of the physical world, not just a part of it, any geometric objects that are not varied will be locally the same throughout space-time (thus having a suitable number of generalized Killing vectors) and the same from model to model (up to gauge equivalence in both cases), the latter sameness being Anderson’s criterion of absoluteness. Thus most likely all nonvariational fields are absolute. The converse requires more careful consideration.

Will absolute objects tend to be nonvariational? I assume that one has already excluded the Rosen-Sorkin trick, so one threat to be absolute and variational is
excluded. Are there comparable theories that merit discussion? The susceptible geometric objects are good candidates to be absolute and yet variational. Perhaps the most important example of a variational and yet absolute (because susceptible) geometric object is $\sqrt{-g}$ in GTR. Exact covectors are another sort of susceptible geometric object; these appear generically in theories with clock fields in the differentiated form $X^A,_{\mu}$. Apart from susceptible fields, likely all absolute objects are nonvariational. No counterexamples come to mind, with the exceptions of the Rosen-Sorkin trick, where the Lagrange multiplier field can bear the burden of satisfying the equation $\frac{\delta S}{\delta \eta^{\mu\nu}} = 0$. Especially for fields with more than $n$ components in $n$ dimensions, introducing the variation of some previously nonvariational absolute object would often be disastrous, bringing in extra field equations and hence leading to a dearth of solutions. For few-component fields, less is at stake in varying a field or not, because the generalized Bianchi identities will tend to make some or all of the potential field equations hold anyway (Bradling and Brown, 2002). For (scalar) clock fields, the new Euler-Lagrange equations are the same as consequences of the old Bianchi identities and field equations, as noted above. The new Euler-Lagrange equation for a scalar density ($w \neq 0$) is only slightly stronger than the consequences of the old Bianchi identities, at least if the scalar density is nowhere vanishing or if $w = 1$. But higher-component fields bring increasing complexity.

Given an action functional $S$ of some fields, all of which are relevant by Anderson’s standards, how does one decide which of them to vary in a principle of least action? Does one make such choices merely on the basis of what gives all and only the desired equations? Anna Maidens holds that this choice is a consequence of physical intuition, and thus does not serve as a source of physical insight regard-
ing which quantities should be regarded as absolute (Maidens, 1998). Although the counterexamples that motivated her doubts have been adequately handled (Pitts, 2006a), doubtless there is some truth in Maidens’ position. However, the new Geroch-Giulini counterexample in GTR shows that our physical intuitions about which structures are physically necessary (and thus not to be varied) can come into tension with the goal of deriving all the field equations from an action principle—at least once our intuitions are duly revised to take into account the sameness-in-all-models of $\sqrt{-g}$, which therefore counts as physically necessary rather than contingent. If “what works” to give the desired equations can disagree with our intuitions, then variational principles cease simply to reflect our intuitions and so might serve as a source of physical insight after all.

It is in fact doubly noteworthy that GTR requires varying the absolute field $\sqrt{-g}$ to get Einstein’s equations. On the one hand, the absoluteness of $\sqrt{-g}$ makes its variationality striking, perhaps puzzling. Perhaps susceptible geometric objects, or scalar densities more specifically, violate the absoluteness-nonvariationality connection for good reasons that remain to be specified. On the other hand, it seems curious to vary an entire field to get $\infty^4$ equations, when the contracted Bianchi identities $\nabla_\mu G^{\mu\nu} = 0$ show that only one equation is actually needed from varying $\sqrt{-g}$ (Henneaux and Teitelboim, 1989). If the reader will pardon the mathematical expression, $\frac{\infty^4}{\infty^4}$ of the variation of $\sqrt{-g}$ is redundant. It appears that if only nonvariational fields count as absolute, then GTR escapes having an absolute object in $\sqrt{-g}$, but perhaps only on a technicality.

The following principle seems worthy of consideration (not to say, endorsement): vary only as many fields (or parts thereof, perhaps nonlocally defined) as it takes to get the full set of desired equations, making maximal use of the general-
ized Bianchi identities. For Maxwell’s electromagnetism, Yang-Mills theories, and Einstein’s GTR, the gauge freedom implies that the Euler-Lagrange equations are not independent. Instead, they are algebraically or (in these cases) differentially connected. For Maxwell’s electromagnetism formulated in terms of the vector potential, in the absence of sources the Lagrangian density is \( \mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \), where \( F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \) is the electromagnetic field strength. The generalized Bianchi identity is the vanishing of the divergence of the Euler-Lagrange derivative \( \partial_\mu \frac{\delta \mathcal{S}}{\delta A_\mu} = \partial_\mu \partial_\nu F^{\mu\nu} = \partial_\mu (\partial^\lambda A_\lambda - \partial_\lambda A^\lambda) \equiv 0 \) as an identity following from \( 1 - 1 = 0 \). If the divergence ‘piece’ of the field equations holds as identity, perhaps one should vary only that portion of \( A_\mu \) that leads to the ‘piece’ of the field equations that are not identities? One can try to use the generalized Bianchi identities to identify the difference-making (as opposed to pure gauge) part of the variation of fields in question; the difference-making variations are those that lead to field equations that otherwise would fail to hold, whereas the gauge variations are those that lead to identities that, being identities, would have held anyway.

Field theorists’ techniques such as Fourier transforms and the inversion of differential operators via Green’s functions (Deser et al., 1977; Jackson, 1975; Kaku, 1993; Sudarshan and Mukunda, 1974) seem to be the most plausible candidates for such a project. One should watch out, however, for complications, such as the Feynman-Faddeev-Popov ghosts in Yang-Mills theory, that arise in the quantum path integral but not in the classical variational principle (Kaku, 1993).\(^{19}\)

Turning to GTR, finding the true degrees of freedom explicitly is a good deal more difficult. While there is gauge fluff in \( \hat{g}_{\mu\nu} \), removing it might be difficult. Perhaps, however, one can hope to address \( \sqrt{-g} \) so that some physical core of it is varied while an external husk of gauge freedom is not varied. Because only one

\(^{19}\)I thank Ikaros Bigi for this point.
equation (not $\infty^4$) is needed, taking the average value of $\sqrt{-g}$ might seem helpful. For a scalar field $\phi$, one expects the average over any region $V$ (or all space-time, if convergence poses no problem) to be given by

$$\bar{\phi} = \frac{\int_V d^4 x \phi(x) \sqrt{-g}(x)}{\int_V d^4 y \sqrt{-g}(y)}.$$  

(2.20)

For a weight $w$ scalar density $\tilde{\phi}(x)$, the natural invariant generalization would seem to be

$$\frac{\int_V d^4 x \tilde{\phi}(x) \sqrt{-g}^{1-w}(x)}{\int_V d^4 y \sqrt{-g}(y)}$$

(2.21)

(a scalar number), or perhaps

$$\sqrt{-g}^w(x) \frac{\int_V d^4 y \tilde{\phi}(y) \sqrt{-g}^{1-w}(y)}{\int_V d^4 z \sqrt{-g}(z)},$$

(2.22)

a field of weight $w$. But for $\tilde{\phi} = \sqrt{-g}^w$, the first average is simply 1 (unless some other field, for which the motivation is not evident, is introduced), while the second average is simply $\sqrt{-g}^w$ itself, neither of which makes progress. While one occasionally sees the four-volume of space-time used as a dynamical variable, that quantity diverges in many interesting contexts. However, it turns out that there has been some work showing how to average geometric objects, such that geometrically meaningful answers might be available more often than one would have expected (Coley and Pelavas, 2006, 2007; Edelen, 1964; Mars and Zalaletdinov, 1997; Paranjape, 2008). Might these works permit defining a nontrivial average of a scalar density that is also a scalar density of the same sort, thereby permitting a splitting of $\sqrt{-g}$ into an average piece that needs variation along with $\hat{g}_{\mu\nu}$ to complete the Einstein equations, and a piece that contributes only field equations that are already entailed by those of $\hat{g}_{\mu\nu}$ via the Bianchi identities? It is
noteworthy that many of these works make important use of volume-preserving coordinates $\sqrt{-g} = 1$. Also it is difficult to imagine what the average value of $\sqrt{-g}$ over all space-time could be (if it exists) except 1 or perhaps $\sqrt{-g}$, so the hope of finding an average portion to vary seems slight. On the other hand, it is not difficult formally to admit variations characterized by a single number (not a function of space or even time) $\delta \sqrt{-g(x)} = \lambda \sqrt{-g(x)}$, where $\lambda$ is real. Whether it seems reasonable to vary a single number pertaining to all space-time, with no dynamics of the ordinary sense, is unclear. The idea of splitting $\sqrt{-g}$ (or its spatial analog) into a piece giving unit volume over a finite region and an overall constant factor has been used profitably (Anderson et al., 2003). While there is no compulsion to regard neighborhoods rather than the (perhaps infinite) whole of space-time as relevant for a variational criterion of absoluteness (in contrast to the nearly obligatory Friedman-Hiskes amendment of Anderson’s criterion), using variational principles on neighborhoods with finite volume does license factoring $\sqrt{-g(x)} = V \mu(x)$, where $V$ is the volume of the neighborhood $U$ and $\int_U d^4 x \sqrt{-g} = V$. It is straightforward to show that all of Einstein’s equations are obtained by varying $\hat{g}_{\mu\nu}$ and $V$, while varying $\mu(x)$ contributes nothing further. Thus for finite regions one can find a partial analog of the Geroch-Giulini counterexample: the variational criterion gives an indeterminate (though not determinately wrong) answer regarding whether $\mu(x)$ is a variational field.

It is well known that complete gauge reduction down to the “true degrees of freedom” (the reduced phase space) can frequently produce nonlocal formulations (Henneaux and Teitelboim, 1992), as it does in the Maxwell, Yang-Mills and Einstein theories. One might understand gauge freedom as the result of requiring local rather than nonlocal theories (Guay, 2008) (though the massive variants of these
theories are local in terms of the true degrees of freedom (Sundermeyer, 1982)). It might be worth exploring whether splitting $\sqrt{-g}$ into some part (perhaps just a constant factor) with non-redundant Euler-Lagrange equations and a background part would be of some use in path integral treatments of quantum gravity. There is an old result to the effect that the conformal factor $\sqrt{-g}$ makes the ‘Euclidean’ (positive-definite) action ill-behaved in the sense of not being bounded below, unlike scalar and Yang-Mills fields (Gibbons et al., 1978). Might recognizing some fixed background in $\sqrt{-g}$ be a solution, or a simpler solution than usual, to this problem? To answer the question would of course require calculation, which I have not done. However, it is noteworthy that this conceptual reflection at least suggests some calculations to do.

2.13 Tetrad-Spinor Counterexample and Resolution via Nonlinear Geometric Objects and the Ogievetsky-Polubarinov Spinor Formalism

It is often asserted that coupling spinor fields, such as represent electrons, to a curved metric tensor, such as represents gravitation in GTR, requires an orthonormal basis of covector fields $f^A_\mu$ (or, equivalently, vector fields $e_\mu^A$) (Bardeen and Zumino, 1994; Deser and Isham, 1976; Fatibene and Francaviglia, 2003; Kaku, 1993; Lawson and Michelsohn, 1989; van Nieuwenhuizen, 1981; Weinberg, 1972, 2000). The basis $e_\mu^A$ and cobasis $f^A_\mu$, considered as matrices of components, are inverses, so it follows that $e_\mu^A f^A_\nu = \delta^\mu_\nu$ and $e_\mu^A f^B_\mu = \delta^B_A$. In four dimensions one calls such a basis a vierbein or tetrad. It is convenient to build a metric out of the covectors $f^A_\mu$ using the definition $g_{\mu\nu} = \text{def} f^A_\mu \eta_{AB} f^B_\nu$, where $\eta_{AB}$ is a signature matrix with 0 off the diagonal and $\pm 1$ on the diagonal; it follows that $e_\mu^C g_{\mu\nu} e^D_\nu = \eta_{CD}$, so the basis is orthonormal (as is the cobasis). From now on $\eta_{AB} = \text{diag}(-1, 1, 1, 1)$
will be employed. The orthonormal tetrad of vector fields typically used to couple
spinor fields, such as represent electrons and the like, to a curved metric tensor,
such as represents gravitation in GTR, introduces a Gerochian vector field (or sev-
eral of them) into the GTR + spinor theory (Pitts, 2006a). Thus while GTR has
been regarded (until the discovery of the Geroch-Giulini $\sqrt{-g}$ counterexample) as
lacking absolute objects, the typical form of GTR + spinor theory counts, oddly
enough, as having an absolute object (Pitts, 2006a). Given both local Lorentz
and coordinate freedom, one can certainly bring the timelike leg into the compo-
nent form $(1, 0, 0, 0)$ at least in a neighborhood about any point. Thus aligning
the tetrad with the simultaneity hypersurfaces is known as imposing the “time
gauge” on the tetrad (Deser and Isham, 1976; Dirac, 1962; Schwinger, 1963). The
tetrad-spinor example seems rather more serious a problem for definitions of ab-
solute objects than the Jones-Geroch cosmological dust example was even before
its resolution in terms of dust holes, because the spinor field is surely closer to
being a fundamental field than is dust or any other perfect fluid. Spinors (actu-
ally vector-spinors for spin $\frac{3}{2}$) are also required in supersymmetric theories such as
supergravity (van Nieuwenhuizen, 1981) and superstring theory, in which internal
and external symmetries are combined. Prior to the appearance of the Geroch-
Giulini $\sqrt{-g}$ counterexample, this tetrad counterexample might have been the
most serious problem facing the Anderson-Friedman project. Given the perhaps
insoluble Geroch-Giulini counterexample for vacuum GTR, finding a formulation
of GTR + spinor lacking absolute objects might not make the difference between
success and failure for the Anderson-Friedman project. Still it is useful to con-
sider whether coupling spinors to gravity actually provides a counterexample of the
sort that dust supposedly posed with a non-vanishing vector field. The ontology
of spinor fields in curved space-time is also at stake.

Before discussing the tetrad-spinor issue, it is worthwhile to consider Anderson’s treatment of spinors satisfying the Dirac equation in a gravitational field (Anderson, 1967, pp. 358-360). Anderson entertains the worry that the Dirac matrices $\gamma^\mu$ satisfying

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}I \tag{2.23}$$

might form an absolute object in flat spacetime, in fact one with a symmetry group smaller than the Poincaré group (though in this context $\gamma^\mu$ is not a vector under arbitrary coordinate transformations—nor even arbitrary transformations near the identity—so it is not eligible to be an absolute object by Anderson’s standards, it would seem). Turning to curved spacetime, Anderson avoids using an orthonormal tetrad by using variable metric-dependent Dirac matrices $\gamma^\mu$ satisfying $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}I$. What follows is a formalism with an (apparently global) symmetry group unrelated to the group of spacetime mappings. However, the implicit relationship between $\gamma^\mu$ and $g^{\mu\nu}$ leaves obscure the coordinate transformation properties of $\gamma^\mu$ and unclear what a suitable action principle might be for deriving the Einstein-Dirac equations—a problematic situation given Anderson’s tentative claim (Anderson, 1967, pp. 88, 89) that absoluteness and nonvariation-ality are coextensive—and what variables it would involve. Thus one can hardly even test Anderson’s formalism for absolute objects; his treatment of spinors is just incomplete. To my knowledge no one else has discussed spinor fields in the context of the absolute objects project until recently (Pitts, 2006a). The tetrad-spinor formalism avoids such difficulties as Anderson’s treatment faces.

The orthonormal tetrad $e^\mu_A$ has two sorts of indices, a world index (here a small Greek letter) that relates to coordinate bases and tensor calculus, and a Lorentz
index (here a large Roman letter); the Lorentz index picks out a particular vector field from a list of four (in four space-time dimensions), the Greek index a particular component. Four vector fields have among them 16 components, rather more than the 10 components of the metric, so there is some redundancy that leaves a new local Lorentz gauge freedom to make arbitrary position-dependent boosts and rotations of the tetrad acting on the Lorentz (Latin) index. This freedom is described by a field of matrices in the group $O(1, 3)$. Given the resulting lack of physical significance of 6 (or $\frac{n(n-1)}{2}$) components of the 16 (or $n^2$) of the tetrad, one can impose a gauge condition to fix some or all of the local Lorentz freedom. When maintaining global Lorentz invariance is desirable in particle physics, it is not uncommon to require the matrix of components $f_{\mu A}$ to be symmetric (with the Lorentz index moved with $\eta_{AB} = \text{diag}(-1, 1, 1, 1)$) (Aladrovandi et al., 1994; Alvarez-Gaumé and Witten, 1984; Bilyalov, 2002; Borisov and Ogievetskii, 1974; Boulware et al., 1979; Cho and Freund, 1975; Choi et al., 1993; Deser and van Nieuwenhuizen, 1974; DeWitt and DeWitt, 1952; Fujikawa et al., 1985; Hamamoto, 1978, 1979; Huggins, 1962; Isham et al., 1971, 1973; Ivanov and Niederle, 1982; Kirsch, 2005; Leclerc, 2006; López-Pinto et al., 1995; Nibbelink et al., 2007; Obukhov and Pereira, 2003; Ogievetskiǐ and Polubarinov, 1965; Passarino, 1984; Tiemblo and Tresguerres, 2004; Tresguerres and Mielke, 2000). A tetrad in a symmetric gauge then counts in some sense as a symmetric square root of the metric. However, most authors see this choice is simply one possible gauge choice for an orthonormal tetrad, not as a means of replacing a tetrad with a fundamentally different entity. Thus the claim that an orthonormal tetrad is required is able to coexistence in the literature with repeated use of techniques that, when fully appreciated, demonstrate the contrary. One finds authors who
both employ the symmetric gauge and assert that a tetrad is required (Aldrovandi et al., 1994; Boulware et al., 1979; Deser and Isham, 1976; Deser and van Nieuwenhuizen, 1974; DeWitt, 1965; DeWitt and DeWitt, 1952; Isham et al., 1971), and so are (assuming no change of views over time) at least implicitly committed to the claim that the symmetric square root of the metric cannot stand on its own in a spinor formalism without a tetrad. However, after thorough searching, I have found no published defenses of the need for an orthonormal tetrad as opposed to the symmetric square root of the metric. It appears that lack of imagination is the main culprit. The symmetric square root of the metric is also not discussed in S. Kichenassamy’s historical survey of Dirac equations in curved space-time (Kichenassamy, 1992). Ne’eman and Sijacki discuss the persistence of erroneous non-existence claims for spinors under the double covering of the general linear group (Ne’eman and Šijači, 1997; Sijacki, 1998). A few authors have explicitly regarded the symmetric square root of the metric \( r_{\mu \nu} \) (not a tensor, despite the notation) as an alternative to a tetrad (Bilyalov, 2002; Borisov and Ogievetskiï, 1974; Ogievetskiï and Polubarinov, 1965). It appears that much of the dispute here is verbal, depending on what counts as coupling directly to the metric (whether the concomitant \( r_{\mu \nu} \) is ‘direct’ enough and whether the author even knows about it or has entertained the idea that it might be a good object in its own right), what counts as a tetrad (whether \( r_{\mu \nu} \), being symmetric and numerically equal to a gauge-fixed tetrad almost everywhere on the manifold, is ‘just’ a tetrad), what counts as a representation (vs. representation up to a sign, as one often sees with spinors), etc. I expect the apparent contradictions to disappear once the ambiguous phrases are disambiguated. Questions about the existence and uniqueness of the symmetric square root in very non-Cartesian coordinates might also play a
role. Regarding that question, some applied mathematics literature on generalized polar decompositions might be relevant (Au-Yeung et al., 2004; Bolshakov and Reichstein, 1995; Bolshakov et al., 1997a, 1996, 1997b,c; Higham, 2003; Lins et al., 2001; van der Mee, 1993; van der Mee and Hovenier, 1992; van der Mee et al., 1999, 2000, 2002).

The supposed need for an orthonormal basis for spinors on curved manifolds appears to create problems for coupling GTR (among other theories) to spinor fields in certain nontrivial topologies, because there are sometimes topological obstructions that prohibit the existence of an orthonormal tetrad on a manifold with a metric. An orthonormal tetrad exists on a four-dimensional manifold if and only if the second Stiefel-Whitney class vanishes (DeWitt et al., 1979). The existence of an orthonormal tetrad in a space-time is closely tied to the existence of a “spinor structure” (Geroch, 1968, 1970). If one wants spinors on such a manifold, then one must hope that the requirement for an orthonormal basis is too strict. Fortunately it has been shown how to have spinors, under a somewhat different but reasonable construal of the term, even without “spinor structure” (Avis and Isham, 1980; Clarke, 1974). The connection between these works and the OP formalism with the square root of the metric is not clear, but they have in common making do with less mathematical structure than has been thought necessary. It would be quite desirable to see a discussion of existence and uniqueness of spinors using the OP formalism.\footnote{I thank Aron Wall for a discussion of this point.} However, every space-time with a Cauchy surface admits an orthonormal tetrad (Geroch, 1970).

