INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK

DISSERTATIONES MATHEMATICAE

ROZPRAWY MATEMATYCZNE

LIII

H. BUSEMANN

Timelike spaces

WARSZAWA 1967
PAŃSTWOWE WYDAWNICTWO NAUKOWE

PRINTED IN POLAND

WROCLAWSKA DRUKARNIA NAUKOWA

CONTENTS

INTRODUCTION

The interest in indefinite metrics derives largely from the theory of relativity. On the other hand (quoting Eddington [5, p. 22]): "Assuming that a material particle cannot travel faster than light the intervals along its track must be timelike. We ourselves are limited by material bodies and can only have direct experience of timelike intervals."

This suggests a study of purely timelike metrics, independently of the question whether they are restrictions of indefinite metrics to timelike intervals. The present paper lays the foundations for such a theory.

The principal property of a timelike space R is this: R is partially ordered ($x \leqslant y$) and a function $\varrho(x, y)$ is defined for $x \leqslant y$ satisfying $\varrho(x, x) = 0$, $\varrho(x, y) > 0$ for $x < y$ and the "*time inequality*" $\varrho(x, y) + \varrho(y, z) \leqslant \varrho(x, z)$ for $x < y < z$. In contrast to the metric case this must be supplemented by other requirements. The function ϱ does not define a topology, which is introduced separately, there must be enough pairs x, y with $x < y$, etc.

In addition, there are interesting spaces where a timelike distance can be defined only locally (they occur also in general relativity.) The basic axioms for timelike and locally timelike spaces and their consequences are given in Section 1.

Our aim is a geometric theory analogous to that of metric G-spaces (see [2]), which proved an adequate basis for many different types of geometric developments. The additional axioms leading to timelike and locally timelike G-spaces are found in Section 2. Although variations of these axioms are possible they will appear quite natural when their meaning in special cases is examined.

However, some remarks on *completeness* are necessary. There are two types of completeness, one concerns small and the other large distances. The former states that $\varrho(x_\nu, y_\nu) \to 0$ if $x_\nu \leqslant y_\nu$, $x_\nu \to x$, $y_\nu \to y$, and not $x \leqslant y$. The latter type proves very elusive. In the metric case the Hopf-

Rinow Theorem in the general version of Cohn-Vossen (see [4]) states that under certain conditions finite compactness, completeness and geodesic completeness are equivalent. This theorem has no analogue in the timelike case and geodesic completeness (i.e. the indefinite prolongability of a geodesic curve) emerges as the most relevant concept.

Section 2 contains, besides the axioms, the existence of segments and geodesics. Section 3 deals with basic topological properties. Unfortunately we could only in relatively few instances refer to the theory of metric G-spaces because just replacing $\leqslant$ by $\geqslant$ in the triangle inequality rarely produces a sufficient argument.

The absence of concrete examples beyond Lorentz spaces is a considerable handicap in the study of timelike spaces. The remainder of the paper is therefore devoted to special cases. In Section 4 we discuss products of timelike and metric spaces which include the Lorentz spaces. We then turn to the most important special class of timelike spaces, namely *timelike Minkowski spaces*, which furnish the local geometries in any differentiable locally timelike G-space.

More specifically, the general theory is discussed in Section 5 and mobility properties in Section 7. There are some unexpected phenomena. For example, the group of stability of a point may (in all dimensions) contain a subgroup which is transitive and abelian on the spheres about the point. There is an interesting open problem, namely the determination of all Minkowskian geometries with pairwise transitive groups of motions. This leads to very simple sounding, but unsolved, problems on projectivities of convex hypersurfaces in affine space, which are of general interest. Section 6 contains all that is known in this respect. The difficulties derive from our lack of information regarding locally compact transformation groups.

We conclude with timelike and locally timelike Hilbert geometries. The above mentioned completeness for small distances is particularly interesting in this case, because it eliminates the timelike form in favor of the locally timelike form, which contains the so-called exterior hyperbolic geometry as a special case.

The present paper is intended as a basis for further investigations. Some subjects in metric G-spaces clearly do not have timelike counterparts, for example the theory of perpendicularity, [2, Chapter II]. But others like mobility, remain and become much more challenging. Of course, there are also entirely novel problems.

Finally we point out that the study of indefinite metrics started in [4] although independent of, and quite different from, the present theory, is closely related to it in purpose.

1. TIMELIKE AND LOCALLY TIMELIKE SPACES

Preorderings, which will turn out to be partial orderings, of the space R, or a neighborhood $U(p)$ of a point p in R, are essential for timelike spaces and will be denoted by $x \leqslant y$ or $x \leqslant_p y$, reserving $x < y$ and $x <_p y$ for the case $x \neq y$. We mention that by definition $x \leqslant x$ or $x \leqslant_p x$ for all $x \in R$ or $x \in U(p)$. The sets of pairs of points x, y in $R \times R$ or $U(p) \times \times U(p)$ for which $x \leqslant y$, $x < y$, $x \leqslant_p y$, $x <_p y$ will be denoted respectively by $(\leqslant)$, $(<)$, $(\leqslant_p)$, $(<_p)$. Similarly, $(>)$ etc. consists of the pairs x, y for which $y < x$, i.e. in the notation used by some authors $(>)^{-1} = (<)$.

The axioms for a *timelike* space are

T_1. *R is a (non-empty) Hausdorff space.*

T_2'. *A preordering $x \leqslant y$ is defined in R. The set $(<)$ is open in $R \times R$ and each neighborhood $W(q)$ of a given point q contains points x, y with $x < q$ and $q < y$.*

T_3'. *A continuous real valued function $\varrho(x, y)$ is defined on $(\leqslant)$ and satisfies*

$$\varrho(x, x) = 0, \qquad \varrho(x, y) > 0 \quad \textit{for} \quad x < y,$$

and the "time inequality"

$$\varrho(x, y) + \varrho(y, z) \leqslant \varrho(x, z) \quad \textit{for} \quad x < y < z.$$

The axioms for a locally timelike space arise from these by localizing T_2' and T_3' and adding a consistency condition.

R is *locally timelike* if it satisfies T_1 and

T_2. *Each point has a neighborhood $U(p)$ in which a preordering $x \leqslant_p y$ is defined. The set $(<_p)$ is open in $U(p) \times U(p)$ (or $R \times R$) and each neighborhood $W(q)$ of a given point $q \in U(p)$ contains points x, y with $x <_p q$ and $q <_p y$.*

T_3. *A continuous real valued function $\varrho_p(x, y)$ is defined on $(\leqslant_p)$ and satisfies*

$$\varrho_p(x, y) = 0, \qquad \varrho_p(x, y) > 0 \quad \textit{if} \quad x <_p y$$

and

$$\varrho_p(x, y) + \varrho_p(y, z) \leqslant \varrho_p(x, z) \quad \textit{if} \quad x <_p y <_p z.$$

T_4. *If $(<_p) \cap (<_q) \neq \emptyset$ then $(<_p) \cap (>_q) = \emptyset$ and $\varrho_p(x, y) = \varrho_q(x, y)$ for $(x, y) \in (<_p) \cap (<_q)$. If $(<_p) \cap (>_q) \neq \emptyset$ then $(<_p) \cap (<_q) = \emptyset$ and $\varrho_p(x, y) = \varrho_q(y, x)$ for $(x, y) \in (<_p) \cap (>_q)$.*

A timelike space may be considered as the special case of a locally timelike space where $U(p) = U(q) = R$ and $(<_p) = (<_q)$ for any two points p, q. This remark frequently obviates separate definitions or proofs for the timelike case which are obtained by simply omitting the subscript p. Where this is clear we will not discuss the timelike space explicitly. The following are two simple but useful examples

(1) *$x \leqslant_p y$ (and $x \leqslant y$) are partial orderings.*

For $x \leqslant_p y$ and $y \leqslant_p x$ give $x \leqslant_p x$ and $\varrho_p(x, y) + \varrho_p(y, x) \leqslant \varrho_p(x, x) = 0$, hence $\varrho_p(x, y) = 0$ and $x = y$.

We say that y lies *between* x and z and write $(x\,y\,z)_p$ (or $(x\,y\,z)$) if $x <_p y <_p z$ and $\varrho_p(x, y) + \varrho_p(y, z) = \varrho_p(x, z)$.

(2) *$(u\,x\,y)_p$ and $(u\,y\,z)_p$ if and only if $(u\,x\,z)_p$ and $(x\,y\,z)_p$.*

Let $(u\,x\,y)_p$ and $(u\,y\,z)_p$; then

$$\begin{aligned}\varrho_p(u, z) &= \varrho_p(u, y) + \varrho_p(y, z) = \varrho_p(u, x) + \varrho_p(x, y) + \varrho_p(y, z)\\ &\leqslant \varrho_p(u, x) + \varrho_p(x, z) \leqslant \varrho_p(u, z),\end{aligned}$$

hence $(u\,x\,z)_p$ and $(x\,y\,z)_p$. The converse is proved in the same way.

If in a locally timelike space we replace, for an arbitrary set of points p, the set $(\leqslant_p)$ by $(\geqslant_p)$ and define a new distance by $\varrho_p'(x, y) = \varrho_p(y, x)$, all axioms will be satisfied.

The space is *consistently ordered* if $(<_p) \frown (>_q) = \emptyset$ for any p, q The space possesses a consistent ordering if by replacing, for a suitable. set of points p, the set $(\leqslant_p)$ by $(\geqslant_p)$ and ϱ_p by ϱ_p', the space can be made consistently ordered. If such a change is possible, we may and will assume that the original space is consistently ordered. There are very interesting spaces which do not possess consistent orderings (see Section 8).

In a consistently ordered space we call *chain* C_{xy} from x to y a set of points $u_0 = x, u_1, \ldots, u_k = y$ with $u_{i-1} \leqslant_{p_i} u_i$ for some p_i. If C_{xy} exists, we write $x \leqslant y$. Then $x \leqslant y$ and $y \leqslant z$ imply $x \leqslant z$. We put

$$\sigma(x, y, C_{xy}) = \sum \varrho_{p_i}(u_{i-1}, u_i).$$

Because of T_4 this is independent of the choice of the p_i for which $u_{i-1} \leqslant_{p_i} u_i$.

(3) *In a consistently ordered locally timelike space, if $xy = \sup_{C_{xy}} (\sigma(x, y, C_{xy}))$ is finite for all x, y with $x \leqslant y$, then xy defines a timelike space.*

The proof is obvious. (3) singles out those locally timelike spaces which can be identified with timelike spaces.

A *local T-curve* in a locally timelike space is a continuous map $x(t)$ of an interval $[\alpha, \beta]$ in $\mathscr{R}$ into a neighborhood $U(p)$ satisfying T_2 such that either (a) $x(t_1) <_p x(t_2)$ for all $t_1 < t_2$ in $[\alpha, \beta]$ or (b) $x(t_2) <_p x(t_1)$

for all $t_1 < t_2$ in $[\alpha, \beta]$. We refer to the two possibilities as cases (a) and (b) respectively. To treat them simultaneously we put

$$x(t_1)x(t_2) = \begin{cases} \varrho_P\big(x(t_1),\ x(t_2)\big) & \text{in case (a)}, \\ \varrho_P\big(x(t_2),\ x(t_1)\big) & \text{in case (b)}. \end{cases}$$

For a partition Δ: $\alpha = t_0 < t_1 < \ldots < t_k = \beta$ of $[\alpha, \beta]$ set

$$L(x, \Delta) = \sum_{i=1}^{k} x(t_{i-1})x(t_i)$$

and define the *length* $L(x)$ of $x(t)$ by

$$L(x) = \inf_{\Delta} L(x, \Delta).$$

Because of the time inequality

(4) $$L(x) \leqslant L(x, \Delta) \leqslant x(\alpha)x(\beta).$$

For any Δ put $\|\Delta\| = \max_i (t_{i+1} - t_i)$. Then we have as for metric spaces:

(5) *If Δ_ν is a sequence of partitions of $[\alpha, \beta]$ with $\|\Delta_\nu\| \to 0$ then $L(x, \Delta_\nu) \to L(x)$.*

Choose Δ_u: $u_0 = \alpha < u_1 < \ldots < u_n = \beta$ such that

$$L(x, \Delta_u) < L(x) + \varepsilon.$$

In Δ_ν choose t_i^ν $(i = 1, \ldots, n-1)$ as the last element smaller than u_i and put $t_0^\nu = \alpha$, $t_n^\nu = \beta$. Then for large ν

$$t_0^\nu = \alpha < t_1^\nu < \ldots < t_n^\nu = \beta, \quad \text{also} \quad \lim_{\nu\to\infty} t_i^\nu = u_i.$$

If Δ_u^ν denotes this partition of $[\alpha, \beta]$, then

$$L(x, \Delta_\nu) \leqslant L(x, \Delta_u^\nu) \quad \text{and} \quad \lim_{\nu\to\infty} x(t_{i-1}^\nu)x(t_i^\nu) = x(u_{i-1})x(u_i).$$

Therefore

$$\limsup L(x, \Delta_\nu) \leqslant \lim L(x, \Delta_u^\nu) = L(x, \Delta_u) < L(x) + \varepsilon$$

and (5) follows from $\liminf L(x, \Delta_u^\nu) \geqslant L(x)$.

If $\alpha \leqslant t' < t'' \leqslant \beta$ we denote by $L_{t'}^{t''}(x)$ the length of $x(t)\,|\,[t', t'']$ (i.e. the restriction of $x(t)$ to $[t', t'']$).

We conclude from (5) that length is additive:

(6) *For any partition Δ: $\alpha = t_0 < t_1 < \ldots < t_k = \beta$*

$$L(x) = \sum_{i=1}^{k} L_{t_{i-1}}^{t_i}(x).$$

(7) $$\textit{If } t_\nu \to t_0 \textit{ then } L_\alpha^{t_\nu}(x) \to L_\alpha^{t_0}(x).$$

For either $t_\nu = t_0$ or $t_\nu < t_0$ and $L_{t_\nu}^{t_0}(x) \leqslant x(t_\nu)x(t_0) \to 0$ or $t_\nu > t_0$ and $L_{t_0}^{t_\nu}(x) \leqslant x(t_0)x(t_\nu) \to 0$. Now the additivity (6) of length yields the assertion.

A continuous map $x(t)$ of $[\alpha, \beta]$ into a locally timelike space is a *T-curve* if a partition $\alpha = t_0 < t_1 < \ldots < t_k = \beta$ exists for which each $x(t)\,|\,[t_{i-1}, t_i]$ is a local T-curve. We define the length of $x(t)$ by

$$L(x) = \sum_{i=1}^{k} L_{t_{i-1}}^{t_i}(x)$$

and conclude from T_4 and (6) that this definition is independent of the partition chosen.

The assertions (5), (6), (7) remain true for T-curves, but $L(x, \Delta_\nu)$ in (5) will, in general, be defined for large ν only.

In timelike spaces there is, of course, no difference between local T-curves and T-curves, moreover always $L(x) \leqslant \varrho\big(x(\alpha), x(\beta)\big)$.

In consistently ordered spaces it suffices to consider the case (a) only. But there are interesting spaces which cannot be consistently ordered with T-curves traversing the same $U(p)$ twice satisfying (a) the first time and (b) the second.

Almost the same proof as for lower semi-continuity of length in the metric case yields upper semi-continuity in the present case (G, p. 20)(¹).

(8) *If $x_\nu(t)$ ($t \in [\alpha, \beta]$, $\nu = 0, 1, 2, \ldots$) are T-curves in a locally timelike space and $x_\nu(t)$ tends uniformly to $x_0(t)$ then*

$$\limsup L(x_\nu) \leqslant L(x).$$

In the timelike case $x_\nu(t) \to x(t)$ for each t suffices.

The length of a T-curve with $\alpha < \beta$ may vanish: If $R = \mathscr{R}$ with the usual order and $\varrho(t_1, t_2) < |t_1 - t_2|^\sigma$, $\sigma > 1$ then R becomes a timelike space in which all local T-curves have length 0. Such curves correspond to non-rectifiable curves in the metric case. Therefore we define:

A T-curve $x(t)$ is *rectifiable* if $L_{t_1}^{t_2}(x) > 0$ for $t_1 < t_2$. On a rectifiable curve $x(t)$, $t \in [\alpha, \beta]$, we can introduce *arclength as parameter*:

$$s(t) = L_\alpha^t(x), \qquad s(\alpha) = 0.$$

The function $s(t)$ is continuous by (7) and strictly increasing by rectifiability. So the inverse $t(s)$ is defined. Then

$$y(s) = x\big(t(s)\big), \qquad 0 \leqslant s \leqslant L(x)$$

is the representation in terms of arclength:

$$L_{s_1}^{s_2}(y) = s_2 - s_1 \quad \text{for} \quad s_1 < s_2.$$

(¹) Paper [2] is quoted as G.

Because of additivity it suffices to prove $L_0^s(y) = s$ and this follows from the uniform continuity of $t(s)$. Also, $L(x) = L(y)$.

