Existence of instantaneous Cauchy surfaces[a)]

Robert Budic, Mark J. Gotay, and Robert H. Gowdy[b)]

Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742
(Received 12 May 1977; revised manuscript received 8 February 1978)

Several properties of instantaneous Cauchy surfaces are obtained. It is shown that a strongly causal spacetime admits an instantaneous Cauchy surface through each of its points, that there is a close and reversible relationship between these surfaces and maximal open globally hyperbolic subsets, that every instantaneous Cauchy surface is contained in a maximal instantaneous Cauchy surface, and that the latter surface is a maximal achronal surface which separates spacetime into past, present, and future. Some other properties of instantaneous Cauchy surfaces are discussed along with a refinement of an earlier topology change property.

I. INTRODUCTION

The preceding paper defined an instantaneous Cauchy surface to be an achronal set whose interior Cauchy development is maximal on the family of all such sets.[1] Several examples were considered, and it was argued that instantaneous Cauchy surfaces may have an important role to play in analyzing the structure of singular spacetimes and quantizing fields on such spacetimes.

Since one can construct spacetimes in which there are no nonempty achronal sets whatsoever, some restriction on the causal structure of spacetime is necessary if the spacetime is to admit the existence of an instantaneous Cauchy surface. The main result of this paper is an existence theorem which shows that instantaneous Cauchy surfaces can be found in all but the most pathological spacetimes.

In Sec. II we establish our notation and collect some well-known facts about globally hyperbolic sets. The main existence theorem is stated and proved in Sec. III. Section IV considers the properties of maximal and minimal instantaneous Cauchy surfaces. It is found, for example, that any instantaneous Cauchy surface is contained in a maximal instantaneous Cauchy surface which is also a maximal achronal set. Section V shows that a maximal instantaneous Cauchy surface is edgeless and that whenever two such surfaces have the same interior Cauchy development, they are homeomorphic.

II. NOTATION AND USEFUL FACTS

The notation of this paper is chosen to be compatible with the monograph by Hawking and Ellis[2] and also with the monograph *Techniques of Differential Topology* by Penrose.[3] One slight departure from the Hawking and Ellis notation is that we only require Cauchy surfaces and partial Cauchy surfaces to be achronal instead of acausal. Domains of dependence, determined by timelike curves, are denoted by $\tilde{D}(S)$ as in Hawking and Ellis. The set that we almost always use is the interior Cauchy development $\mathrm{int}\tilde{D}(S)$ so that it is convenient to denote this set by $D^0(S)$.

Global hyperbolicity can be defined in several ways. We use the definition in Hawking and Ellis[2]:

Definition: A set N is said to be *globally hyperbolic* if the strong causality assumption holds on N and if, for any two points $p,q \in N$, $J^+(p) \cap J^-(q)$ is compact and contained in N.

With this definition, the global hyperbolicity of N is tied to the causal structure of the whole spacetime M. In particular, we will use the following obvious consequence of this definition:

(GH0) If N is a globally hyperbolic subset of a spacetime M and S is achronal relative to N, then S is achronal relative to M.

We list below some additional properties of Cauchy developments and globally hyperbolic sets which will be used in what follows. We assume that M is a time-orientable spacetime.

(GH1) For every achronal set S, $D^0(S)$ is globally hyperbolic.[4]

(GH2) Each open globally hyperbolic set H, considered as a spacetime, contains a Cauchy surface S.[5] In H, $H = D^0(S)$. In M, $H \subset D^0(S)$.

(GH3) Each open globally hyperbolic set can be foliated by the achronal sets of (GH2).[6]

(GH4) If H is open and globally hyperbolic, and S and S' are as in (GH2), then S and S' are homeomorphic.[7]

III. EXISTENCE

Definition: An achronal set S is an *instantaneous Cauchy surface* if and only if, for any achronal set S', $D^0(S) \subseteq D^0(S')$ implies $D^0(S) = D^0(S')$.

Theorem 1: If M is a strongly causal spacetime and p is a point of M, then there is an instantaneous Cauchy surface that includes p.

