

Lorentz-invariant Markov processes in relativistic phase space

By R. M. DUDLEY

1. Introduction

This paper deals with certain “random motions” permitted by the special theory of relativity; that is, with probability measures on sets of trajectories on which speeds are less than or equal to the speed c of light (which we take equal to 1 throughout). We deal with classes of such measures indexed by possible initial states, related to each other by the Lorentz group (implying certain invariance conditions for individual measures).

The definition of “state” as mentioned above is governed by the Markov property, i.e. that given the present state, further knowledge of past states should be irrelevant to the prediction of future states. It is known that position is insufficient for this purpose (see e.g. Theorem 11.3 below), and it seems natural to include the velocity in specifying the state. Indeed, velocities must exist at almost all times since position is a Lipschitzian function of time, and the existence of velocities is generally incompatible with the Markov property of a position process.

We distinguish between speeds strictly less than 1 and those equal to 1, and do not consider processes in which both occur (except in Theorem 11.3). This corresponds to the physical distinction between particles of positive or zero rest mass. Invariant processes of speed 1 turn out to be essentially uninteresting (see § 11) in that they cannot change direction.

We are left, then, with processes of speeds almost always strictly less than 1. For these processes, we can introduce the relativistic “proper time” on each trajectory (see § 6 below). The possible “4-velocities”, i.e. derivatives of space-time position with respect to proper time, then lie in a three-dimensional hyperboloid \mathcal{U} with a symmetric, Lorentz-invariant Riemannian structure (see § 7). Because of our invariance assumptions, the velocity process is itself Markovian (Theorem 3.2) and defined by a “convolution” semigroup on \mathcal{U} .

Such convolution semigroups have fortunately been completely classified by Tutubalin [1]. We thus arrive at an explicit description of all the processes which concern us, in Theorem 8.2, the main theorem of the paper. There are two extreme possibilities: on the one hand, “Brownian motions” in \mathcal{U} (a one-parameter family indexed by a diffusion constant), which yield (the only) processes in which velocity is a continuous (but, of course, not differentiable) function of time; see § 10. On the other hand, there are “Poisson” processes in which the velocity changes only by jumps. Finally, there are (roughly speaking) mixtures of the two.

In order to make the change from coordinate time to proper time as parameter, it seems necessary to make an assumption of continuity in probability. One would like to infer this from the other assumptions through a sort of “infinite divisibility,” but in our situation this seems difficult before the change to proper time, since one has a convolution semigroup on a symmetric space only after the change.

Sections 2 and 3 deal with generalities about Markov processes. Section 4 introduces Lorentz-invariant Markov processes of speeds less than 1. Sections 5–10 carry through the characterization of these processes. In section 11, we establish the triviality of processes with speeds equal to 1 or states given by position only. In section 12 some unsolved problems are mentioned.

For Markov processes I have used specializations of the definitions in Dynkin’s books [1] and [2], since several results are also quoted. It is worth noting, however, that this leads to at least one unsatisfying situation. In the usual approach to the strong Markov property one demands right continuity for sample functions. Here, where a sample function is (in part) a derivative, it becomes a right derivative after imposition of right continuity. It may seem unnatural to consider the right derivative of a function at time t as known at that time, if the left derivative is different. The way out of this difficulty, if it is one, will be left for other researches.

The “diffusion” or “Brownian motion” processes studied in § 10 below have been worked on previously by G. Schay [1] and H. Dinges. Both the latter and R. Hermann advised me against beginning with the proper time (i.e., essentially starting in the middle of § 6 below), as I did in an earlier draft, and I am now glad to have followed their advice. An exchange of correspondence with H. Dinges on this subject in general has been most helpful.

2. Starting probabilities and Markov processes

First we review some measure-theoretic notation. If (X, \mathcal{S}, μ) and (Y, \mathcal{J}, ν) are σ -finite measure spaces (which for our purposes may as well be finite), then $\mathcal{S} \times \mathcal{J}$ denotes the σ -field of subsets of the Cartesian product $X \times Y$ generated by all sets $A \times B$, $A \in \mathcal{S}$, $B \in \mathcal{J}$, and $\mu \times \nu$ the product measure. If F maps X onto another set S , then $F(\mathcal{S})$ denotes the class of all $C \subset S$ such that $F^{-1}(C) = \{x : F(x) \in C\} \in \mathcal{S}$. We let

$$(\mu \circ F^{-1})(C) = \mu(F^{-1}(C)), C \in F(\mathcal{S}).$$

Finally if f is μ -integrable and \mathcal{H} is a sub- σ -field of \mathcal{S} , then $E_\mu(f | \mathcal{H})$ denotes the μ -conditional expectation of f given \mathcal{H} , and if f is the indicator function of a set A , we let

$$E_\mu(f | \mathcal{H}) = \mu(A | \mathcal{H}).$$

To clarify the argument of a function $F = \mu(A | \mathcal{H})$ we may write

$$F(x) = \mu(x : x \in A | \mathcal{H}),$$

where “ $x \in A$ ” will be replaced by a defining condition.

Throughout this paper, R will denote the real line (with its usual topology), R^n its n -fold Cartesian product, and R^+ the nonnegative axis $[0, \infty)$. If S is any topological space, $B(S)$ will denote the Borel σ -field generated by the open sets in S .

We shall consider certain function spaces defined as follows:

Definition. If S is a set, a *path space* for S is a pair (\mathcal{A}, ζ) satisfying the following conditions:

- (1) ζ is a function on \mathcal{A} with $0 \leq \zeta(f) \leq +\infty, f \in \mathcal{A}$.
- (2) Each $f \in \mathcal{A}$ is a function from the interval $[0, \zeta(f))$ to S .
- (3) For each $s \geq 0$ and $x \in S$ there is an $f \in \mathcal{A}$ with $f(s) = x$.

If S is any set, the set \mathcal{A} of all functions f from intervals $[0, \zeta(f))$ to S will be called the *maximal path space* for S .

Given a path space (\mathcal{A}, ζ) and $s \geq 0$, let

$$\mathcal{A}^s = \{f \in \mathcal{A} : \zeta(f) > s\}.$$

Suppose \mathcal{S} is a σ -field of subsets of S . If $0 \leq s \leq t$, $B_t^s(\mathcal{A}, \mathcal{S})$ will denote the σ -field of subsets of \mathcal{A}^s generated by all sets of the form

$$\{f \in \mathcal{A}^s : f(r) \in A\}$$

for $s \leq r \leq t$ and $A \in \mathcal{S}$ (here and throughout $f(r) \in A$ implies $\zeta(f) > r$). $B^s(\mathcal{A}, \mathcal{S})$ will be the σ -field generated by all the $B_t^s(\mathcal{A}, \mathcal{S})$ for $t > s$. Context permitting, $B_t^s(\mathcal{A}, \mathcal{S})$ will be written $B_t^s(\mathcal{A})$ or B_t^s , and likewise for B^s . We call $(\mathcal{A}, \zeta, S, \mathcal{S})$ a *measurable path space*.

