Why Geometry is not Conventional: the Verdict of Covariance Principles

John D. Norton

1 Introduction

The question of which parts of a theory of space or spacetime are conventional and which are factual has dominated foundation studies in space and time for almost a century. The emphasis of Einstein's work in relativity theory has always been on exactly this question. His various principles of relativity, of equivalence of general covariance are all intended to tell us that this or that state of motion cannot be designated as "at rest" factually but only as a convention. A tradition of similar vintage in the philosophy of science literature urges that we cannot choose factually the geometry of a space or which spatially separated events in special relativity are simultaneous. My purpose in this paper is to establish two theses relating the two sets of conventionality claims:

(I). Covariance principles provide a natural way of automatically separating the factual from the conventional components of a theory. This sense of convention is most familiar to us as the sense in which coordinate systems are freely chosen.

(II). Contrary to the claims of Hans Reichenbach and others, the metrical geometry of a space or spacetime is not conventional, or at least not in the sense natural to geometric theories provided by covariance principles. This conclusion follows directly from Reichenbach's "universal forces" construction.

To establish Thesis (II), I will argue that the non-conventional character of metrical geometry is revealed by the need to introduce universal forces as correction factors when we substitute an arbitrary metric for the one revealed by direct measurement. Were metrical structure conventional in the sense of Thesis (I), such correction factors would be unnecessary.

2 The Representation of Physical Spaces and Spacetimes

In constructing a theory of space or spacetime, the principal task of the theorist is to select mathematical structures capable of representing the space or spacetime in question. Virtually all such efforts make use of the manifold of real numbers $\mathbf{R}$ or of ntuples of real numbers $\mathbf{R}^n$. To see how the representation works, consider two simple examples: a two-dimensional Euclidean surface and a Minkowski spacetime of special relativity. They can be represented by $\mathbf{R}^2$ and $\mathbf{R}^4$ respectively.

For all that follows, it is crucial that the reader observe the following. For simplicity of analysis, the $\mathbf{R}^n$ employed here is literally the set of of all ntuples of real numbers (with standard topology assumed) so that points in the set would include $< 0,0,0,0 >$ and $<1,-5,4,1>$. $\mathbf{R}^n$ is not just an n-dimensional manifold with some arbitrary point set and which is "topologically $\mathbf{R}^n$", that is, isomorphic to the manifold of n-tuples of reals with standard topology. The use of R^n in this way corresponds to methods of the older traditions of geometry and spacetime theory, as I have argued elsewhere in NORTON (1989) and (1991).

To represent the properties of the Euclidean surface or Minkowski spacetime, each is coordinated with $\mathbf{R}^2$ or $\mathbf{R}^4$ in the familiar manner. A point of the Euclidean surface is coordinated with a pair $< x, y >$ of reals in $\mathbf{R}^2$. A given straight line in the Euclidean surface is coordinated with the set of pairs $\{< x,y >: Ax = By + C\}$ for suitable real constants A, B and C. Similarly, an event of the Minkowski spacetime is coordinated with a quadruple of reals $< x, y, z, t >$. Other familiar structures of the Minkowski spacetime-worldlines and hypersurfaces-are coordinated with suitable corresponding substructures of R^4.

This coordination of a physical space or spacetime with $\mathbf{R}^n$ is naturally called a coordinate system. The theories of the Euclidean surface and Minkowski spacetime alluded to so far employ just a single coordinate system. Therefore it is natural to call them the "one coordinate system" formulations of these theories.[1]

3 The Problem of Superfluous Structure

Simple as they are, the one coordinate system formulation of these two theories runs into a serious problem. One generally expects that the problem facing a theory is that the physical system under investigation is too rich in structure to be represented easily by readily available mathematical structures. In this case we find precisely the opposite problem. $\mathbf{R}^2$ or R^4 are actually too rich to represent a Euclidean surface and Minkowski spacetime. Both contain superfluous structure to which we cannot allow physical significance.

For example, $\mathbf{R}^2$ contains a preferred origin point $< 0,0 >$ with properties shared by no other pair in $\mathbf{R}^2$.[2] Indeed each of its points is intrinsically distinct from every other. Does this mean that each point of the Euclidean surface represented has unique properties shared by no other point of the surface? It cannot, for the Euclidean surface

[1] For further details, see SALMON (1992, Ch. 5).

[2] For example, it is the only point $< x, y >$ which solves $< x,y >=< 2x, 2y >$.

is by supposition homogeneous, which means that every point is like every other. The surface can have no factually preferred origin. Similarly the x and y axes of $\mathbf{R}^2$ pick out preferred directions in the Euclidean surface. The Euclidean can have no such preferred directions since it is by supposition isotropic.