Proof: Let G_p be the family of all open globally hyperbolic subsets of M containing p. By strong causality, p has an open globally hyperbolic neighborhood so that G_p is not empty. Partially order G_p by set inclusion and let $\{G_\alpha \mid \alpha \in A\}$ be a totally ordered subfamily of G_p. The set $\cup_\alpha \{G_\alpha\}$ is open and globally hyperbolic and is an upper bound for the subfamily $\{G_\alpha\}$. By Zorn's lemma there exists a maximal member H of the family

[a)]Supported in part by the National Science Foundation Grant No. MPS 74-18386-A01.
[b)]Supported in part by the Alfred P. Sloan Foundation.

 0022-2488/78/1907-1580$1.00

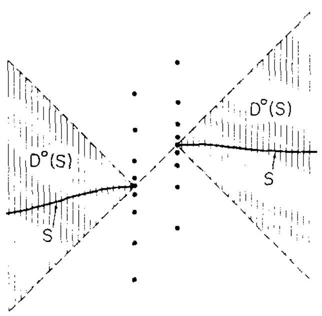

FIG. 1. A nonmaximal instantaneous Cauchy surface. The example consists of two-dimensional Minkowski spacetime with two sequences of points removed so that their limit points (also removed) are null related. The surface S is an instantaneous Cauchy surface because the interior Cauchy development of any other achronal surface will necessarily encounter one of the removed points before it can include the development of S. Larger instantaneous Cauchy surfaces can be obtained by adjoining to S points on the null line (dotted) that connects the limit points. These additional points do not contribute to the interior of the Cauchy development at all. Notice that the resulting larger instantaneous Cauchy surfaces may not be submanifolds because they can include isolated points.

G_p. Now $p \in H$, so by (GH3) and (GH0) we can find an achronal set S through p such that $H \subseteq D^0(S)$. Since H is maximal $D^0(S) = H$ and S is the desired instantaneous Cauchy surface through the point p.

Remark: A slight weakening of both the hypothesis and the conclusion of the existence theorem is possible. In a past(future)-distinguishing spacetime, one can show that an instantaneous Cauchy surface passes through each neighborhood of each point.[8] The method of proof is the same as above except for one point: It must be established that each neighborhood is intersected by an open globally hyperbolic set. This point can be established by a straightforward local construction.

In the process of proving the existence theorem, a connection between instantaneous Cauchy surfaces and maximal open globally hyperbolic sets has been intimated. In fact, for some purposes, one may be more interested in the maximal open globally hyperbolic subsets of a spacetime than in the instantaneous Cauchy surfaces. For this reason, it is useful to state the exact nature of this connection.

Proposition 1: An achronal set S is an instantaneous Cauchy surface if and only if $D^0(S)$ is a maximal open globally hyperbolic set.

Proof: Suppose that H is an open globally hyperbolic set such that $D^0(S) \subseteq H$. By (GH0) and (GH2) one can find some achronal set S' with $D^0(S) \subseteq H \subseteq D^0(S')$. If S is an instantaneous Cauchy surface, then $D^0(S) = H$ and $D^0(S)$ is a maximal open globally hyperbolic set.

From this result and (GH3), we see that the existence of instantaneous Cauchy surfaces is equivalent to the existence of maximal open globally hyperbolic subsets. We now restate the existence theorem in terms of these subsets.

Proposition 2: A strongly causal spacetime is covered by its maximal open globally hyperbolic subsets.

IV. MAXIMAL AND MINIMAL INSTANTANEOUS CAUCHY SURFACES

If one is really interested in the "best possible achronal sets" in a spacetime, then the instantaneous Cauchy surfaces are not the last word. It is possible for one instantaneous Cauchy surface to be a proper subset of another. Figure 1 shows a simple example of this behavior. The interesting achronal sets, we find, are the maximal, $S_{\max}$, and the minimal, $S_{\min}$, instantaneous Cauchy surfaces. These sets are easily constructed from a given instantaneous Cauchy surface.

Proposition 3: If M is a strongly causal spacetime and S is an instantaneous Cauchy surface in M, then

(A) $S_{\min} = \cap_\alpha S_\alpha$, where $\{S_\alpha\}$ is the set of instantaneous Cauchy surfaces contained in S.