While some of the authors cited above have discussed the symmetric square root of the metric $r_{\mu\nu}$ as a nonlinear representation of general coordinate transformations (at least near the identity), and a few have regarded $r_{\mu\nu}$ as conceptually...
independent of a tetrad, the relationship of $r_{\mu\nu}$ to the concept of geometric objects, as discussed by Nijenhuis and others, seems never to have been discussed explicitly. That omission is largely due to the fact that the people who work on geometric objects are mathematicians and general relativists, while the people who use nonlinear group representations are particle physicists; these communities perhaps come closest to overlap in work on supergravity and on gauge theories of gravity. The work of Ranat F. Bilyalov (Bilyalov, 1992, 1998a,b, 2002; Bilyalov and Nikitin, 1998, 2000) is relevant here. Bilyalov in effect makes contact with the notion of geometric objects, re-deriving the “translation equation” (Aczél, 1987; Bilyalov, 2002; Moszner, 1995) (without using the name) that encodes the group property for general coordinate transformations and considering coordinate transformations not close to the identity and coordinates not close to Cartesian. Some recent work in differential geometry has also employed the symmetric square root of the metric (Binz and Pferschy, 1983; Bourguignon, 1993; Bourguignon and Gauduchon, 1992; Iochum and Schücker, 2006; Schücker, 2000), though the positive definite case lacks many of the complexities involved in taking the square root of the metric that appear in cases with indefinite metric signature.

While coordinate transformations near the identity suffice for defining covariant and Lie derivatives (Szybiak, 1963, 1966), the usual geometric objects formalism assumes that large coordinate transformations and very non-Cartesian coordinates are also permitted. This assumption is not, it should be emphasized, obviously a consequence of the equivalence principle, for which approximately Cartesian coordinates with some modest bends and wiggles (more or less as have been used in calculating the energy in asymptotically flat space-times) should suffice. If it turned out that the symmetric square root of the metric $r_{\mu\nu}$ and its
associated spinor formalism failed to exist for some large coordinate transformations, one would have the option of blaming those coordinate systems rather than blaming this spinor formalism. One recalls that Hilbert, Pauli, and Møller already suggested inequalities restricting coordinate systems (Hilbert, 1917; Møller, 1952; Pauli, 1921); one certainly can cover any Lorentzian manifold with coordinate charts with one time coordinate and three spatial coordinates (in the sense of the norms of the coordinate tangent vectors) (Francaviglia, 1988, p. 126). It would be possible (when topology permitted) to introduce an orthonormal tetrad in order to license coordinates that are far from Cartesian. This motivation of convenience is rather less compelling than the usually cited one (that coupling spinors to a curved metric requires a tetrad), however, for introducing a tetrad.

Fortunately Bilyalov finds that all coordinates far from Cartesian are admissible. However, not every ordering of the coordinates is permitted, because not just any coordinate can count as “time” $x^0$ in the sense of being paired with the minus sign in $\text{diag}(-1,1,1,1)$. Suitably reordering the coordinates is in effect the service rendered by Bilyalov’s matrix $T$ (Bilyalov, 1992). Restricting the ordering of the coordinates does alter the geometric objects formalism somewhat, but only in a way that reflects real physics: particle physics routinely attended to time reversals and spatial parity inversions (Kaku, 1993), but these transformations are meaningless without some metric-based distinction among the coordinates. Recently there has been renewed interest in the qualitative physical meaning of coordinate systems (Coll and Morales, 1992); the most detailed reckoning finds 199 options, but some of the distinctions made there are not needed here. While it is generally assumed that arbitrary finite coordinate transformations are permitted, there are not many results dependent on more than infinitesimal transformations and some
finite transformations near the identity (or near the Lorentz group, more generally); thus restricting the coordinate freedom is likely to cause little difficulty in most contexts. A possible drawback to restricting the coordinate order to have time first is that one might like to have the same coordinate serve as a time on one side of a black hole horizon and a spatial coordinate on the other side.\textsuperscript{21} This drawback seems to me not too severe compared to the benefits, however; one can investigate the metric using coordinates that are not licit for spinors, and then transform to other coordinates if spinors are introduced.

It turns out, then, that the symmetric square root of the metric, \( r_{\mu\nu} \), defined from \( g_{\mu\nu} \) with the help of \( \eta_{AB} = \text{diag}(-1,1,1,1) \) as the positive root of

\[
g_{\mu\nu} = r_{\mu\nu} \eta^{NA} r_{\alpha\beta}
\]

forms, with \( \eta_{AB} \), a geometric object (but not a tensor or anything else that is familiar) in its own right, independent of any orthonormal tetrad or even the possibility of one. (I extend the Einstein summation convention using Greco-Roman correspondence of letters when it is obvious.) In fact one might notice that the symmetric square root \( \hat{r}_{\mu\nu} \) of the \textit{conformal metric density} \( \hat{g}_{\mu\nu} \) suffices for this purpose, though Ogievetsky and Polubarinov do not point this out. The geometric object \( r_{\mu\nu} \) is nonlinear, in the sense (Nijenhuis, 1952) that the new components of \( r_{\mu\nu} \) under a coordinate transformation are nonlinear functions of the old components, in contrast to the situation with tensor densities and connections. I find that it has a linear transformation law not only for the Poincaré group of translations, boosts and rotations (Ogievetskiï and Polubarinov, 1965), but also for the rest of the 15-parameter conformal group (on which see (Fulton et al.,

\textsuperscript{21}I thank Charles Misner for this point.)
which is thus the “stability group” for this nonlinear group representation. Thus one notices that the determinant of the metric disappears from \( \Delta \) in equation 22 describing the nonlinear part of the transformation laws for the spinor and the symmetric square root of the metric (Ogievetskiǐ and Polubarinov, 1965).

One way to understand the nonlinearity of the transformation law is to use the component-wise equality of \( r_{\mu\nu} \) with the symmetric-gauged \( f_{\mu N} \) when the latter exists. Consider a coordinate transformation on both entities; sacrificing the local Lorentz freedom in \( f_{\mu N} \) to maintain a symmetric matrix of components (when that is possible) tends to pick out a unique Lorentz boost-rotation in terms of the field and the coordinate transformation, and this boost-rotation multiplies the field itself in the transformation law. The square root of the metric \( r_{\mu\nu} \), however, suffers from no topological obstructions, on account of its being a well defined analytic function of the metric \( g_{\mu\nu} \) and the matrix \( \eta_{AB} \) (or whatever the signature matrix might be, if other signatures or dimensions are used). Thus the square root of the metric is in some respects a hardier creature than an orthonormal basis, evading the hairy ball theorem and the restriction on the second Stiefel-Whitney class (though perhaps more delicate than a tetrad concerning coordinate freedom). A quick argument that works in some coordinate systems follows. At any point \( p \), there is an approximately Cartesian (almost freely falling) coordinate system such that the metric tensor components \( g_{\mu\nu} \) are not very far from \( \eta_{AB} = \text{diag}(-1, 1, 1, 1) \) (or other signature matrix). The binomial series expansion (DeWitt and DeWitt, 1952; Huggins, 1962; Ogievetskiǐ and Polubarinov, 1965) for \( r_{\mu\nu} \) in terms of \( g_{\mu\nu} \) and the signature matrix, which picks out the positive square root of the metric, therefore converges (DeWitt and DeWitt, 1952; Lin and Liu, 2001) in a finite
neighborhood containing \( p \). Thus the components \( r_{\mu\nu} \) of the square root of the metric are well defined in that coordinate system for some finite neighborhood around \( p \). Repeating this process finitely many times about different points on the ball (for the 2-sphere), or as many times as needed to cover the manifold in question, one obtains well behaved components of the square root of the metric, in one or another coordinate system (jumping from patch to patch as needed), everywhere on the manifold. Intuitively, the lack of directionality in \( r_{\mu\nu} \) helps it to be continuous where the (typically numerically equal) symmetric tetrad is discontinuous. The lack of obstructions for \( r_{\mu\nu} \) might not imply any similar hardiness for the pair \( \langle \hat{r}_{\mu\nu}, \psi \rangle \), however.

Relating spinors to geometric objects pays dividends. The existence of a Lie derivative of spinors (with the help of the metric) provides a direct solution to problem of finding external symmetries (translations, rotations and the like) of spinors. The belief that no spinor Lie derivative exists has motivated a less direct solution (Cotăescu, 2000), though it is not impossible that the less direct approach might give easier calculations in some cases. Above it was mentioned that not the full square root \( r_{\mu\nu} \) of the metric \( g_{\mu\nu} \), but only the square root \( \hat{r}_{\mu\nu} \) of the its conformal part \( \hat{g}_{\mu\nu} \), need accompany the spinor \( \psi \) to make a geometric object (or something very similar). Because \( \hat{r}_{\mu\nu} \) and \( \hat{g}_{\mu\nu} \) are equivalent as geometric objects (at least for admissible coordinate systems), one can take the object in question to be \( \langle \hat{g}_{\mu\nu}, \psi \rangle \), thus making part of the object linear. The appearance of only \( \hat{g}_{\mu\nu} \) rather than \( g_{\mu\nu} \) immediately explains some otherwise somewhat obscure facts about conformal Killing vectors and the Lie differentiation of spinors (as in (Benn and Tucker, 1987)). Applying some of Szybiak’s results for nonlinear geometric objects (Szybiak, 1963, 1966) to the object \( \langle \hat{g}_{\mu\nu}, \psi \rangle \) shows for the coordinate
transformation properties of the Lie derivative (the vector field \( \xi \) not being the generator of the coordinate transformation), with irrelevant details suppressed,

\[
(\mathcal{L}_\xi \psi)' \sim \left( \frac{\partial \psi'}{\partial \psi} \mathcal{L}_\xi \psi + \frac{\partial \psi'}{\partial \hat{g}_{\mu\nu}} \mathcal{L}_\xi \hat{g}_{\mu\nu} \right)
\]  

(2.24)

the Lie derivative with respect to conformal Killing vector fields is *nicer*, but Lie differentiation exists in general. For covariant differentiation one has the coordinate transformation property along the lines of

\[
(\nabla \psi)' \sim \left( \frac{\partial \psi'}{\partial \psi} \nabla \psi + \frac{\partial \psi'}{\partial \hat{g}_{\mu\nu}} \nabla \hat{g}_{\mu\nu} \right)
\]  

(2.25)

Using \( \nabla \hat{g}_{\mu\nu} = 0 \), one sees that \( \nabla \psi \) is a spinor-covector and that \( \langle \hat{g}_{\mu\nu}, \nabla \psi \rangle \) is a geometric object without \( \psi \) or \( \nabla \hat{g}_{\mu\nu} \) appearing.

In contrast to the view that an orthonormal tetrad is obligatory for handling spinors in curved space-time, a correct statement is made by some workers in supergravity:

\[\ldots\] in fact it is impossible to define a field, transforming *linearly* under the \( \lambda \) [coordinate] group, which also transforms as a spinor when the \( \lambda \)'s are restricted to represent global Lorentz transformations. It is possible to get around this difficulty by realizing the \( \lambda \) transformations nonlinearly, but this is not a convenient solution\ldots .

To solve these problems we enlarge the gauge group by adjoining to the \( \lambda \) transformations a group of *local* Lorentz transformations, and define spinors with respect to this group. This is a procedure familiar in treatments of nonlinear \( \sigma \) models. Nonlinear realizations of a group are replaced by linear representations of an enlarged (gauge) group. The nonlinearities reappear only when a definite gauge choice is made. Similarly here, by enlarging the gauge group, we obtain linear spinor representations. The nonlinear spinor representations of the general coordinate group reappear only if we fix a gauge for the local Lorentz transformations. (Gates et al., 1983, p. 234)
Inconvenience is rather different from impossibility; the inconvenience mentioned here doubtless refers to the technical complication (involving series expansions and the like) evident in the Ogievetsky-Polubarinov formalism (Ogievetskiǐ and Polubarinov, 1965). Convenience can be goal-relative: Ogievetsky and Polubarinov explicitly presented their perturbative treatment with no gratuitous local Lorentz group as useful for calculations in perturbative quantum gravity. In any case mere technical complication due to nonlinearities is not a philosophically interesting reason to avoid the bloated ontology and unwanted Anderson-Friedman absolute object in the tetrad-spinor formalism. Ivanenko and Sardanashvily discuss matters relevant to the global and strong-field aspects of Ogievetsky-Polubarinov spinors (Ivanenko and Sardanashvily, 1983). At least locally, it seems appropriate to regard the Ogievetsky-Polubarinov formalism as the default formalism, whereas the tetrad formalism serves to linearize the naturally nonlinearity by installing artificial gauge freedom. That this viewpoint has not been the standard one illustrates how a Lakatosian rationally reconstructed history (Lakatos, 1971) can differ from real history due to external factors such as lack of awareness of certain literatures by the actors in question.

2.14 Unwanted Vanishing Absolute Objects: Concomitant Irrelevance

If a philosopher asks what is claimed to exist according to contemporary physical theory, then some criterion for choosing among the many formulations of Einstein’s equations—with many different choices of primitive fields, different numbers of components, etc. (Ashtekar and Lewandowski, 2004; Buchdahl, 1960; Kichenassamy, 1986; Lindström, 1988; Misner et al., 1973; Møller, 1964; Pappas and Stachel, 1978; Peres, 1963; Pons, 2003; Ray, 1975) is needed (Bain,
2004; Jones, 1991; Lyre and Eynck, 2003). One can readily write down formulations in terms of a metric, in terms of a background metric and a potential, in terms of a metric and a connection, in terms of a metric, a connection and a Lagrange multiplier field, in terms of a tetrad of orthonormal vector fields, in terms of a tetrad and a connection, \textit{etc}. At least apart from some degenerate solutions (likely of measure zero) or global issues, these theories have empirically equivalent sets of models. All of these formulations tend to be regarded as ‘just’ GTR in the physics literature, which generally does not seek precise criteria for theory individuation. Clearly they are not strictly the same theory, however, for one theory cannot portray a physical situation as both possible and impossible. Einstein’s ‘theory’ of gravity in these various formulations, then, would be a genuine, un-contrived example of the sort of underdetermination of theories by data that especially afflicts scientific realists, who interpret theories literally, aspire to believe all and only the true ones, perhaps hold to some sort of long-run convergence on truth, and so on. If these various formulations of GTR, or at any rate of Einstein’s field equations, are really distinct theories, then scientific realists can only hope to choose among them using theoretical virtues. Not everyone, it should be noted, despairs of making a rationally motivated choice in such contexts (Musgrave, 1992).

On such matters, the Andersonian tradition, with some friendly amendments partly inspired by Geroch (Pitts, 2006a), is a better guide than most. Anderson insisted on eliminating irrelevant fields. The friendly amendments insist on eliminating locally irrelevant fields and using only irreducible geometric objects.

There are more forms of surplus structure than Anderson’s notion of irrelevance captures. Regions of a fluid’s 4-velocity field are locally irrelevant if the fluid has
holes there. Clock fields are variationally irrelevant because the Euler-Lagrange equations obtainable from a variational principle would hold anyway (albeit not as Euler-Lagrange equations) due to the generalized Bianchi identities. A further sort of excess manifests itself in the so-called ‘Palatini’ formulation\(^{22}\) in which the connection \(\Gamma^\alpha_{\mu\nu}\) is \textit{a priori} independent of the metric and reduces to the metric-based Christoffel symbols \(\{_{\mu\nu}^\alpha\}\) = \(\frac{1}{2}g^{\alpha\sigma}(-g_{\mu\nu,\sigma} + g_{\nu\sigma,\mu} + g_{\sigma\mu,\nu})\) only on-shell. In this case the tensorial difference \(\Gamma^\alpha_{\mu\nu} - \{_{\mu\nu}^\alpha\}\) constitutes an absolute object, albeit one with the value of 0 on-shell.\(^{23}\) Likewise for the potential-field strength electromagnetic variational principle of Julian Schwinger (Schwinger, 1953, 1998), there exists the on-shell vanishing absolute object \(F^\mu_{\nu\nu} - (\partial_\mu A_\nu - \partial_\nu A_\mu)\). At least classically, it is best to eliminate this concomitant irrelevance by reducing the connection \(\Gamma^\alpha_{\mu\nu}\) and field strength \(F^\mu_{\nu\nu}\) to be concomitants of the metric and 4-vector potential, respectively: \(\Gamma^\alpha_{\mu\nu} =_{\text{def}} \{_{\mu\nu}^\alpha\}\) and \(F^\mu_{\nu\nu} =_{\text{def}} \partial_\mu A_\nu - \partial_\nu A_\mu\). If the absence or scarcity of absolute objects means anything interesting, then these formulations yielding the triviality of these vanishing absolute objects are less than ideal. Perhaps one could learn to live with a vanishing absolute object in GTR, but why do so? An attractive principle is the following: if one geometric object is a concomitant of another on-shell, then it should be treated as one off-shell (identically) as well. This principle would exclude the metric-affine variational principle for GTR and also the Rosen-Sorkin Lagrange multiplier trick, the latter because the Riemann tensor on-shell is a concomitant of the zero tensor. It is not impossible that quantization of the metric-affine Einstein-‘Palatini’ variational principle would be inequivalent.

\(^{22}\)The procedure of varying the metric and connection independently is actually due to later work by Einstein (Ferraris et al., 1982; Querella, 1998), whereas Palatini’s contribution was merely to express the variation of the Ricci tensor in terms of the covariant derivative of the variation of the connection (Buchdahl, 1960; Kichenassamy, 1986).

\(^{23}\)Consideration of absolute objects that vanish was inspired by a conversation with Katherine Brading.
to quantization of the metric theory (Querella, 1998). However, this inequivalence would resolve the difficulty: there would be two different theories with different empirical consequences, not just different *prima facie* ontologies.

It turns out that in supersymmetry, however, there might well be a reason to retain fields that vanish given the equations of motion. The auxiliary fields of supersymmetric theories tend to vanish on-shell (Kaku, 1993). While formulations without auxiliary fields are available, formulations with auxiliary fields are preferable for some purposes, as P. K. Townsend explains:

> [f]or many years one of the central problems of supersymmetric field theories has been that of auxiliary fields. Generally, supersymmetry transformations that leave invariant a given action will close to form the usual supersymmetry algebra only when the field equations are used. This is called “on-shell supersymmetry”. If field equations are not needed to close the algebra then the supersymmetry is “off-shell”. For this to happen we must have equal numbers of boson and fermion field components as well as equal numbers of boson and fermion propagating modes (i.e. states). Since fermions always have more components than they propagate modes, a balance in the number of boson and fermion states will usually mean an imbalance in the number of boson and fermion field components. Hence the need for boson “auxiliary” i.e. non-propagating, fields to balance the number of boson and fermion components off-shell while not disturbing the balance of states. Of course equality of the numbers of bosons and fermions is only a necessary condition for off-shell supersymmetry, not a sufficient one. In general, a solution to an auxiliary field problem will require fermion as well as boson auxiliary fields.

There are several reasons why we should want auxiliary fields. The most obvious one is that if the algebra of transformations includes field equations then these transformations are tied to the particular action that yields these equations. The introduction of auxiliary fields frees the transformation rule of reference to a particular action. This allows, inter alia, the addition of invariant actions to produce new invariant actions. Without auxiliary fields this involves a laborious procedure in which terms are added order by order to the action and transformation rules to obtain a new invariant action with new transformation rules.

A second, and more important, motivation for auxiliary fields is that
they allow us to consider superfield perturbation theory in which com-
ponent field Feynman graph calculations can be done together, and
with more ease, as supergraph calculations. This formalism allows
us to deduce immediately, just from the form of the Feynman rules,
many “non-renormalization theorems” which have remarkable conse-
quences for the ultraviolet behavior of the N-extended supersymmetric
theories. (Townsend, 1983, pp. 240, 241)

On the other hand, these advantages appear to be matters of convenience of calcu-
lation or aids in the process of discovery. While easy calculations and discoveries
are desirable, perhaps auxiliary fields that vanish on-shell are nonetheless inappro-
priate for metaphysical questions. It has been suggested before that the context
of discovery (as opposed to justification) is a safe home for various otherwise em-
barrassing phenomena (Reichenbach, 1938). I will not resolve this question here.
I do suggest, however, that supergravity (Boulware et al., 1979; Deser, 1980; van
Nieuwenhuizen, 1981; van Nieuwenhuizen et al., 2006; Weinberg, 2000) is a subject
worthy of philosophers’ attention, for (at least) two reasons. The auxiliary field
issue raises the possibility that a field that vanishes in all physically possible situa-
tions (or dynamical possible trajectories, as Anderson would put it) is nonetheless
physically significant. Moreover, supersymmetry transformations, being in a sense
the square root of translations, constitute an otherwise unimaginable possibility
that stretches the conceptual resources of space-time theorizing. The contingent
empirical fact that supersymmetry either does not occur in nature or is broken
(at least by the vacuum state, if not by the laws) does not deprive the subject of
interest.
2.15 Conclusions about Anderson-Friedman Absolute Objects Program

By now some conclusions can be drawn about the Anderson-Friedman absolute objects program. It is clear from the existence of susceptible kinds of geometric objects (Zajtz, 1988b) that the criterion of sameness in all models picks out some examples that were not intended. The particular example of the Geroch-Giulini $\sqrt{-g}$ counterexample threatens the core of the program by challenging the claim that GTR is generally covariant in the sense of lacking absolute objects, construed in terms of sameness in all models.