We shall need the following fact, proved in Dynkin [1] Lemma 5.9:

Lemma 2.1. *Suppose S is a metric space and (\mathcal{A}, ζ) is a path space for S such that each $f \in \mathcal{A}$ is continuous from the right wherever it is defined. For $0 \leq r \leq s$ let M_s^r be the product σ -field*

$$B_r^s(\mathcal{A}, B(S)) \times B(R^+)$$

in $\mathcal{A} \times R^+$. Let $F(f, t) = f(t)$ for $r \leq t \leq s$. Then the domain D of F belongs to M_s^r and F is M_s^r -measurable on D .

For $h \geq 0$ let θ_h be the transformation of functions defined by

$$(\theta_h f)(t) = f(t+h), \quad t \geq 0.$$

A path space (\mathcal{A}, ζ) will be called *stationary* if for each $h \geq 0$, θ_h takes \mathcal{A} into itself. If $(\mathcal{A}, \zeta, S, \mathcal{S})$ is a stationary measurable path space, clearly θ_h takes B_{t+h}^s onto B_t^s for any $h \geq 0, t \geq s \geq 0$.

A Markov process will be defined below by a class of measures on a path space corresponding to different initial states. The restriction to path spaces is a specialization of the definitions in Dynkin [1] and [2]. The definition in [2] requires stationary transition probabilities, which are sufficient for our purposes. It will be reproduced after our own is completed.

Definition. Suppose $(\mathcal{A}, \zeta, S, \mathcal{S})$ is a measurable path space in which \mathcal{S} contains all one-point sets. Given $x \in S$, a probability measure P on $B^0(\mathcal{A})$ will be called a *starting probability* at x on $(\mathcal{A}, \zeta, S, \mathcal{S})$ if

$$P(f : f(0) = x) = 1.$$

Definition. Suppose $(\mathcal{A}, \zeta, S, \mathcal{S})$ is stationary and for each $x \in S, P_x$ is a starting probability at x on \mathcal{A} . Then $(\{P_x\}, \mathcal{A}, \zeta, S, \mathcal{S})$ or, briefly, $(\{P_x\}, \mathcal{A})$ or $\{P_x\}$, is a *Markov process* if

- (1) For any $t \geq 0$ and $A \in \mathcal{S}, P(t, x, A) = P_x(f : f(t) \in A)$ is \mathcal{S} -measurable in x .
- (2) Whenever $0 \leq t \leq u, x \in S,$ and $A \in \mathcal{S}, P_x(f : f(u) \in A \mid B_t^0) = P(u - t, f(t), A)$ almost everywhere with respect to P_x on \mathcal{A}^s .

The following is known (Dynkin [1], Lemma 2.2):

Lemma 2.2. *If $(\{P_x\}, \mathcal{A}, \zeta, S, \mathcal{S})$ is a Markov process and $A \in B^0(\mathcal{A}, S)$, then $P_x(A)$ is \mathcal{S} -measurable in x .*

The definition of Markov process in Dynkin [2] can be reformulated in our terms as follows:

Definition. Suppose given a stationary measurable path space $(\mathcal{A}, \zeta, S, \mathcal{S})$, a set Ω , a mapping

$$X : \omega \rightarrow x(\cdot, \omega)$$

of Ω onto \mathcal{A} , for each $t \geq 0$ a σ -field \mathcal{M}_t of subsets of

$$\Omega_t = \{\omega \in \Omega : \zeta(x(\cdot, \omega)) > t\},$$

a σ -field \mathcal{M}^0 of subsets of Ω including all the \mathcal{M}_t , and a set $\{Q_x, x \in S\}$ of probability measures on \mathcal{M}^0 . Then $(x(\cdot, \cdot), \zeta, \{\mathcal{M}_t\}, \{Q_x\})$ is a *fields Markov process* if:

- (A) For $0 \leq t \leq u$ and $A \in \mathcal{M}_t$, or $A \in X^{-1}(B_u^0(S, S))$,

$$A \cap \Omega_u \in \mathcal{M}_u.$$

- (B) If P_x is $Q_x \circ X^{-1}$ restricted to $B^0(\mathcal{A}, S)$, then $(\{P_x\}, \mathcal{A}, \zeta, S, \mathcal{S})$ is a Markov process (in the path space sense), and (2) holds with B_t^0 replaced by $X(\mathcal{M}_t)$.

The word ‘‘fields’’ has been added to clarify the difference between the two definitions. If the Q_x are defined only on the minimal σ -field $X^{-1}(B^0)$, we have an isomorphism.

It is clear that a Markov process on a path space defines a fields Markov process, and conversely a fields Markov process defines a path space Markov process. However, non-isomorphic fields Markov processes may define the same path space Markov process.

A Markov process on a maximal path space is called ‘‘canonical’’ by Dynkin [1, 2.11]. Any path space Markov process extends naturally to a canonical one. Conversely, a canonical Markov process $(\{P_x\}, \mathcal{B})$ can be restricted to a path space $\mathcal{A} \subset \mathcal{B}$ if and only if \mathcal{A} has outer measure 1 for all the P_x (Dynkin [1] Theorem 2.5).

We shall also use the strong Markov property, which we proceed to define. Suppose $(\mathcal{A}, \zeta, X, \mathcal{S})$ is a measurable path space. A *stopping time* on the path space is a $B^0(\mathcal{A})$ -measurable function τ from \mathcal{A} to R^+ such that $\tau(f) \leq \zeta(f)$ for all $f \in \mathcal{A}$ and, for any $t \geq 0$,

$$\{f \in \mathcal{A} : \tau(f) \leq t < \zeta(f)\} \in B_t^0(\mathcal{A}).$$

If τ is a stopping time, let

$$\mathcal{A}^\tau = \{f \in \mathcal{A} : \zeta(f) > \tau(f)\}$$

and let B_τ be the σ -field of sets $A \subset \mathcal{A}$ such that for any $t \geq 0$,

$$A \cap \{f : \tau(f) \leq t < \zeta(f)\} \in B_t^0(\mathcal{A}).$$

Then a Markov process $\{P_x\}$ on \mathcal{A} is called *strongly Markovian* if:
 SM(1) Given $B \in \mathcal{S}$,

$$\langle t, x \rangle \rightarrow P_x\{f : f(t) \in B\}$$

is $B(R^+) \times \mathcal{S}$ -measurable,

SM(2) for any stopping time τ , $B \in \mathcal{S}$, and B_τ -measurable function $\eta \geq \tau$,

$$P_x(f : f(\eta) \in B | B_\tau) = P_{f(\tau)}(g : g(\eta - \tau) \in B)$$

almost everywhere on \mathcal{A}^τ with respect to P_x , and $\{x : f(\tau) = x \text{ for some } f \in \mathcal{A}\} \in \mathcal{S}$.

