

Ein Satz über geodätische Verbindbarkeit in normalhyperbolischen Räumen

HANS-JÜRGEN SEIFERT

I. Institut für Theoretische Physik der Universität Hamburg

Eingegangen am 20. Dezember 1966

Abstract. Under certain conditions, two points of a normal hyperbolic manifold joined by a time-like curve can be joined by a time-like geodesic. These conditions are discussed, and some applications are given.

AVEZ [1] behandelt in seiner Doktorarbeit folgendes Problem: In einer normalhyperbolischen Mannigfaltigkeit seien zwei Punkte durch eine zeitartige Kurve verbindbar. Unter welchen Voraussetzungen kann man dann auf die Existenz einer zeitartigen Geodätischen zwischen beiden Punkten schließen? Seine Beweisidee läßt sich unter schwächeren Annahmen über den Raum durchführen (§ 1); seine Behauptung, daß dies in k -vollständigen Räumen (zur Def. vgl. § 2 und KUNDT [4]) allgemein gilt, läßt sich nicht bestätigen, wie durch Gegenbeispiele eingesehen werden kann (§ 2). AVEZ zieht aus der Gültigkeit eines solchen Satzes die bemerkenswerte Folgerung, daß ein k -vollständiger Raum mit nicht verschwindender, nach unten positiv beschränkter Materiedichte zeitlich geschlossen sein muß (§ 5).

§ 1. Ein Satz über geodätische Verbindbarkeit

V_4 sei eine normalhyperbolische Mannigfaltigkeit mit Metrik g_{ab} der Stetigkeitsklasse C^3 , Signatur $(-+++)$, zeitorientiert (es gibt eine globale Zukunftsrichtung); A, B sind Punkte aus V_4 .

$H(A, B)$ sei die Menge aller stetigen, orientierten, zeit- oder lichtartigen Kurvenbögen von A nach B : In einer Umgebung jedes Kurvenpunktes gelangt man mit (ab-)zunehmendem Kurvenparameter in die (Vergangenheit) Zukunft. Im folgenden werden orientierte zeit- oder lichtartige Kurven oder Kurvenstücke kurz als k -Kurven (k -Geodätische) bezeichnet; (k steht für „kausal“).

$\mathcal{C}(A, B)$ sei die Menge aller Punkte, die auf Kurven aus $H(A, B)$ liegen (Durchschnitt der Zukunft von A und der Vergangenheit von B).

Das Funktional

$$F(g) := \int_{g(t)} \operatorname{sgn}(g_{ab} \dot{x}^a \dot{x}^b) \sqrt{|g_{ab} \dot{x}^a \dot{x}^b|} dt$$

sei die (für zeitartige Kurven negative) Kurvenlänge bezüglich der Metrik g_{ab} .

Eine Punktmenge M aus V_4 heie *streng kausal*, wenn es zu jedem Punkt A aus M eine Umgebung U_A gibt, die von den k -Kurven durch A kein zweites Mal getroffen wird.

Es gilt der

Satz. [1]. *Wenn $\mathfrak{S}(A, B)$ nicht leer, kompakt und streng kausal ist, existiert in $H(A, B)$ eine Geodätische, auf der das (unter Umständen schwache) absolute Minimum des Funktionals F auf $H(A, B)$ angenommen wird. Dasselbe gilt für jede k -Homotopieklasse von $H(A, B)$ einzeln.*

Der Beweis besteht aus einer längeren Kette von Standardschlüssen und soll deshalb nur angedeutet werden¹.

$H(A, B)$ ist kompakt. Bew.: Auf $\mathfrak{S}(A, B)$ als einer kompakten Teilmenge einer Mannigfaltigkeit lät sich stets eine (definite) Metrik δ einföhren. Diese induziert eine Metrik Δ auf $H(A, B)$:

$$\Delta(g_1, g_2) := \max\{\sup \delta(P_1, g_2), \sup \delta(P_2, g_1); P_i \in g_i\}; g_i \in H(A, B).$$