Given that something important is wrong with the Anderson-Friedman program, the question arises whether the flaw can be fixed or not. The program is so rich and yields so many of the expected answers that one might well hope that it can be fixed. An attractive candidate for a change is to follow the suggestion that Maidens (Maidens, 1998) ascribes to Hiskes (Hiskes, 1984), namely, adding the requirement of nonvariationality as an additional necessary condition for absoluteness. Thus an absolute object would both satisfy the (local) sameness-in-all-models condition and fail to be varied in an action principle. Such a modification would limit attention to theories that possess an action principle, of course. While most contemporary physicists likely would accept the claim that any fundamental theory of physics—or rather, its classical (non-quantum) variant—should possess an action principle, there are theories of some interest that at least presently do not have one (or did not as recently as 2005). In particular, interacting higher spin (that is, higher than spin 2, which corresponds to gravity) field theories presently tend to lack action principles and are nonlocal in some attractive formulations (Bekaert et al., 2005). This lack of a variational principle is a matter of ongoing research (Barnich and Grigoriev, 2006; Buchbinder and Krykhtin, 2007; Buchbinder
et al., 2007; Fotopoulos et al., 2006) and so might change in the near future. It seems likely that these examples will yield to a variational formulation eventually, so building nonvariationality into the concept of absoluteness likely should not be resisted on their account. Building nonvariationality into the concept of absoluteness would resolve the Geroch-Giulini counterexample, because $\sqrt{-g}$ is varied in the action principle for GTR. Whether this resolution is satisfactory or not will be discussed shortly.

A major problem with fixing the Anderson-Friedman program by building nonvariationality as well as sameness in all models into the concept of absoluteness is that the condition of sameness in all models then seems redundant. It is difficult to imagine a fundamental theory in which one abstained from varying some fields, and yet the field violated local sameness in all models. If nonvariationality entails sameness in all models (at least within the realm of reasonable theories), while sameness in all models also holds in many unwanted cases due to susceptibility, then the condition of sameness in all models does little or no work. Instead it appears to be an accidental feature of nonvariational fields, as well as a surprising feature of the susceptible geometrical objects, whether varied or not, and thus not essentially tied conceptually to absoluteness at all. The lesson here perhaps seems to be that the Anderson-Friedman criterion was off-target rather than incomplete, so attempting a repair would miss the lesson.

If the Anderson-Friedman criterion of sameness in all models does not capture a core feature of absoluteness, perhaps it is best to agree with the various authors, such as Earman (Earman, 2003a) and Wald (Wald, 1993), who take the lack of nonvariational fields to pick out the interesting non-formal general covariance that GTR exemplifies. As Earman notes, this criterion is vulnerable to counterexam-
ples using Lagrange multipliers to derive, for example, the flatness of a flat metric from a variational principle in which the metric is also varied (Rosen, 1966; Sorkin, 2002). Perhaps the solution is to ban from an action principle any geometric objects that are Lagrange multipliers on grounds of irrelevance (Pitts, 2006a), as discussed above. As Earman also notes, clock fields also threaten the criterion of no nonvariational fields as the test for general covariance (Earman, 2006). While Earman declines to seek a more refined criterion to exclude theories with clock fields, it seems to me that the relevant lesson here is that nonvariational fields can be converted into functions of clock fields and their derivatives. If nonvariational fields can be reduced to clock fields, then checking for nonvariational fields while permitting clock fields is a criterion bound to fail: one can simply convert any erst-while nonvariational fields into functions of (derivatives of) clock fields to get an equivalent formulation lacking nonvariational fields, thus satisfying the condition for substantive general covariance even in clearly absurd cases such parameterized field theories in flat space-time. Clock fields will be discussed in more detail in the next section.

Having learned to ban the Rosen-Sorkin trick on grounds of its use of irrelevant fields, and having learned to ban clock fields as inter-convertible with nonvariational fields, one might consider the criterion of no nonvariational fields as the correct notion of substance general covariance. However, it seems to me that this criterion gives the right answer for the Geroch-Giulini case, but for the wrong reason; GTR counts as generally covariant, but not because the correct feature has been identified. As noted above, varying $\sqrt{-g}$ in GTR is doubly odd. It is odd partly because $\sqrt{-g}$ is locally the same in all models, and so seems to be part of the physically necessary furniture of the world, not something that
should be compared between neighboring possible worlds, for reasons discussed by Maidens (Maidens, 1998). Varying $\sqrt{-g}$ in GTR is also odd because almost all of the Euler-Lagrange equations would have held anyway on account of the generalized Bianchi identities; varying an entire field ($\infty^4$ numbers, so to speak) to get a single equation seems like overkill. Why not simply vary one little part of $\sqrt{-g}$ instead? Above I explored some tentative suggestions for finding that one little piece of $\sqrt{-g}$ to vary. Whether those suggestions ultimately bear fruit or not, it appears that we still do not have a fully satisfying answer to the question (discussed with some insight by Maidens) about why we vary exactly the fields that we vary in an action principle.\(^{24}\) Doubtless a rough idea of the desired field equations and a desire (based on convenience or even pragmatic necessity (Guay, 2008)) to work with local field theories are part of the answer, though a further argument for avoiding a mass term might be required. While the availability of a local field theoretic formulation might be ontologically significant, it is difficult to see any ontological significance in our using a local formulation instead of a leaner nonlocal formulation that avoids varying some or all of the (probably nonlocal and perhaps Fourier analysis-related) pieces of the geometric objects that do not contribute essential field equations. (While causality might come to mind, the local formulation might simply make the desired causality properties obvious.)

If, for example, the divergence of the electromagnetic Euler-Lagrange equations would have been true anyway because $1 - 1 = 0$, why vary that part of the vector potential that yields this divergence? And yet relying on the absence of nonvariational objects in (a local field formulation of) GTR to analyze substantive general covariance falls prey to this problem. If varying most of $\sqrt{-g}$ contributes nothing to the field equations that would not have been true anyway (given the variation

\(^{24}\)I thank Ted Jacobson for discussion on this matter.
of $\hat{g}_{\mu\nu}$, then why think that varying $\sqrt{-g}$ instead of some single number (like the average or the space-time volume) reflects anything more than a preference for local field theories on grounds of convenience? I see no answer to this question at present. Thus I conclude that the purely variational criterion for substantive general covariance, though (to my knowledge) extensionally correct in categorizing local field theories as substantively generally covariant or not, likely gives the right answer for a wrong reason. If that is so, then the variational criterion also fails to yield the essence of substantive general covariance.

Perhaps the problem here is that there just isn’t any such thing as substantive general covariance: maybe no single idea captures the crucial novelty of GTR and the various features that Einstein and others have attributed to general covariance.\(^25\) Instead there seems to be a cluster of partially overlapping ideas, many of which admit explicit characterization in terms of differential geometry, the principle of least action, and related tools of contemporary field theory. Much as one can search for the Fountain of Youth without there being any such thing, one can search for the one true General Covariance without there being any such thing. As with some quests for physical objects, the quest for substantive general covariance has been profitable, as I trust has been shown in this chapter, even if the aim of the quest has not been achieved and is perhaps unachievable. But perhaps the next section’s discussion of clock fields will end on a more optimistic note for general covariance.

Background independence is a name widely used in the contemporary canonical quantum gravity project as a synonym for substantive general covariance. Thus the considerations above suggest that there might not be any such concept as

\(^{25}\)Harvey Brown in discussion and John Earman (Earman, 2006) have both suggested that substantive general covariance might be more difficult to analyze than is often thought.
background independence. If so, it can hardly be a virtue of the loop quantum gravity program that it provides a background-independent quantization of gravity or that it preserves background independence as the great lesson about space-time learned from GTR. Of course one could simply stipulate that background independence just is some technical notion to the effect that arbitrary coordinates can be used, none is preferred (as in the case of clock fields), and all fields are varied in the action principle (or path integral, once quantization is in view). But then much of the significance of background independence as a great conceptual lesson of 20th century physics is lost. As I noted above, the technical interest of loop quantum gravity is considerable, so the loss of the concept of background independence would not be a great setback. It might turn out that recognizing a bit of background in $\sqrt{-g}$ would be technically helpful in legitimating the use of some strategies for obtaining a Lie algebra of gravitational constraints (Kuchař and Romano, 1995), despite their implicit reliance on a fixed volume element if deployed for certain purposes (c.f. (Thiemann, 2006)). Concerning another stream of quantum gravity research, perhaps recognizing a bit of fixed background in $\sqrt{-g}$ would help to address the problem that the action is unbounded below in the positive definite case (c.f. (Gibbons et al., 1978)). If even GTR must use a bit of fixed background in $\sqrt{-g}$, then there is little reason to resist the use of a fixed volume element. Of course such suggestions would have to be tested by detailed calculation. The mere possibility of a technical payoff from the conceptual difficulties in analyzing general covariance and the apparent absolute of (most of?) $\sqrt{-g}$ is already interesting, however.
Another promising suggestion for analyzing the assumed substantive general covariance of GTR in contrast to earlier theories can be found in the hope that GTR is “already parameterized” (Arnowitt et al., 1959, 1960, 1962a,b; Dirac, 1964; Kuchař, 1972, 1973, 1976a,b,c, 1978, 1981; Torre, 1992b). Whereas field theories in Minkowski space-time (or other pre- or non-generally covariant theories) admit parameterized formulations with the explicit introduction of clock fields to serve as preferred coordinates in disguise, and nonrelativistic mechanics admits treating time as a dynamical variable in terms of some new parameter (Lanczos, 1949), such theories do not naturally come in this parameterized form. Originally there was hope that GTR was quite literally already parameterized, with clock fields built into the theory in such a fashion that, with sufficient cleverness (likely involving a nonlocal canonical transformation, use of coordinates based on scalar concomitants of the Riemann tensor, etc.) one might explicitly identify the clock fields and deparametrize the theory. Parameterized theories display important formal similarities to GTR in terms of certain Poisson bracket relationships that tend to strengthen the impression that GTR is already parameterized (Dirac, 1964; Isham and Kuchař, 1985). It turns out that there are certain technical obstacles that exclude the strongest versions of the claim that GTR is already parametrized (Torre, 1992b; Westman and Sonego, 2007), but there might be plausible senses in which the claim is true. Given the analogy between parametrized field theories and GTR, there has been considerable work on parametrized field theories as a step towards quantum gravity (Arnowitt et al., 1959, 1960, 1962a,b; Cho and Varadarajan, 2006; Dirac, 1964; Hájíček and Kuchař,
1990; Isham and Kuchař, 1985; Kuchař, 1972, 1973, 1978, 1981; Kuchař and Stone, 1987; Lee and Wald, 1990; Sundermeyer, 1982; Torre, 1992a,b; Varadarajan, 2004, 2007). Some early works in quantum gravity employed clock fields even in GTR (Bergmann and Brunings, 1949; Bergmann et al., 1950), but this approach did not persist. Given the research trends, one might conclude that, whether GTR is already parametrized in some deep sense or not, at least GTR is distinctive in admitting arbitrary coordinates without the need to install clock fields.

In the typical case of a field theory in Minkowski space-time, the parametrization process (in its simpler Lagrangian version) proceeds as follows. One promotes all the (Lorentz) vectors, tensors, etc. to world vectors, tensors, etc. For typical dynamical fields, little change is needed, except perhaps a choice of density weight (that distinction being almost trivial at the Lorentz covariant level because Lorentz transformations have (anti)unit determinant). For spinors, which are almost never discussed in this context, this step is nontrivial; in principle the OP-Bilyalov formalism should suffice, but with strong nonlinearity in the clock field gradients. For the flat background metric, one now has not the matrix $\text{diag}(-1,1,1,1)$ but a Riemann-flat tensor

$$\eta_{\mu\nu} = X^A,_{\mu} \eta_{AB} X^B,_{\nu},$$

(2.26)

where $X^A$ is a set of Cartesian coordinates and $x^\mu$ is a set of arbitrary coordinates. Thus the flat metric is reduced to a function of the partial derivatives of the clock fields. (For metrics with nonzero constant curvature, there will be dependence on one or more of the clock fields themselves as well.) Next one promotes, or perhaps rather demotes (Kuchař, 1973), the preferred coordinates $X^A$ into variational fields in the action principle, which is expressed using the arbitrary coordinates $x^\mu$. While much of the work on parametrized field theories occurs in a Hamiltonian
formalism, the Lagrangian formalism is appropriate for my purposes. One readily demonstrates the claim (made occasionally above) that it does not much matter whether $X^A$ are varied, because the Euler-Lagrange equations from varying $X^A$ are identical to the result of using the other fields’ Euler-Lagrange equations in the generalized Bianchi identities. Let the other fields be represented by a weight $w$ $(1, 1)$ tensor density $\phi^\alpha_\beta$. Because the action $S$ is a scalar, it is unchanged under arbitrary infinitesimal changes of the coordinates $x^\mu$, including the ones generated by an arbitrary vector field $\xi^\mu$. Taking this coordinate transformation to have compact support makes all boundary terms to disappear, including those from pulling off derivatives from $\xi^\mu$ in $L_\xi \phi^\alpha_\beta = \xi^\mu \phi^\alpha_\beta,\mu - \phi^\mu_\alpha,\mu \delta S / \delta \phi^\alpha_\beta + \phi^\alpha_\mu,\beta \phi^\mu_\xi,\mu + w \phi^\alpha_\beta,\mu$ . After making some rearrangements and using the arbitrariness of $\xi^\mu$ to pull off the integral sign, one obtains the generalized Bianchi identity

$$
\phi^\alpha_\beta,\mu \frac{\delta S}{\delta \phi^\alpha_\beta} - w \frac{\partial}{\partial x^\mu} \left( \phi^\alpha_\beta \frac{\delta S}{\delta \phi^\alpha_\beta} \right) + \frac{\partial}{\partial x^\alpha} \left( \phi^\alpha_\beta \frac{\delta S}{\delta \phi^\alpha_\beta} - \phi^\alpha_\mu \frac{\delta S}{\delta \phi^\alpha_\beta} \right) + X^A,\mu \frac{\delta S}{\delta X^A} = 0. \tag{2.27}
$$

Letting $\phi^\alpha_\beta$ be on-shell leaves

$$
X^A,\mu \frac{\delta S}{\delta X^A} = 0. \tag{2.28}
$$

Clock fields have non-vanishing linearly independent gradients, so $\frac{\delta S}{\delta X^A} = 0$, which was to be demonstrated. Here the explicit form of the fields besides $X^A$ ultimately does not matter, but the simple form of the Lie derivative of scalar fields, which is algebraic in the parameters $\xi^\mu$, does important work. Nothing changes if second (or higher) derivatives of $X^A$ are present, as they will be if a background connection is reduced to a function of clock fields and their derivatives.

The functional form of a parametrized theory’s dependence on the clock fields
determines the theory’s symmetry group in a simple way, as a typical example illustrates sufficiently. The translation subgroup of the Poincaré group manifests itself in having only differentiated clock fields in the formulation, because \((X^A + c^A),\mu = X^A,\mu\) for constants \(c^A\). The Lorentz subgroup \(O(1, 3)\) manifests itself in having only the Lorentz-invariant combination of the clock field gradients in the theory formulation, not each gradient by itself. Not coincidentally, the Poincaré group is just the group of Killing vector fields for a flat metric tensor \(\eta_{\mu\nu}\).

Converting a nonvariational object to a function of clock fields and their derivatives is an easy path to the sort of artificial general covariance (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 1998) that presumably would result from a more involved BFT-type procedure applied to massive GTR. The reduction back to the original formulation is accomplished readily by setting, for example, \(X^M = x^\mu\) everywhere. Thus nonvariational fields and clock fields appear to be inter-convertible. However, if one has introduced gauge freedom in order to achieve some other goal, such as a consistent notion of causality (Pitts and Schieve, 2007), then the condition \(X^M = x^\mu\) might not be an allowed gauge condition. Otherwise inter-convertibility seems to hold in general.

Of late philosophers of physics have taken to discussing parametrized theories (Earman, 2003a, 2006; Norton, 2003). As John Norton discusses, Einstein often ascribed to inertial coordinate systems the (objectionable to Einstein) property of acting but not being acted upon; even late in life, when the concept of geometric object would have been available, Einstein persisted in faulting inertial coordinate systems (Norton, 2002). It might be tempting to regard Einstein’s views as outmoded, and presumably surpassed by more sophisticated projects such as that of Anderson and Friedman. As John Earman notes, it is widely held that there
is nothing very interesting about coordinate systems (Earman, 2006). Earman then proceeds to discuss clock fields as an obstacle to defining substantive general covariance, using the example of unimodular gravity. He declines to suggest yet another notion of general covariance to avoid this problem.

It seems to me that noting the inter-convertibility of clock fields with nonvariational fields solves this problem, because then, for example, the formulation of unimodular gravity with clock fields is readily converted to a formulation with a nonvariational background volume element, which by Earman’s standards violates substantive general covariance. Given that clock fields and nonvariational fields are inter-convertible, a criterion for substantive general covariance that bans nonvariational fields can only succeed if it also bans clock fields. Thus one might analyze substantive general covariance as follows.

Definition. A field theory is substantively generally covariant just in case it is formally generally covariant (in the sense of admitting at least arbitrary infinitesimal coordinate transformations and some finite transformations near the Lorentz group), lacks nonvariational fields and lacks clock fields.

To my knowledge this criterion performs as expected, extensionally (assuming that one retains Anderson’s ban on irrelevant fields, of course), with perhaps some ambiguity if the gauge $X^M = x^\mu$ is not permitted, as in massive gravity (Pitts and Schieve, 2007). This criterion of lacking clock fields, however, clearly is at heart much like Einstein’s original claim that GTR was generally covariant in the sense of admitting expression using arbitrary coordinates. To the Kretschmann objection that any theory could be so formulated—one can imagine a counterfactual history in which Kretschmann produces the example of parametrized field theories in Minkowski space-time in response to Einstein’s proposal—one could reply on Einstein’s behalf that Kretschmann’s general covariance was artificial because it
introduced clock fields as a very transparent disguise for preferred coordinates, whereas GTR does no such thing. Thus there is some hope for distinguishing the real general covariance of GTR from the artificial sort.

It is perhaps ironic that so old-fashioned a criterion as the absence of preferred coordinates, with some modern polish, works better than the more sophisticated Anderson-Friedman project and appears to give the desired answers. (Sean Carroll’s view has some similar elements (Carroll, 2004).) The absence of clock fields and the absence of fields not varied in the action principle seem to be not merely coextensive, but very simply related conceptually in light of the inter-convertibility of clock fields and nonvariational fields. The claim that there is nothing much interesting about coordinates is overstated; while it is true for GTR, it is not true in general, and this very fact sheds light on GTR. Of course clock fields $X^A$ have no exemption from the problem that some manifolds cannot be covered with a single coordinate chart; thus one might need to piece together generalized clock fields out of multiple charts.

If there is in fact extensional equivalence between the purely variational definition and the parametrized definition of substantive general covariance, there is no need to choose between them if one merely wants to classify a given theory as substantively generally covariant or not. Above I argued that the variational formulation gives the right answer for perhaps the wrong reason in the case of GTR. Does a similar worry arise here? While I admit surprise that the sophisticated Anderson-Friedman analysis fails and the simple coordinate analysis of the early Einstein essentially succeeds, presently I see no objection to taking the absence of clock fields (and irrelevant fields and nonvariational fields) as an adequate analysis of general covariance. Perhaps the preceding section’s doubts about whether
substantive general covariance exists were premature. Once one bans nonvariational fields and irrelevant fields, formal general covariance without clock fields apparently does imply substantive general covariance.

The parametrized form of a certain massive variant of GTR (Freund et al., 1969; Schmelzer, 2000) in chapter 4 will further illustrate the wisdom of using clock fields as a test for general covariance. Installing artificial gauge freedom by parametrization at least formally restores many features of GTR that one might associate with substantive general covariance: gauge freedom, point individuation questions such as appear in the hole argument, the absence of non-variational fields in the Lagrangian density, a vanishing Hamiltonian apart from boundary terms, etc.