A local T-curve, $x(t) \subset U(p)$ for $t \in [\alpha, \beta]$ is a *segment* if

$$L(x) = x(\alpha)\ x(\beta) = \varrho_p\big(x(\alpha),\ x(\beta)\big) \quad \text{or} \quad = \varrho_p\big(x(\beta),\ x(\alpha)\big).$$

We have (in case (a)) for $\alpha < t_1 < t_2 < \beta$ from (4) and (6)

$$\begin{aligned}\varrho_p\big(x(\alpha),\ x(\beta)\big) &= L_\alpha^{t_1}(x) + L_{t_1}^{t_2}(x) + L_{t_2}^{\beta}(x)\\ &\leqslant \varrho_p\big(x(\alpha),\ x(t_1)\big) + \varrho_p\big(x(t_1),\ x(t_2)\big) + \varrho_p\big(x(t_2),\ x(\beta)\big)\\ &\leqslant \varrho_p\big(x(\alpha),\ x(\beta)\big).\end{aligned}$$

Therefore

$$L_{t_1}^{t_2} = \varrho_p\big(x(t_1),\ x(t_2)\big) \quad \text{for} \quad \alpha \leqslant t_1 < t_2 \leqslant \beta.$$

This shows

(9) *A segment is rectifiable and its restriction to any sub interval is a segment.*

(10) *A rectifiable local T-curve $x(t)$ is a segment if and only if its arclength representation $y(s)$ satisfies*

$$\varrho_p\big(y(s_1),\ y(s_2)\big) = s_2 - s_1 \quad \textit{for} \quad s_1 < s_2$$

or

$$\varrho_p\big(y(s_2),\ y(s_1)\big) = s_2 - s_1 \quad \textit{for} \quad s_1 < s_2.$$

In the first case, if the condition is satisfied, then for any Δ

$$L(\Delta, y) = \varrho_p\big(y(0),\ y(L(x))\big) = \varrho_p\big(x(\alpha),\ x(\beta)\big).$$

The converse is contained in the preceding argument.

$x(t)$ defined on a connected set τ of the t-axis (which contains more than one point) is a *partial geodesic*, if for each $t_0 \in \tau$ an $\varepsilon(t_0) > 0$ exists such that $x(t) \mid [t_0 - \varepsilon(t_0),\ t_0 + \varepsilon(t_0)] \cap \tau$ is a segment. Partial geodesics can be partially ordered: $\{\tau, x(t)\} \leqslant \{\tau', x'(t)\}$ if $\tau \subset \tau'$ and $x(t) = x'(t)$ for $t \in \tau$. Then each well ordered set $\{\tau_i, x_i(t)\}$, $i \in I$ of partial geodesics has an upper bound, namely, the obvious partial geodesic defined on $\bigcup \tau_i$. By Zorn's Lemma each partial geodesic lies on a maximal partial geodesic. Maximal partial geodesics are called *geodesics*. Thus we have

(11) *A partial geodesic, in particular a (proper) segment, can be extended to a geodesic.*

On a partial geodesic $\{\tau, x(t)\}$ we can introduce "arc length" as parameter: select $t_0 \in \tau$ and any real a, put $s(t_0) = a$ and

$$s(t) - a = L_{t_0}^{t}(x) \quad \text{for} \quad t > t_0, \qquad a - s(t) = L_t^{t_0}(x) \quad \text{for} \quad t < t_0.$$

Then $t(s)$ is defined, monotone and continuous and with $y(s) = x\big(t(s)\big)$, $y(a) = x(t_0)$

$$L_{s_1}^{s_2}(y) = s_2 - s_1 \quad \text{for} \quad s_1 < s_2.$$

We will always assume that for a partial geodesic $\{\tau, x(t)\}$ the parameter is arc length, i.e. $L_{t_1}^{t_2}(x) = t_2 - t_1$ for $t_1 < t_2$ and $t_i \epsilon \tau$.

This implies for segments that arc length is parameter with a possible shift of the origin

$$t_2 - t_1 = \varrho_p\big(x(t_1),\ x(t_2)\big) \quad \text{or} \quad = \varrho_p\big(x(t_2),\ x(t_1)\big)$$

if

$$\alpha \leqslant t_1 < t_2 \leqslant \beta = \alpha + x(\alpha)x(\beta).$$

According to our agreement for consistently ordered spaces we have:

(12) *If $x(t)$ is a geodesic in a timelike space then $x(t') < x(t'')$ for $t' < t''$.*

A geodesic $\{\tau, x(t)\}$ in a timelike space is a *line* if $x(t) \mid [t', t'']$ is a segment for any $t' < t''$ in τ.

The axioms T_{1-4} do not contain the existence of segments, which must be derived from additional axioms. The first step of the existence proof works, however, in any locally timelike space.

A *linear set* $\lambda(q, r) = \{\tau, x(t)\}$ is a map $t \to x(t)$ of a subset τ of an interval $[0, \beta]$ in some $U(p)$ with the following properties

$$0 \epsilon \tau \text{ and } x(0) = q, \quad \beta \epsilon \tau \text{ and } x(\beta) = r,$$

either $x(t_1) \leqslant_p x(t_2)$ for all $t_1 \leqslant t_2$ in τ, or $x(t_2)_p \leqslant x(t_1)$ for all $t_1 \leqslant t_2$ in τ, and, with the previous notation,

$$x(t_1)x(t_2) = t_2 - t_1 \quad \text{for} \quad 0 \leqslant t_1 \leqslant t_2 \leqslant \beta;$$

which implies $\beta = qr$.

Since $\tau = \{0\}$ implies $q = r$ we assume $q \neq r$.

The $\lambda(q, r)$ can be partially ordered like partial geodesics and Zorn's Lemma yields the existence of a maximal linear set containing a given $\lambda(q, r)$. Notice that $\{q, r\}$ is a linear set. Thus we have

(13) *A given linear set $\lambda(q, r)$, in particular, $\{q, r\}$ with $q < r$ in $U(p)$, lies in a maximal linear set $\mu(q, r)$.*

So far we only used the existence of the partial ordering $(\leqslant_p)$ and not the additional properties postulated in T_2' or T_2. They serve the purpose of making the theory non-trivial. To see this we consider two examples where R is the (x_1, x_2)-plane. First we define $x < y$ by $x_1 < y_1$ and $x_2 = y_2$; we put $\varrho(x, y) = y_1 - x_1$. Then all axioms are satisfied except that $(<)$ is not open in $R \times R$.

Next define $x < y$ by $x_1 + 1 < y_1$ and $|x_2 - y_2| \leqslant |x_1 - y_1|$ and define

$$\varrho_1(x, y) = y_1 - x_1 - |y_2 - x_2|;$$

then not each $W(q)$ contains points x, y with $x < q < y$.

However, defining $x < y$ by $x_1 < y_1$ and $|x_2 - y_2| \leq |x_1 - y_1|$ and using $\varrho_1(x, y)$, the axioms T_1, T'_2, T'_3 are satisfied. We deduce from T_2 or T'_2 (with $U(p) = R$ in the latter case)

(14) *Given $W(p) \subset U(p)$ then $W'(p) \subset W(p)$ and points u, v in $W(p)$ exist such that $u <_p x <_p v$ for every $x \in W'(p)$.*

For u, v with $u <_p p <_p v$ exist and the openess of $(<_p)$ guarantees that $u <_p x <_p v$ for x in a suitable $W'(p)$.

The most important implication is that any two points x, y in $W'(p)$ have a common predecessor u, and a common successor v.

We observe that in the (x_1, x_2)-plane with the metric ϱ_1 and the second definition of $x < y$ the set $[\tau, x(t)]$ with $\tau = [0, 2]$ and

$$x(t) = (t, 0) \text{ for } 0 \leq t \leq 1, \qquad x(t) = (t+1, 1) \text{ for } 1 < t \leq 2$$

is a maximal linear set $\mu[(0,0), (3,1)]$ for which $x(t)$ is *not* continuous, although $x(t) \to t$ is an isometry, i.e. $\varrho(x(t_1), x(t_2)) = |t_1 - t_2|$ for $t_1 < t_2$.

2. SEGMENTS AND GEODESICS IN TIMELIKE G-SPACES

We now list our additional axioms. Those for a *timelike G-space* are, besides T_1, T'_2, T'_3,

G_1. *The space, R, is locally compact, connected and has a countable base.*

G'_2. *If $x_\nu \leq y_\nu$, $x_\nu \to x$, $y_\nu \to y$, and not $x \leq y$ then $\varrho(x_\nu, y_\nu) \to 0$.*

G'_3. *If $x_\nu \to q$, $z_\nu \to q$ and $(x_\nu\, y_\nu\, z_\nu)$ then $y_\nu \to q$.*

G'_4. *If $q < r$ then points with $(q\,x\,r)$ exist and the closure of all x with this property is compact.*

G'_5. *Each point has a neighborhood $V(p)$ such that for x, y in $V(p)$ with $x < y$ points u and z with $(u\,x\,y)$ and $(x\,y\,z)$ exist.*

G'_6. *If $(u_1\,x\,y)$, $(u_2\,x\,y)$ and $\varrho(u_1, x) = \varrho(u_2, x)$ then $u_1 = u_2$. If $(x\,y\,z_1)$, $(x\,y\,z_2)$ and $\varrho(y, z_1) = \varrho(y, z_2)$ then $z_1 = z_2$.*

The axioms for a *locally timelike G-space R* are T_{1-4}, G_1 and

G_2. *If $x_\nu \leq y_\nu$, $x_\nu \to x \in U(p)$, $y_\nu \to y \in U(p)$ and not $x \leq_p y$ then $\varrho_p(x_\nu, y_\nu) \to 0$.*

G_3. *If $x_\nu \to q \in U(p)$, $z_\nu \to q \in U(p)$ and $(x_\nu\, y_\nu\, z_\nu)_p$ then $y_\nu \to q$.*

G_4. *There is a neighborhood $U'(p) \subset U(p)$ such that for q, r in $U'(p)$ and $q <_p r$ points x with $(q\,x\,v)_p$ exist.*

G_5. *There is a neighborhood $V(p) \subset U(p)$ such that for x, y in $V(p)$ and $x <_p y$ a point u in a given neighborhood of x with $(u\,x\,y)_p$ and a point z in a given neighborhood of y with $(x\,y\,z)_p$ exist.*

G_6. *If* $(u_1\,x\,y)_p$, $(u_2\,x\,y)_p$ *and* $\varrho_p(u_1, x) = \varrho_p(u_2, x)$ *then* $u_1 = u_2$.
If $(x\,y\,z_1)_p$, $(x\,y\,z_2)_p$ *and* $\varrho_p(y, z_1) = \varrho_p(y, z_2)$ *then* $z_1 = z_2$.

The two sets of axioms are strict analogues except for G_4, G_4' and G_5, G_5'. In the latter case we could omit the condition "in a given neighborhood of x (or y)" because the considerations of this section will show that after shrinking $U'(p)$ and $U(p)$ the axiom G_5 will be satisfied with the existence of neighborhoods of x and y when postulated without this existence. We chose the given formulation for emphasis.

As to G_4 and G_4', a local form of the second part of G_4 or G_4' follows from G_2 or G_2' and this suffices for locally timelike G-spaces.

(1) *Given* $W(p)$ *with compact* $\overline{W}(p) \subset U(p)$, *there is a* $W_1(p) \subset W(p)$ *such that* q, r *in* $W_1(p)$ *and* $(q\;x\;r)_p$ *imply* $x \in \overline{W}(p)$.

($U(p) = R$ in the timelike case. We will not mention this modification every time.)

If (1) were false we could find a sequence of point triples q_ν, r_ν, x_ν in $W(p)$ with $(q_\nu\, x_\nu\, r_\nu)_p$, $q_\nu \to p$, $r_\nu \to p$ and $x_\nu \notin \overline{W}(p)$. This contradicts G_3 or G_3'. The compactness of $\overline{W}(p)$ is not used, but this is the most useful case and exhibits the analogy to G_4'. The axioms G_2, G_2' are important completeness requirements which have no counterparts in the metric case.

In the following considerations, in as far as they are local, we will always assume that $qr = \varrho_p(q, r)$ and not $qr = \varrho_p(r, q)$.

G_6 and G_6' play a greater role in timelike spaces than the corresponding axioms for metric G-spaces. To understand this fully we prove

(2) *If* $T_{1\text{-}4}$ *and* $G_{1\text{-}4}$ *hold and the* $\overline{W}(p)$ *of* (1) *lies in the* $U'(p)$ *of* G_4 *then for a maximal linear set* $\mu(q, r) = \{\tau, x(t)\}$ *with* q, r *in* $W_1(p)$ *the set* τ *is the interval* $[0, qr]$.

In the timelike case T_1, T_2', T_3' *and* G_1, $G_{2\text{-}4}'$ *imply the same for any maximal linear set.*

However, $x(t)$ *need not be continuous.*

First we deduce from T_{1-4} and $G_{1\text{-}3}$ that τ is closed for q, r in $W_1(p)$. It follows from (1) that $x(t) \in \overline{W}(p)$. Let $t_\nu \in \tau$ and $t_\nu \to t_0$. We want to show $t_0 \in \tau$ and may assume $0 < t_0 < qr$. Because of (2) a subsequence $\{x(t_k)\}$ of $\{x(t_\nu)\}$ tends to a point x_0.

Let $0 \leqslant t' < t_0 < t'' \leqslant qr$ where t', t'' lies in τ. Then

$$\varrho_p\big(x(t'),\, x(t_k)\big) \to t_0 - t' > 0, \qquad \varrho_p\big(x(t_k),\, x(t'')\big) \to t'' - t_0 > 0$$

hence by G_2

$$x(t') < x_0 \quad \text{and} \quad \varrho_p\big(x(t'), x_0\big) = t_0 - t',$$
$$x_0 < x(t'') \quad \text{and} \quad \varrho_p\big(x_0, x(t'')\big) = t'' - t_0.$$

Therefore, if $\mu(q, r)$ did not contain a point corresponding to t_0 the set $\mu(q, r) \cup x_0$ would still be linear contradicting maximality. The proof for the timelike case uses the second part of G_4' instead of (1).

Assume now that $\tau \neq [0, qr]$. Then α, β in τ with $(\alpha, \beta) \cap \tau = \emptyset$ would exist since τ is closed. By G_4 there is an x_0 with $\big(x(\alpha)\, x_0\, x(\beta)\big)_p$ since $x(\alpha)$, $x(\beta)$ lie in $W(p) \subset U'(p)$. We deduce from (1.2) that

$$\big(x(t')\, x_0\, x(t'')\big)_p \quad \text{for} \quad 0 \leqslant t' \leqslant \alpha \text{ and } \beta \leqslant t'' \leqslant qr.$$

In particular $(q\, x_0\, r)_p$ and $x_0 \in W(p)$. Therefore $\mu(q, r) \cup x_0$ would still be linear.

The distance $\varrho_1(x, y)$ of Section 1 satisfies T_1, T_2', T_3' and $G_1, G_{2\text{-}4}'$ (the latter is proved under more general hypotheses in Section 4) and we noticed that $\mu(q, r)$ with noncontinuous $x(t)$ exist.

We now show

(3) *If* G_6 *or* G_6' *hold in addition to the hypothesis of* (2) *then* $x(t)$ *is continuous, hence* $\mu(q, r)$ *is a segment.*

We must show $x(t_\nu) \to x(t_0)$ for $t_\nu \to t_0$ and $t_\nu \in \tau$. We assume $t_0 > 0$, the case $t_0 = 0$ is treated like $t_0 = qr$. Since each subsequence of $\{x(t_\nu)\}$ has an accumulation point, it suffices to show that $x(t_\nu) \to x_0$ implies $x_0 = x(t_0)$. As in the proof (2) we find for $0 < t' < t_0$ that

$$x(t') < x(t_0) \quad \text{and} \quad \varrho_p\big(x(t'), x_0\big) = t_0 - t' = \varrho_p\big(x(t'), x(t_0)\big).$$

Then $\big(q\, x(t')\, x(t_0)\big)_p$ and $\big(q\, x(t')\, x_0\big)_p$ and G_6 gives $x(t_0) = x_0$.

We will denote a segment from q to r by $\sigma(q, r)$ in the timelike case and $\sigma_p(q, r)$ in the locally timelike case.

$T_{1\text{-}4}$ and G_6 imply

(4) *There is at most one segment* $\sigma_p(q, r)$ *if a point* u *with* $(u\, q\, r)_p$ *or a point* r *with* $(q\, r\, r)$ *exists.*

For let $(u\, q\, r)_p$ and assume two distinct segments $x(t), x'(t)$ $(0 \leqslant t \leqslant \varrho_p(q, r))$ from q to r exist. Then $x(t_0) \neq x'(t_0)$ for some t_0 with $0 < t_0 < \varrho_p(q, r)$ and $\varrho_p\big(q, x(t_0)\big) = \varrho_p\big(q, x'(t_0)\big) = t_0$. On the other hand (1.2) yields $\big(u\, q\, x(t_0)\big)_p$ and $\big(u\, q\, x'(t_0)\big)_p$ contradicting G_6.

We combine our results to obtain the following important facts:

(5) THEOREM. *In a timelike G-space a segment* $\sigma(q, r)$ *exists whenever* $q < r$. *It is unique when either* u *with* $(u\, q\, r)$ *or* v *with* $(q\, r\, v)$ *exists. In particular,* $\sigma(q, r)$ *is unique when* q, r *lie in a* $V(p)$ *of* G_5'.

(6) THEOREM. *In a locally timelike space there is at most one segment* $\sigma_p(q, r)$ *when* u *with* $(u\, q\, r)_p$ *or* v *with* $(q\, r\, v)_p$ *exists, in particular when* q, r *lie in the* $V(p)$ *of* G_5.

Each point has a neighborhood $W_1(p)$ *such that* $\sigma_p(q, r)$ *exists for* q, r *with* $q < r$ *in* $W_1(p)$ *and all these* $\sigma_p(q, r)$ *lie in a compact set* $\overline{W}$.

It should be mentioned that the use of Zorn's Lemma could have been avoided in the existence proof of segments and geodesics, but the arguments would have been longer.

(7) THEOREM. *In a timelike or locally timelike G-space the extension of a partial geodesic (in particular of a proper segment) to a geodesic* $\{\tau, x(t)\}$ *is unique and* τ *is open.*

The uniqueness follows from G_6, G_6' and the openness of τ from G_5, G_5'.

