5 Geometries in collision: Einstein, Klein and Riemann[1]

John D. Norton

General relativity threw into physics and philosophy the antagonism that existed between the two principle directors of geometry, Riemann and Klein. The space–times of classical mechanics and of special relativity are of the type of Klein, those of general relativity are of the type of Riemann. (Cartan, 1927)

Introduction

Of all the achievements of modern theoretical physics, one stands out as unique. Without discernible intermediaries, the general theory of relativity completely transformed our understanding of gravitation. Gravitation ceased to be a field in space and time like the electromagnetic field; it became the geometric curvature of space–time itself. And this transformation was effected essentially by just one scientist, Albert Einstein. Nonetheless, in the eight decades that have passed since Einstein's achievement, a growing chorus of voices have cried that, in spite of his unique success, Einstein never understood the foundations of his own theory. The focus of these problems is Einstein's claims concerning relativity principles, covariance principles and coordinate systems. Specifically, Einstein's theory is generally covariant; its equations remain unchanged in form under arbitrary transformations of the coordinates. Through this property, Einstein claimed, his theory had effected the most complete generalization of the relativity of inertial motion of the special theory to a more general relativity that embraced acceleration. Thus, Fock (1959, p. 368) remarked on this '...point of view that we cannot accept as correct...' and reflected

> The fact that the theory of gravitation, a theory of such amazing depth, beauty and cogency, was not correctly understood by its author, should not surprise us. We should not be surprised at the gaps in logic, and even errors, which the author permitted himself when he derived the basic equations of the theory. In the history of physics we have many examples in which the underlying significance of a fundamentally new physical theory was realized not by its author but by someone else and in which the derivation of the basic equations proposed by the author proved to be logically inconsistent. It is sufficient to point to Maxwell's theory of the electromagnetic field...

1 I am grateful to Jeremy Gray and David Rowe for comments on an earlier version of this chapter.

We might be tempted to dismiss such an appraisal as an ill-tempered fulmination were it not now a mainstream assessment.

My task in this paper is not to exonerate or to censure Einstein. The debate over the significance of covariance principles in relativity theory has grown into many convolutions over the past eight decades and continues today. Elsewhere I have surveyed and discussed these disagreements (Norton 1993, 1995). Rather my purpose here is to explain how Einstein came to make claims that proved so controversial and to understand why they proved so fragile.[2] My account will depend upon an accident of history: general relativity emerged through the collision of two traditions in geometry, that of Felix Klein and that of Bernhard Riemann. The account is based on two claims:

(1) Klein and Riemann employed very different strategies in deciding which of their mathematical structures represented the physically or geometrically real.
(2) Einstein's pronouncements concerning relativity principles in general relativity derive from his importing Klein's strategy, used to good effect in special relativity, into Riemann's geometry, the geometry of general relativity.

I will urge that Einstein's use of Klein's strategy was entirely appropriate within the context of special relativity. For, as Minkowski showed, special relativity provided a beautiful illustration of the power of Klein's approach. Since Einstein saw general relativity as a natural development of special relativity, it was equally natural to retain Klein's strategy of discerning the physically real. The result, however, was an unhappy hybrid interpretation, fated to be ostracized by later generations.

The geometrically real in Klein's Erlangen Program and Riemann's inaugural address

The difference between Klein's and Riemann's approaches is usually understood in terms of the different types of geometries they best addressed. Klein's approach flourished with geometries of uniform spaces, that is (in more modern language), spaces with non-trivial symmetries. These included projective geometries and the geometries of metrical spaces of zero or constant curvature. Riemann's approach extended the methods Gauss developed to deal with surfaces of variable curvature. Such spaces in general admit no

2 This paper provides the historical underpinning for the accounts first developed in Norton (1989) and (1992).

non-trivial symmetries. Beneath this prominent difference lies another. The two approaches employed opposing strategies to determine the geometrically real. Klein employed a subtractive strategy: he would over-describe the space and then direct which parts of the over-description should be accepted as geometrically real. Riemann employed an additive strategy; he would begin with an impoverished description and then only carefully add in further structure in an effort to ensure that all his structures had geometric significance.

Klein did not initiate the use of group and invariant methods in the service of geometry.[3] Their use was widespread in the nineteenth century. I will concentrate on Klein's expressions of these ideas, however, since by the end of the nineteenth century and into the next century, they became widely known and even canonical statements of the approach—as evidenced by the remarks quoted from Cartan (1927) above.