As a further example of the power of clock fields to simulate substantive general covariance in a deceptively plausible way, one might consider GTR as formulated with an ADM split of space-time into space and time (Misner et al., 1973), which introduces a temporal foliation. One can add the time foliation as a new variational field \( T \) at no cost, because the resulting Euler-Lagrange equation is entailed by the others. One now has the ingredients for building lots of new theories using 3-dimensional entities, which in turn are defined using 4-dimensional tensors and \( T \). Of course most of these theories, unlike GTR, will make ineliminable reference to \( T \) and contain absolute simultaneity observably. But the theories’ Lagrangians are built entirely using 4-dimensional tensors and the ‘scalar’ clock field \( T \). Clearly introducing the clock field \( T \) into the theory evacuated the 4-dimensional symmetry of content while preserving it formally. Given that clock fields are such outstanding tools for subverting otherwise promising criteria for substantive general covariance, it seems advisable (assuming that one has already converted all
nonvariational fields into clock fields) to understand substantive general covariance in terms of the absence of clock fields.
CHAPTER 3

INEXACT EMPIRICAL EQUIVALENCE IN ELECTROMAGNETIC,
YANG-MILLS, AND GRAVITATIONAL THEORIES

3.1 Types of Inexact Empirical Equivalence

The most common version of empirical equivalence to be discussed by philosophers is the case of exact empirical equivalence for all models of two theories. The potential interest of such a scenario is evident: obviously there is no chance in any nomologically possible world that experimental progress will resolve the debate, while settling it on theoretical grounds might also be difficult. However, this scenario runs the risk that the two supposedly rival theories are in fact one and the same theory in different guises. Such identification was often made by those influenced by logical empiricism. A related weaker claim is made today by John Norton, namely, that for theories for which the “observational equivalence can be demonstrated by arguments brief enough to be included in a journal article . . . we cannot preclude the possibility that the theories are merely variant formulations of the same theory.” (Norton, 2008, p. 17) Norton evidently has in mind journal articles in the philosophy of science, not physics or some other science (Norton, 2008, p. 33). While Norton aims to deny that the underdetermination of theories by data is generic and that philosophers’ algorithmic rivals carry much force, I aim to show that there are some serious candidates for underdetermination that
arise from within real physics and that have not been discussed much, if at all, by philosophers. I make no inductive claim (which Norton would dispute) that these examples imply that all theories are always underdetermined by evidence. However, the examples available from real physics do seem sufficiently widespread and interesting that it might well frequently be the case that scientific theories are permanently underdetermined by data. It is therefore helpful to observe that the physics literature suggests by example several slightly weaker notions of empirical equivalence that, being weaker, are immune to the strategy of being identified as one and the same theory and hence not rivals, yet strong enough that there is no realistic prospect for distinguishing the two theories empirically.

Among philosophers the question has been raised what to make of the many empirically equivalent (or nearly equivalent) theory-candidates or formulations in gravitation, both for Newtonian gravity and for theories employing Einstein’s equations (Bain, 2004; Jones, 1991; Lyre and Eynck, 2003). Are the several Newtonian (or Einsteinian) theory-candidates just formulations of the same theory, or are they rivals? If they are rivals, are some of these theories better than others? Especially because some versions of both the Newtonian and Einsteinian gravities have flat space-time with absolute objects and a gravitational force, while others employ curved space-time, the theory candidates’ ontologies and explanatory mechanisms vary rather widely, despite the complete or nearly complete empirical equivalence between the two approaches. This choice takes up the issue (discussed by Lotze (Russell, 1897), Poincaré (Poincaré, 1902) and Reichenbach (Reichenbach, 1928)) of the equivalence between curved geometry and flat geometry with universal forces, albeit in a more detailed and vastly more physically plausible way in the context of gravitation (Boulanger and Esole, 2002; Boulware and Deser,
ontologies might embarrass scientific realists, who presumably wish to invest the fields used with (meta)physical significance. An incomplete list of formulations found in the literature on GTR (construed in the physicists’ broad and vague sense) reveals a host formulations in terms of different variables and even different numbers of variables. For curved space-time Einsteinian formulations, there are still many different choices of primitive fields from which to choose, with different numbers of components. To mention just some (excluding spinor formulations, for example), one has the typical metric formulation (itself non-unique in Lagrangian density between, e.g. the Hilbert $R$ Lagrangian density with second derivatives of the metric and the Einstein $\Gamma$ Lagrangian density merely quadratic in the Christoffel symbols, and other choices (Pons, 2003) for example, as well as in the choice of variables between the metric $g_{\mu\nu}$, its inverse $g^{\mu\nu}$, and uncountably many densitized relatives of each); the ADM $3+1$ split with a spatial metric $h_{ij} = g_{ij}$ (lower case Latin indices running from 1 to 3), lapse function $N$ and shift vector $\beta^i$ (Misner et al., 1973); Ashtekar’s “new variables” for the modern canonical quantum gravity project (a densitized triad-connection version of an ADM split) (Ashtekar and Lewandowski, 2004; Jacobson and Romano, 1992); Christian Møller’s orthonormal tetrad formalism (Møller, 1964); the Einstein-‘Palatini’ metric-connection formalism; a metric-connection-Lagrange multiplier
formalism that explains why the metric-connection formalism works (Kichenassamy, 1986; Lindström, 1988; Ray, 1975); a tetrad-connection formalism that derives rather than postulates the vanishing of the connection’s torsion (Papapetrou and Stachel, 1978); and the Peres-Katanaev conformal metric density-scalar density formalism (Katanaev, 2005; Peres, 1963). This last set of variables, though rarely used and little known, turns out to be privileged for Anderson’s absolute objects project because it uses no irrelevant fields (in a fairly well defined sense) and uses only irreducible geometric objects; the failure to use such variables leads to bad performance in inspecting GTR for absolute objects (Pitts, 2006a), as appeared above. Roger Jones identified four formulations of Newtonian gravity; with so many choices of fields from which to choose even within geometrical approaches to Einstein’s equations, as well as (for example) both geometrical and “spin 2” (Lyre and Eynck, 2003; Pitts and Schieve, 2004) options, one sympathizes with his question “realism about what?” (Jones, 1991). In what should the scientific realist believe in order to be a realist about gravitation in light of current physics? Ernan McMullin’s claim that mechanics and theoretical physics generally are anomalously difficult for scientific realism (McMullin, 1991) has some basis.

Apart from some exceptions of perhaps little physical importance (such as solutions of an Ashtekar formulation with a degenerate metric, for example), the various sets of variables for GTR (broadly construed in the fashion of physicists) are empirically equivalent in the sense that all solutions of one set of equations are suitably related (not always one-to-one) with solutions in other sets of variables. Physicists are generally not tempted to regard the resulting theory formulations as distinct theories, partly because their criteria for physical reality are attuned to this mathematical interrelation. Each description comes with an adequate recipe
for distinguishing the physically meaningful from the descriptive fluff, and no further ontological questions are typically asked or answered. Physicists are also quite comfortable with a certain amount of vagueness or merely implicit specificity. For example, is a given energy condition, such as the weak energy condition (Wald, 1984), part of GTR or not? The answer to that question depends, at least, on whether ‘realistic’ matter fields satisfy the condition; but whether a certain kind of matter is realistic is malleable in light of both empirical factors (such as the apparent observation of dark energy in the late 1990s) and theoretical factors (such as recognition that seemingly tame matter fields or quantum fields violate an energy condition hitherto regarded as important) (Barcelo and Visser, 2002). GTR for physicists is in effect a cluster of theories sharing a hard core including Einstein’s equations, while partially overlapping in including or failing to include various additional claims with various degrees of importance, not unlike a Lakatosian research program (Lakatos, 1970, 1971). Perhaps Arthur Fine would commend to philosophers the physicists’ approach, which sounds something like his Natural Ontological Attitude that there is no distictively philosophical question about the real existence of entities employed in scientific theories, so neither realism nor anti-realism is an appropriate doctrine (Fine, 1986). Physicists typically assume some sort of mathematical equivalence as necessary and sufficient for two formulations to be the same theory (though strict equivalence is not always required). Lawrence Sklar discusses a strategy along these lines, which seems not unreasonable if we have no familiarity with the theory’s entities apart from theory itself (Sklar, 1985). If Fine’s call to abstain from metaphysical questions goes unheeded, then the variety in choices of fundamental variables suggests a variety of mutually incompatible ontologies and explanatory mechanisms. Does space-time really
carry a metric only? Does it have a set of four vector fields in terms of which the metric can be defined (and thus make those vectors “orthonormal” at the end of the day)? Does space-time carry a metric and an \textit{a priori} independent connection that happens to match “on-shell” (that is, using some or all of the Euler-Lagrange field equations) the torsion-free Levi-Civita connection determined by the metric? Or is the connection simply defined in terms of the metric, so that the modal force of its metric-compatibility and lack of torsion is logical necessity? Similar questions could be asked about electromagnetism, as Julian Schwinger’s least action principle, formulated in terms of the vector potential $A_\mu$ and \textit{a priori} independent field strength $F_{\mu\nu}$ (Arnowitt et al., 1962b; Schwinger, 1953, 1998), shows. Alan Musgrave, it should be noted, does not despair of answering Jones’s question regarding what realists should be realists about in gravity (Musgrave, 1992), but a full answer will require more detailed treatment than Musgrave gives.

If one does wish to ask the metaphysician’s question about what contemporary physical theories assert to exist, then some criterion for choosing among the many formulations of GTR is needed. On such matters, the Andersonian tradition, with some friendly amendments (Pitts, 2006a), is perhaps the best guide available, as suggested above. Anderson insisted on eliminating irrelevant fields. The friendly amendments insist on eliminating locally irrelevant fields and using only irreducible geometric objects. The resulting collection of fields is such that all geometric objects needed in the theory can be derived from the fundamental fields but not from any smaller set of fields. The outcome for GTR is that the fundamental variables are a conformal metric density (or its inverse) and a scalar density of arbitrary nonzero weight; the metric and connection have no independent existence, but are defined in terms of these two irreducible geometric
objects.

There are various relationships that might obtain between nearly empirically
equivalent theories. The following list is intended to be suggestive rather than
exhaustive, but the variety of options that have genuine physical examples is
already striking. A modal sort of near equivalence is this: theory $T_1$ has all
the models (or “worlds” for variety) of $T_2$, but $T_1$ has some additional models as
well. If expressed in terms of axioms, $T_1$ is logically weaker than $T_2$. An interesting
example would be to consider GTR with the possible further requirement of global
hyperbolicity (Wald, 1984). Consider GTR without the requirement of global
hyperbolicity as $T_1$ and GTR with global hyperbolicity as $T_2$. Clearly there is
no hope for disproving $T_1$ on empirical grounds if one is doing science in a $T_2$-
world. One might also consider ordinary quantum mechanics as $T_2$ and Bohmian
mechanics, which need not enforce the quantum equilibrium condition, as $T_1$. A
third example takes $T_2$ to be Newtonian gravity and $T_1$ to be Cartan’s variant
of it using a space-time with a curved connection (Misner et al., 1973; Norton,
2008); the difference between these theories strikes me as rather more significant
than it seems to Norton. Clearly much depends on how common and interesting
the models in $T_1$ but not in $T_2$ are. If most interesting $T_1$-worlds are also $T_2$-
worlds, then finding oneself in a $T_2$-world will not even probabilistically confirm
$T_2$ much over $T_1$. There might, however, be a sense in which $T_1$ would be noticeably
disconfirmed for a scientist in a $T_2$ world if $T_2$ worlds are only a small portion of
the worlds of $T_1$. Anthropic considerations might also be relevant if embodied
scientists could not exist in some worlds: embodied scientists will certainly not
discover that models incompatible with the existence of embodied scientists are
realized in nature. A second kind of modal near-equivalence could arise if each
theory has some models not in the other theory, along with some shared models. For example, GTR with global hyperbolicity and GTR with asymptotic flatness (such as obtains for localized sources (Wald, 1984)) share some models, while each theory has models that the other lacks. A third kind of near-equivalence arises if every model in $T_2$ is diffeomorphic to part, but perhaps not all, of a model of $T_1$. Some of the most obvious examples, such as might posit that the world began 5 minutes ago (Ellis, 1985; Russell, 1921, 1948) or even 6000 years ago (Clendinnen, 1989; Earman, 1995; Gosse, 1857; Musgrave, 1992; Norton, 2008; Turner, 2007), or that certain objects exist only intermittently (Dejnozka, 1995; Russell, 1948), while otherwise agreeing with conventional history, might seem contrived. In fact examples of this phenomenon need not be of the “ad hoc cut-and-paste variety” (Earman, 1996, p. 630), to use Earman’s phrase. One example is the spin 2 route to Einstein equations with the background metric taken seriously (Pitts and Schieve, 2004) as $T_2$, while geometrical GTR is $T_1$. A second example takes GTR in terms of a metric as $T_2$ and GTR in terms of the Ashtekar variables including a connection and a densitized spatial triad, which permit a degenerate metric (that is, with vanishing determinant) (Jacobson and Romano, 1992), as $T_1$. All of these kinds of empirical near-equivalence have the property that there is no experiment that can be performed in both theories and such that the results disagree, but they give different lists of physically possible worlds. While these sorts of examples merit philosophers’ attention, I will set them aside to focus on a less recondite phenomenon.
3.2 Approximate Empirical Equivalence in Electromagnetism: The Possibility of Massive Photons

There is a kind of empirical near-equivalence that is considerably weaker in some respects, involving differences in occurrent properties in similar events in similar models of the compared theories, and yet implying permanent rather than merely transient underdetermination. It has seemed at least a priori unlikely to some noted physicists that satisfactory physical theories would be isolated, rather than obtainable as limiting cases of a one- (or more) parameter family of theories characterized by various particle masses (Babak and Grishchuk, 2003; Bass and Schrödinger, 1955; Boulware and Deser, 1972). Perhaps more to the point, why should our theorizing about the physical world, given our finite empirical knowledge, single out just one out of a variety of viable theories with differing particle masses? Empirically the photon mass is constrained to be rather small by ordinary standards (Goldhaber and Nieto, 1971; Luo et al., 2003; Tu et al., 2005), less than $10^{-50}$ grams. This bound is close enough to 0 for most practical purposes. However, this mass, not being dimensionless, is not close to 0 in a mathematical sense. The question of the photon mass, arising in real scientific practice, counts as a natural example, rather than cultured or artificial (to use Norton’s classification, based on that for pearls (Norton, 2008)). The contest, however, is not best framed in terms of a pair of theories, because of the infinite possibilities for the value of the photon mass.

This question of the photon mass can be considered at both the classical and quantum levels, giving philosophically interesting test cases for approximate empirical equivalence. While the usual Maxwell electromagnetism has a massless photon (if one may follow the common practice of borrowing quantum terminol-
ogy for classical contexts), it is fairly well known that the massive Proca variants exist and approximate the massless theory for sufficiently small photon mass in both the classical (Jackson, 1975; Sundermeyer, 1982) and quantum contexts (see references in (Boulware and Deser, 1972) or various quantum field theory books, which might speak of a massive vector field or a neutral meson field). Because Maxwell’s electromagnetism is empirically distinguishable from any particular Proca theory (that is, with some given photon mass), and the various Proca theories with different photon masses are also empirically inequivalent, there is no possibility of trivializing the rivalry by regarding the supposed rivals as merely the same theory in different guises. However, for any set of observations with finite precision—which is the only kind that human finitude permits at a given stage of inquiry—there exists a range of sufficiently small photon masses such that the massive electromagnetic theories are empirically indistinguishable from the massless theory. Furthermore, the difference between the massless theory and the massive theories is quite deep conceptually, because only the massless theory has gauge freedom and thus has field equations that mathematically underdetermine the fields’ time evolution (assuming that the potential $A_\mu$ is used rather than the field strength $F_{\mu\nu}$), along the lines of the hole argument in GTR. (In comparison to GTR (Earman and Norton, 1987), the analog of the hole argument for electromagnetism is not a very difficult problem, because space-time point individuation is not at issue due to the purely internal nature of the gauge transformations and consequent ease of finding the gauge-invariant observable field strength $F_{\mu\nu}$.) By the same token, the massive Proca theories merely permit charge conservation (which typically holds as a consequence of the field equations for the charged sources), whereas Maxwell’s theory enforces charge conservation and so can be
coupled only to conserved sources. Thus the contest between Maxwell’s massless electromagnetism and Proca’s massive electromagnetisms provides a paradigm case of approximate empirical equivalence: a contest between (or should one say, among) genuine rivals, which cannot be wholly resolved empirically, and on which matters of considerable interest turn.

The most compact and perspicuous way to begin a technical discussion of a classical field theory is to exhibit its Lagrangian density, a function of some fields and their derivatives, such that the space-time integral of the Lagrangian density $\mathcal{L}$, the “action” $S$ of the theory, satisfies the principle of least (or perhaps merely stationary) action. The source-free Maxwell field equations (in manifestly Lorentz-covariant form) follow from a Lagrangian density of the form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad \text{(3.1)}$$

where the indices are moved using the Lorentz metric $\text{diag}(-1,1,1,1)$ and $F_{\mu\nu} = \text{def} \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength. For Maxwell’s theory, the vector potential $A_\mu$ admits the gauge transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \phi$$

for an arbitrary function $\phi$; this transformation makes no observable difference. This Lagrangian density is manifestly gauge invariant, because it is built from the gauge-invariant field strength only. For the massive Proca electromagnetisms, the Lagrangian density is

$$\mathcal{L}_p = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{m^2}{2} A_\mu A^\mu. \quad \text{(3.2)}$$

Evidently the $A^2$ term breaks the gauge symmetry in the massive case. Whereas
Maxwell’s theory has 2 degrees of freedom at each spatial point (written as $2\infty^3$ degrees of freedom), Proca’s theories have $3\infty^3$ degrees of freedom.\textsuperscript{1} The extra degree of freedom (at each point), however, is weakly coupled for small photon masses and so is not readily noticed experimentally. The treatment of the two theories (or theory types) using the Dirac-Bergmann constrained dynamics formalism is straightforward (Sundermeyer, 1982). The approximate empirical equivalence between Maxwell’s theory and Proca’s theories for small enough photon masses is preserved under quantization: massive quantum electrodynamics (QED) approximates the standard massless QED arbitrarily well (Boulware and Deser, 1972).

It follows that in a world with electromagnetism as the only force, it would be impossible for finite beings to rule out all of the massive electromagnetic theories empirically, and thus impossible to determine empirically whether gauge freedom was a fundamental feature of the electromagnetic laws. Here I am forgetting about the Stueckelberg formulation to be discussed shortly, which shows all the more strongly that gauge freedom \textit{per se} is not even distinctive of massless electromagnetism, unless one bans certain extra gauge compensation fields.

One obvious cause for concern for the empirical equivalence of Maxwell’s theory and the Proca theories pertains to thermal physics, in which various phenomena (such as the equipartition theorem (Huang, 1987) and the thermal history of the universe in Big Bang cosmology (Padmanabhan, 1993)) depend directly on the number of degrees of freedom. Thus one might expect factors of $\frac{3}{2}$ to distinguish predictions of a Proca theory from Maxwell’s theory. The resolution of this worry (Bass and Schrödinger, 1955; Goldhaber and Nieto, 1971) is to argue that the

\textsuperscript{1}The reader will observe the importance of distinguishing $2\infty^3$ from $3\infty^3$, notwithstanding rules for Cantorian transfinite arithmetic (Moore, 1990). The lesson seems to be that physical theories involve continuity properties of sets from which cardinality abstracts. Evidently cardinality does not exhaust the useful notions of “same size” or counting or infinite collections.
third degree of freedom is so weakly coupled (given the small photon mass) that the third degree of freedom (at each point in space) is out of equilibrium. Given how the coupling of the third (longitudinal) degree of freedom scales with the photon mass, the time to reach equilibrium is comparable to or larger than the age of the universe. Thus there is no reason to assume that the longitudinal degree of freedom is at equilibrium with the other electromagnetic degrees of freedom or with the particles, such as electrons, with which photons interact.

One relevant distinction between massive classical electromagnetism and massive quantum electrodynamics pertains to the tendency of classically fixed parameters to acquire quantum corrections.\(^2\) Classically one might take the photon mass to be an arbitrary parameter, handed down from above and not susceptible to explanation, but only to empirical determination. However, in quantum field theory, a small nonzero bare photon mass might acquire large corrections, whereas a vanishing photon mass is forced to stay vanishing by gauge invariance. Thus in massive quantum electrodynamics, the smallness of the photon mass seems to call for explanation, but no explanation (other than fine tuning) is available.