Wegen der δ -Beschräntheit von \mathfrak{S} und damit der Δ -Beschräntheit von H kann aus jeder unendlichen Teilmenge G von H eine cauchykonvergente Teilfolge $\{g_n\}$ ausgewählt werden. Die Menge g aller Grenzwerte von Punktfolgen $\{X_n\}$, ($X_n \in g_n$), ist Element von H : Zum Beweis betrachtet man eine hinreichend kleine Umgebung U eines $P \in g$. Die $g_n \cap U$ sind als k -Kurven *gleichgradig stetig* bezüglich der Zeitkoordinate t in U^2 . Daraus und aus der δ -Vollständigkeit von \mathfrak{S} folgt, daß $g \cap U$ Kurvenstück ist. $g \cap U$ kann als Grenzwert von k -Kurven nirgendwo raumartig sein. g ist kompakt²; nach Überdeckung von g mit endlich vielen U sieht man, daß $g \in H(A, B)$, also Häufungspunkt von G ist. Die Betrachtung benutzt nur die durch die Konformstruktur von V_4 erfolgte Auszeichnung der k -Kurven und die Topologie der punktweisen Konvergenz auf H , erweist sich also als von der Wahl der Metrik δ unabhängig. Die strenge Kausalität erlaubt, konvergente k -Kurven so zu parametrisieren, daß sie auch als Funktionen des Parameters konvergieren.

F ist auf der Menge der *rektifizierbaren* k -Kurven ein *unterhalbstetiges Funktional*. Beweis nach einem Standardverfahren der Variationsrechnung: Man stellt ein g aus $H(A, B)$ durch drei Funktionen g^α eines Zeitparameters dar. Um den Wert $F(g) =: \int f dt$ mit dem auf einer Nachbarkurve $h \in H(A, B)$ zu vergleichen, bettet man g und h in eine

¹ Eine Ausführung der Beweise soll in meiner Diplomarbeit gegeben werden.

² Damit die $[g_n \cap U](t)$ erklärt sind, muß man unter Umständen endlich viele g_n fortlassen und außerdem U auf einen Zylinder mit den Endflächen $t = t^-$ und $t = t^+$ ($t^- < t_P < t^+$) einschränken.

³ Nach dem Diagonalfolgenprinzip: Aus $P = \lim P^i$, $P^i \in g$ (d. h. $P^i = \lim X_n^i$, $X_n^i \in g_n$) folgt $P = \lim X_n^n$, also $P \in g$.

Kongruenz von Kurven aus H ein, die ein Feld $\dot{g}^\alpha(x^\beta, t)$ definieren, und entwickelt den Integranden $f(g^\alpha, \dot{g}^\alpha, t)$ nach den \dot{g}^α . Das Integral über den linearen Anteil f_l längs der geschlossenen Kurve $g - h$ kann nach dem Satz von STOKES umgeformt werden und strebt mit $h \rightarrow g$ gegen Null: $\int_h f_l dt \rightarrow \int_g f_l dt = \int_g f dt$. Im Restglied f_r zweiter Ordnung bilden die zweiten Ableitungen $f_{g^\alpha g^\beta}$ für ein Zwischenwerttripler \dot{g}_z^α eine quadratische Form, die im wesentlichen mit der induzierten Metrik auf den zu g orthogonalen infinitesimalen Hyperflächenelementen übereinstimmt, also für k -Kurven g positiv (semi-)definit ist: $\int_h f_r dt \geq 0$. Für eine Folge $\{g^n\}$ mit $\lim g^n = g$ gilt also $\lim \inf F(g^n) \geq F(g)$. F kann weiterhin auf ganz H , die abgeschlossene Hülle der rektifizierbaren k -Kurven, unterhalbstetig fortgesetzt werden.

Daraus folgt die Existenz einer Kurve g mit $F(g) = \inf F(h)$, $h \in H$. Eine lokale Betrachtung zeigt, daß g geodätisch und $F(g) > -\infty$ ist.

Man unterscheidet zwei Arten von Homotopie (vgl. AVEZ [1]): $g_1, g_2 \in H$ sind homotop bzw. k -homotop, wenn es eine Schar von Kurven bzw. k -Kurven gibt, die eine stetige Deformation von g_1 in g_2 vermitteln. Jede (k -)Homotopieklasse von H ist abgeschlossen und daher ebenfalls kompakt, alle Betrachtungen sind auch für eine einzelne (k -)Homotopieklasse gültig. Insbesondere gilt der Satz für streng kausale k -Homotopieklassen.

§ 2. Gegenbeispiele

Wenn nur die k -Vollständigkeit des Raumes verlangt wird (d. h. alle zeit- und lichtartigen Geodätischen können bis zu jedem Wert eines affinen Parameters fortgesetzt werden), ist der Satz nicht mehr allgemein gültig, wie die folgenden beiden Gegenbeispiele zeigen:

Der de Sitter Kosmos mit negativer Krümmung: Alle zeitartigen Geodätischen durch einen Punkt A schneiden sich im konjugierten Punkt A' und treffen deshalb die raumartige Umgebung von A' nicht. Jeder raumartig zu A' gelegene Punkt B ist aber über eine k -Kurve von A aus erreichbar⁴. Man kann eine Folge von Kurven \widehat{AB} finden, die im Endlichen gegen zwei Lichtstrahlen je durch A und B konvergieren und auf denen F beliebig stark negativ wird. In übersichtlicher Weise läßt sich dieses Verhalten an Hand der Einbettung des de Sitter Kosmos in den E_5 mit der Form $ds^2 = -du^2 - dv^2 + dx^2 + dy^2 + dz^2$ als einschaliges Hyperboloid $u^2 + v^2 - x^2 - y^2 - z^2 = 1$ darstellen (vgl. etwa SYNGE [6]). Die Geodätischen sind die Schnittkurven des Ebenenbüschels durch das Zentrum mit dem Hyperboloid; die zeitartigen Geodätischen durch A sind Ellipsen, die sich im Antipodenpunkt A' wieder treffen, sie über-

⁴ Eine analoge Betrachtung für raumartige Geodätische im positiv gekrümmten de Sitter-Kosmos findet sich bei CALABI und MARKUS [2].

decken genau $\mathfrak{S}(A, A')$ und $\mathfrak{S}(A', A)$, da $(A')' = A$ ist. $\mathfrak{S}(A, A')$ ist der in den Raumrichtungen unbeschränkte, daher inkompakte Streifen zwischen zwei parallelen Erzeugendenbüscheln, dem Zukunftslichtkegel in A und dem Vergangenheitslichtkegel in A' . Über solche Ellipsen und zeitartige Kreise ($x, y, z = \text{const.}$) läßt sich jeder zu A' raumartige Punkt B von A aus erreichen. Da mit $(x^2 + y^2 + z^2) \rightarrow \infty$ die F -Länge dieser Kreise (mit dem Radius $\sqrt{1 + x^2 + y^2 + z^2}$) gegen $-\infty$ geht, gilt: $\inf F(\widehat{AB}) = -\infty$. Die de Sitter-Welt ist konform flach; die Minkowski-Welt kann auf das Innere eines Streifens $\mathfrak{S}(A, A')$ abgebildet werden, innerhalb dessen der Satz noch gilt.

Der Gödel Kosmos⁵. Die Materieweltlinien w bilden eine äquidistante Geodätischenschar. k -Geodätische, die von einem Punkt A auslaufen, fallen auf die Weltlinie w_A durch A zurück, sie beschreiben im Ruhraum des Teilchens in w_A eine Rosettenbahn und können einen endlichen Maximalabstand von w_A , den optischen Horizont, nicht überschreiten. Abgelenkte k -Kurven können bis zum Horizont ihres Ablenkungspunktes gelangen, also Punkte außerhalb des optischen Horizontes von w_A erreichen, genügend oft abgelenkte sogar jedes $P \in V_4$ (s. u.). Weiterhin gibt es geschlossene k -Kurven; diese sind homotop zur vollständig entarteten Kurve (bestehend aus einem Punkt), da V_4 einfach zusammenhängt, aber nicht k -homotop zu ihr, da jede geschlossene k -Kurve einen Durchmesser größer als den optischen Horizont hat. Also haben auch Punktepaare A, B , die k -geodätisch verbindbar sind, wegen des Fehlens von geschlossenen k -Geodätischen nicht in jeder k -Homotopieklasse von $H(A, B)$ eine Geodätische.

Entsprechend ist in beiden Fällen die Voraussetzung über die $\mathfrak{S}(A, B)$ nicht erfüllt; im Gödel Kosmos gilt $\mathfrak{S}(A, B) = V_4$ für beliebige A, B . Denn bekanntlich geht durch jeden Punkt A eine geschlossene zeitartige Kurve; berücksichtigt man, daß der Gödel Kosmos raumzeitlich homogen ist, so folgt, daß von A aus ganz w_A zukünftig zeitartig erreichbar ist. Ferner ist die Menge der von w_A aus zukünftig zeitartig erreichbaren Materieweltlinien sowohl offen als auch abgeschlossen (in der Topologie der lokalen Raumschnitte). Somit lassen sich je zwei Punkte zeitartig verbinden mit verfügbarer Orientierung, und $\mathfrak{S}(A, B)$ ist der ganze Kosmos.