3. Markov processes in product spaces

Suppose $(\mathcal{A}, \xi, X, \mathcal{S})$ and $(\mathcal{B}, \eta, Y, \mathcal{J})$ are stationary measurable path spaces with $\xi \equiv \eta \equiv +\infty$, and for each $x \in X$ ($y \in Y$), P_x (Q_y) is a starting probability on \mathcal{A} (\mathcal{B}).

Let

$$Z = X \times Y, \mathcal{U} = \mathcal{S} \times \mathcal{J}, C = \mathcal{A} \times \mathcal{B},$$

and $\zeta \equiv +\infty$. For $z = \langle x, y \rangle \in Z$ let R_z be the measure $P_x \times Q_y$ on C . Regard $h = \langle f, g \rangle \in C$ as the function

$$h(t) = \langle f(t), g(t) \rangle \in Z, \quad t \geq 0.$$

Theorem 3.1. $(\{R_z\}, C, \zeta, Z, \mathcal{U})$ is a Markov process if and only if both $\{P_x\}$ and $\{Q_y\}$ are Markov processes.

Proof. Clearly $(C, \zeta, Z, \mathcal{U})$ is a measurable path space and each R_z is a starting probability at z . First suppose $\{P_x\}$ and $\{Q_y\}$ are both Markovian. The class of sets $C \in \mathcal{U}$ for which both 1) and 2) in the definition of Markov process hold is closed under finite, disjoint unions and countable increasing unions and decreasing intersections. Thus we can put $C = A \times B$, $A \in \mathcal{S}$, $B \in \mathcal{J}$, and the conclusion follows easily.

Now suppose $\{R_z\}$ is Markovian; let us show that $\{P_x\}$ is Markovian. The measurability condition 1) is clear. Given $0 \leq t \leq u$, $x \in X$, and $A \in \mathcal{S}$, let $z = \langle x, y \rangle$ for some $y \in Y$ and note that the conditional probability of a "rectangle" for a product measure is itself a product, so that if $\mathbf{1}_g = 1$ for all g ,

$$\begin{aligned} & P_x(f : f(u) \in A | B_t^0(\mathcal{A})) \cdot \mathbf{1}_g \\ &= R_z(\langle f, g \rangle : f(u) \in A | B_t^0(C)) \\ &= R_{\langle f(t), g(t) \rangle}(\langle \phi, \psi \rangle : \phi(u-t) \in A) \\ &= P_{f(t)}(\phi : \phi(u-t) \in A) \cdot \mathbf{1}_g, \end{aligned}$$

almost everywhere for R_z . Choosing g suitably, we obtain that $\{P_x\}$ is Markovian, q.e.d.

Now suppose (X, \mathcal{S}) and (Y, \mathcal{J}) are measurable spaces, $Z = X \times Y$, $\mathcal{U} = \mathcal{S} \times \mathcal{J}$, and $(\{R_z\}, C, \zeta, Z, \mathcal{U})$ is a Markov process. Suppose G is a transitive group of automorphisms of the measurable space (X, \mathcal{S}) , and each $\gamma \in G$ acts on Z by $\gamma \langle x, y \rangle = \langle \gamma(x), y \rangle$ and on C in the obvious way. Let K be the natural projection of Z onto Y and $K(\langle f, g \rangle) = g$, $\langle f, g \rangle \in C$, where f and g are both defined on the interval $0 \leq t < \zeta(\langle f, g \rangle)$;

we can let $\eta(g) = \zeta(\langle f, g \rangle)$. Let \mathcal{B} be the set of all $K(h)$ for $h \in \mathcal{C}$. Clearly $(\mathcal{B}, \eta, Y, \mathcal{J})$ is a stationary, measurable path space.

Theorem 3.2. *Suppose $\{R_z\}$ is G -invariant, i.e. for any $z \in Z$, $A \in B^0(\mathcal{C})$, and $\gamma \in G$,*

$$R_z(A) = R_{\gamma(z)}(\gamma(A)).$$

Then for each $y \in Y$ the measure $R_{\langle x, y \rangle} \circ K^{-1}$ is independent of x . Calling it Q_y ,

$$(\{Q_y\}, \mathcal{B}, \eta, Y, \mathcal{J})$$

is a Markov process.

Proof. Each Q_y is well-defined by G -invariance and is clearly a starting probability on \mathcal{B} . For the Markov property, first given $A \in \mathcal{J}$ we have

$$Q_y(g: g(t) \in A) = R_{\langle x, y \rangle}(\langle f, g \rangle: g(t) \in A)$$

for any fixed $x \in X$, and this is \mathcal{J} -measurable in y . Second, suppose $y \in Y$, $0 < t < u$, $A \in \mathcal{J}$. For any $x \in X$,

$$\begin{aligned} R_{x, y}(\langle f, g \rangle: g(u) \in A \mid B_t^0(\mathcal{C})) \\ = R_{f(t), g(t)}(\langle \phi, \psi \rangle: \psi(u-t) \in A) \end{aligned}$$

almost everywhere for $R_{x, y}$ in \mathcal{C}^t . Transforming both sides by K we get

$$Q_y(g: g(u) \in A \mid B_t^0(\mathcal{B})) = Q_{g(t)}(\psi: \psi(u-t) \in A)$$

almost everywhere for Q_y on \mathcal{B}^t , q.e.d.

4. Lorentz-invariant random motions

Let X be a three-dimensional Euclidean space R^3 of ordered triples

$$x = (x_1, x_2, x_3)$$

of real numbers and let

$$|x| = (x_1^2 + x_2^2 + x_3^2)^{1/2}.$$

Let M be the space R^4 of pairs $z = \langle x, t \rangle$, $x \in X$, $t \in R$, with

$$\|z\|^2 = t^2 - |x|^2.$$

\mathcal{L} will denote the group of linear transformations of M into itself which preserve $\|\cdot\|^2$, have determinant 1, and do not change the sign of t (the proper, orthochronous Lorentz group).

M_o will denote the universal "tangent space" of M , or space of derivatives of functions from the real line to M at points of M . M_o is naturally isomorphic to M , but for some purposes the two spaces will be distinguished.

Let \mathcal{A} be the space of functions f from R^+ to X such that

- (1) $|f(s) - f(t)| \leq |s - t|$ for $s, t \geq 0$
- (2) If $f'(s)$ is defined (as it is for almost all s by (1)), $|f'(s)| < 1$.

Let $I(f) \equiv +\infty$. Then $(\mathcal{A}, I, X, B(X))$ is a measurable path space. We shall need the following fact:

Lemma 4.1. *Suppose U is an open subset of R^k , $S = B(U)$, and $(\mathcal{B}, \zeta, U, S)$ is a measurable path space consisting of continuous functions. Let $F = \mathcal{B} \times U$ and*

$$\mathcal{F} = B^0(\mathcal{B}) \times S.$$

Then the function

$$\langle f, t \rangle \rightarrow df(t)/dt$$

is defined on a set in \mathcal{F} and \mathcal{F} -measurable there.