(B) $S_{\max} = \sim I(S)$, where $I(S) := I^+(S) \cup I^-(S)$.

Proof: (A) First we show that $\cap_\alpha S_\alpha$ is an instantaneous Cauchy surface. For each α, $S_\alpha \subseteq S$ and so $D^0(S_\alpha) \subseteq D^0(S)$. Since each S_α and S is an instantaneous Cauchy surface, $D^0(S_\alpha)$ and $D^0(S)$ are maximal globally hyperbolic sets by Proposition 1. Thus $D^0(S_\alpha) = D^0(S)$ for each α. Consequently, if $p \in D^0(S)$, every timelike curve through p must intersect S_α for each α and so every timelike curve through p must intersect $\cap_\alpha S_\alpha$. It follows then that $D^0(S) \subseteq D^0(\cap_\alpha S_\alpha)$ and so from the maximality of $D^0(S)$, $D^0(S) = D^0(\cap_\alpha S_\alpha)$. Thus $D^0(\cap_\alpha S_\alpha)$ is maximal, and since $\cap_\alpha S_\alpha \subseteq S$, the set $\cap_\alpha S_\alpha$ is achronal. A final application of Proposition 1 completes the proof that $\cap_\alpha S_\alpha$ is an instantaneous Cauchy surface. Clearly it is the smallest contained in S.

(B) Since S is achronal, $S \subseteq \sim I(S)$. In fact, one finds that $\sim I(S)$ is itself an achronal set. To show this, suppose that $\sim I(S)$ is not achronal and choose two points, $p \ll q$ in $\sim I(S)$. By strong causality, there exists some open globally hyperbolic subset $N \subseteq I^+(p) \cap I^-(q) \subseteq \sim I(S)$. By (GH0) and (GH2), there exists some achronal set $\Delta S \subseteq N$ such that $N \subseteq D^0(\Delta S)$. Since $\Delta S \subseteq \sim I(S)$, the set $S \cup \Delta S$ is achronal. Moreover, $D^0(S \cup \Delta S)$ contains N and so is strictly larger than $D^0(S)$. This conclusion leads to a contradiction since, by Proposition 1, $D^0(S)$ is maximal. Thus $\sim I(S)$ is an achronal set containing S and so is an instantaneous Cauchy surface. It is not difficult to see that $\sim I(S)$ is the largest achronal set containing S (and hence the largest instantaneous Cauchy surface containing S); for any extension of $\sim I(S)$ would contain points in $I(S)$ and would not be achronal.

Some further properties of $S_{\max}$ and $S_{\min}$ are needed later. It is useful to state these properties explicitly.

Proposition 4: If S is an instantaneous Cauchy surface in a strongly causal spacetime M and $p \in S_{\min}$, then every timelike curve through p intersects $D^0(S_{\min})$.

Proof: First note that for every $p \in S_{\min}$ either $I^+(p) \cap D^0(S) \neq \emptyset$ or $I^-(p) \cap D^0(S) \neq \emptyset$. Otherwise one has $D^0(S) = D^0(S_{\min}) = D^0(S_{\min} - \{p\})$ and $S_{\min} - \{p\}$ is an instantaneous Cauchy surface, contradicting the minimality of $S_{\min}$.

If $I^+(p)\cap D^0(S)\neq\emptyset$, then for any $q\in I^+(p)\cap D^0(S)$, $I^-(q)\cap I^+(p)\subseteq D^0(S)$. Any timelike curve through p must intersect $I^-(q)\cap I^+(p)$ and so must intersect $D^0(S)$. Finally, if $I^-(p)\cap D^0(S)\neq\emptyset$, dual arguments complete the proof.

Proposition 5: $I(S_{\min})\cup S_{\min}=I^+(D^0(S))\cup I^-(D^0(S))$.

Proof: If $p\in I(S_{\min})\cup S_{\min}$, then, for some timelike curve γ through p, $\gamma\cap S_{\min}\neq\emptyset$. By the previous proposition, $\gamma\cap D^0(S)\neq\phi$. Thus, either $p\in I^+(D^0(S))$ or $p\in I^-(D^0(S))$, and consequently $I(S_{\min})\cup S_{\min}\subseteq I^+(D^0(S))\cup I^-(D^0(S))$.