3.3 No Approximate Equivalence in Yang-Mills Theories?

Having pondered the electromagnetic case, one might be tempted to draw three philosophical morals, especially if one is unmoved with surprise by the smallness of the photon mass in massive quantum electrodynamics. One possible moral is the generic availability of rival theories that will remain empirically indistinguishable no matter how far empirical inquiry advances, despite the fact that the rival theories give contradictory answers for the same experiment, because theories with

---

\(^2\)I thank Ikaros Bigi for a discussion of this point.
slightly different particle masses (or perhaps other parameters) are available. A second possible moral is that theories that are nearly empirically equivalent classically remain so under quantization, so empirical equivalence is stable under the change of auxiliary hypotheses from classical to quantum. A third possible moral is that drawing theoretical conclusions about the presence of mathematical indeterminism and gauge symmetries is inadvisable, because contemporary physical theories offer a choice between Maxwell’s electromagnetism, a gauge theory with indeterministic equations for $A_\mu$, and Proca’s theories with deterministic equations and no gauge freedom. While it is inadvisable in general to invest heavily in metaphysical results that are fragile under small changes in physical theory, all three of these lessons are perhaps too hasty, because, among other reasons, it turns out that quantization can be dangerous to the health of massive theories. The preservation of approximate empirical equivalence between the massive and massless cases of electromagnetism under quantization is something of an accident due to the theories’ simplicity, as appears in consideration of the Yang-Mills and gravitational cases from work in the early 1970s. In particular, quantized massive Yang-Mills theory differs from the massless theory even in the limit of vanishing mass (Slavnov, 1972; Slavnov and Faddeev, 1971). Moreover, quantized massive Yang-Mills theory is either non-unitary or not power-counting renormalizable (Delbourgo et al., 1988; Ruegg and Ruiz-Altaba, 2004; Wong, 1971). The massive Yang-Mills theories envisioned are those with a traditional mass term of the form

$$-\frac{m^2}{2}A^i_\mu A^{i\mu},$$

3 I thank Ikaros Bigi for illuminating discussions and advice on these matters.
which breaks the gauge symmetry in an explicit fashion, not the now-standard Higgs mechanism for giving an effective mass and finite range to the weak nuclear force by spontaneous symmetry breaking (by which an interaction between the Yang-Mills vector bosons and the Higgs scalars, after a field redefinition to measure the scalars with respect to a true energy minimum, yields an effective mass term for the Yang-Mills bosons). Non-unitarity is disastrous because negative probabilities seem unintelligible. The lack of power-counting renormalizability seems less disastrous to some authors nowadays (Weinberg, 1995) than it once did: one settles for an effective rather than fundamental theory and thus admits that the theory at hand works only up to some definite energy range, after which further terms would be required. Under quantization, the massive Proca electromagnetic theories escape this painful dilemma merely because their mathematical simplicity as Abelian gauge theories (that is, with a gauge group in which the order of two transformations makes no difference to the result), apart from the mass term that breaks the gauge symmetry, excludes a troublesome term that appears in the non-Abelian Yang-Mills case (Ruegg and Ruiz-Altaba, 2004). Electromagnetism is thus too simple a theory to exhibit the dangers that quantization poses to the health of field theories; its atypical simplicity renders it an unreliable testbed for philosophical morals of the sort suggested above. This phenomenon involving Yang-Mills theories exemplifies or resembles Leplin’s notion of instability of empirical equivalence under change of auxiliary hypotheses (Leplin, 1997). Yang-Mills and massive Yang-Mills field theories are approximately empirically equivalent classically, but this equivalence appears to be violated at the quantum level. If there is an essence of Yang-Mills theories that can be exemplified in either classical or quantum form, then one can take this example as a literal
instance rather than mere analogy to Leplin’s phenomenon, which is cast in terms of logically conjoined theories and hypotheses.

3.4 Approximate Equivalence in Electroweak Theory: Yang-Mills Theory with Essentially Abelian Sector

The underdetermination between the quantized Maxwell theory and the lower-mass quantized Proca theories is permanent. It does not immediately follow that our best science leaves the photon mass unspecified apart from empirical bounds, however. Electromagnetism can be unified with an $SU(2)$ Yang-Mills field describing the weak nuclear force into the electroweak theory (see, for example, (Weinberg, 1996)). The resulting electroweak unification of course is not simply a logical conjunction of the electromagnetic and weak theories; the theories undergoing unification are modified in the process (c.f. (Kukla, 1998), chapter 4). Maxwell’s theory can participate in this unification; can Proca theories participate while preserving renormalizability and unitarity? Probably they can (Barrow and Burman, 1984; Calogeracos et al., 1981; Cornwall et al., 1974a,b; Ignatiev and Joshi, 1996; Ruegg and Ruiz-Altaba, 2004), though a complete demonstration seems still to be lacking.\textsuperscript{4} Thus evidently the underdetermination between Maxwell and Proca persists even in electroweak theory. There is some non-uniqueness in the photon mass term, partly due to the rotation by the weak mixing angle between the original fields in the $SU(2) \times U(1)$ group and the mass eigenstates after spontaneous symmetry breaking. Thus the physical photon is not simply the field corresponding to the original $U(1)$ group, contrary to naive expectations. There are also various empirically negligible but perhaps conceptually important effects that can arise

\textsuperscript{4}I thank Ikaros Bigi for helpful discussion on this issue. Bigi does not incline toward such a theory himself.
in such theories. Among these are charge dequantization—the charges of charged particles are no longer integral multiples of a smallest charge—and perhaps charge non-conservation. Crucial to the possibility of including a Proca-type mass term (as opposed to merely getting mass by spontaneous symmetry breaking) is the non-semi-simple nature of the gauge group $SU(2) \times U(1)$: this group has a subgroup $U(1)$ that is Abelian and that commutes with the whole of the larger group. Were the electroweak theory to be embedded in some larger semi-simple group such as $SU(5)$, then no Proca mass term could be included (Calogeracos et al., 1981). The dependence of this outcome and others recently discussed on involved physical calculations shows that these are not examples of theories for which empirical equivalence can be demonstrated by a brief argument in a philosophy of science paper—examples which have been a target of Norton’s critique (Norton, 2008).

3.5 Is There Approximate Equivalence for Gravity? General Relativity and Its Massive Variants

If Yang-Mills theories qualify the three supposed lessons mentioned above, so that they fail for mass terms for non-Abelian Yang-Mills theories but do hold for the Abelian sector of non-semi-simple Yang-Mills theories such as they electroweak theory, the still greater complication of GTR renders the three lessons even more contingent upon detailed physical calculation. There is a large and growing literature on massive gravities (Boulware and Deser, 1972; Deffayet et al., 2002; Freund et al., 1969; Karch et al., 2001; Ogievetsky and Polubarinov, 1965; Pitts and Schieve, 2007; Vainshtein, 1972; Visser, 1998; Zinoviev, 2007), much of it addressing whether they approach (local) empirical equivalence with Ein-
stein’s equations in the massless limit and are theoretically healthy, even at the
classical level. A widely held view since the early 1970s poses a dilemma (Boul-
ware and Deser, 1972; Tyutin and Fradkin, 1972) asserting that massive gravities
either have $5\infty^3$ degrees of freedom (spin 2) and do not agree with Einstein’s
equations in the massless limit due to the van Dam-Veltman-Zakharov disconti-
nuity (van Dam and Veltman, 1970, 1972; Zakharov, 1970), or they have $6\infty^3$
degrees of freedom (spin 2 and spin 0) and agree empirically with Einstein’s the-
ory in the massless limit (at least classically), but are theoretically unhealthy and
physically unstable because the spin 0 field has negative kinetic energy. Whereas
the Proca theory is the unique local linear massive variant of Maxwell’s elec-
tromagnetism, the most famous massive gravity with $6\infty^3$ degrees of freedom,
the Freund-Maheshwari-Schonberg (FMS) massive gravity (Boulware and Deser,
1972; Freund et al., 1969; Logunov, 1998), is just one member (albeit the best in
some respects) of a 2-parameter family of massive theories of gravity (Ogievetsky
and Polubarinov, 1965), all of which satisfy universal coupling (Pitts and Schieve,
2007). Whether massive gravities are viable even at the classical level remains a
matter of debate in the physics literature. Evidently intuitive expectations about
the ease of constructing approximately empirically equivalent theories to GTR are
threatened by devils in the details.

One possibility worthy of exploration is whether the methods of PT-symmetric
quantization can help. PT-symmetric quantization has exorcised the vicious
ghosts thought to inhabit some theories according to more traditional analyses
(Bender, 2007; Bender and Mannheim, 2008; Mostafazadeh, 2005), though the re-
sulting theories sometimes have surprising phenomenology. Might PT-symmetric
be helpful for massive gravity or for a non-unitary (Delbourgo et al., 1988) mas-
sive Yang-Mills theory? Massive gravities, being bimetric, are also susceptible to causality problems if the relationship between the two metrics’ null cones is not correct; sometimes it is not (Schmelzer, 2000), at least not without help (Pitts and Schieve, 2007). Though matter sees only the effective curved metric and gravity only barely sees the flat background metric due to the smallness of the graviton mass, these theories are only Lorentz-covariant (or covariant under the 15-parameter conformal group in the case of massless spin 0). Thus the usual arguments about superluminality in one frame implying backwards causation in another frame are applicable.

The provisional character of some conclusions in this chapter indicates that in these cases, whatever the ultimate outcomes, the physical details have important philosophical consequences. These examples therefore do not provide strong support for claims that underdetermination is generic. They do, however, provide strong support for the claim that there might well be interesting cases of permanent underdetermination, even in our contemporary best science, such as the electroweak quantum field theory.
CHAPTER 4

EXACT EMPIRICAL EQUIVALENCE AND ARTIFICIAL GAUGE FREEDOM

4.1 Overview

It is well known that many important physical theories have considerable descriptive redundancy. The gauge theories for the four fundamental forces—Maxwell’s electromagnetism, the Yang-Mills theories for the weak and strong nuclear forces, and GTR—display this descriptive redundancy. Though gauge freedom arguably has a deep philosophical significance of some kind, gauge freedom poses something of a challenge technically. In some contexts, especially Maxwell’s electromagnetism or its quantum successor, it can be useful to make a conventional choice of a specific gauge, make calculations in a convenient fashion, and then show that the results did not depend on that specific gauge choice (Weinberg, 1995, p. 345). Unfortunately this simple procedure tends to fail for Yang-Mills theories due to the Gribov ambiguity (Guay, 2008; Kaku, 1993; Sundermeyer, 1982); this failure can be relevant when nonperturbative effects matter. A more elegant, but more abstract and technically difficult procedure, is to work with a space of physically distinct configurations by taking gauge equivalence classes; the resulting reduced phase space (in a Hamiltonian formalism) or other reduction to the “true degrees of freedom” might be difficult to find or use explicitly, however.
These sorts of procedures take gauge freedom to be an obstacle to overcome. Thus Freund, Maheshwari and Schonberg discussed how the absence of gauge freedom in their massive variant of GTR made it easier to quantize than GTR (Freund et al., 1969). One can take the second-class constraints as identities for eliminating unnecessary field variables (Sundermeyer, 1982; Weinberg, 1995). In terms of the Dirac-Bergmann constrained dynamics (Marzban et al., 1989; Pitts, 2006b), massive GTR has only second-class constraints and thus makes the true degrees of freedom available almost immediately in terms of Dirac brackets generalizing Poisson brackets, at least in principle. However, for many theories there are technical challenges involved in working with only the true degrees of freedom (Banerjee et al., 1995; Henneaux and Teitelboim, 1992; Park and Park, 1998), such as that (i) Dirac brackets tend to be nonlocal and field-dependent, (ii) for quantization, finding a suitable factor ordering (recalling that $qp - pq \neq 0$ in quantum mechanics) is often difficult or perhaps impossible, and (iii) finding a complete set of variables to form canonical pairs is often difficult or perhaps impossible.

The pragmatic challenges in dealing with Dirac brackets and the progress in handling gauge theories have led to a change of viewpoint. More recently the view has appeared that gauge freedom is an asset and not so much a liability. Thus technologies have been developed to take theories with constraints but no gauge freedom (of which the Freund-Maheshwari-Schonberg and other massive gravities are examples (Freund et al., 1969; Ogievetsky and Polubarinov, 1965), as are the Proca massive electromagnetisms) and install gauge freedom artificially by adding extra fields and extra symmetries ensuring that the extra fields make no empirical difference. While the ad hoc Stueckelberg trick developed gradually in the middle of the 20th century (Ruegg and Ruiz-Altaba, 2004), and a paper from the 1970s left the
name “Wess-Zumino fields” for certain purposes (Neto, 2006; Wess and Zumino, 1971), the subject of installing artificial gauge freedom reached a mature form in the 1980s (Batalin and Fradkin, 1986, 1987; Batalin et al., 1989; Batalin and Tyutin, 1991; Faddeev and Shatashvili, 1986) and now bears the names of Batalin, Fradkin and Tyutin, or BFT for brevity. This sort of technique has been applied to both Proca’s massive electromagnetism, and for a nontrivial test case, massive Yang-Mills theories (notwithstanding the apparently defective character of most such theories) (Banerjee and Barcelos-Neto, 1997a,b; Grosse-Knetter, 1993; Kim et al., 1997; Kunimasa and Gotô, 1967; Ruegg and Ruiz-Altaba, 2004; Shizuya, 1975a,b; Slavnov, 1972). I anticipate that application of the BFT procedure to massive gravities with \(6\times3\) degree of freedom (Freund et al., 1969; Ogievetsky and Polubarinov, 1965) would reproduce the result of parametrization with clock fields (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 2000), apart from some issues to be discussed below. Such techniques are usually formulated in a Hamiltonian context, where the well-developed Dirac-Bergmann constrained dynamics formalism allows one to speak of “conversion” of second-class constraints into first-class constraints. In the Dirac-Bergmann constrained dynamics technology (Henneaux and Teitelboim, 1992; Sundermeyer, 1982), recently commended to philosophers by John Earman (Earman, 2002, 2003b), second-class constraints are those having nonzero Poisson brackets with some constraint(s), whereas first-class constraints have vanishing Poisson brackets with all constraints, perhaps making use of the constraints themselves.

There has also been a bit of work in a Lagrangian context (Kim et al., 1999; Park and Park, 1998). The Lagrangian formalism provides considerable help in guessing correct answers. Given the conceptual distinction between the context
of discovery and the context of justification (Reichenbach, 1938), one need not be ashamed of a guess-and-check procedure when systematic treatment is difficult. Here there are two advantages. First, the Lagrangian formalism introduces only half as many new fields as the Hamiltonian formalism: new fields but not new conjugate momenta. Having no new momenta can significantly reduce the number of candidate terms in the power series expansion. Second and perhaps even more importantly, the manifest Lorentz covariance of the Lagrangian formulation of relativistic theories still further reduces the number of candidate terms. In general I will limit attention to Hamiltonian formulations of theories, but I prefer Hamiltonians related by Legendre transformations to Lagrangians with the usual symmetries (gauge invariance, Lorentz covariance, etc.) manifest when possible. Below I will make use of a Lagrangian ‘lucky guess’ in installing artificial gauge freedom for FMS massive gravity. The availability of a Lagrangian constraint stabilization algorithm analogous to the Dirac-Bergmann Hamiltonian algorithm (Pons and Shepley, 1995) implies that at least some cases can be treated directly via Lagrangian means.

The typical BFT procedure goes along these lines: start with a Hamiltonian theory (formulation) with \( m \) second-class constraints and no first-class constraints. Introduce \( \frac{m^2}{2} \) pairs of new coordinates \( \theta \) that are canonically conjugate (in a slightly generalized sense). Take the original (second-class) constraints and Hamiltonian as the zeroth order terms in a series expansion involving new fields. Choose the coefficients in the series to make all constraints first-class and make them have vanishing Poisson brackets with the Hamiltonian (Batalin and Fradkin, 1986, 1987; Batalin et al., 1989; Batalin and Tyutin, 1991; Vytheeswaran, 1994, 2002). Given the series expansion, the original formulation is recovered in the limit that the
new fields vanish. One can count the degrees of freedom to show that the new formulation has the same number as the old. Given \( l \) configuration variables (and hence \( 2l \) phase space variables), \( f \) first-class constraints, and \( s \) second-class constraints in a theory formulation, there are \( \frac{2l-2f-s}{2} \) degrees of freedom (Henneaux and Teitelboim, 1992). For the original formulation, there are \( \frac{2l-2a-m}{2} = \frac{2l-m}{2} \) degrees of freedom. For the new formulation, there are \( \frac{(2l+m)-2m-0}{2} = \frac{2l-m}{2} \) degrees of freedom, the same as for the original. The new formulation with first-class constraints presumably allows gauge transformations to achieve \( \theta = 0 \), so the new formulation can be gauge-fixed into the old one by choosing \( \theta = 0 \) (unless some physical principle would be violated thereby, such as causality as in massive gravity (Pitts and Schieve, 2007)).

4.2 Modified Batalin-Fradkin-Tyutin Procedure Gives Stueckelberg Formulation

Directly

Given the starring role playing by Maxwell’s electromagnetism, Yang-Mills theory, and Einstein’s gravity in contemporary physics, and consequently the non-negligible importance of their massive cousins (whether those theories are physically viable or not), it seems reasonable to seek a procedure for converting non-gauge formulations into gauge formulations that is as convenient as possible for massive electromagnetism, massive Yang-Mills theory, and massive GTR. In the electromagnetic and gravitational cases, it turns out that the gauge formulations were found in an \textit{ad hoc} way without the need for such elaborate conversion algorithms as have appeared in recent decades. In the electromagnetic case, the answer is Stueckelberg’s trick, which adds the gradient of a new scalar field to the vector field \( A_\mu \) in the mass term; adding such a term in the kinetic term \(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\)
would make no difference because such a term, like a gauge transformation, has no effect on $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In the gravitational case, the answer is parametrization, whereby preferred coordinates are turned into new fields varied in the action principle. One hopes that parametrization reproduces the outcome from a BFT-type procedure, though a full nonlinear calculation seems never to have been done. (In fact one would expect the relationship to be somewhat indirect, for reasons to be discussed shortly.) For the Yang-Mills works cited above, it turns out that the answer is too complicated to guess easily, but the problem yielded to systematic treatment.

The usual BFT procedure makes no use of a “polarization” of the phase space into configuration and momentum variables, a fact that adds some generality but leaves the results a bit more complicated than is necessary for use in some paradigmatic cases, such as Proca or massive Yang-Mills theories. Whereas one often sees the BFT procedure yield the Stueckelberg formulation of massive electromagnetism only after a rather elaborate calculation involving a canonical transformation to implement a change of variables, path integration, and dropping a boundary term (Banerjee and Banerjee, 1996; Banerjee et al., 1995; Vytheeswaran, 1999), I will now discuss a modified BFT formalism that directly recovers the Stueckelberg version of massive electromagnetism. The modification involves making a distinction between coordinates and momenta in a fundamental way, for both the original field variables and the additional ones added to install gauge freedom. While some generality might be lost thereby, there is a formal simplification that might be important in applications, such as gravitation. It was a considerable achievement in 1958 when it was found how to trivialize the form of the primary constraints in canonical GTR (Anderson, 1958; Dirac, 1958) (and
unpublished work by Bryce DeWitt). The modification made herein to the BFT procedure respects this achievement by leaving the primary constraints alone. It also distinguishes among the new fields a set of new coordinates and new momenta. The result is more Lagrangian-friendly than is the usual BFT procedure, as well as pedagogically simpler. It is optimized for convenience in paradigm cases, but might suffer in generality, however.