Proof. By Lemma 2.1,

$$\langle f, t \rangle \rightarrow f(t) \text{ and } \langle f, t \rangle \rightarrow (f(t+r) - f(t))/r$$

for any $r \neq 0$ have the desired property. For any open set $V \subset R^k$, let

$$Z(V) = \{ \langle f, t \rangle : (f(t+r) - f(t))/r \in V \text{ for small enough rational } r \neq 0 \}. \text{ Clearly } Z(V) \in \mathcal{F}.$$

For each $n = 1, 2, \dots$, let $\{U_{mn}\}_{m=1}^\infty$ be a locally finite open cover of R^k by sets of diameter less than $1/n$. Then

$$A = \{ \langle f, t \rangle : f'(t) \text{ exists} \} = \bigcap_n \bigcup_m Z(U_{mn}).$$

Now any open $V \subset R^k$ is the union of the U_{mn} with $\bar{U}_{mn} \subset V$, where the bar denotes closure. Thus

$$\{ \langle f, t \rangle : f'(t) \in V \} = \bigcup_{m,n} \{ Z(U_{mn}) : \bar{U}_{mn} \subset V \} \cap A.$$

This completes the proof.

Now let V be the open unit ball $\{v: |v| < 1\}$ in R^3 . V is the space of admissible velocities. We shall consider families $\{P_x^v\}$ of measures, where for each $x \in X$ and $v \in V$ P_x^v is a starting probability at x on $(\mathcal{A}, I, X, B(X))$. The family $\{P_x^v\}$ will be assumed to satisfy conditions A) – E) to be formulated below.

(A) For any $n = 1, 2, \dots$, $t_i \geq 0$, $A_i \in B(X)$, $x \in X$, and $v \in V$,

$$P_x^v(f : f(t_i) \in A_i, i = 1, \dots, n) = P_0^v(f : f(t_i) \in A_i - x, i = 1, \dots, n).$$

(B) For any $x \in X$ and $v \in V$,

$$P_x^v(f : f'(0^+) = v) = 1.$$

Since each $f \in \mathcal{A}$ is differentiable almost everywhere and $\{ \langle f, t \rangle : f'(t) \text{ exists} \}$ is product measurable,

$$\{s > 0 : P_x^v(f'(s) \text{ exists}) < 1\}$$

has Lebesgue measure zero by the Fubini theorem for each x and v . We assume this set is actually empty, and formulate a Markov property:

(C) If $t > 0$, $x \in X$, and $v \in V$, then

$$P_x^v(f'(t) \text{ exists}) = 1$$

and for any $A \in \mathcal{B}^t(\mathcal{A})$,

$$P_x^v(A | B_t^0(A)) = P_{f(t)}^{f'(t)}(\theta_t A)$$

almost everywhere for P_x^v (where θ_t is the translation by t , defined in §2 above).

We next formulate a Lorentz-invariance condition. Each $L \in \mathcal{L}$ defines a transformation L^* of functions $f \in \mathcal{A}$ by

$$\begin{aligned} L\langle f(t), t \rangle &= \langle L_{(1)}\langle f(t), t \rangle, L_{(2)}\langle f(t), t \rangle \rangle, \\ L_{(1)}\langle f(t), t \rangle &= (L^* f)(L_{(2)}\langle f(t), t \rangle). \end{aligned}$$

Clearly L^* is measurable from $(\mathcal{A}, B^0(\mathcal{A}))$ to itself. L also defines a transformation V_L of velocities by

$$V_L(f'(t)) = (L^* f)'(L_{(2)}\langle f(t), t \rangle).$$

Our Lorentz-invariance assumption is

(D) $P_0^v \circ (L^*)^{-1} = P_0^{V_L(v)}$ for any $L \in \mathcal{L}$ and $v \in V$.

Let W be the “phase space” $X \times V$. Let d be the metric on W defined by

$$d(\langle x, u \rangle, \langle y, v \rangle) = |x - y| + |u - v|.$$

The last assumption is that derivatives are continuous in probability:

(E) $\lim_{t \downarrow 0} P_0^v(d(\langle f(t), f'(t) \rangle, \langle 0, v \rangle) \geq \varepsilon) = 0$

uniformly in v , for any $\varepsilon > 0$.

I don't know whether (E) follows from the preceding assumptions.

Let \mathcal{Q} be the set of all functions from R^+ to W . Then $(\mathcal{Q}, I, W, B(W))$ is a stationary, measurable path space. Each measure P_x^v on sets

$$\{f: \langle f(t_i), f'(t_i^+) \rangle \in B_i, i = 1, \dots, n\}, t_i \geq 0, B_i \in B(W),$$

extends uniquely by Kolmogorov's theorem (Loève [1] p. 93) to a probability measure $Q_{x,v}$ on $B^0(\mathcal{Q})$. For $t_i > 0$, t_i^+ can be replaced by t_i . Of course, there may exist with positive probability points t , depending on f , at which $f'(t)$ does not exist.

5. Properties of the $Q_{x,v}$.

In this section we infer from (A) – (E) that $\{Q_{x,v}\}$ is a “Fellerian” Markov process, and discuss other continuity properties.

We consider the transition measures

$$Q(h, x, v, A) = Q_{x,v}(f: f(h) \in A)$$

defined for $\langle x, v \rangle \in W, A \in B(W), h \geq 0$. Each $Q(h, x, v)$ is a probability measure on W . We consider weak* convergence of such measures:

$$\mu_n \rightarrow \mu \text{ (weak*)}$$

if for every bounded continuous real-valued function f on W ,

$$\int f d\mu_n \rightarrow \int f d\mu.$$

Let $BL(W)$ be the class of bounded Lipschitzian real-valued functions on W , i.e. functions f with

$$\|f\|_{BL} = \sup_w |f(w)| + \sup_{u \neq w} \frac{|f(u) - f(w)|}{d(u, w)} < \infty.$$

Then $(BL(W), \|\cdot\|_{BL})$ is a Banach space, and any probability measure μ on W defines an element of the dual space $BL^*(W)$, with the dual norm

$$\|\mu\|_{BL}^* = \sup \left(\int f d\mu : \|f\|_{BL} = 1 \right).$$

We have $\mu_n \rightarrow \mu$ (weak*) if and only if

$$\|\mu_n - \mu\|_{BL}^* \rightarrow 0;$$

see e.g. R. R. Rao [1] Theorem 3.1. (For "if", nonnegativity of the measures is essential.)

Theorem 5.1. *Under assumptions (A) – (E) of § 4, $Q(h, x, v, \cdot)$ is jointly weak* continuous in h, x , and v .*

Proof. We shall prove continuity in h and x uniformly on the triple product, then simple continuity in v .