To show the reverse inclusion, suppose $p\in I^+(D^0(S))$. Then for some $q\in D^0(S)$ and some timelike curve γ, p and q are respectively the future and past end points of γ. Let γ' be any inextendible timelike curve containing γ. Since γ' contains $q\in D^0(S_{\min})=D^0(S)$, the set $\gamma'\cap S_{\min}$ must be nonempty. Since p lies on a timelike curve that intersects $S_{\min}$, we have $p\in I(S_{\min})\cup S_{\min}$. A similar argument can be made for $p\in I^-(D^0(S))$ so that $I^+(D^0(S))\cup I^-(D^0(S))\subseteq I(S_{\min})\cup S_{\min}$.

An immediate consequence of Propositions 3 and 5 is

$$S_{\max}-S_{\min}=\sim[I(S_{\min})\cup S_{\min}]=\sim[I^+(D^0(S))\cup I^-(D^0(S))]$$

for any instantaneous Cauchy surface S. It is now easy to see that if S and S' are equivalent in the sense of having the same interior Cauchy development, then $S_{\max}-S_{\min}=S'_{\max}-S'_{\min}$. If S and S' are equivalent instantaneous Cauchy surfaces, then $S_{\max}$ and $S'_{\max}$ are homeomorphic if and only if $S_{\min}$ and $S'_{\min}$ are homeomorphic. We turn to this question in the next section.

V. REFINEMENT OF THE TOPOLOGY CHANGE PROPERTY

Consider the topology change property (GH4) of Cauchy surfaces. In order to apply this result to instantaneous Cauchy surfaces in a straightforward way, the previous paper required one of the surfaces to be acausal so that it would lie in its interior Cauchy development. Here we show how this restriction can be removed.

Theorem 2: If S and S' are instantaneous Cauchy surfaces in a strongly causal spacetime and $D^0(S)=D^0(S')$, then $S_{\min}$ is homeomorphic to $S'_{\min}$ and $S_{\max}$ is homeomorphic to $S'_{\max}$.

Proof: From the comments following Proposition 5, it is sufficient to prove that $S_{\min}$ is homeomorphic to $S'_{\min}$. The property (GH4) cannot be applied directly because of the possibility that $S_{\min}$ and $S'_{\min}$ do not lie entirely within $D^0(S)$. As in the proofs[5-7] of properties (GH2)—(GH4), let γ be a congruence of inextendible timelike curves and use this congruence to produce a map $f:S_{\min}\rightarrow S'_{\min}$.

First, show that the map f is defined on S and is one-to-one and onto. For convenience define $K:=D^0(S)$. It follows from Proposition 4 that any timelike curve $\gamma_p\in\gamma$ through $p\in S_{\min}$ enters K and so intersects $S'_{\min}$ at a point $f(p)$. Thus, γ defines a one-to-one $f:S_{\min}\rightarrow S'_{\min}$. Similarly, γ defines a one-to-one map $g:S'_{\min}\rightarrow S_{\min}$ such that $g=f^{-1}$ so that f is onto.