Im oben beschriebenen de Sitter Kosmos gilt das gleiche. Wegen des zweifachen Zusammenhanges sind die geschlossenen Ellipsen nicht-zusammenziehbar und fallen je nach Zahl der Durchläufe in unendlich viele Homotopieklassen. Das Auftreten geschlossener k -Kurven kann aber verhindert werden durch einen Schnitt z. B. längs eines Hyper-

⁵ Die folgenden Betrachtungen fußen auf Mitteilungen von Herrn Dr. W. KUNDT.

belastet (raumartige Hyperfläche $u = 0, v > 0$) und periodische Fortsetzung. Dann gilt: $\mathfrak{S}(A, B)$ ist genau dann kompakt, wenn B im Inneren von $\mathfrak{S}(A, A')$ oder auf dem Zukunftslichtkegel von A liegt; eine k -Geodätische \overline{AB} existiert genau dann, wenn B im Inneren irgendeines $\mathfrak{S}(A^{(n)}, A^{(n+1)})$ liegt; ($A^{(m)}$ ist m -ter Brennpunkt der k -Geodätischen durch A); für $n > 0$ ist $F(\overline{AB}) \neq \inf F(\widehat{AB}) = -\infty$, da B zeitartig zu einem raumartig zu A' liegenden Punkt liegt.

§ 3. Bemerkungen zur Voraussetzung

Die *Nichtkompaktheit* von $\mathfrak{S}(A, B)$ bedeutet entweder *Unvollständigkeit*: ein Loch in der Zukunft von A (Rand, Singularität) oder *Unbeschränktheit*: Es gibt eine unbegrenzt in die Zukunft fortsetzbare (abgeschlossene, endpunktlose) k -Kurve durch A , deren Punkte durch k -Kurven mit B verbindbar sind; d. h. ein Spiegel in unendlicher Entfernung reflektiert in endlicher Zeit auf die Quelle zurück⁶; es gibt Weltlinien, die vollständig in der Vergangenheit von B liegen.

Die Voraussetzung der Kompaktheit von $\mathfrak{S}(A, B)$ ist für die Gültigkeit des Satzes *nicht notwendig*. Es gibt Räume, in denen je zwei zeitartig zueinander liegende Punkte A, B durch eine kürzeste k -Geodätische verbindbar sind, aber wenigstens ein $\mathfrak{S}(A, B)$ nicht kompakt ist: Beispiele sind *unvollständige*, geodätisch konvexe, irgendwo zeitartigen Rand besitzende Untermengen von Welten, in denen der Satz gilt: etwa ein Vollzylinder mit zeitartiger Achse in der Minkowski Welt.

Es gibt weiterhin Räume⁷ mit einem unbeschränkten $\mathfrak{S}(A, B)$ und k -geodätischer Verbindbarkeit, bei denen aber für ein Punktepaar A, B die Bedingung $\inf F(\widehat{AB}) > -\infty$ nicht mehr erfüllt ist; d. h. die Geodätische \overline{AB} hat nicht mehr den minimalen Wert von F auf $H(A, B)$. Für die Existenz einer minimalen Geodätischen ist die Bedingung $\inf F(\widehat{AB}) > -\infty$ *notwendig*, aber nicht hinreichend, wie im de Sitter Kosmos die Betrachtung zweier Randpunkte eines Streifens $\mathfrak{S}(A, A')$ zeigt.

§ 4. Ein Kriterium

Die Voraussetzung der Kompaktheit von $\mathfrak{S}(A, B)$ ist für je zwei zeitartig zueinander liegende Punkte A, B erfüllt, wenn im V_4 eine *Cauchyhyperfläche* C (zur Def. vgl. PENROSE [5]) existiert: Alle abgeschlossenen k -Kurven ohne Endpunkte treffen C in genau einem Punkt. Denn dann ist die Menge D aller Punkte von C , die durch eine k -Kurve mit A

⁶ Im Gödel Kosmos ist mehrfache (sogar unendlich vielfache) Reflexion nötig, da nur gebrochene Strahlen den jeweiligen optischen Horizont verlassen können; diese k -Kurve kann hier keine Geodätische sein.

⁷ Vgl. Fußnote 1.

verbindbar sind, kompakt: Da D offensichtlich abgeschlossen ist, gäbe es sonst eine Folge von k -Kurven mit divergenten Durchstoßpunkten. In der Einpunkt-Kompaktifizierung $V_4 \cup \{\infty\}$ erhält man eine konvergente Unterfolge $\{f_n\}$, deren Durchstoßpunkte $f_n \cap C$ natürlich gegen ∞ konvergieren. In jeder kompakten Untermenge K von V_4 konvergieren die f_n gegen eine k -Kurve f . Wegen der gleichgradigen Stetigkeit von k -Kurven g_n folgt aus $g = \lim g_n$ sofort: $g \cap C = \lim (g_n \cap C)$. f könnte daher im Endlichen (in beliebigem K) keinen Durchstoßpunkt haben, ein Widerspruch zur Voraussetzung. Daher muß auch das Raumgebiet $\mathfrak{R}(A)$, das von den k -Kurven zwischen A und C erfüllt wird, beschränkt sein, und es gilt: $\mathfrak{S}(A, B) \subset \mathfrak{R}(A) \cup \mathfrak{R}(B)$ und \mathfrak{S} ist kompakt.