We have continuity in h , uniformly in all x, v and h , by assumptions (A) and (E). For continuity in x , let

$$T_y \langle x, v \rangle = \langle x + y, v \rangle,$$

and note that

$$Q(h, x + y, v, \cdot) = Q(h, x, v, \cdot) \circ T_y^{-1}$$

by (A). Letting $\mu = Q(h, x, v, \cdot)$, we have

$$\|\mu - \mu \circ T_y^{-1}\|_{BL}^* = \sup_{\|f\|_{BL}=1} \left| \int f d\mu - \int f \circ T_y d\mu \right|.$$

Since $|(f - f \circ T_y)(w)| \leq |y|$ if $w \in W, y \in X$, and $\|f\|_{BL} = 1$, we have continuity in x , uniformly in h and v .

It remains to prove continuity in v for fixed x and h , and we can take $x = 0$. Suppose $v_n \rightarrow v$ in V . There are $L_n \in \mathcal{L}$ such that $v = V_{L_n}(v_n)$ for all n , with L_n converging to the identity. We also choose numbers $\varepsilon_n \downarrow 0$ such that

$$L_n \langle x, h + \varepsilon_n \rangle = \langle x', t' \rangle \quad \text{and} \quad (t')^2 - |x'|^2 \geq 0 \quad \text{imply} \quad t' > h.$$

For $f \in \mathcal{A}$, $f(0) = 0$, let $t_n(f)$ be the unique value of s such that

$$\langle f(s), s \rangle = L_n \langle x, h + \varepsilon_n \rangle$$

for some x . (Then $t_n(f) > h$ for all $f \in \mathcal{A}$ with $f(0) = 0$.) We need the following:

Lemma 5.2. $f \rightarrow \langle f(t_n(f)), V_{L_n} f'(t_n(f)) \rangle$ is B^h -measurable on \mathcal{A} .

Proof. Clearly $f \rightarrow t_n(f)$ is B^h -measurable, and so is $f \rightarrow \langle f, t_n(f) \rangle$, (from B^h to $B^h \times B((h, \infty))$). The map

$$\langle f, s \rangle \rightarrow \langle f(s), V_{L_n} f'(s) \rangle, \quad s > h,$$

is $B^h \times B(R^+)$ -measurable by Lemmas 2.1 and 4.1 and continuity of V_{L_n} . Composing the last two mappings finishes the proof of the Lemma.

Now we apply assumption (C) to obtain, given $\varepsilon > 0$,

$$\begin{aligned} (*) \quad P_0^v(f : d(\langle f(t_n(f)), V_{L_n} f'(t_n(f)) \rangle, \langle f(h), f'(h) \rangle) \geq \varepsilon \mid B_h^0) \\ = P_{f(h)}^{f'(h)}(g : d(\langle f(h), f'(h) \rangle, \langle g(s_n(g)), V_{L_n} g'(s_n(g)) \rangle) \geq \varepsilon) \end{aligned}$$

almost everywhere for P_0^v , where $s_n(g)$ is the unique value of s such that

$$\langle g(s), s + h \rangle = L_n \langle x, h + \varepsilon_n \rangle$$

for some x . The set of $\langle y, s \rangle \in M$ such that

$$\langle y, s + h \rangle = L_n \langle x, h + \varepsilon_n \rangle$$

for some $x \in X$ is the set of points of the form $L_n \langle z, \alpha_n \rangle$ for some fixed $\alpha_n > 0$ and arbitrary $z \in X$, since L_n^{-1} takes parallel hyperplanes into parallel hyperplanes.

For $y \in X$ the translation

$$T_y : \langle \xi, t \rangle \rightarrow \langle \xi - y, t \rangle$$

of M onto itself takes points of the form $L_n \langle a, \alpha_n \rangle$ onto points of the form $L_n \langle b, \alpha_n(y) \rangle$ for some $\alpha_n(y)$. If $|y| \leq h$, then $\alpha_n(y) > 0$, and as $n \rightarrow \infty$, $\alpha_n(y) \rightarrow 0$ uniformly for $|y| \leq h$.

For any $y \in X$ and $u \in V$, let

$$A_{n,y,u} = P_y^u(g : d(\langle y, u \rangle, \langle g(s_n(g)), V_{L_n} g'(s_n(g)) \rangle) \geq \varepsilon).$$

Applying first T_y and then L_n^{-1} , letting $u_n = V_{L_n}^{-1} u$, we obtain by (A) and (D)

$$A_{n,y,u} = P_0^{u_n}(\phi : d(\langle 0, u_n \rangle, \langle \phi(\alpha_n(y)), \phi'(\alpha_n(y)) \rangle) \geq \varepsilon).$$

Of course, $0 \leq A_{n,y,u} \leq 1$ for all n, y and u . As $n \rightarrow \infty$, since $\alpha_n(y) \rightarrow 0$, and the sequence u_n and its limit u form a compact set for any u , we have by (E) that $A_{n,y,u} \rightarrow 0$ for any u and y with $|y| \leq h$. Integrating (*) with respect to P_0^v over all of \mathcal{A} , we conclude by the bounded convergence theorem that

$$P_0^v(f : d(\langle f(t_n(f)), V_{L_n} f'(t_n(f)) \rangle, \langle f(h), f'(h) \rangle) \geq \varepsilon)$$

tends to 0 as $n \rightarrow \infty$. Now if $B \in B(W)$ let

$$P_n(B) = P_0^v(f : \langle f(t_n(f)), V_{L_n} f'(t_n(f)) \rangle \in B),$$

defined by Lemma 5.2. We have shown that

$$P_n \rightarrow Q(h, 0, v) \text{ weak}^* \text{ as } n \rightarrow \infty. \text{ By (D),}$$

$$P_n(B) = P_0^{v_n}(g : \langle g(h + \varepsilon_n), g'(h + \varepsilon_n) \rangle \in B) = Q(h + \varepsilon_n, 0, v_n, B).$$

We also know that

$$\|Q(h + \varepsilon_n, 0, v_n, \cdot) - Q(h, 0, v_n, \cdot)\|_{BL}^* \rightarrow 0$$

as $n \rightarrow \infty$ since $\varepsilon_n \rightarrow 0$ and the v_n lie in a compact set. Hence $Q(h, 0, v_n, \cdot) \rightarrow Q(h, 0, v, \cdot)$ weak*, and the proof is complete.

It is now easy to prove

Lemma 5.3. *Under assumptions (A) through (E), $(\{Q_{x,v}\}, Q)$ is a Markov process.*

Proof. We know that $(Q, I, W, B(W))$ is a measurable path space. The $Q_{x,v}$ are probability measures on $B^0(Q, B(W))$ and are starting probabilities by assumption (B). For the Markov property, we must first show that given $t \geq 0$ and $A \in B(W)$, $Q(t, x, v, A)$ is jointly measurable in x and v . If A is open, $x_n \rightarrow x$, and $v_n \rightarrow v$,

$$Q(t, x, v, A) \leq \limsup Q(t, x_n, v_n, A)$$

by Theorem 5.1. This implies joint measurability in x and v , and then A can be replaced by an arbitrary Borel set. The Markov property itself then holds by assumption (C), q.e.d.