This implies that a geodesic $\{\tau, x(t)\}$ cannot traverse the same segment in opposite direction (i.e. that $[a, \beta]$ and $[\beta', a']$ in τ exist with $\beta - a = a' - \beta' > 0$ and $x(a+t) = x(a'-t)$ for $0 \leqslant t \leqslant \beta - a$), although allowing reversal of the local partial ordering might seem to make this possible. A geodesic may, of course, be closed and traverse one segment infinitely often in the same direction.

We cannot conclude that for a geodesic τ is the entire real axis, because our axioms do not contain a completeness requirement corresponding to finite compactness postulated for metric G-spaces. The reason for omitting such a postulate will be discussed in the next section.

We now prove various facts concerning the extent of segments and convergence of geodesics.

(8) *If* $x(t)$ *is a geodesic and* $x(t)\,|\,[a, \beta] \subset V(p) \frown W_1(p)$ *then it is a segment* ($W_1(p) = R$ *in the timelike case*).

Assume this is not true. Then there is a $t' \in (a, \beta)$ such that $x(t)\,|\,[a, t']$ is a segment and $x(t)\,|\,[a, t]$ is not for $t > t'$. For if $\varrho_p(x(a), x(t_1)) > t_1 - a$ and $t_2 > t_1$ then

$$\varrho_p(x(a), x(t_2)) \geqslant \varrho_p(x(a), x(t_1)) + \varrho_p(x(t_1), x(t_2)) > t_1 - a + t_2 - t_1 .$$

Since $x(a), x(t') \in V(p)$, there is a v with $(x(a)\, x(t')\, v)_p$ in $W_1(p)$ and hence a segment $\sigma_p(x(t'), v)$. By (1.2) the segment $x(t)\,|\,[a, t']$ continued by $\sigma_p(a(t'), v)$ would be a segment and lead to a geodesic providing a second continuation of $x(t)\,|\,[a, t']$ to a geodesic.

(9) COROLLARY. *In a timelike space each geodesic is a line if and only if for given* $x < y$ *points* u, v *with* $(u\, x\, y)$ *and* $(x\, y\, v)$ *exist.*

The sufficiency follows from (8) and the necessity is obvious.

(10) *Let* $x, y \in W''(p)$ *and* $x < y$ *where* $W''(p) \subset \overline{W}''(p) \subset V(p) \frown W_1(p)$. *Then* $\sigma_p(x, y) \subset \sigma_p(u, v)$ *with* $u <_p x <_p y <_p v$, $\sigma_p(u, v) \subset W''(p) \smile u \smile v$ *and* $u, v \in \overline{W}''(p) - W''(p)$.

We consider the geodesic $x(t)$ for which $x(t)\,|\,[0, \varrho_p(x, y)]$ is $\sigma_p(x, y)$. If $(x\, y\, z)_p$ and $z \in W_1(p)$ then $\sigma_p(y, z)$ is $x(t)\,|\,[\varrho_p(x, y), \varrho_p(x, z)]$. Therefore, if $t_0 = \sup t$ where $x(t)\,|\,[a, t] \subset W''(p)$, it follows that $x(t_0) \in \overline{W}''(p) - W''(p)$.

Similarly for $t_1 = \inf t$ with $x(t)|[t, \varrho_p(x,y)] \subset W''(p)$ then $x(t_1) \in \overline{W}''(p) - W''(p)$ and $x(t)|[t_1, t_0]$ is a segment by (8).

(11) *For a given geodesic* $\{\tau_0, x_0(t)\}$ *and* $t_0 \in \tau_0$ *there is an* $\eta > 0$ *with the following property:*

if $\{\tau_\nu, x_\nu(t)\}$ $(\nu = 1, 2, \ldots)$ *are geodesics,* $[a, t_0 - \varepsilon] \subset \bigcap_{\nu=0}^{\infty} \tau_\nu$ $(a < t_0 - \varepsilon,$ $0 < \varepsilon \leqslant \eta)$, *and* $x_\nu(t) \to x_0(t)$ *for* $t \in [a, t_0 - \varepsilon]$, *then* $[a, t_0 + \varepsilon] \subset \tau_0 \cap \bigcap_{\nu=N}^{\infty} \tau_\nu$ *for a suitable* N *and* $x_\nu(t) \to x_0(t)$ *for* $t \in [a, t_0 + \varepsilon]$.

Put $p = x(t_0)$ and choose $\eta > 0$ such that $x_0(t)|[t_0 - 2\eta, t_0 + 2\eta] \subset W''(p)$. Let $a' = \min(a, t_0 - 2\varepsilon)$, $q_\nu = x_\nu(a')$. For large ν the points q_ν, $x_\nu(t_0 - \varepsilon)$ lie in $W''(p)$. By (10) there is a segment $\sigma_p(q_\nu, r_\nu)$ containing $x_\nu(t)|[a', t_0 - \varepsilon]$ which lies in $W''(p)$ except for $r_\nu \in \overline{W}''(p) - W''(p)$, moreover $\sigma_p(q_\nu, r_\nu)$ is $x_\nu(t)|[a', a' + \varrho_p(q_\nu r_\nu)]$. Any accumulation point of $\{r_\nu\}$ lies in $\overline{W}''(p) - W''(p)$.

An accumulation point of $y_\nu \in \sigma_\nu$ must lie on $x_0(t)$ because of the uniqueness of prolongation. The continuity of the distance $\varrho_p(x, y)$ now shows that $x_\nu(t)$ is defined for $t \in [a, t_0 + \varepsilon]$ and large ν and also that $x_\nu(t) \to x_0(t)$ in this interval.

(12) THEOREM. *Let* $\{\tau_\nu, x_\nu(t)\}$ $(\nu = 0, 1, 2 \ldots)$ *be geodesics and* $[a, \beta] \subset \bigcap_{\nu=0}^{\infty} \tau_\nu$ $(a < \beta)$, *moreover* $x_\nu(t) \to x_0(t)$ *for* $t \in [a, \beta]$. *Let* τ^* *be the set of those* t *which lie in all but a finite number of* τ_ν. *Then* $\tau^* \supset \tau_0$ *and* $x_\nu(t) \to x_0(t)$ *for* $t \in \tau_0$. *The convergence is uniform on any* $[t_1, t_2] \subset \tau_0$.

Put $\tau^a = \tau^* \cap [a, \infty)$ and assume $\tau^a \not\supset \tau_0$. Then $t_0 \in \tau_0$ exists such that $\tau^a \cap [t > t_0] = \emptyset$ but $\tau^* \supset [a, t_0]$. For each $\varepsilon > 0$ there is an $N(\varepsilon)$ with $\tau_\nu \supset [a, t_0 - \varepsilon]$ for $\nu \geqslant N(\varepsilon)$. Choose ε such that (11) is applicable. Then (11) would give $\tau_\nu \supset [a, t_0 + \varepsilon]$ for large ν, so $\tau^* \supset [a, t_0 + \varepsilon]$ which is impossible. With a similar argument for $\tau^\beta = \tau^* \cap (-\infty, \beta]$ this shows $\tau^* \supset \tau_0$.

Convergence on τ_0 also follows from (11) and uniform convergence in closed intervals is equivalent to the easily proved statement $x_\nu(t_\nu) \to x_0(t_0)$ for $t_\nu \to t_0 \in \tau^*$.

A partial geodesic $\{\tau, x(t)\}$ in a timelike space is a *progressing* (*receding*) *ray* if $\tau = [a, \infty)$ $(\tau = (-\infty, a])$, $x(t_1) < x(t_2)$ for $t_1 < t_2$ and

$$\varrho\big(x(t_1), x(t_2)\big) = t_2 - t_1 \quad \text{for} \quad t_1 < t_2.$$

Assume $\{[a, \infty), x(t)\}$ and $\{[\beta, \infty), y(t)\}$ are two progressing rays for which $y(\beta) = x(a')$ with $a' > a$.

If $y(t)$ is a subray of $x(t)$, i.e., $y(t + \beta) = x(t + a')$, we can evidently find sequences $t_\nu \to \infty$, $t'_\nu \to \infty$ such that $x(t_\nu) < y(t'_\nu)$ and, writing *ab* instead of $\varrho(a, b)$, $x(t_\nu)\, y(t'_\nu) \to 0$. In metric G-spaces the existence of such sequences is also sufficient for $y(t+\beta) = x(t+\alpha')$ (see G, p. 137; $\alpha' > \alpha$ is essential). One verifies easily in the Lorentz plane (see Section 4) that in timelike spaces this is not correct even if we require additionally that a sequence $t''_\nu \to \infty$ with $y(t'_\nu) < x(t''_\nu)$ and $y(t'_\nu)\, x(t''_\nu) \to \infty$ exists.

So it looks as though the metric theorem has no timelike analogue; however, it does, and the analogue leads to a generalization of the result in the metric case.

(13) THEOREM. *Let* $\{[\alpha, \infty), x(t)\}$ *and* $\{[\beta, \infty), y(t)\}$ *be progressing rays with* $y(\beta) = x(\alpha')$, $\alpha' > \alpha$. *If there are sequences* t_ν, t'_ν, t''_ν *tending to* ∞ *such that with* $x(t_\nu) = x_\nu$, $y(t'_\nu) = y'_\nu$ *etc. either*

$$x_\nu < y'_\nu < x''_\nu \quad \text{and} \quad x_\nu x''_\nu - x_\nu y'_\nu - y'_\nu x''_\nu \to 0$$

or

$$y_\nu < x'_\nu < y''_\nu \quad \text{and} \quad y_\nu y''_\nu - y_\nu x'_\nu - x'_\nu y''_\nu \to 0$$

then $y(t+\beta) = x(t+\alpha')$ *for* $t \geqslant 0$.

There is, of course, an analoguous theorem for receding rays.

For an indirect proof assume $z = y(\gamma+\beta) \neq x(\gamma+\alpha')$ for some $\gamma > 0$ and put $x(\alpha) = x$, $x(\alpha') = y(\beta) = y$. Then $xz - xy - yz = \delta > 0$. In the first case we have for large ν

$$\begin{aligned} xx''_\nu &= xy + yx_\nu + x_\nu x''_\nu < xy + yy'_\nu - x_\nu y'_\nu + x_\nu x''_\nu \\ &= xy + yz + zy'_\nu + x_\nu x''_\nu - x_\nu y'_\nu, \end{aligned}$$

hence

$$xx''_\nu = xz + zy'_\nu + y'_\nu x''_\nu - \delta + (x_\nu x''_\nu - x_\nu y'_\nu - y'_\nu x''_\nu).$$

This would yield $xx''_\nu < xz + zy'_\nu + y'_\nu y''_\nu$ for large ν.

In the second case, for large ν,

$$xy_\nu < xx'_\nu - y_\nu x'_\nu = xy + yx'_\nu - y_\nu x'_\nu < xy + yz + zy''_\nu - y_\nu x'_\nu - x'_\nu y''_\nu,$$

hence

$$xy_\nu < xz + zy_\nu - \delta + (y_\nu y''_\nu - y_\nu x'_\nu - x'_\nu y''_\nu)$$

and so $xy_\nu < xz + zy_\nu$ for large ν.

The same argument with signs reversed yields for metric G-spaces that $y(t+\beta) = x(t+\alpha')$ if either $x'_\nu y_\nu + y'_\nu x''_\nu - x'_\nu x''_\nu \to 0$ or $y'_\nu x'_\nu + x'_\nu y''_\nu - y'_\nu y''_\nu \to 0$. This contains the statement that $y(t+\beta) = x(t+\alpha')$ if $x'_\nu y'_\nu \to 0$ as the special case $x'_\nu = x''_\nu$ or $y'_\nu = y''_\nu$ which does not make sense in the timelike case.

3. TOPOLOGICAL PROPERTIES. COMPLETENESS

A *line element* L_p at a point p of a locally timelike space is a maximal set of segments $\sigma_p(q, r)$ such that: each of these $\sigma_p(q, r)$ contains p as relative interior point and any two have a further common point. It follows from (2.4) that the intersection of any two segments in L_p is an element of L_p. Each $\sigma_p(q, r)$ with $q <_p p <_p r$ lies in exactly one L_p.

A geodesic H which contains (in an obvious sense) one segment in L_p contains all. Therefore we call L_p a *line element of H at p*. The cardinal number of distinct line elements of H at p is the *multiplicity* of H at p. The point p is a *simple* point of H if its multiplicity is 1, otherwise it is a *multiple* point. H is simple if it has no multiple points. As for G-spaces (G, pp. 44-45) one proves

(1) *The multiplicity of a geodesic at a point is finite or countable. A geodesic has an at most countable number of multiple points.*

(2) *The geodesic $\{\tau, x(t)\}$ is simple if and only if either $x(t_1) \neq x(t_2)$ for $t_1 \neq t_2$ or if $x(t_1) = x(t_2)$ for $t_1 = t_2$ implies $\tau = \mathscr{R}$ and $x(t_1 + t) = x(t_2 + t)$ for all t.*

It follows from (2.5.6) and T_2', T_2 that a (locally) timelike G-space R contains with each point p segments containing this point, therefore, $\dim_p R \geqslant 1$.

Whether a locally timelike G-space is always a topological manifold is not known. This is not even known for metric G-spaces. In all interesting special cases it will be a manifold. Nevertheless there is a certain interest in the topological properties which can be inferred from the axioms. The presently best way for metric G-spaces is not that of G, but proving that the space is locally homogeneous in the sense of Montgomery, see [4, Theorem (3.2)]. This proof cannot be carried over to timelike G-spaces since it uses in an essential way, that there are unique segments connecting any two points in a suitable neighborhood of a given point. However, the approach taken in G can be carried over with some modifications. We will briefly indicate, mostly without complete proofs, how this is done.

In the first place it follows from the remark after (1.14), (2.5.6) and G_1 that

(3) *A (locally) timelike G-space is arcwise and locally arcwise connected.*

Next one observes the following application of (2.6.12).

(4) LEMMA. *Every point p of a (locally) timelike G-space has a neighborhood $W_2(p)$ such that $\sigma_p(q, r)$ exists and is unique for $q < r$ in $W_2(p)$. If q_ν, r_ν $(\nu = 0, 1, 2, \ldots)$ lie in $W_2(p)$, $q_\nu < r_\nu$, $q_\nu \to q_0$, $r_\nu \to r_0$ and $x_\nu(t)$ represents $\sigma_p(q_\nu, r_\nu)$ with $x_\nu(0) = q_\nu$ then $x_\nu(t_\nu) \to x_0(t_0)$ for $t_\nu \to t_0$.*

This is used to prove

(4′) *If two distinct line elements* L_p^1, L_p^2 *exist at* p, *then* $\dim_p R \geqslant 2$.

(The proof produces a set homeomorphic to a 2-simplex with p corresponding to a vertex.)

If $\sigma_p^i \in L_p^i$. Let $\sigma_p(p, q) \subset \sigma_p^1$ and $\sigma_p(p, r) \subset \sigma_p^2$ consist of those points on σ_p^i for which $p \leqslant x$. Using T_2 and the results of the preceding section we can find q' with $(p\, q'\, q)_p$ and r' with $(p\, r'\, r)_p$ such that $q' < r'$ and $\sigma_p(q', r')$ as well as all segments $\sigma_p(p, y)$ with $y \in \sigma_p(q', r')$ lie in the $W_2(p)$ of (4). It then follows from (4′) that $\bigcup_y \sigma_p(p, y)$ is homeomorphic to a 2-simplex.

(5) *A locally timelike G-space of dimension one consists of one simple geodesic. A one-dimensional timelike G-space consists of one line.*

Because of (3) no geodesic can have multiple points, and there si at least one geodesic $\{\tau, x(t)\}$. We want to show that any point v lies on $x(t)$. Let $t_0 \in \tau$ and $q = x(t_0)$. Using the remark after (1.14) we find points $u_0 = q, u_1, \ldots, u_k = v$ such that for each $i = 0, \ldots, k-1$ either a segment $\sigma_{u_i}(u_i, u_{i+1})$ or segments $\sigma_{u_i}(u_i, v)$ and $\sigma_{u_i}(u_{i+1}, v)$ exist. Applying (4′) repeatedly, we see that all these segments must lie on $x(t)$.

The same argument slightly refined yields:

(6) *If* $\dim_p R = 1$ *for some* p, *then* $\dim_p R = 1$ *for all* p.

The second part of (5) is seen as follows: if R is timelike, then $x(t_1) < x(t_2)$ for $t_1 < t_2$ by (1.12) and a segment $\sigma(x(t_1), x(t_2))$ exists. $x(t) | [t_1, t_2]$ is the only arc in R from $x(t_1)$ to $x(t_2)$ and hence is a segment.

(7) THEOREM. *A two dimensional locally timelike G-space is a manifold.*

For simplicity we take a timelike G-space in which all geodesics are lines. The argument is the same in the general case using the uniformity expressed in (2.10) for localization.

Because $\dim_p R = 2$ for all p by (6) there are two distinct line elements at each point p. There are points u, v, r, s with $(u\, p\, v)$, $(r\, p\, s)$, $u < r < s$ but not $(u\, r\, s)$. We prolong each $\sigma(u, x)$, $x \in \sigma(r, s)$ to a $\sigma(u, x^*)$ so that the x^* form an arc B. It suffices to show that any sufficiently small neighborhood U of p is contained in $\bigcup_{x^*} \sigma(u, x^*) = V^*$.

If this were not correct we could, possibly by shrinking $\sigma(r, s)$, produce the following situation. u^* is close to u not in V^*, $(p\, u^*\, c)$ and $y < c$ for $y \in \bigcup_{x \in \sigma(r,s)} \sigma(p, x) = V$.

Put $2\beta = \min xc > 0$ and let π be the map of V in $\bigcup_{x \in V} \sigma(x, c)$ which maps x on x' with $(x\, x'\, c)$ and $x'c = \beta$. Put $V' = \pi V$.

