General Relativity and Gravitation, Vol. 1, No. 4 (1971), pp. 393–400.

Stable and Generic Properties in General Relativity

S. W. HAWKING

Institute of Theoretical Astronomy, University of Cambridge

Read on 18 *May* 1970 *at the Gwatt Seminar on the Bearings of Topology upon General Relativity*

In this paper I want to discuss some aspects of the space of all Lorentz metrics. In particular I shall consider what it means to say that a certain property is stable or generic and shall give some examples.

One would like to have a name for the space L of all Lorentz metrics g on a four-dimensional manifold M. As Professor Wheeler has already introduced the name 'Superspace' for the space of all positive definite metrics on a three-dimensional manifold, I would tentatively suggest 'Metaspace' for L. There are a number of different topologies that can be placed on L; which one uses will depend on the properties one wishes to consider. The topologies differ in how many derivatives of a metric have to be 'near' to those of another metric for the two metrics to be considered 'near' to each other and in what region they are required to be near.

The derivatives of a tensor field (such as a Lorentz metric) on a manifold M are most elegantly described by the bundle of jets over M (see Palais [*1*]). However, I shall use a simple and less sophisticated approach. I put a positive definite metric e on M (this can always be done). This metric can be used to define covariant derivatives of tensor fields on M and also to measure the magnitude of such tensor fields and their derivatives. Thus one can define how near together the derivatives of two metrics are at each point of M. There is also the problem of the regions on which the metrics are required to be near. This is really a question of how the metrics behave near the edge of the manifold, i.e. near infinity.

There are three main possibilities:

1. The metrics can be required to be near only on compact regions of the manifold (Fig. 1). The behaviour near infinity is unrestricted. More precisely, if g is a Lorentz metric, U a compact set of M and ϵ_i $(0 \leqslant i \leqslant r)$ a set of continuous positive functions on M, the neighbourhood $B(U, \epsilon_i, g)$ of g can be defined as the set of all Lorentz metrics whose ith derivatives $(0 \leqslant i \leqslant r)$ differ from those of g by less than ϵ_i on U. The set of all such $B(U, \epsilon_i, g)$ for all U, ϵ_i and g form a sub-basis for the C^r *compact-open* topology for L, i.e. the open sets in this topology are unions and finite intersections of the $B(U, \epsilon_i, g)$.

2. The requirement that the sets U should be compact can be removed

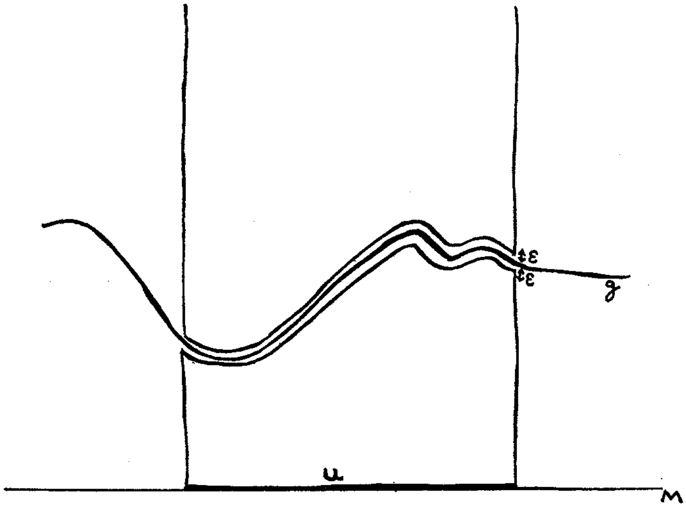

Figure 1—A neighbourhood $B(U, \epsilon, g)$ of the metric g in the compact open topology consists of all metrics which lie within $\epsilon(x)$ of g over the compact set U of M.

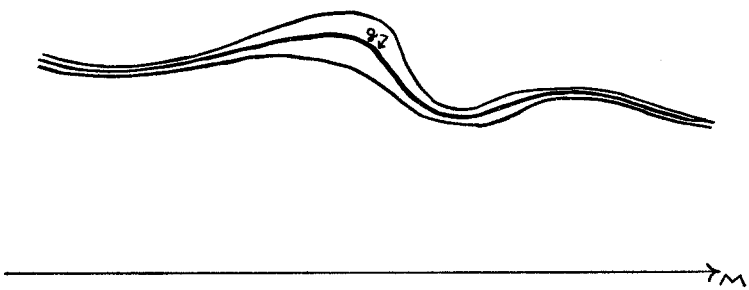

Figure 2—A neighbourhood $B(U, \epsilon, g)$ of the metric g in the open topology consists of all metrics which lie within $\epsilon(x)$ of g over the set U of M. Note that U can equal M, and that ϵ can go to zero at infinity.

and U can be taken to be M. This means that 'nearby' metrics must be nearby everywhere and must have the same limiting behaviour at infinity (Fig. 2). One may call this the *open* topology for L.