(The following section, which presents Lemma 5.4, was received as manuscript added to proof on 15 September 1965.—*Editor*)

We next verify Dynkin's condition "L(Γ)" for compact sets Γ , which says in our case:

Lemma 5.4. For any compact set $K \subset X \times V$ and $u \geq 0$,

$$\lim_{\langle x, v \rangle \rightarrow \infty} \sup_{0 \leq t \leq u} Q(t, x, v, K) = 0.$$

(Here $\langle x, v \rangle \rightarrow \infty$ means $|x| \rightarrow \infty$ and/or $|v| \rightarrow 1$.)

Proof. For $|x| \rightarrow \infty$ the result is clear. Thus it suffices to show that for any compact set $C \subset V$,

$$\lim_{|v| \uparrow 1} \sup_{0 \leq t \leq u} Q(t, 0, v, X \times C) = 0.$$

We may assume that C is a closed ball centered at 0, and then that $v = (z, 0, 0)$ with $0 \leq z \uparrow 1$. For $\langle x, s \rangle \in M$ let

$$L(v) \langle x, s \rangle = \langle (x_1 + zs)/\sqrt{1-z^2}, x_2, x_3, (s + zx_1)/\sqrt{1-z^2} \rangle.$$

Then $L(v) \in \mathcal{L}$ and $V_{L(v)}(0) = v$. Also for any $t \geq 0$, by (D),

$$Q(t, 0, v, X \times C) = P(f : V_{L(v)} f'(s(v)) \in C),$$

where $P = P_0^0$ and $s(v) = s(v, t, f)$ is the unique number s such that

$$\hat{f}(s) = \langle f(s), s \rangle \in H = L(v)^{-1} \{ \langle y, t \rangle : y \in X \}.$$

Let C_δ be the cone $|x| \leq \delta s$. By assumption (B), for any $\delta > 0$ $f(s)$ will P -almost surely lie in C_δ for s small enough. In $H \cap C_\delta$, $s \geq \alpha(v) > 0$, where

$$\alpha(v) = \alpha(v, t, \delta) = t\sqrt{1-z^2}/(1+z\delta),$$

so that $\alpha(v) \rightarrow 0$ as $z \uparrow 1$. Thus if $D(\delta, v)$ is the set of all $\langle x, \alpha(v) \rangle$ with $|x| \leq \delta\alpha(v)$, and

$$S(v) = \{f: f(\alpha(v)) \in D(\delta, v)\},$$

then $P(S(v)) \rightarrow 1$ as $z \uparrow 1$. Also, by (E), for any $\varepsilon > 0$

$$\lim_{z \uparrow 1} P(f: |f'(\alpha(v))| > \varepsilon) = 0.$$

Let $u(f, v) = V_{L(v)}f'(\alpha(v))$. We can apply the Markov property at time $\alpha(v)$ to obtain

$$\begin{aligned} (*) \quad P(f: f \in S(v) \quad \text{and} \quad |V_{L(v)}f'(s(v)) - u(f, v)| \leq \varepsilon | B_{\alpha(v)}^0) \\ = \prod(f, v, \delta) = P_{f(\alpha(v))}^{f'(\alpha(v))}(g: |V_{L(v)}g'(\sigma(v)) - u(f, v)| \leq \varepsilon, f \in S(v)) \\ = 0 \text{ otherwise,} \end{aligned}$$

P -almost everywhere; for $f \in S(v)$ we have let $\sigma(v) = \sigma(v, g, \delta, t)$ be the unique s such that

$$\langle g(s), s + \alpha(v) \rangle \in H.$$

Translating by $-f(\alpha(v))$, using (A), we obtain

$$\prod(f, v, \delta) = P_0^{f'(\alpha(v))}(h: |V_{L(v)}h'(\kappa(v)) - u(f, v)| \leq \varepsilon),$$

where $\kappa(v) = \kappa(v, h, f, \delta, t)$ is the unique κ such that

$$\langle h(\kappa) + f(\alpha(v)), \kappa + \alpha(v) \rangle \in H.$$

Now we transform by $L(v)$, using (D), to get

$$\prod(f, v, \delta) = P_0^{u(f, v)}(j: |j'(\lambda(v)) - u(f, v)| \leq \varepsilon),$$

where

$$\lambda(v) = t - L(v)_{(2)}f(\alpha(v)).$$

Note that $\lambda(v)$ does not depend on j , only on v, t, δ , and $f(\alpha(v))$. Now

$$L(v) \langle x, \alpha(v) \rangle = \langle (x_1 + z\alpha(v))/\sqrt{1-z^2}, x_2, x_3, (\alpha(v) + zx_1)/\sqrt{1-z^2} \rangle = \langle \cdot, x_2, x_3, T \rangle,$$

where

$$T = \frac{t}{1+z\delta} + \frac{zx_1}{\sqrt{1-z^2}}.$$

For $\langle x, \alpha(v) \rangle \in D(\delta, v)$, T is smallest when $x_1 = -\delta\alpha(v)$.

Then

$$T = t(1-z\delta)/(1+z\delta).$$

Thus

$$\lambda(v) \leq 2z\delta t/(1+z\delta) \leq 2\delta t.$$

Hence

$$\prod(f, v, \delta) \geq \inf_{0 \leq s \leq 2\delta t} P_0^{u(f, v)}(j: |j'(s) - u(f, v)| \leq \varepsilon).$$

Thus by (E),
$$\lim_{\delta \downarrow 0} \prod(f, v, \delta) = 1,$$

uniformly in f and v . Thus, integrating our Markov equation (*) with respect to P , we have

$$\lim_{\delta \downarrow 0} \lim_{|v| \uparrow 1} P(f: f \in S(v) \quad \text{and} \quad |V_{L(v)} f'(s(v)) - u(f, v)| \leq \varepsilon) = 1.$$

For $(a, b, c) \in V$ we have

$$\begin{aligned} V_{L(v)}(a, b, c) &= (a + z, b\sqrt{1-z^2}, c\sqrt{1-z^2}) / (1 + za), \\ |V_{L(v)}(a, b, c) - v|^2 &\leq (a^2(1-z^2)^2 + b^2 + c^2) / (1 + za)^2 \\ &\leq (a^2 + b^2 + c^2) / (1 - |a|)^2. \end{aligned}$$

Thus for $\gamma \leq \frac{1}{2}$, $|(a, b, c)| \leq \gamma$ implies

$$|V_{L(v)}(a, b, c) - v| \leq 2\gamma.$$

Hence for any $\varepsilon > 0$ and $0 < \delta < 1$,

$$\lim_{|v| \uparrow 1} P(f: |V_{L(v)} f'(\alpha(v)) - v| \leq \varepsilon) = 1.$$

Thus, letting $\delta \rightarrow 0$, we get

$$\lim_{|v| \uparrow 1} P(f: |V_{L(v)} f'(s(v)) - v| \leq \varepsilon) = 1.$$

Now for any compact $C \subset V$, there is an $\varepsilon > 0$ such that $|v| > 1 - \varepsilon$ and $|w - v| > \varepsilon$ imply $w \notin C$. Then

$$\lim_{|v| \uparrow 1} P(f: V_{L(v)} f'(s(v)) \in C) = 0$$

This limit is uniform in $0 \leq t \leq u$ since

$$\lim_{|v| \uparrow 1} \alpha(v) = \lim_{\delta \downarrow 0} \lambda(v) = 0$$

uniformly for $0 \leq t \leq u$. The proof is complete.