We show $\dim V' \geqslant 2$ following the method of [G, p. 53] where references to the facts used are found. It suffices to prove that any set $F' \subset V'$ which separates $\sigma' = \pi\sigma(r, s)$ from $u' = \pi u$ has at least dimension 1. The set $\pi^{-1} F' = F$ is closed and separates $\sigma(r, s)$ from p, so does a continuum F_0 in F, hence F_0 contains a point r_0^* of $\sigma(p, r)$ and a point s_0^* of $\sigma(p, s)$. Then $F_0' = \pi F_0 \subset F'$, moreover πr_0, πs_0 lie in F_0' and are distinct. Since F_0' is a continuum, $\dim F_0' \geqslant 1$. So $\dim V' \geqslant 2$. Now $\bigcup_{x' \in V'} \sigma(x', c)$ contains the product of V' and a segment, and has at least dimensions 3.

For our next topic, completeness, as well as for later purposes, we need some auxiliary facts.

In a timelike space R the set of those points whose distance from a given point exceeds a certain number, plays in many respects the rôle of open ball in metric spaces. We, therefore, introduce the notation ($\sigma \geqslant 0$)

$$F(x, \sigma) = \{y\colon\ x < y \text{ and } xy > \sigma\},$$
$$P(x, \sigma) = \{u\colon\ u < x \text{ and } ux > \sigma\}.$$

We put $F(x, 0) = F(x)$, $P(x, 0) = P(x)$. These sets consist respectively of all points which follow or precede x and are called the *future* and the *past* of x. The closures of $F(x, \sigma)$, $P(x, \sigma)$, $F(x)$, $P(x)$ are denoted by $\overline{F}(x, \sigma)$, $\overline{P}(x, \sigma)$, $\overline{F}(x)$, $\overline{P}(x)$. We prove:

(8) *In a timelike G-space (for $\sigma \geqslant 0$) if $x < y$ then $F(x, xy + \sigma) \supset \overline{F}(y, \sigma) - y$; if $u < x$ then $P(x, ux + \sigma) \supset \overline{P}(u, \sigma) - u$.*

For let $z \in \overline{F}(y, \sigma) - y$. For $z \in F(y, \sigma)$ the assertion follows from $xz \geqslant xy + yz > xy + \sigma$.

Let $z_\nu \in F(y, \sigma)$ and $z_\nu \to z \neq y$. From $xz_\nu \geqslant xy + yz_\nu > xy + \sigma$ follow $x < z$ and $xz \geqslant xy + \sigma$.

Choose w with $(x\, w\, y)$ so close to x that $w < z$ and $w < z_\nu$ for large ν. This is possible because $(<)$ is open. We have

$$wz_\nu \geqslant wy + yz_\nu, \qquad \text{hence} \qquad wz > wy + \sigma.$$

If $wz > wy + \sigma$ then $xz \geqslant xw + wz > xw + wy + \sigma = xy + \sigma$. Let $wz = wy + \sigma$. Then $(x\, w\, z)$ is impossible. This is clear for $\sigma = 0$ because of G_6'. If $\sigma > 0$ choose v with $(w\, v\, z)$ and $wy = wy$. Then $(x\, w\, v)$ and we have again a contradiction to G_6'.

Therefore in either case

$$xz > xw + wz \geqslant xw + wy + \sigma = xy + \sigma.$$

The following slightly stronger statement for $\sigma = 0$ is also needed.

(9) *If* $x < y_\nu < z_\nu$, $y_\nu \to y \neq x$, $z_\nu \to z \neq y$ *and* $x < y$ *then* $xz > xy$. *If* $v < x$, $u_\nu < v_\nu$, $v_\nu \to v \neq x$, $u_\nu \to u \neq v$ *and* $v < x$ *then* $ux > vx$.

For we have $xy_\nu \to xy > 0$ and $xz_\nu \geqslant xy_\nu + y_\nu z_\nu \geqslant xy_\nu \to xy$.

Choosing w as before we have $w < y_\nu$, $w < z_\nu$ for large ν, hence

$$wz_\nu \geqslant wy_\nu + y_\nu z_\nu \geqslant wy_\nu \to wy$$

therefore $wz = wy$ and the proof proceeds as before.

In locally compact metric spaces satisfying postulates analogous to $G'_{4,5}$ completeness, geodesic completeness ($\tau = \mathscr{R}$ for a geodesic $\{\tau, x(t)\}$) and finite compactness (bounded infinite sequences have accumulation points) are equivalent. In Riemannian geometry this is essentially the content of the Hopf-Rinow Theorem.

This theorem does not seem to have an analogue in timelike spaces. Apparently there is no concept of completeness or finite compactness which is shared by all spaces which we want to admit. Examples will be found in the next two sections.

No really interesting timelike space has the property that a sequence of points x_ν with $p < x_\nu$ and $0 < \gamma \leqslant px_\nu \leqslant \beta$ has an accumulation point. However Lorentz spaces are finitely compact with the following definition (see Section 5):

A timelike space is *finitely compact* if a sequence of points x_ν with $p < q \leqslant x_\nu$ and $px_\nu \leqslant \beta$, or $x_\nu \leqslant q < p$ and $x_\nu p \leqslant \beta$, has an accumulation point.

A timelike space is *complete* if $x_\nu < x_{\nu+\mu}$ $(\nu, \mu = 1, 2, \ldots)$ and $\varrho(x_\nu, x_{\nu+\mu}) \leqslant \gamma_\nu$ (or $x_{\nu+\mu} < x_\nu$) and $\varrho(x_{\nu+\mu}, x_\nu) \leqslant \gamma_\nu$) and $\gamma_\nu \to 0$ imply that $\{x_\nu\}$ converges.

This could be formulated locally with $x_\nu \in U(p)$, $x_\nu <_p x_{\nu+\mu}$, etc. However, we are interested in G-spaces where $\overline{U}(p)$ may be taken as compact; then $\{x_\nu\}$ has at least one accumulation point and it can be proved that it has only one.

We prove

(10) *A finitely compact timelike G-space is complete.*

For let $x_\nu < x_{\nu+\mu}$ and $x_\nu x_{\nu+\mu} \leqslant \gamma_\nu$, $\gamma_\nu \to 0$. Then $x_1 x_\nu \geqslant x_1 x_2 > 0$ for $\nu \geqslant 2$ and $x_1 x_\nu \leqslant \gamma_1$. Therefore finite compactness yields the existence of an accumulation point for every subsequence. We must show that only one accumulation point exists. If there were two, q and q', then we would have subsequences $\{i_\nu\}$ and $\{j_\nu\}$ of $\{\nu\}$ with $i_\nu < j_\nu$, $x_{i_\nu} \to q$, $x_{j_\nu} \to q'$. Also let $\{k_\nu\} \in \{i_\nu\}$ with $k_\nu > j_\nu$. Then

$$x_1 x_{j_\nu} \geqslant x_1 x_{i_\nu} + x_{i_\nu} x_{j_\nu}, \qquad x_1 x_{k_\nu} \geqslant x_1 x_{j_\nu} + x_{j_\nu} x_{k_\nu}.$$

Now G_2 implies $x_1 < q$, $x_1 < q'$ and the inequalities yield $x_1 q = \lim x_1 x_{i_\nu} = \lim x_1 x_{j_\nu} = xq'$ contradicting (9).

For comparison with the metric case note that we did not use all axioms for a G-space, only the validity of G_2' and (9), (the proof of (9) uses G_2').

The converse which is valid under certain conditions in the metric case, makes no sense here, because completeness deals only with sequences satisfying $x_\nu < x_{\nu+\mu}$.

A (locally) timelike G-space is *geodesically complete* if $\tau = \mathscr{R}$ for every geodesic $\{\tau, x(t)\}$.

This definition could be applied to any locally timelike space, but would often lead to absurdities, for example, if no proper segments exist.

Assume that for a geodesic $\{\tau, x(t)\}$ in a timelike G-space the set τ has a finite upper bound β. If $t_1 < t_2 < \ldots$ and $t_\nu \to \beta$ then $x(t_\nu) < x(t_{\nu+\mu})$ but—in contrast to the metric case—$x(t_\nu)\, x(t_{\nu+\mu}) \geq t_{\nu+\mu} - t_\nu$. Therefore we cannot conclude that $\{x(t_\nu)\}$ converges if the space is complete. The only result in this direction is:

(11) *For a timelike G-space in which all geodesics are lines completeness implies geodesic completeness. The converse does not hold.*

That the converse is false follows from (4) in Section 5. The following sections will show that geodesic completeness is the most relevant concept. It should be remembered that G_2 has the nature of a completeness postulate but without an analogue in metric spaces.

We conclude this section by defining motion. This is simple for timelike spaces:

A motion of a timelike space R is a topological map Φ of R on itself which maps $(<)$ on itself and preserves distances, i.e. if $x < y$ if and only if $\Phi x < \Phi y$ and $\varrho(\Phi x, \Phi y) = \varrho(x, y)$.

The definition is necessarily more involved for locally timelike spaces owing to the largely arbitrary choice of the $U(p)$ and the possibility that the space cannot be consistently ordered.

A motion of a locally timelike space R is a topological map Φ of R on itself with the following property: Each point has a neighborhood $N(p) \subset U(p)$ such that $\Phi N(p) \subset U(\Phi_p)$ and either $\Phi x <_{\Phi_p} \Phi y$ and $\varrho_{\Phi_p}(\Phi x, \Phi y) = \varrho_p(x, y)$ for all pairs x, y with $x < y$ in $N(p)$ or $\Phi y <_{\Phi_p} \Phi x$ and $\varrho_{\Phi_p}(\Phi y, \Phi x) = \varrho_p(x, y)$ for all pairs x, y with $x < y$ in $N(p)$.

Since this definition is cumbersome we observe:

(12) *A topological map of a (locally) timelike G-space on itself is a motion, if and only if it maps T-curves on T-curves and preserves the lengths of all T-curves.*

4. PRODUCTS OF TIMELIKE AND METRIC SPACES

Our next aim is significant examples to elucidate the theory. In this section S_1 is a timelike space (the results all carry over to the case where S_1 is locally timelike), S_2 is a metric space and $R = S_1 \times S_2$. The distance in S_i is denoted by σ_i, points in S_1 by a, b, c (with sub- and superscripts), points in S_2 by x, y, z and points in R by $p = (a, x)$, $q = (b, y)$, $r = (c, z)$, etc. The set $(<)$ is defined by $p = (a, x) < q = (b, y)$ meaning $a < b$ and $\sigma_1(a, b) > \sigma_2(x, y)$. Transitivity follows for $r = (c, z)$ from

$$(1) \qquad \sigma_1(a, c) \geqslant \sigma_1(a, b) + \sigma_1(b, c) > \sigma_2(x, y) + \sigma_2(y, z) \geqslant \sigma_2(x, z).$$

Moreover, T_2' for S_1 implies T_2' for R.

We impose different distances ϱ_a $(a \geqslant 1)$ on R

$$\varrho_a(p, q) = [\sigma_1^a(a, b) - \sigma_2^a(x, y)]^{1/a}.$$

For $\varrho_1(p, q) = \sigma_1(a, b) - \sigma_2(x, y)$ we find from (1) that $\varrho_1(p, q) + \varrho_1(q, r) \geqslant \varrho_1(q, r)$ with equality only if $\sigma_1(a, b) + \sigma_1(b, c) = \sigma_1(a, c)$ and $\sigma_2(x, y) + \sigma_2(y, z) = \sigma_2(x, z)$.

To prove that the other ϱ_a satisfy the time equality we observe:

(2) *If* $h_1 > k_1 \geqslant 0$, $h_2 > k_2 \geqslant 0$ *and* $a > 1$ *then*

$$(h_1^a - k_1^a)^{1/a} + (h_2^a - k_2^a)^{1/a} \leqslant [(h_1 + h_2)^a - (k_1 + k_2)^a]^{1/a}$$

with equality only when either $k_1 = k_2 = 0$ *or* $h_1 : k_1 = h_2 : k_2$.

Put $\varepsilon_i = (h_i^a - k_i^a)^{1/a}$ *or* $h_i^a = \varepsilon_i^a + k_i^a$. By Minkowski's Inequality [8, p. 31],

$$[(\varepsilon_1 + \varepsilon_2)^a + (k_1 + k_2)^a]^{1/a} \leqslant (\varepsilon_1^a + k_1^a)^{1/a} + (\varepsilon_2^a + k_2^a)^{1/a}$$

with equality only if $k_1 = k_2 = 0$ or $\varepsilon_1 : k_1 = \varepsilon_2 : k_2$.

This is equivalent to the assertion, from which we conclude:

(3) *If* $a > 1$ *then*

$$\varrho_a(p, q) + \varrho_a(q, r) \leqslant \varrho_a(p, r)$$

with equality for $p < q < r$ *only if* $(a\, b\, c)$, $\sigma_2(x, y) + \sigma_2(y, z) = \sigma_2(x, z)$ *and either* $x = y = z$ *or* $\sigma_1(a, b) : \sigma_2(x, y) = \sigma_1(b, c) : \sigma_2(y, z)$.

We denote the space R with the distance ϱ_a by R_a and have proved that R_a is timelike for all a. We turn to the axioms G_1, $G_{1\text{-}6}'$. Obviously

(a) G_1 *holds in* R_a $(a \geqslant 1)$ *if it holds for* S_1 *and* S_2.

(b) G_2' *holds in* R_a $(a \geqslant 1)$ *if it holds in* S_1.

For if $p_\nu = (a_\nu, x_\nu) \to (a, x) = p$, $q_\nu = (b_\nu, y_\nu) \to (b, y) = q$, $p_\nu \leqslant q_\nu$ and $p \leqslant q$ does not hold, then $a \leqslant b$ does not hold, hence $\sigma_1(a_\nu, b_\nu) \to 0$

and

$$\sigma_1(a_\nu, b_\nu) \geqslant \sigma_2(x_\nu, y_\nu) \quad \text{gives} \quad \varrho_a(p_\nu, q_\nu) \to 0.$$

(c) G_3' *holds in* R_a $(a \geqslant 1)$ *if it holds in* S_1.

This follows from the condition for equality in (3). To discuss G_4' we define $(x\,y\,z)$ in S_2 as for S_1, i.e. x, y, z are distinct and $\sigma_2(x, y) + \sigma_2(y, z) = \sigma_2(x, z)$. If S_2 is finitely compact then the points y satisfying $(x\,y\,z)$ lie (for fixed x, z) in a compact set and if for given x, z a point y with $(x\,y\,z)$ exists then any two points of S_2 can be joined by a segment, [G, p. 29]. Therefore the conditions for equality on (3) yield

(d) G_4' *holds in* R_a $(a \geqslant 1)$ *if it holds in* S_1 *and* S_2 *is finitely compact and contains with* $x \neq z$ *a point* y *with* $(x\,y\,z)$.

Similarly one deduces from (3)

(e) G_5' *holds in* R_a $(a \geqslant 1)$ *if it holds in* S_1 *and* S_2 *and either any two points* a, b *in* S_1 *with* $a < b$ *or any two points in* S_2 *can be joined by a segment.*

So far no difference between the cases $a = 1$ and $a > 1$ has appeared. However

(f) *If* S_1 *and* S_2 *satisfy* G_6' *then* R_a *does for* $a > 1$ *but in general not for* $a = 1$.

For, assume $(p\,q\,r_i)$ with $\varrho_a(q, r_1) = \varrho_a(q, r_2)$ $(a > 1)$. Then, if $r_i = (c_i, z_i)$ the conditions for equality in (3) give

$$\sigma_1(a, b) + \sigma_1(b, c_i) = \sigma_1(a, c_i), \quad \sigma_1(a, b) > 0, \quad \sigma_1(b, c_i) > 0,$$

$$\sigma_2(x, y) + \sigma_2(y, z_i) = \sigma_2(x, z_i)$$

and either $x = y = z_i$ or x, y, z_i are distinct and

$$\sigma_1(a, b) : \sigma_2(x, y) = \sigma_1(b, c_i) : \sigma_2(y, z_i).$$

If $x = y$ then $x = y = z_1 = z_2$ and $\sigma_1(b, c_1) = \sigma_1(b, c_2)$ so $c_1 = c_2$ and $r_1 = r_2$. If x, y, z_i are distinct then the above relations imply $\sigma_1(b, c_1) = \sigma_1(b, c_2)$ and $\sigma_2(y, z_1) = \sigma_2(y, z_2)$ and G_6' in S_i gives $c_1 = c_2$ and $z_1 = z_2$ hence $r_1 = r_2$.

The negative assertion for $a = 1$ is one of several enunciated in:

(4) *If* S_1 *is a timelike G-space and* S_2 *is a metric G-space, then* R_1 *satisfies neither* G_6' *nor* (3.8) *or* (3.9). *Also,* R_1 *contains maximal linear sets* $\mu(q, r) = \{[0, \varrho_1(q, r)], p(t)\}$ *for which* $p(t)$ *is not continuous.*

There are points a, b, c_1, c_2 in S_1 and x_1, x_2 in S_2 such that $(a\,b\,c_1)$, $(a\,c_1\,c_2)$ and $\sigma_2(x_1, x_2) = \sigma_1(c_1, c_2)$. Putting $p = (a, x_1)$, $q = (b, x_1)$, $r_i = (c_i, x_i)$ we find

$$\varrho_1(p, q) = \sigma_1(a, b), \qquad \varrho_1(q, r_1) = \sigma_1(b, c_1),$$
$$\varrho_1(q, r_2) = \sigma_1(b, c_2) - \sigma_2(x_1, x_2) = \sigma_1(b, c_1),$$
$$\varrho_1(p, r_2) = \sigma_1(a, c_2) - \sigma_2(x_1, x_2) = \sigma_1(a, c_1) = \varrho_1(p, q) + \varrho_1(q, r_2).$$

So $(p\,q\,r_i)$ and $\varrho_1(q, r_1) = \varrho_1(q, r_2)$ but $r_1 \neq r_2$.