3. Define the set $F(U, \epsilon_i, g)$ as the set of all metrics whose ith derivatives differ from those of g by less than ϵ_i and which coincide with g outside the compact set U. The neighbourhood $B(\epsilon_i, g)$ is then defined as the union of the $F(U, \epsilon_i, g)$ for all compact sets U. The neighbourhoods $B(\epsilon_i, g)$ form a sub-basis for the *fine topology* on L. As its name suggests, the fine topology is finer than the open topology which in turn is finer than the compact–open topology. In other words, there are more open sets in the fine topology than in the open topology and still more than in the compact–open topology.

As well as the C^r topologies, one can use the Sobolev W^r topologies, these differ from the C^r topologies in that instead of requiring the difference between the derivatives to order r of two nearby metrics to be pointwise small, they require the integrals of the squares of these differences to be small (the squares and the integrals are here defined with respect to the positive definite metric e on M). Clearly a C^r tensor field is also a W^r field and it follows from a fundamental lemma of Sobolev that a W^{r+3} field in four dimensions is a C^r field. This means that a W^{r+3} topology is finer than the corresponding C^r topology which in turn is finer than the W^r topology. The W^r topologies play a fundamental role in the Cauchy problem in general relativity.

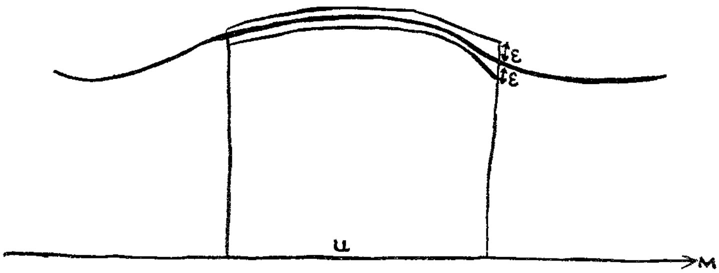

Figure 3—A neighbourhood $B(\epsilon, g)$ of the metric g in the fine topology consists of all the metrics which lie within ϵ of g over some compact set U and coincide with the metric g outside U.

Having put those topologies on the space of all Lorentz metrics, one can define what it means to say that a property of a metric is stable: a property P of a metric g is *stable* in a given topology on L if in that topology there is an open neighbourhood of g, every metric of which has the property P, i.e. if every sufficiently nearby metric has the property P. The reason for considering stable properties is this. A physical theory is a correspondence between certain physical observations and a mathematical model (in this case a manifold with Lorentz metric). The accuracy of the observations is always limited by practical difficulties and by the uncertainty principle. Thus the only properties of space-time that are physically significant are those that are stable in some appropriate topology. Other unstable properties will not have any physical relevance but may be of mathematical inconvenience in that they may provide counter examples to general theorems one would like to prove about *all* metrics in a certain region of L, i.e. the theorem may hold for *almost all* metrics in the region but fail for some particular metrics. One can say that such a theorem holds *generically* or that a property is *generic* in a region of L if it holds *almost everywhere* on that region (by 'almost everywhere' I mean that it holds on an open dense outset of the region of L). For physical purposes it is sufficient to prove that a theorem holds generically because the metric of the mathematical model for space-time is defined with only limited accuracy.

A given property may be stable or generic in some topologies and not in others. Which of these topologies is of physical interest will depend on the nature of the property under consideration as will be seen in the examples I shall give. Roughly speaking, if one is concerned with structure in a bounded region of space-time then the appropriate topology is the compact open but if one is interested in statements about the existence or non-existence of something everywhere in space-time, one should use the open or the fine topology in order to restrict the behaviour of the metric near infinity. As the open topology is coarser than the fine topology it is a stronger requirement on a property for it to be stable in the open topology than in the fine topology and still stronger for it to be stable in the compact–open topology.