Under assumptions (A)–(E), each measure $Q_{x,v}$ gives outer measure 1 to the class \mathcal{G} of functions in Q having limits from the left and continuous from the right for each $t \geq 0$ (Dynkin [2], Theorem 3.6).

The set of $\langle x(\cdot), v(\cdot) \rangle \in \mathcal{G}$ such that $|x(s) - x(t)| \leq |s - t|$ for all rational $s, t \geq 0$ is $B^0(\mathcal{G})$ -measurable and has $Q_{y,u}$ -measure 1 for each $\langle y, u \rangle \in W$ since P_y^u is concentrated in \mathcal{A} . Thus by definition of \mathcal{G} , $|x(s) - x(t)| \leq |s - t|$ for all $s, t \geq 0$ with $Q_{y,u}$ -probability 1 on \mathcal{G} . The map

$$\langle t, \langle x(\cdot), v(\cdot) \rangle \rangle \rightarrow x'(t)$$

is $B(R^+) \times B^0(\mathcal{G})$ -measurable by Lemma 4.1. For each t , $x'(t) = v(t)$ with $Q_{y,u}$ -probability one. Hence by the Fubini theorem, for $Q_{y,u}$ -almost all $\langle x(\cdot), v(\cdot) \rangle \in \mathcal{G}$ we have $x'(t) = v(t)$ for almost all t , so that since $v(\cdot)$ is locally bounded and $x(\cdot)$ is Lipschitzian, $x(\cdot)$ is an indefinite integral of $v(\cdot)$ and $x'(t^+) = v(t)$ for all $t \geq 0$. Let \mathcal{W} be the set of $\langle x(\cdot), v(\cdot) \rangle \in \mathcal{G}$ with $x'(t^+) = v(t)$ for all $t \geq 0$. Then we have a Markov process

$$(\{Q_{x,v}\}, \mathcal{W}, I, W, B(W)).$$

Let $\overline{\mathcal{W}}$ be the set of functions $t \rightarrow \langle c + t, f(t) \rangle, f \in \mathcal{W}, c, t \in R$. Let \overline{W} be the product space $R \times W$, and let R_c be the unit mass concentrated in the function $t \rightarrow c + t$. Let $Q_{c,x,v}$ be the product measure $R_c \times Q_{x,v}$. Then by Theorem 3.1,

$$(\{Q_{c,x,v}\}, \overline{\mathcal{W}}, I, \overline{W}, B(\overline{W}))$$

is a Markov process. This process is ‘‘Fellerian’’ by Theorem 5.1 and right continuous, hence strongly Markovian (Dynkin [1], Theorem 5.10).

6. The proper time

For any $f \in \mathcal{A}$ (defined early in §4) the proper time $\tau(s, f)$ is defined (cf. Møller [1] §§ 36, 37) by

$$\tau(s, f) = \int_0^s \left(1 - \left| \frac{df}{dt} \right|^2 \right)^{1/2} dt$$

where the integrand is clearly bounded and measurable. By assumption on \mathcal{A} , it is strictly positive almost everywhere (Lebesgue measure). For $F \in \mathcal{W}, F(t) = \langle f(t), v(t) \rangle$, or for $F \in \overline{\mathcal{W}}, F(t) = \langle f(t), c + t, v(t) \rangle$, we let $\tau(s, F) = \tau(s, f)$. Then let

$$\varphi_s^i(f) = \tau(t, f) - \tau(s, f)$$

for $0 \leq s \leq t, f \in \overline{\mathcal{W}}$. Clearly $\varphi_s^i(f) = 0$ for any $s \geq 0$, and $\varphi_s^i(f) > 0$ if $0 \leq s < t$. We have

$$\varphi_s^i(f) + \varphi_u^t(f) = \varphi_u^s(f)$$

for $0 \leq s \leq t \leq u, f \in \overline{\mathcal{W}}$, and $\varphi_s^i(f)$ is continuous in s and t . The functional φ is stationary in the sense that

$$\varphi_s^i(\theta_h f) = \varphi_{s+h}^{s+h}(f)$$

for $h \geq 0, 0 \leq s \leq t, f \in \overline{\mathcal{W}}$.

The strong Markov process $(\{Q_{c,x,v}\}, \overline{\mathcal{W}}, I, \overline{W}, B(\overline{W}))$ is ‘‘strongly’’ measurable as defined by Dynkin [2], 3.17, i.e. the conclusion of Lemma 2.1 above holds, since the hypothesis (right continuity) does. We can extend each of the σ -fields $B_i^s(\overline{\mathcal{W}})$ by the sets which are subsets of sets of measure zero for all the $Q_{c,x,v}$, thus obtaining a ‘‘complete’’ fields Markov process as defined by Dynkin [2], 3.6.

Thus we have established all the hypotheses of a theorem on ‘‘random change of time’’: Dynkin [2], Theorem 10.10. To state the theorem in our case, we let $\tau \rightarrow t(\tau, f)$ be the inverse of the increasing function $t \rightarrow \tau(t, f)$ for any $f \in \overline{\mathcal{W}}$. Then for any fixed $\tau > 0, t(\tau, \cdot)$ is a stopping time. For $f \in \overline{\mathcal{W}}$ let

$$(Af)(\tau) = f(t(\tau, f)), \zeta'(Af) = \sup \{ \tau(t, f) : t \geq 0 \}, x(\tau, f) = (Af)(\tau).$$

Theorem 6.1. $(x(\tau, \cdot), \zeta' \circ A, B_{t(\tau, \cdot)}, Q_{c,x,v})$ is a fields Markov process.

Corollary. $(\{Q_{c,x,v} \circ A^{-1}\}, A(\overline{\mathcal{W}}), \zeta', \overline{W}, B(\overline{W}))$ is a (path space) Markov process.

We have the inclusion

$$A(B_{t(\tau, \cdot)}) \supset B_{\tau}^0;$$

I don’t know whether the converse inclusion holds (given right continuity, in general).