That (3.8) and (3.9) do not hold is seen in a similar way: There are points a, b, c in S_1 and y, z, z_ν in S_2 with $(a\,b\,c)$ and $\sigma_1(b, c) = \sigma_2(y, z)$, $z_\nu \to z$ and $\sigma_2(y, z_\nu) < \sigma_2(y, z)$. Then with $p = (a, y)$, $q = (b, y)$, $r_\nu = (c, z_\nu)$, $r = (c, z)$ the hypotheses $p < q$, $q < r_\nu$, $r_\nu \to r \neq q$ of (3.8) are satisfied but

$$\varrho_1(p, r) = \sigma_1(a, b) + \sigma_1(b, c) - \sigma_2(y, z) = \sigma_1(a, b) = \varrho_1(p, q)$$

hence $F(p, \varrho_1(p, q))$ does not contain $\overline{F}(q) - q$. Putting $q_\nu = q$ yields a negative answer to (3.9).

To construct a discontinuous $p(t)$ consider a segment $\{[0, \sigma_1(a, b)], a(t)\}$ from a to b $(a < b)$ in S_1. Choose values $0 < t' < t' + \varepsilon < \sigma_1(a, b)$ and points x, z in S_2 with $\sigma_2(x, z) = \varepsilon$. Define

$$p(t) = \big(a(t), x\big) \ \text{ in } [0, t'], \qquad p(t) = \big(a(t+\varepsilon), z\big) \ \text{ in } [t', \sigma_1(a, b) - \varepsilon].$$

One readily verifies

$$\varrho_1\big(p(t_1), p(t_2)\big) = t_2 - t_1 \quad \text{for} \quad 0 \leqslant t_1 \leqslant t_2 \leqslant \sigma_1(a, b) - \varepsilon,$$

so $p(t)$ defines a maximal linear set but is not continuous.

Combining our results we have

(5) THEOREM. *If S_1 is a timelike G-space and S_2 is a metric G-space, then R_a is a timelike G-space if $a > 1$. The space R_1 satisfies all axioms but G_6' and has the other properties listed in* (4).

Note. Minor modifications of our arguments show that R_a is for $a > 1$ a locally timelike G-space if S_1 is a locally timelike G-space and S_2 is a metric G-space.

Now we turn to completeness:

(6) THEOREM. *If S_1 is a complete timelike space and S_2 is a complete metric space then R_a is a complete timelike space for $a > 1$, but in general not for $a = 1$.*

Let $p_\nu < p_{\nu+\mu}$ and $\varrho_a(p_\nu, p_{\nu+\mu}) \leqslant \gamma_\nu$ and $\gamma_\nu \to 0$, $a > 1$. Then

$$\varepsilon_\nu = \sigma_1(a_{\nu-1}, a_\nu) - \sigma_2(x_{\nu-1}, x_\nu) > 0.$$

Using the Mean Value Theorem and (1) we find

$$\varrho_a^a(p_\nu, p_{\nu+\mu}) = \sigma_1^a(a_\nu, a_{\nu+\mu}) - \sigma_2^a(x_\nu, x_{\nu+\mu})$$
$$\geqslant \Big[\sum_{\nu+1}^{\nu+\mu} \sigma_1(a_{i-1}, a_i)\Big]^a - \Big[\sum_{\nu+1}^{\nu+\mu} \sigma_2(x_{i-1}, x_i)\Big]^a = \Big(\sum_{\nu+1}^{\mu} \varepsilon_i\Big)^a a c_{\nu,\mu}^{-1}$$

where

$$\sum_{\nu+1}^{\nu+\mu} \sigma_2(x_{i-1}, x_i) \leqslant c_{\nu,\mu} \leqslant \sum_{\nu+1}^{\nu+\mu} \sigma_1(a_{i-1}, a_i).$$

Therefore

$$\sum_{\nu+1}^{\nu+\mu} \varepsilon_i \Big(\sum_{\nu+1}^{\nu+\mu} \sigma_2(x_{i-1}, x_i)\Big)^{a-1} \leqslant a^{-1}\gamma_\nu^a \to 0.$$

There are two cases:

1) All $\sigma_2(x_{i-1}, x_i) = 0$ then $\varrho_a(p_\nu, p_{\nu+\mu}) = \sigma_1(a_\nu, a_{\nu+\mu})$ and the assertion follows from the completeness of S_1.

2) Not all $\sigma_2(x_{i-1}, x_i)$ vanish. The last inequality with $\nu = 1$ shows that both series

$$\sum_1^\infty \varepsilon_i \quad \text{and} \quad \sum_1^\infty \sigma_2(x_{i-1}, x_i)$$

converge. The latter and

$$\sigma_2(x_\nu, x_{\nu+\mu}) \leqslant \sum_{\nu+1}^{\nu+\mu} \sigma_2(x_{i-1}, x_i)$$

shows that $\{x_\nu\}$ is a fundamental sequence in S_2, so that x_ν tends to a point x in S_2. Moreover,

$$\sigma_1^a(a_\nu, a_{\nu+\mu}) = \varrho_a(p_\nu, p_{\nu+\mu}) + \sigma_2^a(x_\nu, x_{\nu+\mu}) \leqslant \gamma_\nu'$$

with $\gamma_\nu' \to 0$, hence a_ν converges to a point a in S_1 and $(a_\nu, x_\nu) \to (a, x)$.

The space L_1^{n+1} which will be defined presently is an example for the negative part of (6) as well as of (7) and (8).

(7) *If S_1 and S_2 are finitely compact then so is R_a for $a > 1$, but not necessarily for $a = 1$.*

For, let $p < q < r_\nu$ and $\varrho_a(p, r_\nu) \leqslant \beta$ $(a > 1)$. Then

$$\sigma_1(a, b) - \sigma_2(x, y) = \gamma > 0, \qquad \sigma_1(b, c_\nu) \geqslant \sigma_2(y, z_\nu)$$

and

$$\beta^a \geqslant \sigma_1^a(a, c_\nu) - [\sigma_1(a, c_\nu) - \gamma]^a + [\sigma_1(a, c_\nu) - \gamma]^a - \sigma_2^a(x, z_\nu),$$
$$\sigma_2(x, z_\nu) - \sigma_2(x, y) \leqslant \sigma_2(y, z_\nu) \leqslant \sigma_1(b, c_\nu) \leqslant \sigma_1(a, c_\nu) - \sigma_1(a, b)$$

or

$$\sigma_2(x, z_\nu) \leqslant \sigma_1(a, c_\nu) - \gamma .$$

Therefore

$$\beta^\alpha \geqslant \sigma_1^\alpha(a, c_\nu) - \big(\sigma_1(a, c_\nu) - \gamma\big)^\alpha ,$$

which implies $\sigma_1(a, c_\nu) \leqslant \beta'$ with a suitable $\beta' < \infty$ and hence $\sigma_2(x, z_\nu) \leqslant \beta'$. The finite compactness of S_1 and S_2 now yields a subsequence $\{k\}$ of $\{\nu\}$ for which c_k and z_k converge to points c, z; so that $r_k \to (c, z)$.

An immediate consequence of (3) is

(8) *If S_1 is a geodesically complete timelike G-space and S_2 is a metric G-space, then R_α is geodesically complete for $\alpha > 1$ but not necessarily for $\alpha = 1$.*

We now consider the most important special case where S_1 is the real axis with the usual order and distance $\sigma_1(u_1, u_2) = u_2 - u_1$ and S_2 is the n-dimensional ($n \geqslant 1$) euclidean space. We introduce coordinates $x^1, \dots, x^n$ so that the distance takes the standard form

$$\sigma_2(x, y) = e(x, y) = \Big[\sum (x^i - y^i)^2\Big]^{1/2}.$$

In this case we denote ϱ_α by λ_α and the space by L_α^{n+1}. Then for $p = (u, x)$ and $q = (v, y)$ the relation $p < q$ means $v - u > e(x, y)$ and

$$\lambda_\alpha(p, q) = [(v-u)^\alpha - e^\alpha(x, y)]^{1/\alpha}, \qquad \alpha \geqslant 1.$$

In particular, $\lambda_2(p, q)$ defines the $(n+1)$-dimensional Lorentz space L_2^{n+1} [2].

(9) *In L_α^{n+1} ($\alpha \geqslant 1$) the affine lines*

$$r(t) = [1 - \lambda_\alpha^{-1}(p, q)t]p + t\lambda_\alpha^{-1}(p, q)q, \qquad p < q, \ -\infty < t < \infty$$

are geodesics and they are lines, i.e.

$$\lambda_2\big(r(t_1), r(t_2)\big) = t_2 - t_1 \quad \textit{for} \quad t_1 < t_2 .$$

They are the only geodesics when $\alpha > 1$, there are others if $\alpha = 1$, in particular such for which the corresponding set τ is bounded.

This implies that L_α^{n+1} is geodesically complete for $\alpha > 1$ but not for $\alpha = 1$.

All statements in (9) except the very last follow from (3). Cosider $x(s) = (s, 0, \dots, 0)$ and $u(s) = s + \operatorname{arc\,tan} s$. Then $r(s) = \big(u(s), x(s)\big)$ is a line (but s is not arc length) because

$$r(s_1) < r(s_2) < r(s_3) \qquad \text{for} \qquad s_1 < s_2 < s_3$$

---

(2) Strictly speaking L_2^{n+1} is the restriction to the pairs $(u, x) \leqslant (v, y)$ of the indefinite metric $|(v-u)^2 - e^2(x, y)|^{1/2}$ defined for all pairs which is usually denoted as Lorentz space.

and

$$\lambda_1(r(s_1), r(s_2)) = \arctan s_2 - \arctan s_1 \quad \text{for} \quad s_1 < s_2.$$

However, the range of the length t if $s = 0$ corresponds to $t = 0$, is $\tau = (\pi/2, \pi/2)$.

The sequence $r(\nu)$ $(\nu = 1, 2, \ldots)$ also shows that L_1^{n+1} is neither complete nor finitely compact, confirming the last statement in (6) and (7). Using the preceding results we have

(10) *The spaces* L_a^{n+1} $(a > 1)$ *are finitely compact and geodesically complete G-spaces in which all geodesics are lines.*

The spaces L_1^{n+1} *satisfy all axioms but* G_6'. *They are neither complete, nor finitely compact nor geodesically complete.*

However, L_1^{n+1} has the property that any segment can be extended to a geodesic $\{\tau, x(t)\}$ for which $\tau = \mathscr{R}$. This is easily verified.

We mentioned that L_2^{n+1} does not have the property that every sequence of points q_ν with $p < q_\nu$ and $0 < \alpha \leqslant \lambda_2(p, q_\nu) \leqslant \beta$ has an accumulation point. This is true for all a: If $e(x, x_{\nu+1}) = \nu^{1/a}$, $q = (0, x_1)$, $p_\nu = ((\nu+1)^{1/a}, x_{\nu+1})$ then $\lambda_a(q, p_\nu) = 1$ but no subsequence of $\{p_\nu\}$ converges.

Finally we mention the obvious fact

(11) *If* Φ_1 *is a motion of the timelike space* S_1 *and* Φ_2 *is a motion of the metric space* S_2, *then* $(a, x) \to (\Phi_1 a, \Phi_2 x)$ *defines a motion of* R_a $(a \geqslant 1)$.

5. TIMELIKE MINKOWSKI SPACES

We now discuss the type of timelike G-space which is basic in the sense that under differentiability hypotheses any timelike G-space behaves locally like a space of this type.

We call these spaces Minkowski spaces with a twofold justification: they are the obvious analogue to the metric Minkowski spaces and they comprise the Lorentz space L_2^4 which in relativity is often called the Minkowski space.

A *timelike Minkowski space* is defined by the following properties: it is a timelike G-space for which the underlying space R is the n-dimensional affine space A^n $(n \geqslant 2)$ with the usual topology and the translations of A^n are motions.

In order to discribe these spaces we need certain functions: In terms of affine coordinates the function $f(x) = f(x', \ldots, x^n)$ is a *gauge function* if

(a) $f(x)$ *is defined on an open convex set* $D\ (\neq \emptyset)$ *which is a cone with the origin as apex, i.e.* $\lambda D = D$ *for* $\lambda > 0$,

(b) $f(\lambda x) = \lambda f(x)$ *for* $\lambda > 0$,

(c) $f(x) > 0$,

(d) $f(x_\nu) \to 0$ *if* $x_\nu \to x \notin D$,

(e) $f\big((1-\theta)x+\theta y\big) > (1-\theta)f(x)+\theta f(y)$ *for* $0 < \theta < 1$ *unless* $x = \lambda y$ *with* $\lambda > 0$.

It follows from (b) and (c) that the origin 0 does not belong to D. The property expressed by (e) may be called *strong concavity* of $f(x)$ in analogy to strong convexity, see [G, p. 99]. It has two important equivalent forms:

(e′) $f(x+y) > f(x)+f(y)$ *unless* $x = \lambda y$ *with* $\lambda > 0$.

(e″) *The set* $f(x) \geqslant \varrho > 0$ *is strictly convex, i.e.* $f(x) \geqslant \varrho$, $f(y) \geqslant \varrho$ *and* $x \neq y$ *imply* $f\big((1-\theta)x-\theta y\big) > \varrho$.

This equivalence is established as for convex $f(x)$ [G, pp. 99, 100] and can, in fact, be reduced to the convex case by observing that $-f(x)$ is convex. Therefore $f(x)$ is continuous, and if we put $f(0) = 0$ then (d) implies that $f(x)$ is continuous at 0.

Denote the boundary of D by C.

(1) *The set* $D \cup C$ *possesses at* 0 *a supporting hyperplane which intersects* $D \cup C$ *only at* 0.

For, denote by E the intersection of all closed half spaces bounded by hyperplanes through 0 and containing the set $\{x, f(x) \geqslant 1\}$. Since this set is strictly convex, there is a hyperplane through 0 intersecting E only at 0. But any half space whose boundary contains 0 and which contains $f(x) \geqslant 1$ also contains $f(x) \geqslant \varrho$ by (b) for any ϱ, so $E \supset D$ and $E \supset D \cup C$. It is easily seen that actually $E = D \cup C$.

We now establish the relation between Minkowski spaces and gauge functions.

(2) THEOREM. *Let* R *be a timelike Minkowski space with affine coordinates* $x = (x^1, \dots, x^n)$ *and distance* $\varrho(x, y)$. *Then* $\varrho(x, y) = f(y-x)$, *where* $f(x)$ *is a gauge function and* $x < y$ *is equivalent to* $f(y-x) > 0$.

Conversely, if a gauge function $f(x)$ *in* A^n *is given and* $x < y$ *is defined by* $f(y-x) > 0$, *then* $\varrho(x, y) = f(y-x)$ *defines a timelike Minkowski space.*

We prove the second part first. The topological properties $\mathbf{T}_1$, $\mathbf{G}_1$ are trivial. The continuity of $\varrho(x, y)$ follows from that of $f(x)$, and so does the openness of $(<)$. The existence of x, y with $x < q < y$ in a given

$W(q)$ is clear from (b), the time inequality is contained in (e′). G_2' follows from (d).

Let $(x\,y\,z)$ or $f(y-x)+f(z-y) = f(z-x)$; then (e′) gives $(y-x) = \lambda(z-y)$ with $\lambda > 0$ or

$$y = \frac{1}{1+\lambda}x + \frac{\lambda}{1+\lambda}z, \qquad z = -\frac{x}{\lambda} + \frac{1+\lambda}{\lambda}y,$$

which proves G_2' and G_6' and reversing the argument gives G_3', G_4'.

Now let $R = A^n$ be a timelike Minkowski space with $x' = x+a$ as motions. Then $x < y$ implies $x+a < y+a$, in particular $0 < y-x$ and $\varrho(x, y) = \varrho(0, y-x)$. Put $f(x) = \varrho(0, x)$. Then $\varrho(x, y) = f(y-x)$ for $x < y$. Let D be the set $f(x) > 0$. If $0 < x$, $0 < y$ then $x < x+y$ and $\varrho(0, x+y) \geqslant \varrho(0, x) + \varrho(x, x+y)$.

(α) *If $x \in D$, $y \in D$ then $x+y \in D$ and $f(x+y) \geqslant f(x)+f(y)$.*

We call z a *midpoint* of x and y if $x < z < x$ and $\varrho(x, z) = \varrho(z, y) = \varrho(x, y)/2$.

(β) *If z is a midpoint of x and y then so is $u = x+y-z$.*

For $u-x = y-z$, $y-u = z-x$ give $f(u-x) = f(y-z) > 0$ hence $x < u$, similarly $u < y$ and $f(u-x) = f(y-z) = f(x-y)/2 = f(z-x) = f(y-u)$. This implies

(γ) *If a midpoint z of x and y exists and is unique then $z = (x+y)/2$.*

If the segment $\sigma(x, y)$ is unique $(x < y)$ then any two points v, w with $v < w$ on $\sigma(x, y)$ have a unique midpoint because of (1.2) and (2.5). We conclude from (γ) that $\frac{3}{4}x+\frac{1}{4}y$ and $\frac{1}{4}x+\frac{3}{4}y$ are the midpoints of x, z and z, x respectively. Using the continuity of $f(x)$ we find that $\sigma(x, y)$, if unique, is the affine segment consisting of the points $x_\theta = (1-\theta)x + \theta y$ $(0 \leqslant \theta \leqslant 1)$ and that

$$\varrho(x, x_\theta) = \theta\varrho(x, y) \qquad \text{for} \qquad 0 \leqslant \theta \leqslant 1.$$

Now segments are locally unique, see (2.5) and hence locally affine segments. This shows that any geodesic curve $x(t)$, $a \leqslant t \leqslant b$, is an affine segment, hence the only geodesic curve from $x(a)$ to $x(b)$. On the other hand, there is a $\sigma(x(a), x(b))$ and this is a geodesic curve. It follows that the geodesics are affine lines and that for $x < y$ the geodesic through x and y has the form

(δ) $z(t) = \left(1-t\varrho^{-1}(x, y)\right)x + t\varrho^{-1}(x, y)y, \quad -\infty < \varepsilon < \infty, \; t$ arclength.