Klein's subtractive strategy

Klein's Erlangen Program was an attempt to bring systematic order to the proliferation of different geometries emerging in the nineteenth centuries. The key was to seek the group characteristic of each geometry. Once this group had been found, the real entities of the geometry could then be recovered as the invariants of the group. Such is the theme of Klein's (1872, p. 463) canonical statement of his Erlangen Program:

> There are spatial transformations which leave the geometric properties of spatial structures completely unchanged.... We designate the intension of all these transformations as the *principal group*[4] of spatial alterations; *geometric properties are not altered by transformations of the principle group.* Also conversely one can say: *geometric properties are characterized by their invariability under transformations of the principle group.* (Emphasis in original)

Klein's approach provides a means of discerning geometric properties. Implicit in this statement is the assumption that there are properties that are not geometric and that one needs to proceed carefully in order not to confuse the two. This was a commonplace already of the geometry of Klein's time. In projective geometry, one learned to discount familiar geometric notions such as lengths on a Euclidean surface in order to attend to the properties peculiar to the geometry, those that were unchanged under projective transformation. Klein canonized these notions into a powerful general method. We are able to

3 While I never claimed otherwise, I am grateful to David Rowe and Jeremy Gray for emphasizing the pervasiveness of these group and invariant theoretic approaches in the nineteenth century.

4 [Klein's footnote] 'That these transformations form a group is conceptually necessary'.

over-describe in our geometrical treatises. His Erlangen Program showed us how to subtract the non-geometric properties from these descriptions to leave the geometric properties exposed.

Klein noted in his conclusion that his method performed a special service in the context of analytic geometry. Here the danger of improper over-description was greatest. His method provides a systematic escape from this danger. In analytic geometry, one employs the resources of algebra to the ends of geometry. The geometric space is represented by algebraic variables—the coordinates of the space—and geometric structures in the space are represented by algebraic expressions. Klein (1872, p. 488) explained:

> It has often been objected that analytic geometry privileges arbitrary elements through the introduction of coordinate systems, and this objection applies to every way of treating extended manifolds in which individuals are characterized by the values of variables. While this objection was only too often justified with the defective way in which coordinate methods were used earlier, it vanishes with the rational treatment of the method. The analytic expressions that can arise in the investigation of a manifold in the sense of a group, must be, in accord with their meaning, independent of the coordinate system, in so far as it is arbitrarily chosen, and it is worthwhile also to make this independence *formally* evident. That this is possible and how it has to happen is shown by modern algebra in which the formal concept of the invariant, which is at issue here, is manifest most clearly. It possesses a general and exhaustive law of formation for invariant expressions and operates on principle only with them. One should also place the same demand on formal treatments if groups other than the projective are used as a basis. (Emphasis in original)

This case of analytic geometry is of special interest to us, for the geometric methods that Einstein will come to use lie fully within this analytic tradition. Klein here poses a problem central to Einstein's work: how are we to know if some structure described analytically in terms of coordinates has geometric significance? Klein assures us that the method of his Erlangen Program supplies the answer: we need only check whether the expression is an invariant of the group of the geometry.

Riemann's additive strategy

The project of Riemann (1854), his widely celebrated inaugural address, is immediately familiar to modern readers. He wishes to give a description of metrical spaces with variable curvature. This description inheres in two notions: what he calls the 'n-fold extended manifold' and a 'relation of measure'. They correspond directly to a modern n-dimensional manifold on which metrical structure is defined by means of a quadratic differential form. Familiar as these notions are, what is striking in the address is the great difficulty Riemann finds in describing his n-fold extended manifold. The difficulty is quite explicitly addressed by Riemann. In introducing the concept of the n-fold extended manifold, he begins with a plea for indulgence from the reader (p. 412)

> ... I think myself the more entitled to ask considerate judgement inasmuch as I have had little practise in such matters of a philosophical nature, where the difficulty lies more in the concepts than in the construction, and because I have not been able to make use of any preliminary studies whatever aside from some very brief hints which Privy Councillor Gauss has given on the subject in his second essay on biquadratic residues and in his Jubilee booklet, and some philosophical investigations of Herbart.

The ensuing discussion does not disappoint the reader expecting a laboured development. The notion of manifold is built from general concepts allowing 'modes of determination' and Riemann labours mightily to convey his intent. Typical of his efforts is his explication of what we now recognize as the dimensionality of the manifold: (p. 413)

> In a concept whose modes of determination form a continuous manifold, if one passes in a definite way from one mode of determination to another, the modes of determination which are traversed constitute a simply extended manifold and its essential mark is this, that in it a continuous progress is possible from any point only in two directions, forward or backward. If now one forms the thought of this manifold again passing over into another entirely different, here again in a definite way, that is, in such a way that every point goes over into a definite point of the other, then will all the modes of determination thus obtained form a doubly extended manifold. In a similar procedure one obtains a triply extended manifold when one represents to oneself that a double extension passes over in a definite way into one entirely different, and it is easy to see how one can prolong this construction indefinitely.