The situation is similar for the one coordinate system formulation of special relativity. The preferred origin of $\mathbf{R}^4$, $< 0,0,0,0 >$ cannot represent a factually preferred event in the spacetime, for the spacetime is postulated to be homogeneous. Again, the distinct set of straights parallel to the t-axis of $\mathbf{R}^4$ picks out a preferred state of inertial motion in the Minkowski spacetime. Yet the principle of relativity postulates that no such state of inertial motion can be preferred other than through a conventional stipulation.

This problem of superfluous structure is solved routinely by expanding the set of coordinate systems used in formulating the theory. These coordinations can usually be mapped onto one another by a set of transformations that form a group. The group of these transformations is known as the covariance group of that formulation of the theory. The crucial postulate is that each coordination represents the physical properties of the space or spacetime equally well. It follows immediately that the only aspects of the coordination that can have physical significance are those that are the same in all the coordinate systems of the formulation. In other words, physical significance can only be attributed to the invariants of the formulation's covariance group.

To employ this solution of the problem in the case of the Euclidean surface, we proceed from the one coordinate system formulation to the "standard formulation" of the theory. We allow any coordinate system produced from the original by any translation, rotation or reflection of the original coordinate system. These transformations form the covariance group of the formulation. We call the coordinate systems of the standard formulation, standard coordinate systems. It now follows automatically that the Euclidean surface is homogeneous. Consider for example some point p of the Euclidean surface that is coordinated with the preferred origin $< 0,0 >$ of $\mathbf{R}^2$. Is p thereby picked out as a preferred origin of the surface factually distinct from any other point? It is not. The covariance group of the theory includes arbitrary translations. Therefore there will exist a transformation within the group to a new coordinate system in which any other nominated point–say "q"–is coordinated with $< 0,0 >$. By supposition, each of these coordinations represent the Euclidean surface equally well. Therefore whatever physical properties are conferred onto p by virtue of the fact that is can be coordinated with $< 0,0 >$ in some coordinate system must also be conferred onto any other nominated point q of the surface. That is, the inhomogeneity of $\mathbf{R}^2$ does not confer inhomogeneity onto the Euclidean surface. The use of a particular coordinate system in which p in the surface is coordinated with $< 0,0 >$ does not indicate that p is factually distinct from all other points; it merely reflects a conventional choice of that particular coordinate system.

A similar argument establishes that the anisotropy of $\mathbf{R}^2$ does not confer anisotropy onto the Euclidean surface. Under a coordinate transformation within the covariance group, any nominated direction in the Euclidean surface can be transformed to be parallel to the x-axis, for example, of the coordinate system.

The implementation of the solution in the case of special relativity is analogous. The standard formulation of the theory allows all coordinate systems generated from the original by a transformation in the extended Lorentz group. In particular, the transition to this standard formulation automatically deprives special physical significance from states of inertial motion represented by curves parallel to the t-axis in some coordinate system. The argument proceeds as above. Consider any inertial state of motion. There will always be a Lorentz transformation to a coordinate system in which that motion is represented by curves parallel to the t-axis. Since all coordinate systems are postulated to represent the physical spacetime equally well, then every inertial state of motion can be designated equally properly as "at rest". The decision to use one or other coordinate system and thereby privilege one set of inertial motions does not reflect a factual property of the Minkowski spacetime, but is merely a conventional choice of coordinate system. In this way, the principle of relativity of special relativity is expressed by the Lorentz covariance of the theory.[3]

If preferred origins and states of rest do not have physical significance in Euclidean geometry and special relativity, what does? We answer by selecting invariants. The straightness of a line in a Euclidean space remains invariant under rotations, transformations and reflections.[4] Therefore this straightness (but not direction) is physically significant. Similarly inertial states of motion remain inertial under Lorentz transformation and are physically significant. Most fundamentally, we assign physical significance to distance ℓ in Euclidean geometry and interval s in special relativity. The two quantities are given by the fundamental forms

$$d\ell^2 = dx^2 + dy^2 + dz^2 \tag{1}$$

which gives the distance $d\ell$ between two points in a Euclidean surface whose coordinates differ by dx, dz and dz; and

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

which gives the interval between two events in a Minkowski spacetime whose coordinates differ by dx, dy, dz, dt. (c is the speed of light.) Both of these quadratic differential forms remain invariant under the transformations of the covariance group of the standard formulation of the theory in question.