Next, show that f is continuous by expressing it locally as a composition of continuous maps. Choose any $p\in S_{\min}$ and choose neighborhoods U_p of p and U_q of $q:=f(p)$ with compact closures. Choose a point $\tilde p\in\gamma_p\cap U_p$ and a point $\tilde q\in\gamma_p\cap U_q$. From (GH3) we can pick achronal sets $\tilde S$ and $\tilde S'$ in K such that $\tilde p\in\tilde S$, $\tilde q\in\tilde S'$, and $D^0(\tilde S)=D^0(\tilde S')=K$. The congruence γ defines the one-to-one and onto maps $f_{\min}:S_{\min}\rightarrow\tilde S$, $\tilde f:\tilde S\rightarrow\tilde S'$, and $f'_{\min}:\tilde S'\rightarrow S'_{\min}$ so that $f=f'_{\min}\circ\tilde f\circ f_{\min}$. The function $f_{\min}$ will be continuous at p if the sequence $\{\tilde q_n|\tilde q_n=f_{\min}(p_n)\}$ converges to $\tilde q$ whenever the sequence $\{p_n\}$ in $S_{\min}$ converges to p. If any subsequence of $\{\tilde q_n\}$ converges to a point $q'\in U_p$ but $q'\neq\tilde q$, then $q'\in\gamma$ and either $q'\in I^+(\tilde q)$ or $q'\in I^-(\tilde q)$. Suppose $q'\in I^+(\tilde q)$. But then $I^+(\tilde q)$ is a neighborhood of q' and must contain $\tilde q_n$ for sufficiently large n. One then has $\tilde q_n\in I^+(\tilde q)$ which contradicts the achronality of $\tilde S$. The same argument holds if $q'\in I^-(\tilde q)$. If $\{\tilde q_n\}$ contains a subsequence $\{q_{n'}\}\subseteq M-U_p$, then the timelike curves of γ which join $p_{n'}$ and $\tilde q_{n'}$ must intersect $\dot U$. As $\overline{U}$ is compact, the boundary $\dot U$ is also compact and this sequence of intersections will have a cluster point q'. One can then apply the previous argument to contradict the achronality of $\tilde S$. Thus $f_{\min}$ is continuous. This same argument can also be used to show that $f'_{\min}$ is continuous. Since (GH4) implies the continuity of $\tilde f$, we have established that f is continuous.

A time reversal of the preceding argument shows that f^{-1} is also continuous so that f is continuous and open and therefore a homeomorphism.

VI. DISCUSSION

Proposition 1 establishes a close and reversible connection between instantaneous Cauchy surfaces and maximal open globally hyperbolic sets. Theorem 1 shows that instantaneous Cauchy surfaces are plentiful in most of the spacetimes that one would wish to consider. Proposition 3 connects these instantaneous Cauchy surfaces to minimal and maximal instantaneous Cauchy surfaces.

For most purposes, it is the minimal and maximal instantaneous Cauchy surfaces that are of interest. Propositions 3–5 spell out a variety of properties of these surfaces. From part (B) of Proposition 3, a maximal instantaneous Cauchy surface can be thought of as a "global instant of time" because it divides spacetime into past, present, and future (compact partial Cauchy surfaces in a causally continuous spacetime share this same property[9]). Such a surface is an achronal boundary and is therefore a closed, edgeless, imbedded C^{1-} three-dimensional submanifold of spacetime—a partial Cauchy surface. Proposition 4 shows that a minimal instantaneous Cauchy surface has an important global property in common with a spacelike hypersurface. However, it should be noticed that $S_{\min}$ is not, in general, an acausal or spacelike surface and can have null generators. Proposition 5 and the comments that follow it can be used to deduce the properties of $S_{\max}-S_{\min}$. The property that has been used in this paper is the fact that this set depends only on the maximal open globally hyperbolic set $D^0(S)$ and not on the particular instantaneous Cauchy surface S that generates it. It is also quite easy to show that $S_{\max}-S_{\min}$ is generated by a congruence of null geodesics.

Theorem 2 requires considerable work to extend the property (GH4), but this work is necessary because one often prefers to give data on null hypersurfaces which need not be contained in their own Cauchy developments. As in the previous paper,[1] the most natural way to interpret Theorem 2 is to give various negative statements of it. Thus, we find that topology changes in maximal instantaneous Cauchy surfaces cannot occur through the regular evolution of hyperbolic field equations. If two such surfaces are not homeomorphic, then they must have distinct Cauchy developments and the spacetime cannot be globally hyperbolic.

[1]R.H. Gowdy, J. Math. Phys. **18**, 1798–1801 (1977).
[2]S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge U.P., Cambridge, 1973).
[3]R. Penrose, *Technique of Differential Topology in Relativity* (SIAM, Philadelphia, 1972).
[4]Reference 3, p. 48, Propositions 5.22, 5.23.
[5]Reference 3, Theorem 5.25.
[6]Reference 3, Theorem 5.26.
[7]This result follows directly from the theorem cited above.
[8]We are using the definition of past-distinguishing which is given in Hawking and Ellis, Ref. 2, p. 192.
[9]R. Budic and R.K. Sachs, J. Math. Phys. **15**, 1302–9 (1974).