I need hardly point out just how much this discussion leaves for the reader to make precise. Just what is a 'mode of determination'? What is 'passing over into another entirely different'? It seems not that difficult to answer these questions very precisely. Indeed Felix Klein (1927, p. 289), in reporting Riemann's work, did exactly that:[5]

> Riemann laid at the foundations of his investigations n variables $x_1, x_2, \ldots, x_n$ each of which can take all real values. Riemann denoted the totality of their systems of values as a *manifold of n dimensions*; by a fixed system of values $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$, he meant a point in this manifold. Riemann hereby avoided expressing the word space, since he later wanted to introduce the space given to us intuitively as a special case of a three fold extended manifold. (Emphasis in original)

5 It has been pointed out to me by David Rowe that a literal reading of Klein's remarks suggests that Riemman's manifolds are topologically $\mathbf{R}^n$ (in the modern sense). That would automatically prevent a single space of Riemann from representing a complete spherical space. Since Riemann's scheme clearly included such spaces, I cannot believe that Klein intended this implication. I presume that Klein was not attending to the issue of global versus local topology and that his remarks were intended to apply to local patches of Riemann's spaces.

Drawing on the resources of analytic geometry, Klein identified Riemann's manifolds with the n-dimensional number manifolds $\mathbf{R}^n$—the manifold whose points are all n-tuples of reals.[6] At a stroke, Klein had eliminated the vagueness of Riemann's exposition. Riemann's 'mode of determination' is simply an n-tuple of reals, understood as a set of values of the variables $x_1, x_2, \ldots, x_n$. A curve in $\mathbf{R}^n$ is Riemann's simply extended manifold. Its extension to a surface gives the doubly extended manifold; and so on.

Klein is correct, of course, in noticing that Riemann avoided use of the term space so that ordinary physical space could be introduced as an instance of his notion of manifold. That Riemann intends the notion of manifold to cover more than ordinary spaces is made clear by his mention (p. 413) of the colours as forming a continuous manifold. But if Klein is right and Riemann intended his manifolds to be number manifolds, we must marvel at Riemann's obstinacy in avoiding their mention in his long discussion of Section I of his address on the concept of manifold. The notion of a number manifold was quite familiar to Riemann. He introduces it later (p. 416) in his address as a way of fixing locations in manifolds and to enable him to avail himself of the resources of algebra in the treatment of his relations of measure.

Why, then, did Riemann, unlike Klein, avoid use of number manifolds in explicating his concept of manifold? The answer is given by their differing strategies of discerning the geometrically real. Number manifolds over-describe Riemann's manifolds. Riemann clearly did not want there to be a notion of length intrinsic to his manifolds. His task was to *add* that in with the introduction of his relations of measure. A number manifold, however, is very rich in structure. There is a kind of notion of length built into it, for example, that of differences of coordinate values. And there is much more. There are preferred positions; the origin $(0, 0, \ldots, 0)$ is unique, for example. And there are preferred directions; for example, that of the distinctive coordinate axes.

For Klein the use of number manifolds was natural because of the immediate precision they brought. Since he favoured the subtractive strategy, he is untroubled by the extra structure. The subtractive strategy told him how to ignore all the structure that is deemed non-geometric. One must find the appropriate group—in this case the group of all transformations[7] —and work solely with its invariants. These are the geometrically real structures of the manifold.

6 In order to avoid confusion over my very specific use of term 'number manifold', I stress that a number manifold is a manifold whose point set literally *is* a set of n-tuples of numbers (real or complex). It is *not* merely a manifold that is topologically $\mathbf{R}^n$, that is, a topological space that can be coordinatized locally or globally by $\mathbf{R}^n$.

7 What determines the extent of 'all' is notoriously vague. Are continuity and differentiability conditions imposed? Are the transformations C^0 transformations?

Riemann's project was devoted to a different strategy of defining the geometrically real; I have called it his 'additive strategy'. Its whole point was to built up the geometry in two steps. The first defined the notion of n-fold extended manifold whose intrinsic properties were just continuity and dimensionality *and nothing more*. In particular, there is no notion of length intrinsic to these manifolds. The second step then added this geometric notion of length, which was introduced on a highly localized basis. It would be antithetical to this project to begin its first phase with a number manifold so rich in superfluous structure. The starting point of the project is a manifold free of metrical notions. Thus Riemann could not avail himself of the group and invariant strategy of discarding superfluous structure, because he was pursuing the implications of a vastly general vision of geometry.[8] He was reduced to the tortured gropings of his efforts to describe manifolds with just the amount of geometric structure he wanted and no more.