4 General Covariance: Making Geometric Structure Explicit

While the standard formulations of Euclidean geometry and special relativity suffice for most practical purposes, this is not the case for other theories of space and time. Most notable is general relativity, which requires a generally covariant formulation. It is now customary to employ generally covariant formulations of all theories of space and spacetime. This custom ensures that comparative studies of these theories reflect actual

[3]Readers concerned that I am disagreeing with the now widespread view that covariance principles are physically vacuous should consult my paper NORTON (1989) and (1991).

[4]More precisely, of course, the linearity of the relation between coordinates representing the line is preserved.

differences between the theories and not accidents of different types of formulation. However if we continue to employ covariance principles as a means of separating factual from conventional content, it seems, at first blush, that we deny physical significance to too much when we move to generally covariant formulations of Euclidean geometry and special relativity.

Pursuing only the example of Euclidean geometry, let us expand the covariance group to the general group of transformations, that is, to the group of all C^∞ transformations on $\mathbf{R}^2$. Thus we allow an arbitrary C^∞ transformation from a standard coordinate system (x, y) to the coordinate system (x_1, x_2). Consider for example the transformation

$$x = (x_1)^3, \quad y = (x_2)^3$$

for which we have transformation coefficients

$$p_{11} = \frac{\partial x}{\partial x^1} = 3(x_1)^2, \ p_{22} = \frac{\partial y}{\partial x_2} = 3(x_2)^2, \ p_{12} = \frac{\partial x}{\partial x_2} = 0 = \frac{\partial y}{\partial x_1} = p_{21}.$$

In general, under such a transformation, the fundamental form (1) of a Euclidean surface fails to remain invariant. That is,

$$d\ell^2 = d(x_1)^2 + d(x_2)^2 \tag{2}$$

is not invariant.

It follows immediately that distances as determined by the form (2) have no physical significance. The distance $d\ell$ between two infinitesimally close points will vary according to the coordinate system chosen. That distance will therefore not be invariant. Notice that were the space in question a topological space in which distance along curves has no physical significance, then this outcome would not trouble us. Indeed we would choose to expand the covariance of the theory to general covariance precisely in order to display the lack of physical significance of distance.

Of course the problem of giving a generally covariant formulation of Euclidean geometry is routinely solved in the following way. Within the generally covariant formulation, we retain a subset of coordinate systems corresponding to the standard coordinates. These coordinates are specially adapted to the geometry of the surface in the sense that the form (1) gives the physically significant distances of the surface. To complete the formulation of the theory, we replace the form (2) by a form that preserves the invariance of dl under arbitrary C^∞ transformation. The form automatically corrects for the deviations of the new coordinate system (x_1, x_2) from a standard coordinate system $(x, y) = (X_1, X_2)$ in such a way that dl remains invariant. We write (with summation over repeated indices implied)

$$d\ell^2 = (dX_m)(dX_m) = p_{mi}\, p_{mk}\, dx_i\, dx_k = g_{ik}\, dx_i\, dx_k \tag{3}$$

where $g_{ik} = p_{mi}\, p_{mk}$ will be recognized as the coefficients of a metric tensor. In practical terms, these coefficients g_{ik} encode the deviations of the relevant coordinate system from a standard coordinate system and enable us to preserve the invariance of the distance ℓ. We now also think of the metric tensor as giving the surface the physically significant property of distance via the form (3). That is, with the transition to general covariance, that aspect of the surface's geometric structure, its metrical structure, has been made explicit as the metric tensor.

Because it will parallel what is to follow, I should like to review the overall procedure that led to the appearance of the metric tensor. We began with the one coordinate system formulation of Euclidean geometry and proceeded to expand the covariance group of the formulation through successively larger groups. The effect was to deny physical significance to successively more structure. With the transition to the generally covariant formulation, this process had proceeded too far. We had denied physical significance to distance. We preserved the physical significance of distance by rescuing the relevant aspect of geometric structure, the metrical structure, as the correction coefficients $g_{ik} = p_{mi}p_{mk}$. These coefficients were re-introduced as a new and distinct component of the theoretical structures of the theory, the metric tensor, which encapsulates the physically significant distance properties of the surface.

5 Reichenbach's Relativity of Geometry and the Universal Forces Argument

While the thesis has been advanced in various forms by Einstein, Poincaré and others, the notion that there is an important convention in metrical geometry is best known from the work of Hans Reichenbach.[5] Reichenbach observed that we cannot investigate the metrical geometry of a surface until we have some system for determining metrical distances on the surface. Such a system, he argued, required what he called an arbitrarily chosen "coordinative definition," which coordinates some particular thing with some concept. For example, we may decide to define distance on the surface (a concept) as coinciding with the results of measuring operations using a movable rod (a thing) taken by definition to be of unit length. Reichenbach did not feel that metrical geometry was physically vacuous, as may be suggested by the use of the label "conventionalism." Thus he elected to use the term "relativity of geometry" to indicate that his view held to the dependence of the results of measurement on a definition.