Applying (δ) with $x = 0$ we find $f(\lambda y) = \lambda f(y)$ for $\lambda > 0$.

This and (α) prove (e) via (e′), and (d) follows from G_2'. Because $0 < a$, $0 < b$ imply $0 < a+b$, hence $0 < (a+b)/2$, the set D is convex.

A consequence of this discussion is:

(3) *A timelike Minkowski space is geodesically complete and all its geodesics are lines.*

In general the space will not be complete, but there is a simple criterion for completeness. To formulate it we remember that $f(x-a) > 0$ is the future $F(a)$ of a.

$F(a, \sigma)$ denotes the set of x satisfying $\varrho(a, x) > \sigma$, in our case $f(x-a) > \sigma$. The closure $\bar{F}(a, \sigma)$ is $\{x, f(x-a) \geqslant \sigma\}$. The boundary of $F(a)$ is, according to the language of relativity, the light cone $C(a)$. With the previous notation $F(0) = D$, $C(0) = C$.

(4) *A timelike Minkowski space is complete or finitely compact if and only if no hyperplane exists which separates a generator of $C(0)$ from $\bar{F}(0, 1)$.*

We could, of course, have used any point a and any $\sigma > 0$ instead of 0 and 1. A generator G of $C(0)$ is a ray with origin 0 lying on $C(0)$. It is clear that a hyperplane separating $C = C(0)$ from $\bar{F} = \bar{F}(0, 1)$ must be parallel to G. Because of (4.10) we must prove:

1) *If the condition in (4) is satisfied then the space is finitely compact.*

2) *If it is not satisfied then the space is not complete.*

To show 1) let $0 < y$ or $f(y) > 0$ and $y \leqslant z_\nu$ with $f(z_\nu) = \varrho(0, z_\nu) \leqslant \beta$. No generator G' of $C(y)$ lies outside $F(0, \beta)$ because then a suitable hyperplane through G' would separate $F(0, \beta)$ from the generator of C parallel to G'. Therefore the intersection of $\{x, f(x) \leqslant \beta\}$ and $\bar{F}(y)$ is compact and the sequence $\{z_\nu\}$ has an accumulation point.

For 2) let H be a hyperplane separating the generator G of C from $\bar{F}$. Let G_1 be any other generator of C and in the plane determined by G and G_1. Let G, G_1 be the non-negative x- and y-axes of an affine coordinate system such that $G_1 \cap H$ is the point $(0, 1)$. Put $p_i = (x_i, y_i)$ $(i = 0, 1, 2, \ldots)$ with $x_i = i$ and $y_i = 1 - 2^{-i}$. Then $\varrho(p_0, p_i) < 1/2$ and more generally $\varrho(p_k, p_{k+j}) < 2^{-k}$, but the sequence $\{p_i\}$ does not converge.

An explicit example of a non-complete timelike Minkowski space is given by $D = F(0) = \{x\colon x^1 > 0, x^2 > 0\}$ and $f(x) = x^1 x^2 (x^1 + x^2)^{-1}$, where $f(x) = 1$ is the branch $x_1 > 0$ of the hyperbola $(x^1 - 1)(x^2 - 1) = 1$. Such examples show that a geodesically complete timelike space with a transitive group of motions need not be complete, whereas any locally compact metric space with a transitive group of motions is complete.

The significance of G_2' as a completeness condition may be seen from the following observation: In a timelike Minkowski space with gauge function $f(x)$ restrict $f(x)$ to an open convex subcone of $F(0)$ with apex 0. Then all axioms except G_2' are satisfied.

Let A^n be the underlying space of a locally timelike space G-space for which the translations are motions. The crucial arguments in the preceding proof were (β) and (γ) and both are local. We conclude therefore that the affine lines are geodesics and that (δ) holds. The space satisfies the hypotheses of (1.3) and nothing new is gained.

However, consider a locally timelike space R which is a manifold of dimension $n \geqslant 2$ and possesses a transitive abelian group of motions. Then no motion except the identity has fixed points. (For if $\varphi a = a$ and $b = \psi a$ then $\varphi b = \varphi\psi a = \psi\varphi a = \psi a = b$.) The group is therefore simply transitive and the space can be identified with the group space. According to a theorem of Pontrjagin [14, p. 170] the group is the product of n groups isomorphic either to the real numbers or to the circle group. The universal covering space R is the affine space A^n and the lifted group is the group of translations. So we have the previous case. R possesses a consistent ordering, but the hypothesis of (1.3) will not be satisfied unless R is simply connected. We may express this as follows:

(5) *A locally timelike G-space which is a topological manifold and possesses a transitive abelian group of motions is locally Minkowskian.*

The Möbius strip and the one-sided torus can also be provided with locally timelike Minkowski metrics, but not with arbitrary ones, because the reflection in some line must exist (as for example in L_2^2). This leads to spaces which cannot be consistently ordered.

6. PROJECTIVITIES OF CONVEX HYPERSURFACES

Sections 7 and 8 lead to problems concerning convex hypersurfaces which are of considerable independent interest, but have only been partly solved, although they sound very simple. What is known concerning these problems is the content of the present section. In order not to interrupt the discussion later we first establish a lemma of a general nature which is due to Montgomery:

(1) *If a Lie group Γ acts transitively on the manifold M then so does its identity component Γ_0.*

The number of components of Γ is finite or countable and they have the form $\varphi_i\Gamma_0$, $\varphi_i \in \Gamma$. For any point $p \in M$, let $\Gamma_0(p)$ be the orbit of p under Γ_0. Because Γ is transitive on M, $\bigcup_i \varphi_i\Gamma_0(p) = M$ and each $\varphi_i\Gamma_0(p)$ is homeomorphic to $\Gamma_0(p)$, which is the union of a countable number of compact sets. Therefore

$$\dim \varphi_i\Gamma_0(p) = \dim M \quad \text{for all } i.$$

It follows that $\varphi_i \Gamma_0(p)$ contains an open subset of M, see [9, p. 46], and hence is open. If $\Gamma_0(p) \neq M$ then $M - \Gamma_0(p)$ would be the union of some of the $\varphi_i \Gamma_0(p)$ and M would not be connected.

Consider now a closed convex hypersurface K in A^n $(n \geqslant 2)$ and let I be its interior, put $K^0 = K \cup I$. We complete A^n to the n-dimensional projective space P^n by adding a hyperplane $x^{n+1} = 0$ and will be concerned with the group Γ_K of projectivities which map K on itself. These take also I and $E = P^n - K^0$ into themselves. Conversely, a projectivity which maps I or E on itself also maps K on itself. We show first:

(2) *If Γ_K is transitive on K then K is an ellipsoid.*

The theorem is elementary for $n = 2$, so we may assume $n > 2$. Then K is simply connected and by (1) the identity component Γ_0 of Γ_K acts transitively on K. According to a theorem of Montgomery [11, p. 226] Γ_0 contains a compact subgroup Γ_c acting transitively on K. Now a result due to the collective effort of several mathematicians, see Nagano [13], implies that Γ_c is isomorphic to a subgroup of the orthogonal group $O(n)$ and this gives readily the assertion.

A much more elementary method of deducing the assertion from the existence of Γ_c is found at the end of this section.

For $n = 3$ we could have referred to Lie [10] who determined all surfaces possessing transitive groups of projectivities onto themselves. However, he always assumes that the surface is analytic, and in this area innocent looking smoothness hypotheses may change completely the character of a problem, see Theorem 7.

The principal unsolved problem is finding all K for which Γ_K is transitive on I. The problem would be quite accessible if it were known that under this hypothesis different orbits of Γ_K have different closures or that the number of distinct orbits is finite. We solve the problem only in special cases.

For this purpose we consider the metrization of I as a Hilbert geometry (for details see [G, Section 18]).

We put $h(a, a) = 0$ and if a, b are distinct points of I let the projective line $\overline{ab}$ through a and b intersect K in x and y. Remembering that for any permutations (i_1, i_2) and (j_1, j_2) of $(1, 2)$

$$R(a_{i_1}, a_{i_2}, x_{j_1}, x_{j_2}) = [R(a_1, a_2, x_1, x_2)]^{\pm 1},$$

where $R(\ )$ denotes the crossratio, we put

$$h(a, b) = |\log R(a, b, x, y)|. \tag{3}$$

Then $h(a, b)$ satisfies the axioms for a metric space. With this metric I is finitely compact, the intersections of the projective lines with I are isometric to R and hence are geodesics. They are the only geodesics, unless there are two proper segments on K whose convex hull contains points of I. (The proof of [G, (18.5)] leads to this result, but the formulation of (18.5) does not quite express it.)

We prove first

(4) *$K^0 = K \cup I$ is a simplex if and only if Γ_K has a subgroup Γ_a which is abelian and transitive on I.*

The set $x^i > 0$ $(i = 1, \ldots, n)$ in A^n may be considered as the interior I of a simplex and the maps $\bar{x}^i = \beta^i x^i$, $\beta^i > 0$ $(i = 1, \ldots, n)$ provide a transitive abelian group of projectivities of I.

For a proof of the necessity we observe that for $\varphi \in \Gamma_a$ the distance $h(x, \varphi x)$ is independent of x. For if $y \in I$ then $\psi \in \Gamma_a$ with $\psi x = y$ exists and $h(y, \varphi y) = h(\psi x, \varphi\psi x) = h(\psi x, \psi\varphi x) = h(x, \varphi x)$. Let $u \in K$, $\varphi \in \Gamma_a$ and $u \neq \varphi u = v$. We will show that u is not an extreme point of K^0.(3) Let $a_\nu \in I$ with $a_\nu \to u$. Then $b_\nu = \varphi a_\nu \to \varphi u = v$. If $\overline{a_\nu b_\nu}$ intersects K in x_ν and y_ν and the names are such that the order is $y_\nu, a_\nu, b_\nu, x_\nu$ then

$$h(a_\nu, b_\nu) = \log R(a_\nu, b_\nu, x_\nu, y_\nu).$$

For a subsequence $\{i\}$ of $\{\nu\}$ we have $x_i \to x^*$, $y_i \to x^*$. Since $h(a_\nu, b_\nu)$ is independent of ν, it follows that $x^* \neq v$ and $y^* \neq u$. Therefore u lies in the interior of a segment on K and hence is not an extreme point of K^0.

Thus every element of Γ_a leaves each extreme point of K fixed. Since K^0 is the closure of the convex hull of the set of extreme points and $\dim K^0 = n$, there are at least $n+1$ extreme points which do not lie in a hyperplane. There cannot be more since the elements of Γ_a are projectivities.

A hyperplane contains at most n extreme points no $n-1$ of which lie in an $(n-2)$ flat; otherwise each element of Γ_a would leave the hyperplane pointwise fixed and Γ_a would not be transitive on I. Proceeding in this way we see that K^0 has precisely $n+1$ extreme points and hence is a simplex.

A point p of K is an *Euler point* if the following holds: K is differentiable at p, all sections of K by two-flats through p which do not lie in the tangent hyperplane have curvatures at p, and these obey the classical relations of Meusnier and Euler. Almost all points of K are Euler points (see [3, p. 23]). The Gauss curvature at an Euler point can be defined in the usual way as product of the principal curvatures.

Our aim is to show that K is an ellipsoid if it possesses an Euler point with non-vanishing Gauss curvature and Γ_K is transitive on I.

(3) The definition and the properties of extreme points used here can be found in [1, Section 10].

The proof is based on a lemma whose explicit formulation would be very long.

In the s-dimensional euclidean space E^s with Cartesian coordinates $x^1, \ldots, x^s$ let two closed continuous hypersurfaces which are starshaped with respect to the origin be given by

$$|x| = r(u) \quad \text{and} \quad |x| = \varrho(u), \quad u \in S$$

where S is the unit sphere. The numbers

$$\min r(u) = r_m, \quad \max r(u) = r_M, \quad \min \varrho(u) = \varrho_m, \quad \max \varrho(u) = \varrho_M$$

are finite and positive.

Let φ be a projectivity of the E^s completed to P^s which leaves O fixed and takes the first hypersurface into the second:

$$\varphi\big(r(u)u\big) = \varrho(v)v, \quad v = v(u), \quad |v| = 1.$$

Then φ possesses a unique representation of the form

$$y^i = \frac{\sum_{j=1}^{s} a^i_j x^j}{\sum_{j=1}^{s} b_j x^j + 1} \quad \text{or} \quad y = \frac{Ax}{b\cdot x + 1} \quad \text{with } \det A \neq 0.$$

The lemma states that

$$(5) \qquad |a^i_j| \leqslant 2\varrho_M r_m^{-1}, \quad |\det A| \geqslant r_M^{-s} \varrho_m^s, \quad |b_j| < r_m^{-1}.$$

The last relation is true if $b = (b_1, \ldots, b_s) = 0$. Assume $b \neq 0$, and put $u_b = b|b|^{-1}$. We conclude from

$$\varrho(v)v = Ar(u)u\big(b\cdot r(u)u + 1\big)^{-1} = Au\big(b\cdot u + r^{-1}(u)\big)^{-1}$$

that $b\cdot r(u)u + 1 \neq 0$ for all u. Since $b\cdot u_b > 0$, we see that

$$br(-u_b)(-u_b) + 1 = -|b|r(-u_b) + 1 > 0,$$

hence

$$|b_j| \leqslant |b| < r^{-1}(-u_b) \leqslant r_m^{-1}.$$

Next put $u_k = \varepsilon(\delta^1_k, \ldots, \delta^s_k)$, $v(u_k) = v_k$ where $\varepsilon = 1$ if $b_k \geqslant 0$, $\varepsilon = -1$ if $b_k < 0$. Then

$$\varrho^2(v_k) = \sum_i \big((v_m)v^i_k\big)^2 = \sum_j (a^i_k)^2\big(|b_k| + r^{-1}(u_k)\big)^{-2}$$

and from $|b_k| \leqslant r_m^{-1}$

$$\varrho_M^2 \geqslant (a^j_k)^2 4^{-1} r_m^2$$

which gives the first inequality in (5).

Finally let A_i^k be the cofactor of a_i^k in the matrix A so that $(A_i^k \det^{-1} A) = A^{-1}$. Put $A_i = \left(\sum_k (A_i^k)^2\right)^{1/2} > 0$. Choose $\varepsilon = \pm 1$ such that $\varepsilon \sum_k b_k A_i^k \geqslant 0$ and put

$$w_i = \varepsilon A_i^{-1}(A_i^1, \ldots, A_i^s).$$

Then

$$\varrho^2\big(v(w_i)\big) = \sum_j \Big(\sum_k a_k^j A_i^k A_i^{-1}\Big)^2 \Big(\varepsilon \sum_k b_k A_i^k A_i^{-1} + r^{-1}(w)\Big)^2 \leqslant A_i^{-2} \det^2 A\, r_M^2$$

and

$$\varrho_m^{2s} r_M^{-2s} \leqslant \prod_i A_i^2 \det^{2s} A \leqslant \det^2(A_i^k) \det^{2s} A = \det^2 A,$$

where the second estimate follows from Hadamard's Inequality, [8, p. 34].

As an application we have

(6) *Let $\{K_\nu^1\}$ and $\{K_\nu^2\}$ be sequences of closed convex hypersurfaces in E^s tending to closed convex hypersurfaces K^1, K^2 containing the origin in their interiors. If a projectivity of E^s (completed to P^s) exists which leaves 0 fixed and takes K_ν^1 into K_ν^2, then there is a projectivity φ taking K^1 into K^2 (leaving 0 fixed).*

It is easy to verify with examples that assuming 0 to remain fixed is essential.

Let K_ν^1 and K_ν^2 be given by

$$|x| = r_\nu(u) \quad \text{and} \quad |x| = \varrho_\nu(u), \quad u \in S.$$

Then for large ν and suitable $r_1, r_2, \varrho_1, \varrho_2$

$$0 < r_1 \leqslant r_\nu(u) \leqslant r_2, \quad 0 < \varrho_1 \leqslant \varrho_\nu(u) \leqslant \varrho_2.$$

The projectivity φ_ν has the form

$$y = A_\nu x (b_\nu \cdot x + 1)^{-1}, \quad \det A_\nu = \det(a_{i\nu}^k) \neq 0$$

and we conclude from (4) that

$$|a_{i\nu}^k| \leqslant 2\varrho_2 r_1^{-1}, \quad |\det A_\nu| \geqslant r_2^{-s} \varrho_1^s, \quad |b_{j\nu}| < r_1^{-1}.$$

For a suitable subsequence $\{\mu\}$ of $\{\nu\}$ we have $a_{i\mu}^k \to a_i^k$, $b_{j\mu} \to b_j$ with $\det A \neq 0$ and

$$\varphi:\ y = Ax(bx + 1)^{-1}$$

defines a projectivity satisfying the assertion.

Consider now a closed convex hypersurface K in E^n which possesses an Euler point p with non-vanishing Gauss curvature. Denote by H the tangent hyperplane of K at p and by H_σ the hyperplane parallel to H at distance $\sigma > 0$ intersecting K, let p_σ be the intersection of H_σ with the normal to K at p. Project $C_\sigma = K \cap H_\sigma$ parallel to this normal on H

and dilate it in the ratio $1 : \sqrt{\sigma}$ obtaining $\bar{C}_\sigma$. Then $\bar{C}_\sigma$ tends for $\sigma \to 0+$ to an ellipsoid, namely to the Dupin Indicatrix of K at p, see [3, Section 3].