Special relativity and Klein's Erlangen Program

The primary burden of Einstein's 1905 special theory of relativity was to deny physical significance to the cosmic state of rest presumed by then current theories of the electromagnetic aether. To do this he showed in his celebrated 1905 'On the electrodynamics of moving bodies' that Maxwell's electrodynamics did satisfy a principle of relativity, as long as one recognized the group of transformations of the coordinates of space and time that properly represented the relativity of inertial motion. That group proved to be the Lorentz group and not the Galilean group assumed in classical physics. Einstein showed that the equations of Maxwell's electrodynamics retained their form under Lorentz transformation and embarked on the project of modifying the remainder of physics so that it too would satisfy this requirement of Lorentz covariance. As Einstein (1940, p. 329) later summarized it, this requirement was the essence of the theory:[9]

> The content of the restricted relativity theory can be summarized in one sentence: all natural laws must be so conditioned that they are covariant with respect to Lorentz transformation.

8 Editor's remark: It is also worth noting that the first rigorous analysis of the simplest type of magnitude, length on a line, and the identification of the line with the real number continuum, date from the work of Cantor and Dedekind. This work, let alone its general acceptance, dates from after Riemann's death; Riemann could reasonably have rejected the glib identification of the two current in his day.

9 An almost identical formulation appears in Einstein (1952, p.148), an appendix to his popular text on relativity theory: 'The whole content of the special theory of relativity is included in the postulate: The laws of Nature are invariant with respect to the Lorentz transformations.'

This covariance entailed that no state of rest could have physical significance, for no law embodying such a state of rest could retain covariance under the Lorentz transformation.

Einstein's strategy lay close to that of Klein's Erlangen Program. Einstein reserved physical significance for laws that remained unchanged by the transformations of the relevant group. Correspondingly Klein distinguished the geometric properties as those that remained unaltered by the transformations of the group of the geometry. This much of the similarity was evident in 1905. But there was a deeper connection to be found. It was already foreshadowed by Klein's pronouncements decades earlier on the Erlangen Program. In his original manifesto, Klein had all but reduced geometry to the study of the invariants of groups. He wrote (1872, p. 463):

> The following all embracing problem arises as the generalization of geometry: *Let there be given a manifold and a group of transformations on it; one should investigate the structure belonging to the manifold with regard to such properties as are not changed by the transformations of the group.* (Emphasis in original)

By giving the Lorentz group such privileged position in his theory, Einstein had brought it to the threshold of Klein's geometry. It was Hermann Minkowski, mathematician and colleague of Klein at Göttingen, who then took the theory past the threshold. In his famous popular lecture of 1908, he sought to motivate his new space–time approach to special relativity by reflecting on Newtonian mechanics. That theory's equations, he noted, exhibit a twofold invariance; they are unaltered under arbitrary change of position and under transition to arbitrary states of uniform motion. He continued (1909, p. 1)

> We are accustomed to look upon the axioms of geometry as finished with, when we feel ripe for the axioms of mechanics, and for that reason the two invariances are probably rarely mentioned in the same breath. Each of them by itself signifies, for the differential equations of mechanics, a certain group of transformations. The existence of the first group is looked upon as a fundamental characteristic of space. The second group is preferably treated with disdain, so that we with untroubled minds may overcome the difficulty of never being able to decide, from physical phenomena, whether space, which is supposed to be stationary, may not be after all in a state of uniform translation. Thus the two groups, side by side, lead their lives entirely apart. Their utterly heterogeneous character may have discouraged any attempt to compound them. But it is precisely when they are compounded that the complete group, as a whole, gives us to think.

Minkowski proceeded to treat both groups on a par by compounding them. The group of spatial translations was associated with the geometry of space. That group was expanded by the group of transformations between uniform states of motion. The result defined, in accord with Klein's prescription, a new geometry—but now it was not the geometry of space, but of space–time. The concern with groups and geometry led directly to a new notion of fundamental importance, space–time.[10]

Writing shortly afterwards, Klein was clearly delighted that his geometric methods had found application at the forefront of physics. To draw Minkowski's space–time formulation closer to his geometry, Klein (1910, p. 539) provided a brief scheme for translating between the relativity talk of the physicists and his preferred geometric language:

> What the modern physicists call 'relativity theory' is the theory of invariants of a four dimensional space–time-region, x, y, z, t (the Minkowski 'world') with respect to a particular group of collineations, namely the 'Lorentz group'; —or more generally, and turned round the other way:
> If one wants to make a point of it, it would be all right to replace the phrase 'theory of invariants relative to a group of transformations' with the words 'relativity theory with respect to a group.'

The generalization of special relativity: the taming of coordinate systems

The special theory of relativity was the first step for Einstein. He turned almost immediately to the task of finding a new theory that would extend the principle of relativity to acceleration. His first published step came in a 1907 review article (Einstein, 1907/08); after much labour and misadventure, the theory reached its final form in November 1915 and was summarized in a well-known review article the following year (Einstein, 1916). The nature and significance of the generalization achieved has been the subject of continued debate. Thus we must attend carefully to how Einstein understood his generalization.