As an illustration I shall discuss *stable causality*. Ordinary causality can be defined as the absence of closed timelike curves. If there were such curves, one could in theory travel round them and arrive in one's past. The logical difficulties that could arise from such time travel are fairly obvious: for example, one might kill one of one's ancestors. These difficulties could be avoided only by an abandonment of the idea of free-will: by saying that one was not free to behave in an arbitrary fashion if one travelled into the past. This is not something which it is very easy to accept, however, and it seems more reasonable to believe that there are no closed timelike curves. As well as actually closed timelike curves, it would seem reasonable to exclude 'almost closed' timelike curves, i.e. to require that there should be no point p such that every small neighbourhood of p intersects some timelike curve more than once. A metric with this property is said to satisfy *strong causality*. However, even strong causality is not enough to ensure that space-time is not on the verge of violating causality as is shown by the example of Fig. 4 in which a strip of two-dimensional Minkowski space has been identified along the edges to form a cylinder and three 'baffles' have been cut out of the space to prevent there from being any closed or almost closed timelike curves. Nevertheless there are timelike curves which pass arbitrarily close to other timelike curves which then come arbitrarily close to the first curves. In fact Brandon Carter has shown that there is a whole hierarchy of higher causality conditions corresponding to different numbers of baffles and to different numbers of limiting processes. This hierarchy is more than countably infinite: one can define an $(\infty+1)$th causality condition, and $(\infty+2)$th condition and so on. However, one can define an ultimate causality condition which is stronger than all this hierarchy and which corresponds to space-time not being on the verge of violating causality: a metric g is said to satisfy the stable causality condition if, in the C^0 open topology on L, there is an open neighbourhood of g no metric of which has closed timelike curves. In other words one can vary g by a small amount everywhere, without introducing closed timelike curves. One has to use the open rather than the compact open topology in the definition since in the compact open topology any open neighbourhood of any metric g contains a metric g in which there are closed timelike curves. This is because a neighbourhood of g consists of all the metrics which are near g on a compact set U. However, outside U they can differ by an arbitrary amount and so can admit closed timelike curves.

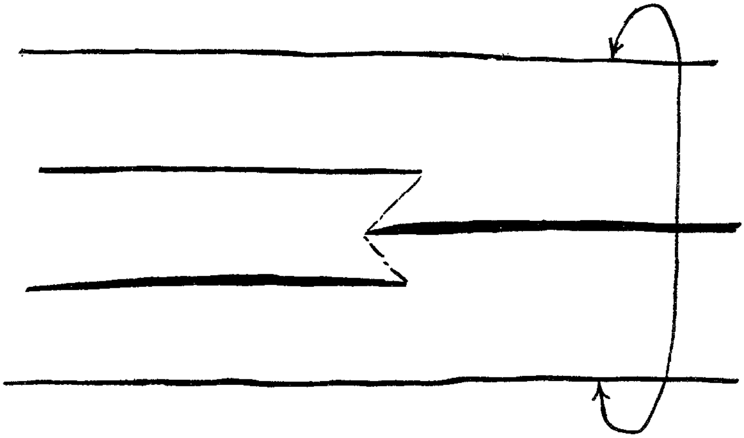

Figure 4.—A space which is on the verge of violating causality but contains no closed timelike curves and no almost closed timelike curves.

One could also define stable causality using the fine topology. I think that this would be definitely weaker than using the open topology, that is, I think that a metric which satisfied Carter's ∞th causality condition but not (say) the $(\infty+2)$th condition would be stably causal in the fine but not in the open topology. The open topology is probably more physical than the fine topology since to establish that two metrics actually coincide outside some compact set would require an exact measurement which is not physically possible. The definition of stable causality with the open topology has the further advantage of being related to another physically significant property, the existence of cosmic time functions. By a *cosmic time function* I mean a smooth function t which increases along every future directed timelike or null curves. The spacelike surfaces of constant value of such a function can be regarded as surfaces of simultaneity in the universe though, of course, they are not unique. One can show (Hawking [*2*]) that a metric admits such cosmic time functions if and only if it is stably causal in the C^0 open topology. It follows incidentally from this result that stable causality defined with the C^∞ open topology is equivalent to that defined with the C^0 open topology.

The region in L on which stable causality holds is the interior of the region on which ordinary causality holds. Since the region on which ordinary causality is violated is open, the union of this region with the region on which stable causality holds is an open dense set in L. It thus is generic for a metric either to violate causality or to be stably causal.

I would conjecture (but have not proved) that it is generic for a metric satisfying ordinary causality also to satisfy stable causality, i.e. that the stably causal metrics are dense in the causal metrics.

A rather different kind of stability occurs in the Cauchy Problem for General Relativity. If the solutions of the Einstein equations did not depend continuously in some sense on the initial data, it would be physically useless to make predictions since one could never know the initial data exactly. It turns out that there is a sense in which this dependence is continuous but there seem to be one or two slightly odd features which I shall try to describe.