Nonetheless the term conventionalism continues to be used of the view championed by Reichenbach – and properly so. For, according to Reichenbach's view, differing choices of coordinative definitions will yield differing geometries for the surface in question. In particular, whether the geometry is Euclidean or non-Euclidean will depend upon a definition. Thus whether the surface is or is not Euclidean must be described as a matter of convention, for its Euclidean character can be eradicated, for example, by a conventionally chosen definition.

This role of coordinative definitions emerged clearly in Reichenbach's discussion of universal forces. Universal force fields, by definition, affect all material in the same way and admit no insulating walls. Let us imagine that we use a unit measuring rod to map out the geometry of a surface. If a universal force field which distorts distances is present, our results will be affected by these distortions. However because universal forces affect all materials alike and cannot be shielded, we have no independent way to correct for these distortions. Measurements conducted with another measuring rod

[5] See especially REICHENBACH (1958, Ch. 1) upon which the discussion of this section is based.

of different composition will be distorted equally. Thus, Reichenbach concludes, we can only accommodate universal forces if we set by definition that they vanish or have some other value.[6]

As an aside from my main argument, I must note that the notion of a universal force, as a genuine, physical force, is an extremely odd one. They are constructed in such a way as to make verification of their existence impossible in principle. The appropriate response to them seems to me not to say that we must fix their *value* by definition. Rather we should just ignore them and for exactly the sorts of reasons that motivated the logical positivists in introducing verificationism. Universal forces seem to me exactly like the fairies at the bottom of my garden. We can never see these fairies when we look for them because they always hide on the other side of the tree. I do not take them seriously exactly because their properties so conveniently conspire to make the fairies undetectable in principle. Similarily I cannot take the genuine physical existence of universal forces seriously. Thus to say that the values of the universal force field must be set by definition has about as much relevance to geometry as saying the the colors of the wings of these fairies must be set by definition has to the ecology of my garden.

Undeterred by such quibbles Reichenbach[7] used the notion of universal force to state a theorem from which it followed that a surface could at least locally be held to have any nominated metrical geometry provided we were prepared to stipulate the presence of an appropriate universal force. In particular, imagine that the uncorrected results of measurement yield a metric g'_{ik} for some surface, then we can infer that the geometry of that surface is actually given by the metric g_{ik} if we stipulate that our measuring rods were under the influence of a universal force field F_{ik} where all three quantities are related by the equation

$$g'_{ik} = g_{ik} + F_{ik}. \tag{4}$$

Clearly if g'_{ik} is Euclidean, g_{ik} need not be, according to the value of F_{ik}, which is set by stipulation. That is, whether a surface is Euclidean depends on a conventional stipulation.

6 Universal Forces Construction Displays Non-Conventionality of Geometry

The central move of the theorem discussed immediately above is to add another transformation to the groups of transformations used in formulating the theory. These transformations are best described as gauge transformations on the metric g_{ik}; that is, the gauge transformations map g'_{ik} onto any arbitrary g_{ik} differing from it by an arbitrary additive tensorial factor F_{ik}, according to equation (4).

[6]Differential forces, however, affect different materials differently. Thus their presence can be revealed by comparing the behavior of different materials. Heat is an example of a differential force. One can reveal the presence of thermal expansions in regions of varying temperature by comparing the differential expansion of copper and wooden measuring rods.

[7]REICHENBACH (1958), p. 33.

This addition of a new group of transformations is strongly analogous to the expansions of the covariance groups seen above. If we assume that each transformation relates structures equally able to represent the factual nature of the geometric surface, then the addition of this gauge transformation deprives more of the theory's structure of physical significance. In particular, the distance ℓ, as given by the quadratic differential form (3) now ceases to be an invariant. Adding an arbitrary gauge term F_{ik} to the coefficients g_{ik} in (3) will in general alter the value of $d\ell^2$. If the surface in question were merely a topological space with no factual metric properties, this circumstance would be exactly what we want. The lack of invariant character of $d\ell^2$ would merely be reflecting that the space has no factual metrical properties.

However, if the surface does have factual metrical properties, then our addition of the gauge transformation has clearly gone too far. Moreover its does so in exactly the same way as did the expansion of the covariance group of the theory to the arbitrary C^∞ transformations of the generally covariant formulation. In the latter case we rescued the structure with physical content by preserving certain coordinate systems as specially adapted to the geometry and introducing correction factors into the form (2) to correct for distortions introduced when employing coordinate systems other than the standard coordinate systems.