Einstein's methods in special relativity—both before and after Minkowski—were algebraic: in geometry they were the methods of analytic geometry. His spaces or space–times are represented by algebraic variables, the coordinates of space, time or space–time. Thus his approach risked the danger that Klein had discussed decades before when introducing his Erlangen Program. Each coordinate system introduces arbitrary elements; we must be alert not to accord improper significance to them. In Einstein's coordinate formulations of special relativity, each coordinate system tacitly defined a state of rest-points whose spatial coordinates are constant are at rest according to that

10 In a recent study of Minkowski's work in relativity physics, Leo Corry (manuscript) has emphasized that the Erlangen Program played only a small part in Minkowski's efforts. His deeper motivation lay in the desire to axiomatize and thereby clarify physical theory. Corry also suggests that, by Minkowski's time, the connection of groups and geometry was a topic of general interest, not so strongly connected with the specifics of Klein's Erlangen Program. I will persist here, however, with special attention to Klein, since he was one of the most visible proponents of this viewpoint and he provides us very clear statements of the connection between groups and geometry and of the subtractive strategy.

coordinate system. The content of the principle of relativity is to deny physical significance to this coordinate-based notion of rest. Thus Einstein's formulations of the principle of relativity are propositions *about coordinate systems*. For example, his popular text, Einstein (1917, p. 13), puts it so:[11]

> If, relative to [Galileian coordinate system] K, K' is a uniformly moving co-ordinate system devoid of rotation, then natural phenomena run their course with respect to K' according to exactly the same general laws as with respect to K. This statement is called the *principle of relativity* (in the restricted sense). (Emphasis in original)

He proceeded to explain (p. 14) that, were this principle to fail, then all Galileian coordinate systems would cease to be equivalent for the description of natural phenomena.[12]

> In this case we should be constrained to believe that natural laws are capable of being formulated in a particularly simple manner, and of course only on condition that, from amongst all possible Galileian co-ordinate systems, we should have chosen *one* (K_0) of a particular state of motion as our body of reference. We should then be justified (because of its merits for the description of natural phenomena) in calling this system 'absolutely at rest', and all other Galileian systems K 'in motion'. (Emphasis in original)

Einstein used Klein's strategy to deny physical reality to this coordinate-based state of rest. Physical reality accrues only to that which is common to all Galileian coordinate system, that is, to the invariants of the Lorentz transformation.

In pursuing his quest for a generalized principle of relativity, Einstein's strategy remained exactly the same. The coordinate systems of special relativity still harboured illegitimate elements. In particular, each Galileian coordinate system defined a state of rest. The class of all such states of rest formed the class of inertial motions. These inertial motions were accorded physical reality in special relativity since the class as a whole remained invariant under Lorentz transformation. An extension of the principle of relativity to acceleration must deprive these inertial motions of their preferred status. This would be achieved by an expansion of the covariance group of his theory beyond the Lorentz group. Einstein (1919, p. 230) described his program in impassioned rhetorical questions:

> Should the independence of physical laws of the state of motion of the coordinate system be restricted to the uniform translatory motion of coordinate systems in respect to each other? What has nature to do with our

11 A 'Galileian' coordinate system is an inertial coordinate system.
12 Pedants will note that Einstein's argument is fallacious, but inessentially so for our purposes. The failure of the principle of relativity does not entail the existence of a single state of rest. The relativity of inertial motion may prevail in just one direction of space, for example.

> coordinate systems and their state of motion? If it is necessary for the purpose of describing nature, to make use of a coordinate system arbitrarily introduced by us, then the choice of its state of motion ought to be subject to no restriction; the laws ought to be entirely independent of this choice (general principle of relativity).

Success in the quest for the generalization of principle of relativity lay in the achievement of general covariance. Having introduced the notion of the arbitrary coordinate system of Gauss' theory of surfaces, Einstein (1917, p. 97) stated the principle that was satisfied by his general theory of relativity:

> The following statement corresponds to the fundamental idea of the general principle of relativity: 'All Gaussian coordinate systems are essentially equivalent for the formulation of the general laws of nature.'

Through its general covariance, Einstein's general theory of relativity had succeeded in denying physical reality to the arbitrary elements of Einstein's space–time coordinate systems. Only invariants of Einstein's arbitrary transformations[13] may have physical reality. The only elements of the coordinate systems that may have physical reality are their topological properties, such as their dimensionality and continuity.