In order to discuss the stability of the Cauchy Problem I shall adopt the approach of Choquet-Bruhat and Geroch [3]. An *initial data set* (S, W) is defined to be a three-dimensional manifold S together with two symmetric tensor fields h^{ab} and χ^{ab} on S. The fields h and χ are required to obey four constraint equations which can be expressed as tensor equations in S with covariant derivatives defined by the three-dimensional metric h. The Cauchy Problem for empty space then consists of finding a development (M, g, θ) for (S, W). A development (M, g, θ) is a four-dimensional manifold M, a Lorentz metric g which satisfies the empty space Einstein equations and an imbedding $\theta : S \to M$ which is such that $\theta(S)$ is a Cauchy surface for M in the metric g and such that the first and second fundamental forms of $\theta(S)$ are h and χ respectively. One can define initial data sets and developments for the Einstein equations with matter in a similar way. The manifold M is diffeomorphic to $S \times R^1$ and the imbedding θ can be chosen to identify S with $S \times 0$. The metric g of the development can be unique only up to isometries which leave $\theta(S)$ pointwise fixed, i.e. two metrics are equivalent if there is a diffeomorphism $\theta : M \to M$ which carries one metric into the other. The Cauchy Problem thus becomes a problem of finding a map from the set of initial data sets to the set of equivalent classes under diffeomorphisms of metrics on $S \times R^1$. One would like to know whether this map is unique and continuous in an appropriate topology. In fact Choquet-Bruhat and Geroch have shown that there is a unique maximal development of Sobolev class W^r if the initial data is of class W^r and $r \geqslant 5$ (by the initial data being of class W^r I mean that h is a W^r tensor field on S, and χ, which represents first derivates of g, is a W^{r-1} field). I have slightly improved this result by lowering r to 4. I discovered something rather curious, however: I was not able to prove that a C^∞ initial data set had a C^∞ development. I could prove that it had a W^r and therefore a C^{r-3} development for every r but it seemed possible that these developments might get smaller as r got larger (Fig. 5), and that there might not be any region on which g was C^∞.

To describe the sense in which the Cauchy Problem is stable I shall define $\tilde{L}$ to be the space of equivalent classes of metrics on $S \times R^1$, that is, $\tilde{L}$ is the quotient of L by the group of diffeomorphisms of $S \times R^1$ which leave $S \times 0$ fixed. The space $\tilde{L}$ inherits topologies from L in a natural way under the quotient map $\pi : L \to \tilde{L}$ which assigns a metric to its equivalence class, i.e. a set V is open in $\tilde{L}$ if and only if $\pi^{-1}(V)$ is open in L. The topologies in the space I of pairs (h, χ) of initial data fields on S can be defined in a manner similar to those for L. Since χ represents first derivatives, the derivative conditions on it should be one order lower than those on h. Thus by two sets of initial data being near in the W^r topology, I shall mean that the h's are near in the W^r sense and the χ's are near in the W^{r-1} sense.

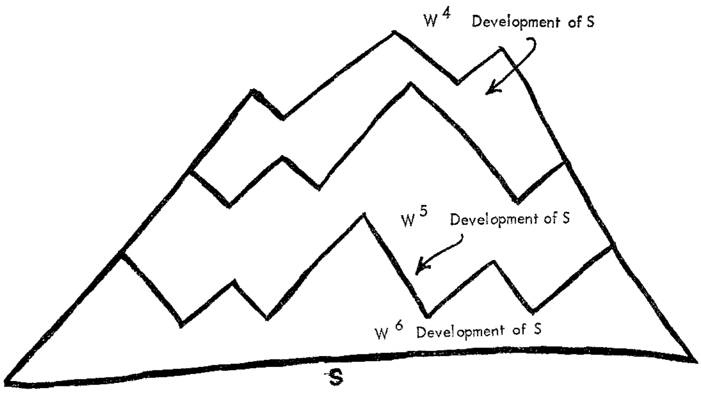

Figure 5.—The Cauchy development of a surface S with C^∞ Cauchy data.

One would not expect the map from the initial data I to the metrics $\tilde{L}$ to be continuous in general in the open or fine topology since one would expect that a small change in the initial data might produce a small change in the limiting value at infinity of the metric of the development. I have been able to prove that the map from the W^r initial data I to the W^r metrics $\tilde{L}$ is continuous in the W^{r-1} compact–open topologies in both spaces. It may be continuous in the W^r topologies but I have not been able to prove this.

One can also ask whether various properties of the developments are stable or generic in terms of the topologies on the space of initial data I. To conclude I shall give some results for the case where S is compact. These follow from theorems in Hawking [4], and Hawking and Penrose [5]. Since S is compact, it has no edge and so the compact–open, the open and the fine topologies on I are all equivalent.

1. It is generic for the maximal development to be geodesically incomplete.

2. If the maximal development can be imbedded in a larger space-time for which S is not a Cauchy surface, then it is generic for all such extensions either to be geodesically incomplete or to violate causality.

3. There is an open region of I for which all such extensions are geodesically incomplete.

References

1. Palais, R. (1968) *Foundations of Global Non-linear Analysis*, Benjamin.
2. Hawking, S. W. (1968) *Proc. Roy. Soc*, A., **308**, 433.
3. Choquet-Bruhat, Y. and Geroch, R. (1969) *Comm. Math. Phys.*, **14**, 329
4. Hawking, S. W. (1967) *Proc. Roy. Soc*. A., **300**, 187.
5. Hawking, S. W. and Penrose, R. (1970) *Proc. Roy. Soc*. **314**, 529.