From a Lagrangian point of view, one wants to recover the second-class formulation from the first class formulation by imposing the vanishing of new coordinates (such as the Stueckelberg scalar) and their time derivatives ("velocities"), rather than the vanishing of new fields and momenta. The mathematical details will make this discussion clearer. For the massive Proca electromagnetism, the Lagrangian density is

$$\mathcal{L}_p = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{m^2}{2} A_\mu A^\mu.$$  

Here as elsewhere I used the $-+++$ signature, letting Greek indices run from 0 to 3 and Latin indices from 1 to 3. The primary constraint (at each point) is

$$\pi^0 = \frac{\partial \mathcal{L}_p}{\partial A_{0,0}} = 0.$$  

Performing the Legendre transformation to get the canonical Hamiltonian density (not a redundant term in the constrained dynamics literature) gives

$$\mathcal{H}_{pc} = \pi^\sigma A_{\sigma,0} - \mathcal{L}_p = \frac{1}{2} (\pi^\sigma)^2 + \pi^\sigma A_{0,a} + \frac{1}{4} F_{ij} F_{ij} + \frac{m^2}{2} (A_i)^2 - \frac{m^2}{2} (A_0)^2.$$  

Using the canonical Hamiltonian $\int d^3 x \mathcal{H}_{pc}$ (or the primary Hamiltonian—it does not matter) to find the time evolution of $\pi^0$ and to enforce its continued vanishing
gives

\[ \{ \pi^0(y), \int d^3x \mathcal{H}_{pc}(x) \} = \pi^a \cdot a(y) + m^2 A_0(y) \approx 0. \]

Thus the secondary constraint is \( \pi^a \cdot a + m^2 A_0 \) everywhere. Demandig that the secondary constraint be preserved by the time evolution requires the use of the primary Hamiltonian \( H_{pp} = H_{pc} + \int d^3x v(x) \pi^0(x) \), where \( v \) is a new Lagrange multiplier. Preserving the secondary constraint (with the help of an arbitrary test function to smear the Dirac delta functions from the fields’ Poisson brackets as needed) gives

\[ \{ \pi^a \cdot a(y) + m^2 A_0(y), H_{pp} \} = -m^2 A_1 \cdot i + m^2 v \approx 0, \]

fixing \( v \) and leaving no arbitrariness in the evolution of the system. The Poisson brackets of the constraints among themselves are

\[ \{ \pi^0(x), \pi^0(y) \} = 0, \]

\[ \{ \pi^0(x), \pi^a \cdot a + m^2 A_0(y) \} = -m^2 \delta(x, y), \]

and

\[ \{ \pi^a \cdot a + m^2 A_0(x), \pi^1 \cdot i + m^2 A_0(y) \} = 0. \]

The vanishing Poisson brackets here vanish without the use of the constraints themselves. The matrix of Poisson brackets of constraints has non-vanishing determinant, so the theory is indeed second-class as advertised.

How does the constrained dynamics treatment of the Stueckelberg gauge for-
mulation of massive electromagnetism differ? The Lagrangian density is

\[ \mathcal{L}_s = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{m^2}{2} (A_\mu + \partial_\mu \phi) (A^\mu + \partial^\mu \phi). \]

As with clock fields, it does not matter if the Stueckelberg scalar \( \phi \) is varied or not. The reason is that the gauge transformation formula for the Stueckelberg field \( \phi \) is algebraic in the gauge parameters (as is also the case for massive Yang-Mills (Banerjee and Barcelos-Neto, 1997b) and for massive gravity (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 2000)). If \( \phi \) is not varied, then its erstwhile equation of motion will still follow from the Euler-Lagrange equation for \( A_\mu \), either using the generalized Bianchi identity or by simply taking the divergence. Turning to the Hamiltonian formalism, the primary constraint is \( \pi^0 = \frac{\partial \mathcal{L}_s}{\partial A_{0,0}} = 0 \), as before; the momenta for \( A_i \) are also unchanged. The new momentum for the new field \( \phi \) is

\[ P = \frac{\partial \mathcal{L}_s}{\partial \dot{\phi}_{i,0}} = m^2 A_0 + m^2 \phi_{i,0}, \]

so the vanishing of the new momentum and the vanishing of the new field’s time derivative are in general incompatible. Performing the Legendre transformation to get the canonical Hamiltonian density gives

\[ \mathcal{H}_{sc} = \frac{1}{2} (\pi^a)^2 + \pi^a A_{0,a} + \frac{1}{4} F_{ij} F_{ij} + \frac{m^2}{2} (A_i)^2 + \frac{P^2}{2m^2} - A_0 P + m^2 A_i \phi_{i,i} + \frac{m^2}{2} (\phi_{i,i})^2. \]

Preserving the primary constraint yields the secondary constraint

\[ \pi^a_{i,a} + P \approx 0. \]
Preserving the secondary constraint gives neither a fixed Lagrange multiplier \( v \) nor a tertiary constraint. The primary constraint, being unchanged, still Poisson-commutes with itself. The new secondary constraint also Poisson-commutes with itself. Most importantly, \( \{ \pi^0(x), \pi^a, P(y) \} = 0 \), so the Stueckelberg formulation is first class, as advertised. This is the same Poisson bracket algebra as in the Maxwell theory. One advantage of the Stueckelberg formulation of massive electromagnetism over the Proca formulation is the ease of taking the massless limit (Zinoviev, 2007); the gauge freedom remains, while the extra degree of freedom decouples as \( m \to 0 \).

Comparing the secondary constraints of the two formulations is illuminating. For the Proca formulation, the secondary constraint is \( \pi^a, m^2 A_0 \), whereas the secondary constraint for the Stueckelberg formulation is \( \pi^a + P \). If one pretends that the Stueckelberg formulation has been produced by the usual BFT procedure with the boundary condition of getting the old formulation when the new fields all vanish, then one should expect the new constraints to reduce to the old ones once the new fields \( \phi \) and \( P \) are set to 0. The primary constraint, being unchanged, satisfies this expectation, but the secondary constraint does not: \( \pi^a + P \) becomes \( \pi^a + 0 \neq \pi^a, m^2 A_0 \). This result is disappointing, if one hoped that the BFT procedure would reproduce conveniently the Stueckelberg formulation. On the other hand, the relation \( P = m^2 A_0 + m^2 \phi,0 \) already told us that the vanishing of the new momentum \( P \) generally is inconsistent with the vanishing of the new velocity \( \phi,0 \), as was noted above. Perhaps one could circumvent this problem with a canonical transformation, but why work so hard? Another response to this difficulty would be to replace the BFT boundary condition that new constraints reduce to the old ones when the new fields (here \( \phi \) and \( P \) vanish, with the
condition that the new constraints reduce to the old ones when *those new fields that appear in the Lagrangian density* vanish (here \( \phi \)) and the appropriate relation between the new momenta and the new velocities holds. The relation between the new momenta and the new velocities arises, from the Hamiltonian angle, from the equation of motion \( \phi_{,0} = \frac{\delta H}{\delta P} \): the functional derivative of the new Hamiltonian with respect to the new momenta is made to vanish. Thus I propose imposing \( \phi = 0 \) and \( \frac{\delta H}{\delta P} = 0 \), rather than \( \phi = 0 \) and \( P = 0 \), as the means to recover the Proca formulation by gauge-fixing the Stueckelberg formulation. Imposing my conditions makes the Stueckelberg secondary constraint and Hamiltonian reduce to the Proca secondary constraint and Hamiltonian, as desired. One can also consider whether a set of extra conditions suffices to fix the gauge by forming, along with the primary and secondary constraints, a matrix of Poisson brackets with non-vanishing determinant (Sundermeyer, 1982). It is easy to show that the conditions \( \phi = 0 \) and \( P = 0 \) fail to gauge-fix the Stueckelberg formulation because the resulting four conditions \( \pi^0, \pi^a_{,a} + P, \phi = 0 \) and \( P = 0 \) fail to form a matrix of Poisson brackets with non-vanishing determinant. By contrast, my conditions \( \phi = 0 \) and \( \frac{\delta H}{\delta P} = 0 \) do fix the gauge. In general my proposal involves making a fundamental distinction among the new fields between new coordinates and new momenta, unlike the usual BFT procedure.

One might not like the fact that the BFT procedure sometimes changes the primary constraints. In the Stueckelberg formulation (not resulting from the BFT procedure), it was clear above that all the old momenta are unchanged relative to the Proca formulation, and indeed relative to the massless Maxwell case. Given that primary constraint \( \pi^0 \) has vanishing Poisson bracket with itself, all the non-vanishing Poisson brackets that make the Proca formulation second-class in Hamil-
tonian form evidently involve the secondary constraint $\pi^{\alpha}_{\gamma \alpha} + m^2 A_0$. Clearly it is no fault of the primary constraints that the Proca theory has second-class constraints, then. So why modify the primary constraints, if one can avoid it? Leaving the primary constraints unchanged should also remove some of the arbitrariness inherent in the BFT procedure. If the primary constraints are left unchanged, then the change in the Hamiltonian entails a definite change in the secondary (and higher order, if any) constraints.

One can derive, honestly and simply (with no dropping of boundary terms, field redefinitions through canonical transformations, or functional integration), the Stueckelberg formulation starting with the Proca formulation. Retaining the old primary constraints and introducing one new field and corresponding new momentum, one makes an unknown change in the Hamiltonian $\Delta H$, and hence a corresponding change in the secondary constraints. Then one seeks the following: the dynamical preservation of the new secondary constraints by the new Hamiltonian, the first-class nature of all constraints in the new formulation, and the boundary condition that the vanishing of the new coordinate only (not the new momentum) and the vanishing of the derivative of the new Hamiltonian with respect to the new momentum gives the Proca Hamiltonian. The resulting Poisson bracket algebra is the same as in Maxwell’s theory. Making an inverse Legendre transformation from the gauged Hamiltonian formulation yields precisely the usual Stueckelberg Lagrangian. It would be of interest to apply the modified BFT procedure to Yang-Mills theories as well.
4.3 Gauge Freedom in Massive Gravity: An External Symmetry

Whereas the previous section considered internal symmetries, for which each point lives a life by itself, there are other symmetries in which gauge transformations link what goes on at different space-time points. The most famous of these are the “external” space-time diffeomorphisms and subgroups thereof such as rigid translations. Less famous but quite important are the transformations of supersymmetry, in which internal and external symmetries are combined. Also of interest are the gauge transformations of Einstein’s equations on a background, which are a symmetry distinct from diffeomorphisms but sharing some of their mathematical features (Grishchuk et al., 1984; Pitts and Schieve, 2004). While “internal” and “external” symmetries are mutually exclusive, evidently they are not exhaustive. Therefore I call the contrast class to (purely) internal symmetries, “non-internal symmetries.” The most general symmetry can have internal and non-internal parts.

It is of considerable interest to add gauge freedom artificially (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 2000) to massive theories of gravity (Freund et al., 1969; Ogievetsky and Polubarinov, 1965) in order to compare Einstein’s GTR, massive theories, and massive theories with gauge freedom. This comparison, the analog of Maxwell vs. Proca vs. Stueckelberg comparison, juxtaposes the paradigmatic generally covariant theory (GTR), a theory formulation in which general covariance is broken totally but as simply as possible (massive GTR), and a formulation in which general covariance is in many respects restored artificially (gauged massive GTR). If the massive theories prove theoretically healthy and empirically viable, then they also pose a case of underdetermination, as in electromagnetism. Whether these theories are physically viable,
however, is not crucial for present purposes.

The gauge freedom to be installed here is external, involving the crucial transport term

\[ \xi^\alpha \frac{\partial}{\partial x^\alpha} g_{\mu\nu} \]

or the like (Sundermeyer, 1982, p. 258), which has no analog for electromagnetic or Yang-Mills transformations, in which the derivative of the potential does not appear. This sort of term is responsible for issues of space-time point individuation in the hole argument. There have been various studies that introduce gauge compensation or so-called Stueckelberg fields for linear theories of spin 2 (and perhaps also spin 0) fields (Delbourgo and Salam, 1975; Duff et al., 2001; Hamamoto, 1979, 1996, 1997; Zinoviev, 2007). Current work on higher-spin (that is, higher than the spin 2 of gravity) fields often takes place in a “gauge invariant” formalism making use of gauge compensation fields (Buchbinder and Krykhtin, 2005; Buchbinder et al., 2007). However, it is not at all clear what form the generalized Stueckelberg trick should take for nonlinear theories involving spin 2. Indeed one might worry that it would be horribly complicated and involve derivatives to all orders (thus being nonlocal), given the exponential form of the finite Yang-Mills gauge transformations and of presence of derivatives of all orders in finite gauge transformations in gravitation (Grishchuk et al., 1984). Fortunately the answer is available, painlessly, in a simple form via parametrization (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 2000).

Making use of parametrization, this section will show in effect what obtains from the application of a BFT-type procedure (modified as above for Lagrangian-friendliness) to a nonlinear massive theory of gravity, the Freund-Maheshwari-Schonberg theory (Freund et al., 1969; Logunov, 1998). (I expect that there would
be equivalence between the modified Lagrangian-friendly procedure outlined above and parametrization.) This theory is in some respects privileged over the other massive gravities (Ogievetsky and Polubarinov, 1965), though not by virtue of the universal coupling derivation, which applies to all of the Ogievetsky-Polubarinov theories (Pitts and Schieve, 2007) (pace (Boulware and Deser, 1972)). Making contact with parametrization, along with some known results involving the Poisson bracket algebra for GTR (Sundermeyer, 1982), obviates difficult calculations. For present purposes the fact that the spin 2-spin 0 theories are related to Einstein’s in the same way that Proca’s is to Maxwell’s trumps any concern (noted above) that spin 2-spin 0 theories might not be healthy theories of gravity.

An additional philosophical reason to consider versions of gravity with a non-zero ‘graviton rest mass’ is the way that they exemplify the importance, as discussed recently by Harvey Brown (Brown, 2005b), of giving physical explanations not (or not simply) in terms of space-time structure, but in terms of detailed physical laws. Massive GTR (Freund et al., 1969; Ogievetsky and Polubarinov, 1965) (forgetting its problems with positive energy and causality (Pitts and Schieve, 2007) for the moment) has the Poincaré group for its symmetry group, and yet matter ‘sees’ only the effective curved metric $g_{\mu\nu}$, because the metric appearing in the matter action is the curved effective metric, not the flat background $\eta_{\mu\nu}$. For sufficiently small graviton rest masses, in practice the flat background metric is very difficult to observe even in gravitational experiments in these theories, though it is certainly present in the field equations. The fact that the symmetry group is the Poincaré group of STR might lead one to anticipate typical STR phenomenology, in which rods and clocks see the flat metric, but this anticipation is clearly seen to be false on examination of the matter action. The “space-time structure”
of these theories involves both a curved metric $g_{\mu\nu}$ and a flat metric $\eta_{\mu\nu}$, but this mere listing of geometric objects leaves one with almost no idea what to expect from experiments. Why, as Brown might ask, does matter see the curved metric and not the flat metric? Appeal to space-time structure here gives no explanation, but a glance at the matter action $S_{\text{matter}}[g_{\mu\nu}, u]$ (with matter denoted by $u$) gives a partial answer immediately: $\eta_{\mu\nu}$ is absent from the matter action and appears only in the gravitational action. (There remains the further question of why matter couples to $g_{\mu\nu}$ in such a fashion that material objects behave as rods and clocks for $g_{\mu\nu}$, rather than relating to $g_{\mu\nu}$ in some more complicated fashion (Brown, 2005b).\textsuperscript{1}) These same points apply to massive scalar gravity (Freund and Nambu, 1968; Pitts, 2009), which generalize Nordström’s theory (Deser and Halpern, 1970), with flat metric $\eta_{\mu\nu}$ and the merely conformally flat metric $g_{\mu\nu}$ conformally related to it. Matter sees only the conformally flat $g_{\mu\nu}$; gravity alone involves the flat metric $\eta_{\mu\nu}$, because $\sqrt{-\eta}$ appears in the mass term for gravity, but nowhere else. Massive scalar gravity turns out to be wrong empirically, predicting no bending of light, but that does not matter for present purposes; unlike massive GTR, massive scalar gravity is free of worries about positive energy or causality. The fact that explanations in terms of space-time structure appear to work so well for STR and GTR is due to the scarcity of geometric objects arguably pertaining to geometry (as opposed to matter) in the theory, yielding (almost) unique results for matter coupling and the behavior of the metric. But given the possibility of writing theories such as massive gravities, or even theories in which different kinds of matter ‘see’ radically different space-time structures, explaining physical phenomena in terms of mere space-time structure, without detailed recourse to the Lagrangian density, is typically an unsuccessful strategy.

\textsuperscript{1}I thank Katherine Brading for discussing this matter.
The treatment of the FMS theory in Hamiltonian form follows ([Boulware and Deser, 1972; Pitts and Schieve, 2007]). The Lagrangian density is (apart from divergences and other unimportant terms)

$$\mathcal{L} = \sqrt{-g} R(g) - m^2 \left( -\sqrt{-g} - \sqrt{-\eta} + \frac{1}{2} \sqrt{-g} g^{\mu\nu} \eta_{\mu\nu} \right).$$  \hspace{1cm} (4.1)

This Lagrangian density includes a formal cosmological constant term \(\sqrt{-g}\) and an unimportant constant term \(\sqrt{-\eta}\), but the term \(-\frac{m^2}{2} \sqrt{-g} g^{\mu\nu} \eta_{\mu\nu}\) breaks the gauge symmetry by introducing preferred coordinates implicitly. This term cancels the linear term in the gravitational potential (construed as something like the difference between the curved and flat metrics), which is responsible for the peculiar cosmological constant behavior with a potential growing with distance (Freund et al., 1969). The distinctive mass term phenomenology, with Yukawa exponential decay of the potentials, then arises from the quadratic part of \(\sqrt{-g}\).

Making the usual ADM (3+1)-dimensional split (Misner et al., 1973) of the curved metric \(g_{\mu\nu}\), one uses for dynamical variables the lapse function \(N\) relating the proper time to coordinate time, the shift vector \(\beta^i\) expressing how the spatial coordinate system moves over time, and a curved spatial metric \(h_{ij}\) with the inverse \(h_{ij}\) and determinant \(h\). Letting \(g^{\mu\nu}\) be the inverse curved metric as usual, one has \(g^{00} = -N^{-2}\), \(g_{ij} = h_{ij}\), and \(g_{0i} = h_{ij} \beta^j\). For temporary convenience I partly fix the coordinates to have \(\eta_{00} = -1\) and \(\eta_{0i} = 0\). The above Lagrangian density, after dropping a divergence, becomes

$$\mathcal{L} = N \sqrt{h} \left[ R + K_{ab} K^{ab} - K^2 + m^2 \left( 1 - \frac{h^{ij} \eta_{ij}}{2} \right) \right] + m^2 \left[ \sqrt{-\eta} + \frac{\sqrt{h}}{2N} \left( \eta_{ij} \beta^i \beta^j - 1 \right) \right].$$  \hspace{1cm} (4.2)
The canonical momenta, as in GTR, are

\[ \pi^{ij} = \frac{\partial L}{\partial h_{ij,0}} = \sqrt{h}(K^{ij} - h^{ij} K), \quad P_1 = \frac{\partial L}{\partial \beta^i_0} = 0, \quad P = \frac{\partial L}{\partial N,0} = 0. \]  

(4.3)

The four vanishing canonical momenta are primary constraints in constrained dynamics (Sundermeyer, 1982).

Performing the generalized Legendre transformation and using the primary constraints gives the canonical Hamiltonian density

\[ \mathcal{H} = N \left[ \mathcal{H}_0 + m^2 \sqrt{h} \left( \frac{1}{2} h^{ij} \eta_{ij} - 1 \right) \right] + \beta^i \mathcal{H}_i - m^2 \sqrt{-\eta} + \frac{m^2 \sqrt{h}}{2N} (1 - \eta_{ij} \beta^i \beta^j), \]  

(4.4)

where, as usual,

\[ \mathcal{H}_0 = \frac{1}{\sqrt{h}} \left( \pi^{ij} \pi_{ij} - \frac{1}{2} \pi^2 \right) - \sqrt{h} R, \quad \mathcal{H}_i = -2D_j \pi^j_i, \]

and \( D_j \) is the three-dimensional torsion-free covariant derivative compatible with \( h_{ij} \). Setting \( m = 0 \) recovers the usual GTR form that is purely a sum of constraints, but \( m \neq 0 \) destroys that form and leads to six, not two, degrees of freedom at each point in space. The zeroth-order term \(-m^2 \sqrt{-h} \eta\) has been retained to give Minkowski space-time zero energy. The secondary constraints are obtained in effect by varying the lapse \( N \) and shift vector \( \beta^i \), yielding the modified Hamiltonian constraint

\[ \frac{\partial \mathcal{H}}{\partial N} = \mathcal{H}_0 + m^2 \sqrt{h} \left( -1 + \frac{1}{2} h^{ij} \eta_{ij} \right) - \frac{m^2 \sqrt{h}}{2N^2} (1 - \eta_{ij} \beta^i \beta^j) = 0 \]  

(4.5)
and the modified momentum constraint

\[
\frac{\partial \mathcal{H}}{\partial \beta^i} = \mathcal{H}_i - \frac{m^2 \sqrt{\eta}}{N} \eta_{ij} \beta^j = 0.
\]  

(4.6)

These constraints are second-class (Pitts, 2006b). Thus this theory is a gravitational analog of Proca’s massive electromagnetism.