Success and failure: Klein meets Riemann

While Einstein had been euphoric over his success in formulating a generally covariant gravitation theory, his interpretation of its general covariance was almost immediately challenged. The *locus classicus* of this challenge is Kretschmann (1917). It was the first major statement of a critical position that grew slowly from a minority opinion to a mainstream judgement. I have reviewed this critical tradition elsewhere in detail (Norton, 1993, 1995), so I need only mention here the two themes that pervade it. First, the achievement of general covariance is typically regarded as failing to contribute to the physical content of a theory. It proved easy to find generally covariant formulations of commonly known theories of space and time, both special relativity and Newton's, for example. Thus whatever physical content could be associated with general covariance would have to be a part of all these theories as well. Second, the tradition sought to dissociate covariance principles from relativity principles. The latter are symmetry principles expressing a uniformity of space–time structure. In standard formulations of special relativity, the covariance group just happens to be the same as the symmetry group of the Minkowski space–time. But this coincidence no long occurs in the case of general relativity. While the covariance group has grown to hold arbitrary transformations, the space–times of general relativity typically admit no

13 Einstein does not make clear precisely which group he intends.

symmetries at all. In many space–times, the symmetry group will just be the identity group.

So how are we to understand Einstein's achievement? Did his attainment of general covariance serve the sort of purpose he intended. In one sense, it did—and completely so. To see that sense, we need to recall the focus of Einstein's work, the coordinate system. I have gone to pains to show that Einstein's target was illegitimate elements in these coordinate systems. In attaining general covariance, Einstein did succeed in his aim of denying these elements physical reality. But that success was a narrow one. The inertial structure encoded in Einstein's coordinate systems is essential for his general theory of relativity; it defines the free fall trajectories of the theory. Had Einstein been successful in eliminating this structure from his theory, he would merely have a topological space as his space–time—too impoverished a structure to stand as a gravitation theory.

Einstein did not eliminate this inertial structure in his transition from special to general relativity. To see why, we need to review how Einstein proceeded from special to general relativity. We now tend to think of special relativity as the theory of a Minkowski space–time—a space–time endowed with a flat, Lorentz signature metric. But Einstein's 1905 Lorentz covariant formulation of special relativity made no *explicit* mention of such a metric. Its structures were there implicitly, of course. Indeed they could be recovered directly from the covariance group: the Minkowski line element of Minkowski's space–time formulation (where the coordinates (t, x, y, z) have the usual meaning)

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \tag{5.1}$$

is an invariant of the Lorentz group and thus automatically accorded physical reality in the theory. In the transition to a generally covariant formulation, this Minkowski line element (5.1) would lose physical reality, since it would cease to be an invariant of the allowed coordinate transformations. But then the metrical structure it represents would be lost. So Einstein replaced the Minkowski line element (5.1) with the familiar

$$ds^2 = g_{ik}\, dx^i\, dx^k \tag{5.2}$$

where the coefficient g_{ik} of the space–time metric transform covariantly accordingly to the rules for the transformation of covariant, second rank tensors. This insertion of an explicit space–time metric preserves the inertial structure (and more) that a transition to general covariance would otherwise have denied physical reality.

However, in formulating a space–time theory based on an invariant line element (5.2), Einstein had left the geometric tradition of Klein and entered that of Riemann. Klein's group-based method of discerning the physically real no longer functions without significant modification. Therein lie the interpretative woes that Einstein brought on his head. Of course Klein's subtractive strategy still works as far as the coordinate systems are concerned. Their physically real properties are just the invariants of the theory's group. Since that group admits arbitrary transformations, the properties are merely topological. But in the Riemannian program, this topological space is the first step in the construction of the space. Physically or geometrically real structure is then added to this space by means of a quadratic differential form—in the case of general relativity, the line element (5.2). Because of the way this line element behaves under coordinate transformation, it is immune to the subtractive strategy.

One way to see this breakdown of Klein's subtractive strategy is to compare two versions of special relativity each formulated in the manner of Riemann. The first is Lorentz covariant; its metrical structure is introduced by positing the line element (5.1). The second is generally covariant; its metrical structure is introduced by positing the line element (5.2). (The g_{ik} must here represent a flat metric, of course.) In Klein's approach, an expansion of the group of a geometry has the effect of subtracting from the physically real. But the physical content of the line elements (5.1) and (5.2) remain unaltered by the expansion of the theory's group from the Lorentz group in (5.1) to the general group in (5.2); the alteration of the line element by the introduction of the matrix of coefficient g_{ik} has the effect of protecting the content of (5.1) from the subtractive stratagem. A simpler way to see the failure of Klein's strategy is this. In a Lorentz covariant theory, in Klein's approach one need not posit the line element (5.1). If the geometric structure of the theory is the invariants of the group, then one can derive this line element. The same thinking is of no avail in general relativity. That theory's group is the general group; but that group does not define for us a unique line element (5.2).