Assume that Γ_K is transitive on I and give $q \in I$. A projectivity which maps H_σ on a hyperplane H'_σ through q sends C_σ into $C'_\sigma = H'_\sigma \cap K$. Also, $\varphi_\sigma | H_\sigma$ is a projectivity of H_σ on H'_σ.

For a suitable sequence $\sigma_\nu \to 0+$ the H'_{σ_ν} converge to a hyperplane H' and $C'_{\sigma_\nu} \to C' = H' \cap K$. A motion of E^n takes C'_{σ_ν} into C^2_ν in H such that q goes into p and C^2_ν converges to an image C^1 of C' under a motion. With $\bar{C}_{\sigma_\nu} = C^1_\nu$ we have $\psi_\nu C^1_\nu = C^2_\nu$ where ψ_ν is a projectivity of H which leaves p fixed. It follows from (5) that a projectivity ψ of H will take C^1 into C^2. Therefore C^1 and hence C' is an ellipsoid.

This discussion also shows that all sections of K by hyperplanes parallel to H^1 on at least one side of H' must be ellipsoids homothetic to C'. Since q was arbitrary in I, this yields readily that K is an ellipsoid. Thus we proved:

(7) THEOREM. *Let K be a closed convex hypersurface in E^n which possesses an Euler point with non-vanishing Gauss curvature. If the interior I of K possesses a transitive group of projectivities then K is an ellipsoid.*

The simplex and other examples show that the hypothesis concerning the Euler point is essential.

(7) not only substantiates our earlier statement that very harmless looking conditions may prove highly restrictive, but (7) will most probably also be essential in determining all K for which Γ_K is transitive on I by allowing inductive reduction of the dimension.

In addition the arguments leading to (7) provide an elementary proof for the fact (deduced under (2) from deep results on transformation groups) that K is an ellipsoid if Γ_K possesses a compact subgroup Γ_c which is transitive on K. Every point is an Euler point hence the Gauss curvature cannot vanish. Let $p \in I$. The orbit $\Gamma_c(p)$ of p under Γ_c is compact, hence stays away from K. The previous arguments show that through some point q of $\Gamma_c(p)$ there passes a hyperplane H such that the intersections of K with the hyperplanes parallel to H on at least one side of H are homothetic ellipsoids. The same must hold for every point of $\Gamma_c(p)$ and hence for every point of I.

Concerning the K with Γ_K transitive on I the following conjecture seems reasonable:

The convex bodies $K^0 = K \cup I$ for which Γ_K is transitive on I are convex hulls of a finite number of points and solid ellipsoids of dimension $\geqslant 2$ with the property that the dimension of the hull decreases if one of the points is omitted or one of the ellipsoids is replaced by a lower dimensional convex subset.

7. MINKOWSKI SPACES WITH PAIRWISE TRANSITIVE GROUPS OF MOTIONS

Consider again an n-dimensional timelike Minkowski space with affine coordinate $x^1, \ldots, x^n$ and the translations $x' = x + a$ as motions. In this section we are interested in additional motions and show first:

(1) *A motion φ of a timelike Minkowski space is an affinity.*

The proof uses

(2) *Given a finite number of points $a_1, \ldots, a_k$ then p with $p < a_i$ $(i = 1, \ldots, k)$ exists.*

This is obvious because $F(0)$ and hence $F(p) = F(0) + p$ is an open cone.

Putting generally $\varphi x = x'$ we must show that $a_2 = (1-t)a_1 + ta_3$ $(0 < t < 1)$ implies $a_2' = (1-t)a_1' + ta_3'$. Let $p < a_i$ $(i = 1, 2, 3)$; then $p' < a_i'$. Since translations are motions, we may assume $p = p' = 0$. Because the set $(<)$ is open in $A^n \times A^n$ we have $b_1 = \mu_1 a_1 < a_3$ for small $\mu_1 > 0$. The ray λa_2 $(\lambda \geqslant 0)$ intersects the affine segment from b_1 to a_3 in a point b_2, so with suitable λ, μ_2 between 0 and 1

$$b_2 = \mu_2 a_2 = (1-\lambda) b_1 + \lambda a_3.$$

It follows from (δ) in Section 5 that $b_1' = (\mu_1 a_1)' = \mu_1 a_1'$ and similarly $b_2' = \mu_2 a_2' = (1-\lambda) b_1' + a_3'$.

The following lemma is needed:

(3) *In the (x, y)-plane let $y = g(x)$ be a strictly convex decreasing and differentiable curve in (γ, ∞), where $\gamma \geqslant 0$, $g(x) > 0$ and $g(x) \to \infty$ for $x \to \gamma$. Let the tangent at $q = (x, g(x))$ intersect the x-axis at q_x and the y-axis at q_y. If*

$$\beta \geqslant |q - q_y| : |q - q_x| \geqslant \alpha > 0$$

then $\gamma = 0$ and $\lim\limits_{x \to \infty} g(x) = 0$.

If $|q - q_y| : |q - q_x| = \alpha > 0$ then $g(x) = kx^{-\alpha}$ $(0 < x < \infty)$.

For $q_y = (0, g(x) - xg'(x))$ hence

$$\beta \geqslant |q - q_y| : |q - q_x| = -xg'(x)/g(x) \geqslant \alpha,$$

whence

$$\frac{g(x)}{g(\gamma+1)} \leqslant \left(\frac{\gamma+1}{x}\right)^{\beta} \quad \text{for} \quad x \leqslant \gamma + 1,$$

$$\frac{g(x)}{g(\gamma+1)} \leqslant \left(\frac{\gamma+1}{x}\right)^{\alpha} \quad \text{for} \quad x \geqslant \gamma + 1.$$

A group Γ of motions of a timelike space is called *pairwise transitive* if given $p < q$ and $p' < q'$ with $\varrho(p, q) = \varrho(p', q')$ a motion in Γ exists which takes p into p' and q into q'. A *triplewise transitive* group is defined analogously for $p < q < r$, $p' < q' < r'$ with $\varrho(p, q) = \varrho(p', r')$, $\varrho(p, r) = \varrho(p', r')$, $\varrho(q, r) = \varrho(q', r')$.

A metric Minkowski space with a pairwise transitive group of motions is euclidean, [G, p. 101]. A timelike Minkowski space with a pairwise transitive group of motions need not be a Lorentz space. It is when the group of motions is triplewise transitive. The problem of finding all timelike Minkowski spaces with pairwise transitive groups of motions leads to difficult unsolved problems on convex bodies and hypersurfaces, including the one concerning the K with Γ_K transitive on I. The remainder of this section discusses the facts which are known in this direction.

(4) *A timelike Minkowski space with a pairwise transitive group of motions is finitely compact.*

If $p < q$, $p < q'$ and $\varrho(p, q) = \varrho(p, q') = \Delta > 0$ then a motion exists which leaves p fixed and takes q into q'. Applying this to $q = 0$ and using (1) we see that the convex hypersurface $K_\Delta, f(x) = \Delta$ possesses a transitive group of central affinities. Since any convex hypersurface is almost everywhere differentiable (even twice differentiable), see [3, p. 23], K_Δ is everywhere differentiable.

Let $q \in K_\Delta$. The tangent plane T_q of K_Δ at q intersects the light cone $C(0)$ in a set S_q homeomorphic to S^{n-2}. A line L in T_q through q intersects S_q in two points s, s'. The ratio $|q-s| : |q-s'|$ of the cartesian distances is an affine invariant. As S traverses S_q this ratio attains a maximum β and minimum $\alpha = \beta^{-1}$, which are independent of q.

According to (5.4) it suffices to show that no hyperplane separates a generator of $C(0)$ from K_Δ and this follows at once from (3).

The two dimensional case is easily handled.

(5) *A timelike Minkowski plane with a pairwise transitive group of motions can in suitable affine coordinates x^1, x^2 be represented as follows: $x < y$ means $x^1 < y^1$ and $x^2 < y^2$ and $\varrho(x, y) = (y^1 - x^1)^{1-\mu}(y^2 - x^2)^\mu$ with $0 < \mu < 1$.*

Note. L_2^2 corresponds to $\mu = 1/2$ and is characterized geometrically by the fact that it can be reflected in each geodesic.

For a proof we choose x^1, x^2 such that $F(0)$ is the first quadrant $x^1 > 0$, $x^2 > 0$. A centro-affinity mapping $q \in K_\Delta$ on $q' \in K_\Delta$ leaves the ratio $|q - q_{x_2}| : |q - q_{x_1}|$ of (3) fixed hence $x^2 = k(x^1)^{-\alpha}$ with $\alpha > 0$.

Because $f(\lambda x) = \lambda f(x)$ for $\lambda > 0$ the corresponding function is

$$c(x^1)^{1-\mu}(x^2)^\mu, \quad 0 < \mu < 1, \ c > 0,$$

where c may be omitted because different c give isometric spaces.

To discuss the cases $n > 2$ we imbed A^n in the n-dimensional projective space P^n by adding a hyperplane H_∞ and a coordinate x^{n+1} so that H_∞ is $x^{n+1} = 0$. The cone $C(0)$ intersects H_∞ in C_∞ which, because of (5.1), is a closed convex hypersurface in H_∞ relative to a suitable $(n-2)$-flat as ideal locus. C_∞ bounds a set F_∞, the intersection of $F(0)$ with H_∞.

A motion of the Minkowski space induces a projectivity of H_∞ on itself which takes C_∞ and F_∞ into themselves.

Denote as *stability group* of a point p in a timelike space the group of all motions which leave p fixed. If the space has a pairwise transitive group of motions then this stability group is transitive on each sphere $S(p, \Delta) = \{x; p < x \text{ and } \varrho(p, x) = \Delta\}$. Using (6.3) we prove

(6) THEOREM. *A timelike Minkowski space for which the stability group of a point p possesses a subgroup which is transitive and abelian on the spheres $S(p, \Delta)$ is in suitable affine coordinates given by: $x < y$ meaning $x^i < y^i$ $(i = 1, \ldots, n)$ and*

$$\varrho(x, y) = \prod_{i=1}^{n} (y^i - x^i)^{a_i}, \qquad a^i > 0, \ \sum a^i = 1.$$

For we conclude from (6.3) that $F_\infty \cup C_\infty$ is a simplex in H_∞. Therefore we can choose affine coordinates in A^n such that $F(0)$ is the set $x^i > 0$ $(i = 1, \ldots, n)$. Put $S_\Delta = S(0, \Delta)$. Since S_Δ is almost everywhere differentiable and the stability group of 0 is transitive on S_Δ, the sphere S_Δ is everywhere differentiable and being convex it is of class C^1 (see [3, p. 6]). For $y \in S_\Delta$ the tangent hyperplane of S_Δ is with $\partial f/\partial x^i = f_i$ given by

$$\sum x^i f_i(y) = f(y).$$

Denote its intersection with the x^i-axis by $\bar{a}_i$ and let a^i be the i-th coordinate of $\bar{a}_i$. Then $a^i = f(y)/f_i(y)$. The line through $\bar{a}_i$ and y intersects $x^i = 0$ at a point $\mu_i \bar{a}_i + (1 - \mu_i) y$ with $\mu_i < 0$ and

$$\frac{\mu_i}{\mu_i - 1} \cdot \frac{f(y)}{f_i(y)} = y^i.$$

Then $0 < a_i = \mu_i(\mu_i - 1)^{-1} < 1$ and a_i is invariant under the given motions. Therefore $f(x)$ satisfies the differential equations $f_i(x)/f(x) = a_i/x^i$ which yields

$$f(x) = k \prod_{i=1}^{n} (x^i)^{a_i} \qquad \text{with} \qquad \sum a^i = 1,$$

because $f(\lambda x) = \lambda f(x)$ for $\lambda > 0$. The factor k has no significance since different k lead to isometric spaces. This proves (6). We observe that

$$\bar{x}^i = \beta^i x^i + \Delta^i, \quad \beta^i > 0, \quad \prod (\beta^i)^{a_i} = 1 \tag{7}$$

is the identity component of the group of motions. This may not be the entire group, which may contain elements interchanging the coordinate axis depending on the values of the a^i. The Lorentz space is for $n > 2$ not a special case because C_∞ is not an ellipsoid.

If a connected metric space possesses a group of motions which has on orbit containing an open set ($\neq \emptyset$) then the group is transitive on R. The group (7) shows that the corresponding statement is *not* true for timelike spaces. It contains the subgroup

$$\bar{x}^h = \beta^h x^h + \sigma^h, \quad \bar{x}^j = \beta^j x^j, \quad 1 \leqslant h \leqslant k \leqslant n, \; k+1 \leqslant j \leqslant n$$

which is transitive on $\{x\colon x^j > 0\}$.

Since, according to (6), the existence of a pairwise transitive group of motions does not characterize Lorentz spaces we now prove

(8) THEOREM. *A timelike Minkowski space with a triplewise transitive group of motions is a Lorentz space.*

The hypothesis implies that for any two rays in H_∞ issuing from a given point r in F_∞ a projectivity of F_∞ induced by a centro-affinity of S_Δ exists which leaves r fixed and sends the first ray into the second. This implies that C_∞ possesses a transitive group of projectivities and hence is, by (6.1), an ellipsoid.

This alone does not imply that the space is a Lorentz space, for C_∞ is an ellipsoid for all the spaces L_a^n of Section 4 (see below). Consider a point $q \in S_\Delta$. The tangent hyperplane T_q of K_Δ at q intersects $C(0)$ in an ellipsoid E_q with a center c_q. If $q = c_q$ for some q on S_Δ, then this holds for all q on S_Δ and S_Δ is a branch of a hyperboloid, so that the space is Lorentzian. We prove that $c_q \neq q$ is impossible.

If q_∞ and c_∞ are the projections of q and c_q form 0 on H_∞, then a projectivity of F_∞ corresponding to an element of the stability group of 0 which leaves q_∞ fixed also leaves c_∞ fixed, which contradicts the transitivity of the group of projectivities of F_∞ on the rays with origin q_∞.

As gauge function of L_a^n we may take

$$f(x) = \Big[(x^n)^a - \Big(\sum (x^i)^2\Big)^{a/2}\Big]^{1/a}, \qquad x^n > \Big[\sum_{i=1}^{n-1} (x^i)^2\Big]^{1/2}.$$

The light cone is the same for all a and $f(x)$ is invariant under the affinities

$$\bar{x}^i = \sum_{k=1}^{n-1} a_k^i x_k \quad (i = 1, \ldots, n-1), \; (a_k^i) \text{ orthogonal}, \; \bar{x}^n = x^n.$$

Therefore the stability group of 0 is (to a high degree) transitive on the generators of $C(0)$, but the group of motions is not pairwise transitive.

As an immediate consequence of (6.2, 7) we have:

(9) THEOREM. *The time cone* $C(0)$ *of a timelike Minkowski space is elliptic if one of the following conditions* a), b) *is satisfied.*

a) *The stability group of* 0 *is transitive on the generators of* $C(0)$.

b) *The group of motions is pairwise transitive and a cross-section of* $C(0)$ *by a hyperplane not through* 0 *possesses an Euler point with non-vanishing Gauss curvature.*

Under either condition the space need not be Lorentzian.

We know the last assertion to be true for a) and establish it for b) as follows. Let $C = C(0)$ be given by $(x^n)^2 - \sum_{i=1}^{n-1}(x^i)^2 = 0$, $x \geqslant 0$ and consider the group Γ^* of central affinities which maps C on itself and leaves the generator G: $x_n = x_1, x_2 = \ldots = x_{n-1} = 0$ of C fixed. A sphere $S_\Delta = \{x, f(x) = \Delta > 0, x_n > 0\}$ must have the property that a variable segment tangent to S_Δ, cut out by C and with one endpoint on G be divided in a constant ratio by the point of contact with S_Δ. In each plane through G the metric must be of the type described in (5) with the same μ for all planes. A simple calculation shows that S_Δ must have the form

$$(10) \qquad (x^n - x^1)^{2\mu-1}\Big((x^n)^2 - \sum_{i=1}^{n-1}(x^i)^2\Big)^{1-\mu} = c, \qquad 0 < \mu < 1.$$

The group Γ^* considered as group of projectivities of H_∞ is the group of those hyperbolic motions of $\sum_{i=1}^{n-1}(x^i)^2 < (x^n)^2$ which leave the point $x^n = x^1 = 1$, $x^2 = \ldots = x^{n-1} = 0$ fixed. They take the hypersurfaces (10) into themselves. For $n = 3$ they are found in Lie [10, p. 221] in the form obtained from (10) by the coordinate transformation $x^n - x^1 = u_1$, $x^n + x^1 = u_n$, $x^i = u_i$ for $i = 2, \ldots, n-1$, i.e. $u_1^{2\mu-1}\big(u_n u_1 - \sum_{i=2}^{n-1}(u_1^2)\big)^{1-\mu} = c$ or, for $n = 3$, $u_3 = u_2^2 u_1^{-1} + c^1 u_1^{\mu(\mu-1)^{-1}}$ where $c^1 > 0$.

One checks readily that the Gauss curvature is positive. The fact that $u_2, \ldots, u_{n-1}$ occur for $n > 3$ only in the combination $\sum_{i=2}^{n-1} u^2$ then shows that (10) is a convex hypersurface in $x^n > [\sum(x^i)^2]^{1/2}$.