Eine Abschwächung des Kriteriums ist möglich: Für je zwei Punkte einer Untermenge M von V_4 ist \mathfrak{S} kompakt, wenn alle k -Kurven, die M treffen, in der Vergangenheit (bzw. der Zukunft) auf eine abgeschlossene Punktmenge M' zurückverfolgt werden können und M' in je genau einem Punkt treffen. Denn nur diese Eigenschaften von C werden beim obigen Beweis benutzt. Ein Beispiel im de Sitter Kosmos: M ist das Innere eines Streifens $\mathfrak{S}(A, A')$, M' die Symmetriehyperbene (Mittelhyperbel zwischen A und A').

§ 5. Anwendungen

Dieser Satz hat eine Reihe von topologischen und differentialgeometrischen Konsequenzen.

Zwei *Cauchyhyperflächen* C_1, C_2 haben die gleiche *Zahl von Zusammenhangskomponenten* und sind gleichzeitig (in-)kompakt. Denn die Teilmenge D^A von C_2 , die von den k -Kurven durch ein $A \in C_1$ getroffen wird, ist ein (durch die Geodätischen vermitteltes) stetiges Abbild der Menge S_3 aller zeit- und lichtartigen Richtungen nach Identifizierung einer Zukunftsrichtung mit der entgegengesetzten Vergangenheitsrichtung; also ist die Teilmenge D^* von C_2 , die von den k -Kurven durch eine Zusammenhangskomponente $C^* \subset C_1$ getroffen wird, stetiges Abbild von $S_3 \times C^*$: D^* ist zusammenhängend und kompakt, wenn C^* es ist.

In einem Raum gelte für alle zeitartigen Einheitsvektoren u^a : $R_{ab}u^a u^b \leq -d^2 < 0$: Die *Materiedichte* hat eine *positive untere Schranke*, oder die kosmologische Konstante bewirkt Entsprechendes. Dann folgt (AVEZ [1]) wegen des Vorzeichens der zweiten Variation der Kurvenlänge F für eine minimale k -Geodätische g : $|F(g)| < \sqrt{3} \pi/d$. Bei Gültigkeit des Satzes ist der Raum also *k-unvollständig* oder besitzt *geschlossene k-Kurven*. HAWKING [3] erhält ebenfalls aus dem Ausdruck für die zweite Variation weitere Kriterien für das Auftreten von Singularitäten in k -vollständigen Räumen mit Cauchyhyperfläche. Die Ricci-Tensoren des Gödel Kosmos und des negativ gekrümmten de Sitter Kosmos erfüllen obige Bedingung, ohne daß eine Singularität auftritt. Im letzteren

gibt es sogar abgeschlossene, randfreie Raumschnitte, die allerdings nicht alle Eigenschaften einer Cauchyhyperfläche haben.

Bei Gültigkeit des Satzes wird die Zukunft eines Ereignisses von einer Schar raumartiger *Hyperflächen schlicht überdeckt*. Die Pseudosphären konstanter Eigenzeit (längs den Geodätischen gemessen) überdecken im allgemeinen die Zukunft stellenweise mehrfach. Obige Hyperflächen bekommt man durch die Vorschrift, die Eigenzeit längs der kürzesten Geodätischen zu messen. Sie sind abgeschlossene, randfreie Teile der entsprechenden Pseudosphären, sind stückweise glatt, und stehen senkrecht auf den Geodätischen. Sie sind bis auf den Lichtkegel (Abstand = 0) raumartig.

Ich danke den Mitgliedern des Hamburger Seminars für Allgemeine Relativitätstheorie, besonders Herrn Dr. W. KUNDT, für viele Anregungen und Hinweise und vor allem Herrn Professor P. JORDAN für seine großzügige Förderung.

Literatur

- [1] AVEZ, A.: Thèses, Paris 1963.
- [2] CALABI, E., and L. MARKUS: Ann. Math. **75**, 63 (1962).
- [3] HAWKING, S. W.: Singularities in universes (1966 preprint). Singularities in closed universes (1966 preprint).
- [4] KUNDT, W.: Z. Physik **172**, 488 (1963).
- [5] PENROSE, R.: Phys. Rev. Letters **14**, 57 (1965).
- [6] SYNGE, J. L.: Relativity: The general theory. Amsterdam: North Holl. Publ. 1960.