In sum, the transition from special to general relativity is not marked by the overall denial of physical reality to theoretical structures; rather it is characterized by a re-organization in the methods used to represent the physically real. Coordinate systems are deprived of all but topological properties as physically real—this according to the subtractive strategy of Klein; but many of these properties are reintroduced by the explicit introduction of a quadratic differential form—this according to the additive strategy of Riemann.

Number manifolds and Einstein's causal argument

This last summary is an unforgiving over-simplification of Einstein's treatment of the physically real in relativity theory. It ignores at least two circumstances: the first is a matter of historical contingency in the history of mathematics, the second is what would probably be Einstein's own response to this summary.

Riemann's spaces have two parts: a manifold and a quadratic differential form. The manifold is required to have just topological properties. We have seen how Riemann laboured to explain the notion of a manifold with just these properties. As late as the early 1910s, when Einstein developed general relativity, a mathematically clean description of such manifolds was not part of the standard repertoire of applied mathematics. When a precise mathematical representation of a manifold was needed, one proceeded as Riemann eventually had to in his inaugural address: one introduced a number manifold such as $\mathbf{R}^n$. It was quite common as late as the 1920s to define a manifold as $\mathbf{R}^n$.[14] This was certainly the practice that Einstein found in Minkowski's geometrical formulation of special relativity, for he would have heard in Minkowski's (1908, p. 76) famous popular lecture:[15]

> We will try to visualize the state of things by the graphic method. Let x, y, z be rectangular co-ordinates for space and let t denote time.... A point of space at a point of time, that is, a system of values x, y, z, t, I will call a 'world-point'. The manifold of all thinkable x, y, z, t systems of values we will christen 'the world'.

Minkowski's 'world' is the system of thinkable values of the four real-valued coordinates x, y, z, t; that is, it *is* $\mathbf{R}^4$.

This use of number manifold brought the familiar risk of arbitrary elements. Klein's subtractive strategy controlled them for Minkowski: only those elements of his number manifolds that were invariant under Lorentz transformation could claim physical reality. This problem persisted as Einstein proceeded to general relativity. He needed to supply his theory with a manifold suitable for the construction of a space–time by Riemann's methods. That is, the manifold had to have only topological properties so that an appropriate line element could be introduced to carry the metrical structure. Since Einstein, like Minkowski, continued to use number manifolds—in effect his coordinate systems—when a manifold was needed, he had to be sure that he denied physical reality to their arbitrary elements. Coordinate differences could not represent physical times elapsed or distances traversed; straights in the number manifolds could not be free fall trajectories. Such times, distances and free fall trajectories were all to be defined by the line element. So it was entirely appropriate for Einstein to invoke Klein's subtractive strategy and deny physical reality to all those elements of his number manifolds that were not invariant under the arbitrary transformations of his theory.

In hindsight, it seems so clear that the structure denied physical reality in the number manifolds was merely reintroduced in the line element (5.2). They had not been eliminated; they had been relocated. I am fairly sure how Einstein would respond to this assessment. He would object that the structure reintroduced in the line element has been decisively altered in its physical

14 Such, for example, is the definition given in Levi-Civita (1925, p. 1). For a more extensive discussion of the development of the notion of manifold see Norton (1989, Section 3).
15 I have replaced the Perrett and Jeffrey translation of Minkowski's *Mannigfaltigkeit* as 'multiplicity' with 'manifold'.

properties—and in this alteration we find the true advance of general relativity. He explained this in a 1954 appendix to his text on relativity theory (Einstein, 1953, pp. 139–40):

> It is the essential achievement of the general theory of relativity that it has freed physics from the necessity of introducing the 'inertial system' (or inertial systems). This concept is unsatisfactory for the following reason: without any deeper foundation it singles out certain co-ordinate systems among all conceivable ones. It is then assumed that the laws of physics hold *only* for such inertial systems (e.g. the law of inertia and the law of the constancy of the velocity of light). Thereby, space as such is assigned a role in the system of physics that distinguishes it from all other elements of physical description. It plays a determining role in all processes, without in its turn being influenced by them. Though such a theory is logically possible, it is on the other hand rather unsatisfactory. (Emphasis in original)

That is, he would agree that the transition to general covariance had deprived certain properties of coordinate systems of physical reality, most notably any association with coordinate straights and the free falls of inertial motion. But he would insist that the inertial structure restored through the line element was of a quite different kind. It responds dynamically through the Einstein field equations to the physical content of space–time. Thus this inertial structure is deflected towards the Sun by the Sun's mass. The inertial structure associated with coordinate systems is not influenced causally by the processes occurring within space–time; the inertial structure of the line element (5.2) of general relativity is so influenced. This change, Einstein assures us, is the essential achievement of general relativity.