For $n = 3$ one easily sees that the metrics given by (5) and (10) exhaust the timelike Minkowski spaces with pairwise transitive groups of motions either by examining Mostow [12] for the groups acting as a plane which can be groups of motions of a Hilbert geometry or by proving that the triangle and the ellipse are the only closed convex curves in the plane whose interiors possess transitive groups of affinities, compare also [G, p. 370]. Thus

(10) *The three dimensional timelike Minkowski spaces with pairwise transitive groups of motions are in suitable affine coordinates given by the gauge functions*

$$\prod_{i=1}^{3} (x^i)^{a_i}, \qquad a_i > 0, \qquad \sum a_i = 1, \; x^i > 0$$

or

$$(x^3 - x^1)^{2\mu-1}\left((x^3)^2 - (x^1)^2 - (x^2)^2\right)^{1-\mu}, \quad 0 < \mu < 1, \quad x^3 > [(x^1)^2 + (x^2)^2]^{1/2}.$$

Finally we observe that the group Γ^* as subgroup of the group Γ of the motions of a Lorentz space is transitive on the side containing $F(0)$ of the hyperplane tangent to $C(0)$ along G, so that we have again a subgroup whose orbit is a proper open subset of the space.

8. TIMELIKE HILBERT GEOMETRIES

The analogues to the metric Hilbert geometries encountered in Section 6 are of special interest because they provide a test for our axioms. Two possibilities present themselves, one for a timelike space, the other for a locally timelike space which cannot be consistently ordered. The latter seems more natural, because it includes as special case the hyperbolic geometry defined in the exterior of an ellipsoid. In fact, in the timelike case our axiom G_2' cannot be satisfied if all other axioms are.

Let K_1, K_2 be complete convex hypersurfaces in A^n such that the closed convex sets K_1^0, K_2^0 bounded by K_1 and K_2 are disjoint. (If K_1 or K_2 is a hyperplane this condition defines K_1^0 or K_2^0.) We consider all open oriented affine segments $L(a_1, a_2)$ where $a_i \in K_i$ and $a_\theta = (1-\theta)a_1 + + a_2 \notin K_1^0 \cup K_2^0$ for $0 < \theta < 1$.

The space R is the union of all $L(a_1, a_2)$ with the topology of A^n. The relation $p < q$ is defined to mean that p and q lie on an $L(a_1, a_2)$ and q follows p. Then $p < q$ and $q < r$ imply $p < r$. For $p < q$ on $L(a_1, a_2)$ we define

(1) $$h(p, q) = \tfrac{1}{2}k|\log R(p, q, a_1, a_2)| = \tfrac{1}{2}k \log R(p, q, a_2, a_1).$$

(For the following compare [G, Section 18].) With this distance $L(a_1, a_2)$ is isometric to the real axis.

Assume that $p < q < r$ are not collinear (see Figure). Let $p, q \in L(a_1, a_2)$; $q, r \in L(b_1, b_2)$; $p, r \in L(c_1, c_2)$ the affine lines through c_i and b_i $(i = 1, 2)$ intersect at a point t (possibly at ∞) and the line through t and q intersects the segment from p to r in a point s. Finally, the line containing $L(c_1, c_2)$ intersects the line through a_i and b_i in a point c_i'. Then

$$R(p, q, a_2, a_1) = R(p, q, c_2', c_1'), \qquad R(q, r, b_2, b_1) = R(s, r, c_2', c_1'),$$

$R(p, q, a_2, a_1) R(q, r, b_2, b_1) = R(p, r, c_2', c_1') \leqslant R(p, r, c_2, c_1)$ with equality only when $c_1 = c_1'$ and $c_2 = c_2'$. Therefore $h(p, q) + h(q, r) < h(p, r)$ for non-collinear $p < q < r$, if and only if no proper affine segments $T_i \subset K_i \cap \bar{R}$ exist which are coplanar. We will assume that this condition is satisfied.

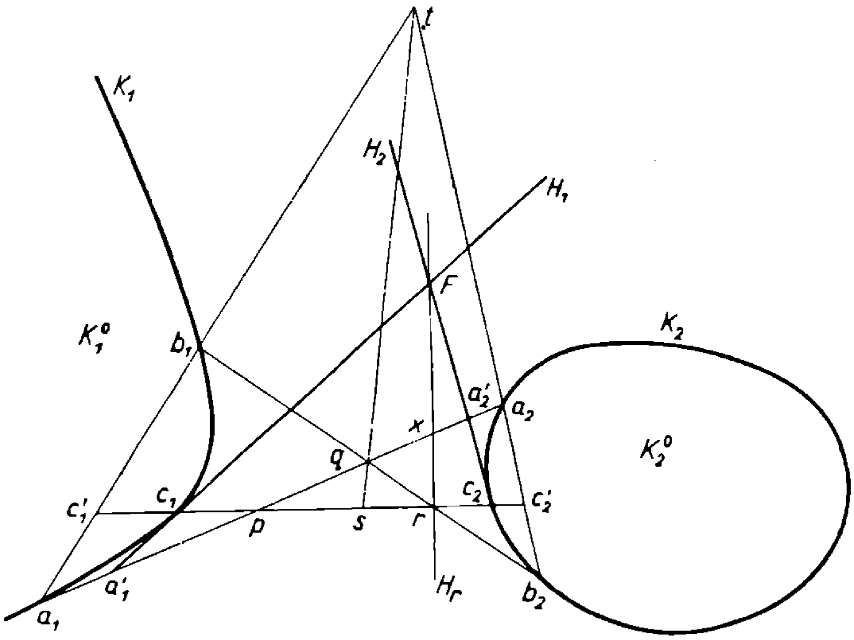

Fig. 1

The sphere $h(p, x) = \sigma > 0$ lies on a strictly convex hypersurface; it possesses at a given point r a supporting hyperplane H_r touching it only at r; and H_r separates p from the sphere (except for r).

To see this let H_i be a supporting hyperplane of K_i at c_i and H_r the hyperplane through r and the $(n-2)$-flat $F = H_1 \cap H_2$ (possibly at ∞). For convenience consider the point $x = H_r \cap L(a_1, a_2)$. Then H_i intersects $L(a_1, a_2)$ in a_i' and according to our assumption either $a_1 \neq a_1'$ or $a_2 \neq a_2'$ or both. Therefore

$$R(p, x, a_2, a_1) < R(p, x, a_2', a_1') = R(p, r, c_2, c_1).$$

Examining the space with respect to our axioms $T_1, T_{2,3}'$ one finds that $(<)$ is open in $R \times R$ if and only if no $L(a_1, a_2)$ lies on a supporting line (or a supporting hyperplane) of K_1 or K_2. It is clear that $(<)$ cannot be open if $L(a_1, a_2)$ exist which lie on supporting lines of K_1 or K_2 (or both) which contain interior points of R. If $L(b_1, b_2)$ lies on the boundary of R then an $L(a_1, a_2)$ of the previous type exists. Thus for $(<)$ to be open it is necessary and sufficient that no $L(a_1, a_2)$ lying on a supporting of K_1 or K_2 exists.

If we assume this, then G_2' is not satisfied. For then there are points $p \in R$ with a sequence $L(a_1^\nu, a_2^\nu)$ containing p and such that one of the sequences $\{a_i^\nu\}$, say $\{a_2^\nu\}$, tends to a point a_2, whereas $\{a_1^\nu\}$ diverges, so that $L(a_1^\nu, a_2^\nu)$ tends to a ray through p with a_2 as endpoint. (Actually the points p for which such $L(a_1^\nu, a_2^\nu)$ do not exist are exceptional.) If we choose x_ν on $L(a_1^\nu, a_2^\nu)$ such that $p < x_\nu$ and $h(p, x_\nu) = \sigma > 0$, then the definition of $R(p, x_\nu, a_2^\nu, a_1^\nu)$ yields that x_ν tends to a point x on R different from p and a_2. But $p < x$ is not true since R does not intersect K_1. This with $h(p, x_\nu) = \sigma$ contradicts C_2'.

All other axioms for a timelike G-space are satisfied. The space is geodesically complete and all geodesics are lines.

We summarize our results without repeating all details:

(2) *The Hilbert metric outside two convex hypersurfaces K_1, K_2 satisfies all axioms for a timelike G-space with the exception of G_2' if and only if no proper segments $T_i \subset K_i \cap \bar{R}$ exist which are coplanar and no $L(a_1, a_2)$ lies on a supporting line of K_1 or K_2. Under these conditions the axiom G_2' does not hold.*

The $L(a_1, a_2)$ are isometric to the real axis and are the geodesics, so that the space is geodesically complete.

The case where K_1 is a hyperplane and K_2 is a strictly convex hypersurface is the timelike analogue to Funk's "*Geometrie der spezifischen Massbestimmung*", see [7]. If K_1 is considered as hyperplane at infinity of a euclidean space with distance $e(x, y)$ then (1) becomes

$$h(p, q) = \tfrac{1}{2} k \log \frac{e(p, a_2)}{e(q, a_2)}.$$

Notice that the spheres $h(p, x) = \sigma > 0$ are homothetic to K_2. But $(<)$ is not open and G_2' does not hold.

To define the locally timelike analogue to Hilbert's Geometry let K be a closed convex hypersurface in the n-dimensional projective space P^n and I its interior. This means that hyperplanes not intersecting K exist and that K is convex with interior I relative to such a hyperplane considered as locus at infinity. Put $K^0 = K \cup I$.

For $p \notin K^0$ take the union $C(p)$ of all supporting lines of K^0 through p, i.e. the lines through p intersecting K but not I. Then $C(p) - C(p) \cap K$ has two nappes $C_1(p), C_2(p)$, and $C_i(p)$ bounds together with K an open set $F_i(p)$. Each point p has a neighborhood $U(p)$ such that with a proper choice of the notations $F_1(q), F_2(q)$ the set $F_i(q)$ depends continuously (in an obvious sense) on q for $q \in U(p)$. Such a choice of notation is not possible in all of $R = P^n - K^0$, which will be our space.

In $U(p)$ define $q <_p r$ by $r \in F_1(q)$. Then $q <_p r$ and $r <_p s$ imply $q <_p s$. Any projective line which intersects I intersects K in two points

a_1, a_2 and for any two points q, r in R on this line we define

$$h(q, r) = \tfrac{1}{2}k|\log R(p, q, a_1, a_2)|.$$

If $q <_p r$ then the line through q and r intersects I and we define

$$\varrho_p(q, r) = h(p, q).$$

Then T_4 is satisfied because $r \in F_1(q)$ is equivalent to $q \in F_2(r)$. The time inequality

$$\varrho_p(q, r) + \varrho_p(q, s) \leqslant \varrho_p(p, s) \qquad \text{for} \qquad q <_p r <_p s$$

follows from the previous discussion, which also shows that equality holds only if q, r, s are collinear, unless certain pairs of segments lie on K. This need not be made precise because G_2 is valid if and only if no segment of a projective line lies on K.

For, if L is a supporting line of K containing a proper segment with endpoints b_1, b_2, let $p \in L - L \cap K$ and L_ν a line through p intersecting I and whose intersections with K tend to b_1 and b_2. If q is a point of L in $U(p)$ and different from p and $q_\nu \in L_\nu$ tends to q, then

$$\lim h(p, q_\nu) = \tfrac{1}{2}k|\log R(p, q, b_1, b_2)| > 0$$

although neither $p <_p q$ nor $q <_p p$ which contradicts G_2.

On the other hand $h(p, q_\nu) \to 0$ if $L \cap K$ is a point.

The remaining axioms are easily verified. So we have:

(3) THEOREM. *Let K be a strictly convex hypersurface in P^n with interior I. If $R = P^n - K \cup I$ and $U(p)$, $\varrho_p(q, r)$ are defined as above, then R is a geodesically complete locally timelike G-space which cannot be consistently ordered.*

The geodesics lie on the projective lines intersecting I. The length of the arc of a geodesic from p to q is $h(p, q)$.

The motions of R are the restrictions to R of the projectivities of P^n which map K on itself.

That such projectivities are motions is evident and that every motion has this form is seen in a similar way as (7.1).

For a motion φ of R we have $\varphi C(p) = C(\varphi p)$. Therefore, if the group of motions is transitive on R, any $C(p)$ can be moved into any $C(q)$. We note that in this case the strict convexity of K can be proved:

(4) *If for a closed convex surface K with interior I in P^n the group of projectivities mapping K on itself is transitive on $R = P^n - K \cup I$ then K is strictly convex.*

According to a theorem by Ewald and Rogers [6], if H is a hyperplane not intersecting the closed convex hypersurface K then for almost all $C(p)$ with $p \in H$ the generators of $C(p)$ touch K in one point only. Our hypothesis implies that this is true for every $C(q)$.

If the group Γ of motions of R is transitive on R then for no $x \in K$ the orbit $\Gamma(x)$ can consist of x alone since Γ would then map the union of all supporting hyperplanes of K at x on itself and hence would not be transitive on R.

Let $y \in \Gamma(x) - \{x\}$. Give a point z on K and a neighborhood $V(z)$ of z on K. Choose q in R so close to z that one of the sets on K bounded by $C(q) \frown K$ lies in $V(z)$. For any point p in R collinear with x and y let $\varphi \in \Gamma$ take p into a point of $F_1(q) \smile F_2(q)$. Then one of the points x, y goes into a point of $V(z)$. Therefore $\Gamma(x) = \Gamma(x) \smile \Gamma(y)$ is dense in K. If $\Gamma(x)$ contains an open subset of K no other orbit $\Gamma(u)$ $(u \in K)$ can be dense in K, hence $\Gamma(x) = K$ and K is an ellipsoid by (6.2).

The only other alternative is that $\dim \Gamma(x) < n-1$, [9, p. 46]. Since $\Gamma(x)$ is the countable union of compact sets, the number of distinct orbits $\Gamma(x)$ $(x \in K)$ must be non-countable. So we see: Unless K is an ellipsoid it is the union of a non-countable number of distinct orbits $\Gamma(x)$ each of which is dense on K. It is very probable that this cannot happen, but the absence of information on locally compact transformation groups seems to make a proof difficult at the present time except for $n = 3$. For, $\dim \Gamma \geqslant n$ and $\dim \Gamma(x) \leqslant n-2$ for $x \in K$ imply that the stability group of $x \in K$ has at least dimension 2 and Γ acts effectively on K.

A group of motions Γ of a locally timelike G-space R is *transitive on the line elements of* R if for two given line elements L_p and $L_{p'}$ a motion in Γ exists which maps p on p' and the geodesic containing L_p on the geodesic containing $L_{p'}$ (or equivalently, which maps all sufficiently small segments in L_p on elements of $L_{p'}$).

It is clear that: in a locally timelike Hilbert space K must be an ellipsoid if the group of motions is transitive on the line elements for then $\Gamma(x)$ $(x \in K)$ is open on K.

If K is an ellipsoid we speak of a *locally timelike hyperbolic geometry*. It is the restriction to the pairs $x \leqslant_p y$ of the so-called exterior hyperbolic geometry, see [15], which is an indefinite metric defined for pairs in any position (compare footnote (²) on p. 28). Thus we may say

(5) *A locally timelike Hilbert geometry is hyperbolic if its group of motions is transitive on the line elements.*

We repeat that transitivity on the points suffices for $n = 2, 3$ and probably for all n.

For $n = 2$ this geometry gives rise to a second locally timelike G-space. The geodesics through a point p are the projective lines which do not intersect or touch the ellipse K (see [15]). Whether there is analogue to this for arbitrary strictly and/or differentiable closed convex K

is not known. The standard definition of distance when K is an ellipse uses the conjugate complex intersections of the projective lines with K and cannot be generalized.

If in (1) we choose K_1 and K_2 as the branches of the hyperboloid $(x^n)^2 - \sum_{i=1}^{n-1} (x^i)^2 = 1$, the resulting timelike space is, essentially, a restriction of the locally timelike hyperbolic geometry. By changing coordinates it can be given a form analogous to the Poincaré model of hyperbolic geometry in $y^n > 0$, namely

$$ds = \left[(dy^n)^2 - \sum_{i=1}^{n-1} (dy^i)^2\right]^{1/2} (ky^n)^{-1}$$

and $x < y$ means $y^n - x^n > [\sum_{i=1}^{n-1} (y^i - x^i)^2]^{1/2}$. From this form it would be hard to guess that the space, which is not geodesically complete, is part of a geodesically complete space. We have again the phenomenon of a subgroup of the group of all motions (of the locally timelike hyperbolic geometry) which has an open orbit.

REFERENCES

[1] T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Berlin 1934 and New York 1948.

[2] H. Busemann, *The geometry of geodesics*, New York 1955, quoted as G.

[3] — *Convex surfaces*, New York 1958.

[4] — and J. K. Beem, *Axioms for indefinite metrics*, Rend. Circolo Mat. Palermo.

[5] A. S. Eddington, *The mathematical theory of relativity*, 2nd ed., Cambridge 1930.

[6] G. Ewald and C. A. Rogers, *The faces of a convex body*, Israel J. Math.

[7] P. Funk, *Über Geometrien, bei denen die Geraden die Kürzesten sind*, Math. Ann. 101 (1929), pp. 226-237.

[8] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed., Cambridge 1952.

[9] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton 1941.

[10] S. Lie, *Bestimmung aller Flächen, die eine kontinuirliche Schar von projektiven Transformationen gestatten*, Ber. Sächs. Akad. Wiss. Leipzig 47 (1895), pp. 209-260.

[11] D. Montgomery and L. Zippin, *Topological transformation groups*, New York 1955.

[12] G. D. Mostow, *The extensibility of local Lie groups of transformations and groups on surfaces*, Ann. of Math. 52 (1950), pp. 606-636.

[13] T. Nagano, *Differential geometry and analytic group theory methods in foundations of geometry*, Algebraic and Topological Foundations of Geometry, Oxford 1962.

[14] L. Pontrjagin, *Topological groups*, Princeton 1939.

[15] И. М. Яглом, Б. А. Розенфельд и Е. У. Ясинская, *Проективные метрики*, Усп. Мат. Наук 19 вып. 5 (1964), стр. 51-113.

English translation:

I. M. Yaglom, B. A. Rozenfel'd and E. U. Yasinskaya, *Projective metrics*, Russ. Math. Surveys 19 (1964), No 5, pp. 49-107.