With the BFT procedure as amended above regarding the primary constraints and the boundary conditions on the gauge compensation fields, one should be able to install artificial gauge freedom in the FMS theory in much the same fashion as in Proca’s electromagnetism, apart from computational difficulty. The task is easy, however, if one makes use of parametrization. One need only make the transformation \( \eta_{\mu\nu} \rightarrow X^A{}_{;\mu} \eta_{AB} X^B{}_{;\nu} \) in a single piece of the mass term, \(-\frac{m^2}{2} \sqrt{-g} g^{\mu\nu} \eta_{\mu\nu}\), which breaks the gauge symmetry, and the work is nearly finished. By leaving \( \sqrt{-\eta} \) alone, one obtains a Lagrangian formally the same as GTR with four minimally coupled massless scalar fields and a cosmological constant. This Lagrangian density is manifestly coordinate-invariant and has no fields unvaried in the action principle, so GTR-type generalized Bianchi identities relating the Euler-Lagrange equations follow. When the clock fields are expanded in a background piece and perturbing gauge compensation field, the result at lowest order resembles the Stueckelberg trick. The fact that the clock fields look formally like matter fields in the Lagrangian density is a reminder of the flexibility in deciding which fields pertain to gravity and space-time and which do not.

While the success of the installation of gauge freedom is already now evident at the Lagrangian level, it can be displayed perspicuously using the Hamiltonian also. Applying the Dirac-Bergmann constrained dynamics procedure to the result, one readily sees that all constraints are now first-class. The cosmological constant
term does not affect the Poisson bracket algebra of constraints; neither do the
minimally coupled scalar fields (Sundermeyer, 1982, p. 253). That one of the
scalar fields has negative energy is relevant to the viability of the theory, but not
to the Poisson bracket algebra. Contrary to the usual BFT procedure, I require
the primary constraints \( P \) and \( P_i \) and the other old momenta to suffer no change
while gauge freedom is installed. The new gauge compensation fields, which are
clock fields \( X^A \), have canonical momenta \( \pi_A = -m^2 \sqrt{-g} g^{0\mu} X^B_{,\mu} \eta_{BA} \). Gauge
freedom is installed by adding a term to the Hamiltonian density,

\[
\Delta \mathcal{H} = N \mathcal{H}_{0s} + \beta^i \mathcal{H}_{is} - \frac{m^2}{2} \sqrt{-g} g^{\mu \nu} \eta_{\mu \nu} , \quad (4.7)
\]

where

\[
\mathcal{H}_{0s} = \pi_A \eta^{AB} \pi_B \frac{2m^2}{\sqrt{\hbar}} + \frac{m^2}{2} X^A_{,i} \eta_{AB} X^B_{,j} h^{ij} \sqrt{\hbar}
\]

and

\[
\mathcal{H}_{is} = X^A_{,i} \pi_A.
\]

The altered Hamiltonian with the unchanged primary constraints gives the new
secondary constraints. They have the same Poisson bracket algebra as vacuum
GTR, one sees without calculation, because neither the cosmological constant,
nor the minimally coupled scalar fields, nor the constant term \( \sqrt{-\eta} \) changes the
algebra.

Studying the FMS massive theory of gravity and its parametrized variant
has illuminated questions involving empirical equivalence in much the way that
studying the Proca and Stueckelberg massive electromagnetisms illuminated the
question of empirical equivalence for electromagnetism. This parallel reflects and
illustrates and deep technical and conceptual similarities between Maxwell’s elec-
tromagnetism and Yang-Mills theories on the one hand, and Einstein’s GTR on the other. The FMS theory is distinguished among the 2-parameter family of massive gravities (Ogievetsky and Polubarinov, 1965) by its containing the flat metric $\eta_{\mu\nu}$ to the first power only, not higher powers or the inverse metric $\eta^{\mu\nu}$ or the determinant $\sqrt{-\eta}$ in the crucial symmetry-breaking term. It is this fact that, after parametrization, yields the Lagrangian that is formally just four massless scalar fields (one with the “wrong” sign) with the standard kinetic term. Parametrization of any of the other OP massive gravities will result in higher powers of $\partial X$. Whether the resulting exotic kinetic terms fall within the scope of minimally coupled matter (as understood at the time of (Sundermeyer, 1982, p. 253)) is unclear. Moreover, given the many ways that the lapse $N$ and shift $\beta^i$ and the field derivatives $\partial X$ enter the Lagrangian densities for most OP theories, finding the Hamiltonian could be challenging. Thus ascertaining by direct calculation whether the Poisson bracket algebra has changed could be difficult. Fortunately it is clear on Lagrangian grounds that the replacement $\eta_{\mu\nu} \rightarrow X^A,_{\mu} \eta_{AB} X^B,_{\nu}$ has the desired effect in these theories also. The manifest coordinate covariance of the Lagrangian density, a weight 1 scalar density with no nonvariational fields such as $\eta_{\mu\nu}$ in the Euler-Lagrange equations, indicates that there will be generalized Bianchi identities (Sundermeyer, 1982) much as in GTR. One also knows how to count gauge invariances and degrees of freedom directly from the Lagrangian (Henneaux et al., 1990; Pons and Shepley, 1995); doing so gives the expected results.

Studying the FMS massive theory of gravity and its parametrized variant has shed light on general covariance. Massive variants of GTR are interesting because their Lagrangian densities are, except for one term key term, those of GTR. Thus
one can isolate most or all the phenomena of substantive general covariance or its violation; the absence or presence of a mass term implies the presence or absence of substantive general covariance, respectively (assuming that GTR exemplifies it, an assumption that is not quite so obvious after the Geroch-Giulini $\sqrt{-g}$ counterexample to the Anderson-Friedman absolute objects program). Installing artificial gauge freedom by parametrization at least formally restores many features of GTR that one might associate with substantive general covariance: gauge freedom, first-class constraints, a Hamiltonian that is a sum of constraints (perhaps apart from a constant in this case), point individuation questions such as appear in the hole argument, the absence of non-variational fields in the Lagrangian density (apart from a constant term), etc. However, parametrized GTR’s gauge freedom is clearly artificial. Apart perhaps from causality worries that appear to problematize gauge-fixing (Pitts and Schieve, 2007) to recover the original FMS formulation, it would be exceedingly natural to regard the parametrized and non-gauge formulations as the same theory, much as one might identify the Proca and Stueckelberg formulations of massive electromagnetism. That parametrization of a non-gauge theory such as FMS massive gravity can mimic so many features of GTR that one associates with substantive general covariance tends to confirm that the presence or absence of clock fields is perhaps the best criterion for deciding whether a formally generally covariant theory is merely formally generally covariant, or is substantively so.

4.4 Generalized BFT Conversion, Unconstrained Theories and Parametrization

For philosophical and physical purposes, there might be considerable interest in various generalizations of the BFT-type procedure. Why must one be interested
only in converting a theory formulation with only second-class constraints into a theory with that same number of first-class constraints and no second-class constraints? There are physically interesting cases of theories naturally formulated with first- as well as second-class constraints (Kim et al., 1999; Monemzadeh and Shirzad, 2005; Park and Park, 1998); clearly a generalization of the BFT conversion algorithm is required to accommodate such cases. This generalization, however, is hardly ground-breaking conceptually. There might also be reasons, however, to perform only a partial conversion, leaving some second-class constraints and some first-class constraints. Such a procedure might be worthwhile in cases, perhaps such as massive gravity, where installation of gauge-freedom is made to satisfy an important physical principle such as null-cone causality, but one wants to stay as close to the true degrees of freedom as possible. If some 1960s writers on massive gravity (Freund et al., 1969; Ogievetsky and Polubarinov, 1965) had noticed that their theories were acausal (Pitts and Schieve, 2007; Schmelzer, 2000), perhaps they, writing before turning second-class constraints into first-class constraints came to seem like a good idea in general, would have sought such a procedure. There are also reasons for studying unconstrained theories parametrized with clock fields, such as a massive scalar field satisfying the Klein-Gordon equation in Minkowski space-time (Kuchař, 1973). The BFT conversion algorithm takes for granted that there are second-class constraints that one wants to turn into first-class constraints, as with massive gravity (Arkani-Hamed et al., 2003; Pitts and Schieve, 2007; Schmelzer, 2000); the BFT procedure seems simply inapplicable to unconstrained theories. One can imagine even more general possibilities, such as leaving some second-class constraints untouched while adding gauge freedom elsewhere in a theory, such as by parametrizing the Proca
theory. Kuchař and C. L. Stone have taken Maxwell’s electromagnetism, a theory with first-class constraints, and parametrized it (Kuchař and Stone, 1987). The general idea of adding artificial gauge freedom in fact is not tied essentially to the presence of second-class constraints, or of any constraints at all, in the initial theory formulation.

A sufficiently general sort of procedure that adds gauge freedom seems to be the following. Take a theory formulated in terms of a Hamiltonian (perhaps starting from a Lagrangian); there need not be any constraints of any kind. Add a new term (chosen by an educated guess or with sufficient generality) to the Hamiltonian and perhaps some new constraints (perhaps obtained by starting with the Lagrangian). Run the Dirac-Bergmann stabilization algorithm to find all the constraints. Gauge freedom has indeed been added if there are at least some first-class constraint(s) that have no first-class ancestors in the original formulation (though there might be second-class ancestors), while retaining (perhaps under modification) whatever first-class constraints were present in the initial formulation as well. (The second clause is to ensure that one does not lose any gauge symmetries in the process, because that seems to violate the spirit of adding gauge freedom.) Confirm that the new formulation has the same number of degrees of freedom as the original. Verify that the new formulation can be gauge-fixed into the original one when the new fields reduce to some trivial configuration. The successful implementation of this sort of procedure installs gauge freedom even in theories that perhaps had no constraints initially, such as if one parametrizes the theory of a massive or massless scalar field. Parametrizing Proca’s electromagnetism, which starts with second-class constraints in the electromagnetic sector and then acquires first-class constraints pertaining to space-time, should also fit
within this framework. In this fashion one can subsume both the parametrization process and the BFT conversion algorithm into a more general procedure for adding gauge freedom. Parametrization adds external gauge symmetry to theories that might or might not have any constraints. The BFT procedure adds gauge freedom, whether internal or external, to theories with second-class constraints. The more general procedure just outlined adds gauge freedom, whether internal or external or both, to theories that might or might not have any constraints. Both the BFT procedure and parametrization are applicable in the case of massive gravity (Freund et al., 1969; Ogievetsky and Polubarinov, 1965), which can be viewed as having a geometrical symmetry broken by the mass term. The demonstration above that parametrization of FMS massive gravity (Freund et al., 1969) yielded just the sort of result that one expects from BFT (modified concerning the boundary conditions and the distinction between coordinates and momenta) confirms the utility of subsuming both the BFT and parametrization procedures into a more general field-theoretic technique.

4.5 Artificial Gauge Freedom and Recent Views on Underdetermination: Glymour and Quine

Relatively recent (post-positivist) discussions of underdetermination and empirical equivalence are not entirely unaware of the possibility of adding descriptive redundancy to theories. According to Sklar,

[t]ypically one can generate alternative theories saving the phenomena by some process which introduces into a theory otiose elements whose place in the theory “cancels out.” Most interesting, of course, are the historical cases where the theory with the otiose elements came first and where it was an important scientific discovery that one could eliminate them by a conceptual revision. (Sklar, 1985, p. 62)
As Sklar’s comment suggests, there is no expectation that introducing such “otiose” elements might be scientifically productive. Van Fraassen’s versions of Newtonian mechanics with different velocities for the center of mass of the solar system (re)introduce a quantity that is unobservable due to the relativity of motion in the theory (Laudan and Leplin, 1991; van Fraassen, 1980). While one might regard the Stueckelberg, BFT and parametrization technologies as introducing a generalized relativity of motion—that is, of the evolution of fields—these philosophical discussions do not recognize the technical sophistication that such introductions might involve, and indeed already had involved, in the physics literature. Recent discussions of empirical equivalence often imply one or more claims which now seem doubtful in certain important physical cases. First, the formulation with the extra entities is often taken to be a different theory from the original leaner formulation. Second, the formulation with the extra entities is thought to have no practical advantage over the original leaner formulation. Third, the formulation with the extra entities is thought to have no conceptual advantage over the original leaner formulation. After commenting on the presence of some or all of these themes in works by Glymour and Quine, I will recall how each of these claims is at variance with the beliefs and practices of contemporary physicists working on certain problems.

Quine mentions briefly a possibility that might accommodate the transition from Proca theories to Stueckelberg theories:

...suppose we had an adequate theory of nature, and then we were to add to it some gratuitous further sentences that had no effect on its empirical content. By ringing changes on these excrescences we might get alternative theories, logically incompatible, yet always empirically equivalent. This gratuitous branching of theories would be of no interest to the thesis of under-determination, since the adequate original theory was itself logically compatible with each one of these
taking some Proca theory for the moment as the “adequate theory of nature”—obviously not a realistic claim, but it does even quantize well—can one take the Stueckelberg formulation merely to add “some gratuitous further sentences”? If one makes the field redefinition
\[ \tilde{A}_\mu = \text{def} \ A_\mu + \partial_\mu \phi, \]

one can rewrite the Stueckelberg Lagrangian density
\[ \mathcal{L}_s = -\frac{1}{4} F_{\mu \nu} F^{\mu \nu} - \frac{m^2}{2} (A_\mu + \partial_\mu \phi) (A^\mu + \partial^\mu \phi). \]

in terms of \( \tilde{A}_\mu \) and \( \phi \), but \( \phi \) no longer appears in the Lagrangian. While field redefinitions with derivatives are not to be trusted without question (c.f. the proof for algebraic redefinitions (Schouten, 1954)), this one seems harmless. The Euler-Lagrange equations for \( \tilde{A}_\mu \) make no reference to \( \phi \) (because the expression \( A_\mu + \partial_\mu \phi \) in the mass term yields the same field strength \( F_{\mu \nu} \) as results from \( A_\mu \)) and are identical to those of the Proca theory (apart from the typography of \( \tilde{A}_\mu \)). This field redefinition can be taken as a reconstrual of predicates that Quine allows in testing for theoretical equivalence. The new field equation for \( \phi \) is vacuous, so one can prescribe \( \phi \) freely: the extra field \( \phi \) appears solely in “some gratuitous further sentences” with no empirical content. Thus the application of Stueckelberg’s trick to a Proca formulation fits within Quine’s discussion.\(^2\) While

\(^2\)An analogous procedure appears to work for Yang-Mills theories (Ruegg and Ruiz-Altaba, 2004), though that is not always obvious (Banerjee and Barcelos-Neto, 1997b). One should distinguish between an Abelian-like conversion \( A^a_\mu \to A^a_\mu + \partial_\mu \phi^a \) of the Yang-Mills mass term used by some authors (Salam, 1962; Umezawa and Kamefuchi, 1961) on the one hand, and
for Quine the Stueckelberg formulation is logically compatible with the Proca formulation, the Stueckelberg formulation (even after the field redefinition that reconstrues the predicates) is a different theory from Proca’s: the Stueckelberg theory posits the existence and certain behavior of the field $\phi$, whereas the Proca theory does not. The connotations of “gratuitous further sentences” and “ringing changes on these excrescences” indicate that Quine would have anticipated nothing practically useful or conceptually insightful to result from the Stueckelberg trick.\(^3\)

Clark Glymour’s discussion of the underdetermination of geometry (or otherwise) can be substantially carried over to the case of artificial gauge freedom. Glymour, responding to Reichenbach, invites the reader to

> [s]uppose you find yourself teaching high school physics, Newtonian mechanics in fact. Suppose further than a bright and articulate student named Hans one day announces that he has an alternative theory which is absolutely as good as Newtonian theory, and there is no reason to prefer Newton’s theory of his. According to his theory, there are two distinct quantities, gorce and morce; the sum of gorce and morce acts exactly as Newtonian force does. Thus the sum of gorce and morce acting on a body is equal to the mass of the body times its acceleration, and so on. Hans demands to know why there is not quite as much reason to believe his theory as to believe Newton’s. What do you answer?

I should tell him something like this. His theory is merely an extension of Newton’s. If he admits that an algebraic combination of quantities is

\[
A_\mu^a \rightarrow A_\mu^a + D_\mu^a \phi^a
\]

resembling a non-Abelian gauge transformation on the other (Grosse-Knetter, 1993). The latter has been criticized for not actually converting all second-class constraints into first-class constraints (Banerjee et al., 1995), contrary to claims made about it. Given the consequences for counting degrees of freedom and given the interrelation between Hamiltonian and Lagrangian symmetries (Henneaux et al., 1990), this controversy is surprising, but I have not attempted to resolve it. For parametrized massive GTR, the relevant field redefinition is $g^{AB} = \text{def} g^{\mu\nu} X^A_{\mu} X^B_{\nu}$ and the like; using the preferred coordinates $X^A$ to define the action integral makes the arbitrary label coordinates $x^\mu$ disappear from the action.

\(^3\)Quine’s relativism about ontology and general pro-science naturalistic form of philosophy (Quine, 1969) are more helpful than the remarks quoted above, as viewed from particle physics. Confronted with the discussion to follow, he might conclude that at the quantum level, it is Stueckelberg rather than Proca that gives an adequate theory, in which case my criticisms become largely inapplicable. I thank Don Howard for discussion.
a quantity, then his theory is committed to the existence of a quantity, the sum of gorce and morce, which has all of the features of Newtonian force, and for which there is exactly the evidence there is for Newtonian forces. But in addition his theory claims this this quantity is the sum of two distinct quantities, gorce and morce. However, there is no evidence at all for this additional hypothesis, and Newton’s theory is therefore to be preferred.

The gorce plus morce theory is obtained by replacing “force” wherever it occurs in Newtonian hypotheses by “gorce plus morce”, and by further claiming that gorce and morce are distinct quantities neither of which is always zero. The computations which give values for force will not give values for gorce or for morce, but only for the sum of gorce and morce. Indeed, in general if we have a set of simultaneous equations such that using these equations, values for some of the variables in the equations may be determined from values of other variables, then values for the new variables will not be determined. Implicit in the discussion is a certain articulation of the principle that we prefer a theory with fewer untested hypotheses to one with more untested hypotheses. (Glymour, 1977).

Whether this analogy accurately captures the main issues involved in the question of the conventionality of geometry may be doubted (Pitts and Schieve, 2004), but that need not concern us now. It fits well enough the transition from a Proca formulation to a Stueckelberg formulation of massive electromagnetism (apart from the assumption of a merely algebraic relation among the fields), and mutatis mutandis the general process of installing artificial gauge freedom. It is evident that Glymour considers the gorce plus morce theory distinct from Newton’s theory, practically in no way advantageous to Newton’s theory, and conceptually inferior to Newton’s theory. By parity of reasoning, one would expect a similar verdict from him on the Stueckelberg formulation of massive electromagnetism and on the results of BFT conversion or parametrization.

Quine and Glymour (and they seem not to be unusual among philosophers) take the formulations with additional unobservable entities not only to be differ-
ent theories from the leaner originals, but also to be practically and conceptually inferior to the leaner originals. By contrast, it is overwhelmingly taken for granted in the physics literature that the Stueckelberg formulation is a formulation of the very same theory as the Proca formulation. More generally, it is overwhelmingly taken for granted that formulations obtained using the BFT algorithm or parametrization are formulations of the same theory after the surgery as before. With the possible exception of A. A. Logunov and his school, no physicist (to my knowledge) regards gauge-fixing as producing a numerically distinct theory. When gauge freedom has been installed artificially, as in the Stueckelberg, BFT and parametrization cases, one can eliminate the extra fields, the “otiose elements” or “excrescences,” by gauge-fixing. Whereas physicists agree that the result is the very same theory as before gauge-fixing, notable philosophers of science are committed to the view that a different and superior theory is produced by gauge-fixing. Whereas philosophers of science would expect the formulations with artificial gauge freedom to be practically inferior to their leaner ancestors, physicists find artificial gauge freedom useful in taking the limit of vanishing photon or other particle mass (Zinoviev, 2007), often use artificial gauge freedom in studying higher spin fields, avoid the technical challenges sometimes present with using the true degrees of freedom by installing gauge freedom artificially (Banerjee et al., 1995; Henneaux and Teitelboim, 1992; Park and Park, 1998), and prove desirable properties of a physical theory in different gauges, such as unitarity in one gauge and renormalizability in another (Weinberg, 1996, chapter 21)(Kaku, 1993, chapter 10). On the other hand, if one individuated theories as strictly as do Quine and Glymour, one would have (at least) a theory that had gauge freedom but was not obviously unitary or renormalizable, an empirically equivalent but dis-
tinct theory that had no gauge freedom and that was unitary but not obviously renormalizable, and a third empirically equivalent but distinct theory that lacked gauge freedom and that was renormalizable but not obviously unitary. If one can regard the tetrad-spinor formalism as obtained from the OP spinor formalism by the installation of artificial gauge freedom (Gates et al., 1983, p. 234) (perhaps in a Lakatosian rational reconstruction of the history of spinors in curved space-time), then the linearity of the tetrad-spinor formalism is another practical benefit from artificial gauge freedom. Parametrizing special relativistic theories has been fruitful conceptually in revealing similarities between GTR and parametrized theories, as the discussion of the suggestion that GTR is “already parametrized” demonstrated. Parametrizing massive GTR also has the advantage—it is unclear to what degree it is practical vs. conceptual—of taking acausal theories and giving them some hope of causality (Pitts and Schieve, 2007). While this last application does not strictly contradict Quine or Glymour, who assumed that one had an adequate theory at the beginning, it does exemplify the importance of artificial gauge freedom, contrary to what they presumably would have expected.