These causal considerations were not the last-minute excuses of an Einstein seeking to repair an ailing interpretation. They persist throughout the corpus of his writing on general relativity. (See Norton, 1993, 3.9). In so far as we can find a satisfying account of Einstein's causal concerns and its relation to a generalized principle of relativity, it is through Anderson's notion of the 'absolute object'. However, the interpretative program based on this notion remains controversial. (See Norton, 1993, Section 8; 1995, Section 6).

Conclusion

In his general theory of relativity, Einstein bequeathed us an uncomfortable mix of strategies for discerning the physically real. He used Klein's subtractive strategy to deprive his coordinate systems of all but topological properties. He used Riemann's additive strategy to locate metrical properties in a quadratic differential form of his invariant line element. His proclamations on the physical content of the theory, however, seem only to acknowledge the import of the first strategy: the expansion of the covariance group deprives much mathematical structure of physical reality. But they seem to ignore that much of what was denied physical significance is restored by its incorporation into the line element. This tension has defined the ensuing debate over the correct understanding of Einstein's relativity principles.

References

Cartan, E. (1927) *L'Ensignment Mathématique*, **26**, 200–225. Reprinted (1952) *Oeuvres Complétes* Part 1, Vol. II, pp. 841–866, Gauthier-Villars, Paris.

Corry, L. (manuscript) Hermann Minkowski and the postulate of relativity.

Einstein, A. (1907, 08) Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen, *Jahrbuch der Radioaktivitt und Elektronik*, **4** (1907), 411–462; **5** (1908), 98–99.

Einstein, A. (1916) Die Grundlage der allgemeinen Relativitätstheorie, *Annalen der Physik*, **49**, 769–822. Translated (1952) without p. 769 as 'The foundation of the general theory of relativity,' pp. 111–164 in H.A. Lorentz *et al.*, *The principle of relativity*, Dover, New York.

Einstein, A. (1917) *Relativity: the special and the general theory*. Translated (1977) by R.W. Lawson, Methuen, London.

Einstein, A. (1919) What is the theory of relativity. In *Ideas and opinions*, pp. 227–232, Bonanza, New York.

Einstein, A. (1940) The fundamentals of theoretical physics. In *Ideas and opinions*, pp. 323–335, Bonanza, New York.

Einstein, A. (1952) Relativity and the problem of space. Appendix V to Einstein (1917).

Einstein, A. (1953) *The meaning of relativity*, Princeton University Press, Princeton.

Fock, V. (1959) *The theory of space time and gravitation*, Translated by N. Kemmer, Pergamon, New York.

Klein, F. (1872) *Programm zum Einstritt in die philosophische Facultät und den Senat der k. Friedrich-Alexanders-Universität zu Erlangen*. (Erlangen, A. Deichert). Presented (1921) as 'Vergleichende Betrachtungen über neuere geometrische Forschungen,' in F. Klein, *Gesammelte Mathematische Abhandlungen: Vol. 1 Linien Geometrie, Grundlegung der Geometrie zum Erlanger Programm*, pp. 460–497, R. Fricke and A. Ostrowski (eds) Julius Springer, Berlin. Citations to and quotes from the latter.

Klein, F. (1927) *Vorlesungen Über Nicht-Euklidische Geometrie*, Chelsea, New York.

Kretschmann, E. (1917) Über den physikalischen Sinn der Relativitätspostulat, A Einsteins neue und seine ursprünglische Relativitätstheorie, *Annalen der Physik*, **53**, 575–614.

Levi-Civita, T. (1925) *Calcolo differenziale absolute*, Bologna Zanichelli. English translation by E. Persico, *The absolute differential calculus*, Blackie, London and Glasgow. Reprint, Dover, New York, 1977.

Minkowski, H. (1909) Raum und Zeit, *Jahresbericht der Deutschen Mathematiker-Vereinigung*, **18**, 75–88; (1908) Space and time. In A. Einstein *et al.*, *The principle of relativity*. Dover, 1952.

Norton, J.D. (1989) Coordinates and covariance: Einstein's view of spacetime and the modern view, *Foundations of Physics*, **19**, 1215–1263.

Norton, J.D. (1992) The physical content of general covariance. In J. Eisenstaedt and A. Kox (eds), *Studies in the history of general relativity: Einstein studies*, (Vol. 3), Birkhauser, Boston.

Norton, J.D. (1993) General covariance and the foundations of general relativity: eight decades of dispute, *Reports on Progress in Physics*, **56**, 791–858.

Norton, J.D. (1995) Did Einstein stumble: the debate over general covariance, *Erkenntnis*, **42**, 223–245. Volume reprinted (1995) as *Reflections on spacetime: foundations, philosophy, history*, U. Maier and H.-J. Schmidt (eds), Kluwer, Dordrecht.

Riemann, B. (1854) On the hypotheses which lie at the foundations of geometry. Translated (1959) by H.S. White as pp. 411–425 in D.E. Smith (ed.), *A source book in mathematics*, Dover, New York.