We can do precisely the same thing here. Let us single out as preferred a metric g'_{ik}, the one revealed by uncorrected distance measurements. Whenever we transform to a new metric g_{ik} we will preserve our path by noting the additive term F_{ik} used in the transformation. We now define the distance ℓ to be given by the form

$$d\ell^2 = g'_{ik}\ dx_i\ dx_k = (g_{ik} + F_{ik})\ dx_i\, dx_k. \tag{5}$$

The coefficients F_{ik} are "correction factors," to use Reichenbach's[8] own term. The quantity $d\ell^2$ as defined in (5) is once again an invariant under all the transformations of the theory, both coordinate transformations and gauge transformations of the metric. Since the values of $d\ell^2$ must coincide with the factual results of uncorrected measuring operations, this transition to the form (5) must be taken by the Reichenbachian conventionalist.

How are we to interpret the need for the factors F_{ik} in the form (5)? Their presence serves the same function as the presence of the factors $g_{ik} = p_{mi}\ p_{mk}$ in the form (3), where the factors g_{ik} protect the metrical structure against loss of physical significance in the transition from the standard to the generally covariant formulation. Indeed the correction factors g_{ik} themselves become symbolic of the metrical structure of the surface. Correspondingly the presence of the factors F_{ik} in (5) serves the function of preserving the metrical structure against loss of physical significance under the introduction of the gauge transformation (4). Indeed we might well say that the structure F_{ik} now represents that physically significant metrical structure (in conjunction with the coefficients g_{ik}) in the same way as the coefficients g_{ik} (in conjunction with the coordinate system (x_1, x_2)) represents the metrical structure in the generally covariant formulation.

We can choose freely between different sets of coefficients g_{ik} in the form (5) by means of the gauge transformation. However the crucial point is that this does not

[8] REICHENBACH (1958), p 33.

reveal a convention in metrical structure, for all such choices are subject to correction by the factors F_{ik}. In precise analogy, we can choose an arbitrary coordinate system in the generally covariant formulation of the theory. This does not mean that metrical properties of the surface are conventional. For all such coordinate systems the coordinate differentials are corrected by the factors g_{ik} in the form (3) when recovering judgements of distance.

7 Conclusion

Reichenbach's universal forces construction demonstrates that the metrical geometry of a surface is not conventional in the sense natural to geometry, the one provided by covariance principles and exemplified in our conventional freedom to choose coordinate systems.[9] The obvious question that goes beyond this paper is whether there is some other interesting sense in which metrical geometry is conventional. There is, of course, a trivial sense in which the truth of any factual statement is dependent on the definitions of its terms. We can conventionally choose to alter the truth value of any factual statement by arbitrarily changing the meanings of its terms. Salmon[10] correctly rules out this type of convention as "trivial semantic conventionalism" because the existence of the convention does not depend upon any physical facts of the world. Under this criterion, our freedom to stipulate that any state of inertial motion in special relativity is the "rest" state is a non-trivial convention, for it depends the physical fact of the principle of relativity. Can we find physical facts upon which to base the conventionality of metrical geometry?[11] I do not think that we can look to universal forces as a source of some appropriate physical fact. As I indicated briefly above, I cannot take seriously the postulation of a physical force so contrived by definition as to be beyond detection.

[9]The use of the term "coordinate" and its adjectival form "coordinative" in "coordinate system" and "coordinative definition" suggests a close connection between the conventional freedom in choosing coordinate systems and in setting coordinative definitions. While I do think that there is a reasonably strong suggestion of a connection between the two conventions, it should be noted that this suggestion is not as strong in Reichenbach's original German. For example, in REICHENBACH (1928), the German text of REICHENBACH (1958), for "coordinate system" Reichenbach used the standard "Koordinatensystem" (e.g. p. 279). For "coordinative definition," however, he used "Zuordnungsdefinition" (e.g. p. 23), which could be translated as "associative definition" or "correspondence definition," if one wished to avoid the term "coordinative."

[10]SALMON (1969), p. 61).

[11]This is not the place to consider GRÜNBAUM's (1973, Part I) claim that space is metrically amorphous and has no intrinsic metrical properties, so that these properties must be provided conventionally by us.

Semantical Aspects of Spacetime Theories

Edited by
Priv.-Doz. Dr. Ulrich Majer
Universität Göttingen
and
Prof. Dr. Heinz-Jürgen Schmidt
Universität Osnabrück

B·I·
Wissenschaftsverlag
Mannheim · Leipzig · Wien · Zürich

1994