How is it that philosophers of science disagree with physicists on the value of artificial gauge freedom? One explanation is that the relevant physics has become prominent only rather recently, perhaps too recently to have influenced the philosophical works discussed. A deeper reason, however, is that physicists, often tacitly, ascribe to physical theories an ontology that differs from the prima facie ontology indicated by the fields in the Lagrangian density. To be specific, physicists take for granted, and sometimes assert, that the real is the invariant, where the relevant symmetry group(oid) is picked out (primarily) by the symmetries of the action. That only gauge- and coordinate-invariant entities are considered real,

---

4Don Howard has called attention to this theme (Howard, 1998).
and thus part of a theory’s ontology, is a limited application of the principle of the identity of indiscernibles. Though that principle is rarely invoked by name among physicists, the idea is ubiquitous. Thus the real ontology of the Stueckelberg formulation does not include the extra scalar field $\phi$, because $\phi$ is “mere gauge” (indeed gauge-equivalent to 0) and hence unreal or “unphysical.” While the identity of indiscernibles is far from obscure among philosophers, recent philosophers of science who are inclined toward scientific realism have avoided invoking the principle in some contexts where physicists almost unanimously invoke it. One might take the modern reemergence of the hole argument in GTR (Earman and Norton, 1987) as evidence of this super-realism among philosophers of science, though doubtless modern mathematicians’ fondness for active diffeomorphisms (which seem to require mathematical haecceities) has also contributed. To a large degree, modern physical theories dictate their own ontologies, indicating clearly by the presence of gauge freedom—especially artificial gauge freedom, which typically can be removed without spoiling manifest locality—when the naive ontology is not correct. The venerable tradition of trying to express scientific theories in terms of logic and set theory, as is evident in Quine (Quine, 1975) among others, is rather remote from the language of contemporary physics (Ladyman et al., 2007). Thus Quine regards as equivalent two formulations that are logically equivalent after reconstrual of predicates, but a formulation that adds superfluous entities with arbitrary values produces, in his view, a different and inferior theory. While explicitly eliminating superfluous fields is important for some tasks in the philosophy of physics, such as the Anderson-Friedman absolute objects program discussed in chapter 2, to some degree Lagrangian field theory takes care of itself ontologically with the help of the principle that the real is the invariant.
CHAPTER 5

CONCLUSION

5.1 Insights for the Philosophy of Physics

Consideration of a variety of theories (and formulations thereof) of contemporary physical relevance has yielded a variety of insights pertaining to both the philosophy of physics and to general philosophy of science. Concerning the philosophy of physics, a detailed exploration of general covariance, especially involving the Anderson-Friedman absolute objects program, has disposed of some old difficulties while occasioning the discovery of a new and more challenging problem. Concerning the general philosophy of science, the question of empirical equivalence and underdetermination of theories by data proves to have highly nontrivial examples, as well as surprising failures, when one looks into quantum field theory. Lessons for both the philosophy of physics and the general philosophy of science pertain not only to the standard questions in those disciplines, but also to the methods used for answering them effectively.

Concerning the philosophy of physics, the longstanding question of general covariance proves, upon detailed examination of the Anderson-Friedman absolute objects formalism, to be somewhat elusive. Various old counterexamples have been resolved, but the new Geroch-Giulini counterexample in GTR itself becomes more visible once the metric $g_{\mu \nu}$ is broken into its irreducible parts, the conformal
metric density $\hat{g}_{\mu\nu}$ and $\sqrt{-g}$, which are a tensor density (of fractional weight) and a scalar density, respectively. The general phenomenon causing difficulty here I have called the susceptibility to absoluteness of certain kinds of geometric objects, including scalar densities, tangent vectors, and most kinds of tangent vector densities as well. As it stands, the Anderson-Friedman analysis counts susceptible geometric objects as absolute. One might consider modifying the definition of absoluteness to exclude susceptible geometric objects, but thus far I can find no compelling reason either for retaining the usual definition or for modifying it to exempt the susceptible objects. If the usual definition is retained, it follows that GTR has an absolute object, so either GTR is not generally covariant or general covariance must be redefined as something other than the lack of absolute objects. Partly because Bach-Weyl conformal gravity lacks any Anderson-Friedman geometric object, one should not conclude immediately that the usual definition captures nothing of interest. If the susceptible geometric objects are exempted from absoluteness, then both the Geroch-Giulini counterexample involving $\sqrt{-g}$ and the tetrad-spinor counterexample are resolved (even without the alternative OP spinor formalism).

Rather than exempting susceptible geometric objects, one might envision modifying the definition of absoluteness to involve a variational principle. One way to do this is by introducing a criterion involving a variational principle as a further necessary condition for absoluteness. However, it appears that the variational criterion would then do all the work, while the Anderson-Friedman condition of sameness in all models does none. One might then consider using a purely variational criterion, largely abandoning the Anderson-Friedman project (apart from insights such as removing irrelevant fields), but the redundancy of the field equa-
tions in important physical theories makes one wonder if the variational criterion gives the right list of generally covariant theories for the wrong reason. It seems evident that the preference for local field theories plays an important role here, but whether that preference is merely pragmatic or involves genuine insight is unclear. The fact that susceptible geometric objects’ Euler-Lagrange equations say little or nothing about the susceptible objects themselves also casts doubt on the variational analysis of general covariance.

Converting fields not varied in an action principle to functions of (derivatives of) clock fields, on the other hand, appears to give the expected list of substantively generally covariant theories for plausibly the right reason. This proposal, which is largely inspired by the claim that GTR is already parametrized, bears considerable resemblance to the pre-Kretschmann views of Einstein on GTR’s distinctively lacking preferred coordinates. One might then view clock fields as a stratagem to deploy in a Kretschmann-type objection that general covariance is trivial or purely formal, while taking the absence of clock fields in GTR to show that it, unlike STR, has no preferred coordinates and ipso facto is substantively generally covariant. Thus both trivial and substantive general covariance admit clear analyses in terms of clock fields. The study of massive GTR in chapter 4 confirms the utility of using the absence or presence of clock fields as the criterion for substantive general covariance.

A second possible lesson, in light of the OP treatment of spinors in chapter 2, is that requiring covariance with respect to arbitrary finite coordinate transformations (including those sufficiently near Lorentz transformations), not just arbitrary infinitesimal coordinate transformations and some finite transformations, might be inadvisable. This sort of weakened formal general covariance suffices for Lie
and covariant differentiation mathematically, and for the equivalence principle physically, while avoiding the supposedly necessary inflation from a metric with 10 components to an orthonormal tetrad with 16. While a tetrad treatment of spinors might retain advantages for calculation in some contexts (especially when a diagonal metric tensor is unavailable), philosophers strive to separate descriptive fluff from physics with ontological significance. In the tetrad formalism, one could formally lessen the ontological commitment to a tetrad by taking the equivalence class with respect to position-dependent $O(3,1)$ transformations, but it is not very clear what it is to believe in such an equivalence class of tetrads. Ceteris paribus it is obviously better to have only 10 fields (besides the spinor) rather than 16; of course other things are not equal, but the benefits of the OP spinor formalism might well be worth the costs, at least for philosophical purposes. It is highly desirable to have a modern global treatment of the existence and uniqueness of OP spinors in the context of nontrivial topologies.

Besides answers regarding standard questions in the philosophy of physics, the study of general covariance has also yielded some methodological lessons. First, much of value for space-time theory has happened in physics after 1915-16, when the completed GTR appeared. Spinor fields, clock fields, and massive versions of Maxwell’s and Einstein’s theories are just a few examples.

Second, important mathematical tools involving geometric objects appeared in the middle of the 20th century and must be retained and in some cases recovered. The transition to so-called ‘coordinate-free’ differential geometry, while no doubt essential for a proper treatment of global issues, has led to unfortunate omissions and distortions relevant to the Anderson-Friedman project on account of the merely partial translation of old results into the new vocabulary(s). Ten-
sor densities, for example, are less well known today than in the 1960s, but they are crucial for resolving the Torretti and Norton counterexamples and for setting up the Geroch-Giulini counterexample. Besides tensors, connections, and tensor densities, one needs the general theory of geometric objects (and their kin, such as Siwek's pseudo-objects), some of which are nonlinear. Especially after the switch from global to local equivalence, it is more important to be thorough than to adopt a modern style, if one is forced to choose one or the other. The modern mathematical literature on natural bundles is indeed relevant if a global treatment of geometric objects is desired. Properly including non-coordinate types of covariance requires the generalization to gauge-natural bundles, though even this formalism might not be strictly adequate due to its involving only groups (not groupoids) and, perhaps, due to its assuming covariance under finite coordinate transformations (not just infinitesimal ones and some finite ones), depending on the ultimate resolution of the spinor issue. Spivak notes at the start of his pentalogy on differential geometry that

> no one denies that modern definitions are clear, elegant, and precise; it's just that it's impossible to comprehend how anyone ever thought of them. And even after one does master a modern treatment of differential geometry, other treatments often appear simply to be about totally different subjects. (Spivak, 1979, p. v)

Because finding modern-style definitions is not easy, scholars who wish to write in a modern style but lack the skills of a professional mathematician to invent their own tools, might be tempted to force-fit problems in the philosophy of physics into pre-existing conceptual boxes that mathematicians have already supplied (or have supplied in the most widely read texts). The discussions of densities, spinors, and groupoids in chapter 2 illustrate how physically and philosophically important concepts might not (now or in the recent past) yet have an adequate (and widely
known) modern mathematical treatment. Component-based differential geometry has a certain vagueness that is often considered a liability. That vagueness, however, can be an asset by avoiding premature precision when a problem has not yet been perfectly formulated. It might suffice to make a small change in component notation, such as by supplementing some equations with some inequalities, while the corresponding modern techniques might look very different. Component notation’s ambiguity might also be helpful in addressing the hole argument by avoiding smuggling gratuitous space-time point haecceities into the formalism.

A third methodological lesson for space-time philosophy is the importance of particle physics. Much of the particle physics literature consists of relativistic classical field theory that serves as preparation for quantization. Thus one need not be especially interested in quantum mechanics to benefit from particle physics literature. The neglect of spinor fields in space-time philosophy is not justified, not only because (quantum analogs of) spinors exist and indeed compose the bulk of ordinary matter, but also because spinors impose their own demands concerning global topology and/or perhaps coordinate freedom, demands quite dissimilar to those imposed by tensorial/bosonic fields. Until and unless an adequate quantum treatment of space-time and/or gravity is available, philosophers of space-time ought to strive to make their works at least quantization-ready.

5.2 Insights for the General Philosophy of Science

Developments in 20th century physics, especially in particle physics but also in gravitation, shed light on empirical equivalence and the underdetermination of theories by data, a standard issue in the general philosophy of science. Leaving aside some modal varieties of underdetermination with genuine physical examples,
particle physics reveals some interesting cases of permanent underdetermination between theories with different properties for the ‘same’ or corresponding models. The availability of massive Proca variants of Maxwell’s electromagnetism shows that underdetermination can be permanent rather than transient, even without there being any two empirically equivalent theories. Instead the competition is between Maxwell’s massless theory and the one-parameter family of massive Proca theories with different photon masses, which tend in the massless photon limit to reproduce the predictions of Maxwell’s electromagnetism. The difference between massless and massive theories is very important conceptually, because the massless theory is a gauge theory with mathematically indeterministic field equations and guaranteed charge conservation; by contrast, the massive theory has no gauge freedom and hence has deterministic field equations, but does not enforce charge conservation apart from the matter equations of motion. Given the empirical inequivalence between every pair of theories employed in the comparison, there is no possibility of identifying the ostensible rivals as merely linguistic variants of the same theory. This classical approximate empirical equivalence for electromagnetism proves to persist under quantization. Yang-Mills theories offer a precise classical analog of the massive Proca vs. massless Maxwell competition, with a smooth limit of vanishing mass. This equivalence, perhaps surprisingly, is broken under quantization, however, except in special cases involving an invariant Abelian subgroup of the gauge group. Massive variants of GTR provide an analogous competition to Einstein’s theory; the massive theories violate substantive general covariance in the mass term within the gravitational Lagrangian density. In this case there are various problems, some of which would profit from further attention in the physics literature. It is controversial whether the massive grav-
ities are healthy theories; the most widely held view is that they are unhealthy even classically, though various counterarguments suggest that this conclusion is not yet demonstrated. Massive variants of GTR also suffer from a causality problem, though that seems to be fixable by introducing artificial gauge freedom and distinguishing the (in principle) observable flat metric in the gravitational field equations from the flat background metric. Electromagnetic underdetermination between Maxwell and low-mass Proca theories evidently persists even into $SU(2) \times U(1)$ electroweak unification, because the $U(1)$ subgroup probably permits a Proca mass term while preserving unitarity and renormalizability. The dependence of the results on the details of quantum field theories suggests caution in either affirming or denying general theses about the existence of approximately empirically equivalent theories. Introducing descriptive redundancy deliberately in the form of artificial gauge freedom also pays off for general philosophy of science. It yields either underdetermination between Proca and Stueckelberg or the conclusion that the amount of convention in the form of gauge freedom is itself conventional, a conclusion analogous to one already suggested by Einstein regarding coordinate freedom in GTR. It is also noteworthy that different desiderata might be demonstrable features of one theory only (or most readily) if the theory is formulated in different ways for each demonstration, as in the case of unitarity and renormalizability of quantum field theories.

The discussion of clock fields in chapter 2 on general covariance turns out to involve a special case of artificial gauge freedom, so the same concept sheds light on general covariance in the philosophy of physics and on underdetermination in the general philosophy of science. A Lagrangian-friendly modification of the BFT formalism for converting second-class constraints into first-class constraints was
suggested above; then one immediately arrives at the Stueckelberg formulation starting from the Proca theories, and presumably directly arrives at parametrized massive GTR when starting with massive GTR. The physics literature on artificial gauge freedom also proves to admit considerable generalization to cases with no second-class constraints and perhaps no constraints at all; one need not convert all second-class constraints into first-class constraints, either, depending on one’s goals. These techniques provide algorithms for making generalized Kretschmann-type objections, to the effect that gauge symmetries are trivial or merely formal. Induction over special features of massive electromagnetism, Yang-Mills theories, and GTR in the formulations with artificial gauge freedom suggests some criteria for distinguishing artificial gauge freedom from the ‘natural’ gauge freedom in the massless cases (apart from the obvious presence of a mass term, which might not generalize). For the massive variants with artificial gauge freedom, the gauge parameters appear algebraically, so the new gauge compensation fields have redundant field equations. For the theories with natural gauge freedom, eliminating gauge freedom tends to produce nonlocal formulations, whereas the massive theories with artificial gauge freedom can be gauge-fixed (apart from causality worries for massive gravity) back into local field theories with no gauge freedom. Whether these differences are the correct criteria or not, they suggest that there might exist relevant differences that permit distinguishing artificial gauge freedom from the more familiar natural sort, in much the way that the suggestion that GTR is already parametrized suggests that substantive general covariance is formal general covariance achieved without clock fields. Descriptive redundancy need not be a nuisance. In the present contexts it is a resource: artificial gauge freedom sheds light not only on the question of empirically equivalent theories,
but also on general covariance.

Relating the manifest and the scientific images of nature (Eddington, 1928; Ladyman et al., 2007; Sellars, 1963) occasionally suggests that some features of the common-sense world are less fundamental than they seem. The example of the Bach-Weyl theory of gravity suggests that the laws of nature possibly lack the resources to define the lengths of timelike curves and hence ages, including the age of the universe. Theories with more than one space-time metric, including scalar-tensor theories, suggest the possibility that ages are radically ambiguous, perhaps to the point that even finitude vs. infinitude is ill-defined. Both the Bach-Weyl theory and some multi-metric theories seem to be metaphysically possible. Thus longstanding arguments about the necessity or impossibility of an infinite past, going back at least to Aristotle and Philoponus, respectively, thus require reformulation even to be intelligible (to say nothing of persuasiveness) in light of the possibilities disclosed by modern physics. Three possible strategies for reformulation involve the existence (or otherwise) of a first moment, the extensibility or inextensibility of the space-time toward the past, and conditionalization on the existence and uniqueness of sufficient temporal metrical structure.

This study suggests some methodological insights for the general philosophy of science as well. First, studying contemporary physical theories readily provides genuine or plausible examples of interesting kinds of underdetermination of theories by data. Thus philosophers need not rely on thin contrived examples; physics is important even for general philosophers of science. Second, comparing some contemporary physical theories suggests apparently novel sorts of underdetermination, as far as philosophical discussions are concerned. Some of these are modal. Another kind involves two theories that differ in that each model of one theory is
diffeomorphic to a proper part of a model of another theory. While this idea is not new, the fact that there are serious physical examples makes it more worthy of attention than when it savored of a version of radical skepticism. Approximate underdetermination, such as can arise between massive and massless theories, involves differences in occurrent properties between the two worlds, but has the surprising feature that underdetermination is permanent even though empirical equivalence is only approximate. Third, there is a surprising degree of dependence on the physical details, including difficult calculations in quantum field theory, for approximate underdetermination. Fourth, underdetermination can be broken in surprising ways when auxiliary hypotheses are changed, much as Leplin has anticipated. Fifth, the kinds of underdetermination discussed here are immune to trivialization as just linguistic variants of the same theory. Underdetermination might not be ubiquitous in contemporary physics, but there are enough nontrivial examples and types of it, apparently including electroweak theory, that the subject remains of considerable interest.


Alexander, H. G., editor (1956). *The Leibniz-Clarke Correspondence*. Manchester University, Manchester.


203


213


diffeomorphisms for the parametrised Maxwell field. *Classical and Quantum
Gravity*, 4:319.

Kuchowicz, B. (1973). Conformally flat space-time of spherical symmetry in

Kniper, N. H. and Yano, K. (1955). On geometric objects and Lie groups of
transformations. *Koninklijke Nederlandse Akademie van Wetenschappen, A,

*Analysis*, 53:1.

Kukla, A. (1996). Does every theory have empirically equivalent rivals? *Erkenn-
ntnis*, 44:137.

York.

to the massive Yang-Mills field. *Progress of Theoretical Physics*, 37:452.


Lakatos, I. (1970). Falsification and the methodology of scientific research pro-
grammes. In Lakatos, I. and Musgrave, A., editors, *Criticism and the Growth
of Knowledge*. Cambridge University Press.

Lakatos, I. (1971). History of science and its rational reconstruction. In Buck,
R. C. and Cohen, R. S., editors, *PSA: Proceedings of the Biennial Meeting of
the Philosophy of Science Association, 1970*, Boston Studies in the Philosophy

Lanczos, C. (1949). *The Variational Principles of Mechanics*. University of
Toronto, Toronto.


third edition.


Luchter, J. (1973). Determination of the linear non-homogeneous geometric objects of the type \([m, n, 2]\), where \(m \leq n\). *Commentationes Mathematicae Prace Matematyczne*, 17:137.
the photon rest mass with a rotating torsion balance. Physical Review Letters,
90:081801.

Lusanna, L. and Pauri, M. (2006). The physical role of gravitational and gauge
degrees of freedom in general relativity - I: Dynamical synchronization and
generalized inertial effects. General Relativity and Gravitation, 38:187. gr-
qc/0403081v2.

underdetermination in gravitational theories. Journal for General Philosophy
of Science, 34:277.

Magnano, G. and Sokolowski, L. M. (1994). Physical equivalence between nonlin-
ear gravity theories and a general-relativistic self-gravitating scalar field. Physical
Review D, 50:5039. gr-qc/9312008.

Maidens, A. (1998). Symmetry groups, absolute objects and action principles in
general relativity. Studies in the History and Philosophy of Modern Physics,
29:245.


gravity and volume-preserving coordinates. Journal of Mathematical Physics,
38:4741–4757. dg-ga/9703002.

Philosophy of Science, 69:S221.

massive fields coupled to sources. Journal of Mathematical Physics, 30:1877.

thesis, University of Southampton. supervised by James A. Vickers;


Smolin, L. (2002). Quantum gravity at the background independent level. In _PSA_.