The "velocity component" of a function in $A(\overline{\mathcal{W}})$ is not the derivative of the "space component". To remedy this situation, let U be the transformation of V into M_0 defined by

$$U(v) = \langle v, 1 \rangle / \sqrt{1 - |v|^2}.$$

Then $\|U(v)\|^2 = 1$ for any $v \in V$, and the last component of $U(v)$ is positive. $U(v)$ is the four-velocity associated with the "velocity" v (cf. Møller [1] § 37). Let \mathcal{U} be the set of all $U(v)$, $v \in V$. \mathcal{U} is one nappe of a three-dimensional hyperboloid. For

$$\xi = \langle x, c, v \rangle \in M \times V,$$

let $U_2(\xi) = \langle x, c, U(v) \rangle \in M \times \mathcal{U}$. Then U_2 is a homeomorphism of $M \times V$ onto $M \times \mathcal{U}$, and it defines a transformation U_2 of $A(\overline{\mathcal{W}})$ onto another function space \mathcal{V} . For $\langle x, c, U \rangle \in M \times \mathcal{U}$, $z = \langle x, c \rangle$, $U = U(v)$, let

$$\prod_{z,U} = Q_{c,x,v} \circ A^{-1} \circ U_2^{-1}, \zeta(U_2 f) = \zeta'(f)$$

(U_2 is one-to-one). Then by a standard result on transformation of phase spaces (Dynkin [2] Theorem 10.13) we have

Theorem 6.2. ($\{\prod_{z,U}\}$, \mathcal{V} , ζ , $M \times \mathcal{U}$, $B(M \times \mathcal{U})$) is a Markov process.

We now have that for $f = \langle \phi, \psi \rangle \in \mathcal{V}$, where ϕ has values in M and ψ in \mathcal{U} , $\phi'(\tau^+) = \psi(\tau)$ for all $\tau < \zeta(f)$, since the derivative with respect to τ is the 4-velocity (Møller [1] § 37, equation (38)). Also ψ has a limit from the left for all $\tau > 0$. Thus ψ is bounded on any bounded closed subinterval of $[0, \zeta(f))$, and ϕ is locally Lipschitzian and an indefinite integral of ψ .

Assumption (A) of § 4 (spatial homogeneity) and the same condition for the measures R_c of § 5 imply the corresponding condition for the $\prod_{z,U}$:

(I) If $y, z \in M$, $U \in \mathcal{U}$, $\tau \geq 0$, and $A \in B(M)$,

$$\prod_{z,U}(\langle \phi, \psi \rangle : \phi(\tau) \in A) = \prod_{z+y,U}(\langle \phi, \psi \rangle : \phi(\tau) \in A + y).$$

Also, assumption (D) of § 4 yields

(II) For any $L \in \mathcal{L}$, $A \in B^0(\mathcal{V})$, and $U \in \mathcal{U}$,

$$\prod_{0,U}(A) = \prod_{0,L(U)}(L(A)).$$

An $L \in \mathcal{L}$ acts on the tangent space M_0 of M through the natural isomorphism of M and M_0 , or, equivalently here, through its differential dL , so that if $U \in \mathcal{U}$ and $\phi'(\tau) = U$, then

$$L(U) = dL(\phi(\tau))/d\tau.$$

Now let P be the natural projection of $M \times \mathcal{U}$ onto \mathcal{U} . By (I) and Theorem 3.2,

$$(\{\prod_{z,U} \circ P^{-1}\}, P(\mathcal{V}), \eta, \mathcal{U}, B(\mathcal{U}))$$

is a Markov process independent of z , where $\eta(P(f)) = \zeta(f)$ for any $f \in \mathcal{V}$. P is one-to-one on \mathcal{V} , so that this process determines $\{\prod_{z,U}\}$ (by integration).

For $\sigma \geq 0$, $U \in \mathcal{U}$, and $A \in B(\mathcal{U})$, let

$$P_\sigma(U, A) = \prod_{z,U}(\langle \phi, \psi \rangle \in \mathcal{V} : \psi(\sigma) \in A)$$

(which is independent of z). By the Markov property, we have the Chapman–Kolmogorov equation (semigroup property)

$$P_{\sigma+\tau}(U, B) = \int_{\mathcal{U}} P_{\sigma}(U, dV) P_{\tau}(V, B)$$

for any $\sigma, \tau \geq 0, U \in \mathcal{U}, B \in \mathcal{B}(\mathcal{U})$. By (II), we have

$$P_{\sigma}(U, B) = P_{\sigma}(L(U), L(B)), L \in \mathcal{L}.$$

In § 8, we shall reproduce the known classification of all semigroups $\{P_{\sigma}\}$ satisfying our conditions.

Lemma 6.3. *For any $\{\prod_{z, U}\}$ satisfying our conditions and $\tau > 0$,*

$$\prod_{z, U} \{ \langle \phi, \psi \rangle : \phi'(\tau) \text{ exists} \} = \prod_{z, U} \{ f : \zeta(f) > \tau \}.$$

Proof. Let \mathcal{N} be the set of all $\tau > 0$ for which the conclusion does not hold. For $f = \langle \phi, \psi \rangle \in \mathcal{V}, \phi'(\tau)$ exists for almost all τ in the interval $(0, \zeta(f))$. Thus by Lemma 4.1 and the Fubini theorem, \mathcal{N} has Lebesgue measure 0. Thus for any $\tau > 0$ there exist $t > 0$ and $u > 0$, neither of them in \mathcal{N} , such that $t + u = \tau$. For any $\sigma > 0$,

$$\prod_{0, U} \{ \langle \phi, \psi \rangle : \phi'(\sigma) \text{ exists} \}$$

is independent of U by condition (II). Letting $\sigma = u$ and applying the Markov property at time t , we conclude that $\tau \notin \mathcal{N}$, q.e.d.

Of course, for $\langle \phi, \psi \rangle \in \mathcal{V}, \phi'(\tau)$ is equal to $\psi(\tau)$ if the former is defined.

7. The space \mathcal{U}

\mathcal{U} is well-known as “hyperbolic space” or “Lobachevsky space” (Gelfand and Berezin [1]). It is acted on transitively by the proper Lorentz group \mathcal{L} . The subgroup K_p of \mathcal{L} leaving a point p of \mathcal{U} fixed is isomorphic to the orthogonal group K on three-dimensional Euclidean space, and is a maximal compact subgroup of \mathcal{L} . \mathcal{U} can be regarded as the homogeneous space \mathcal{L}/K_p of right cosets of K_p in \mathcal{L} . (Of course, K_p is not a normal subgroup of \mathcal{L} .)

\mathcal{U} has a natural \mathcal{L} -invariant Riemannian structure, inherited from the “pseudo-Riemannian” or Lorentz form $\|\cdot\|^2$ on M_o by restriction to the tangent spaces of \mathcal{U} .

We put geodesic polar (or spherical) coordinates on \mathcal{U} (see Helgason [1] p. 401) as follows: we take p as $(0, 0, 0, 1)$ for a given Lorentz coordinate system (x, y, z, t) on M_o . Given any point U in \mathcal{U} we let $\varrho(U)$ be the Riemannian distance from p to U in \mathcal{U} . Now each surface $\varrho = \varrho_0 > 0$ in \mathcal{U} is isometric to a Euclidean sphere of some radius $f(\varrho_0)$. (K_p acts transitively on these surfaces, i.e. \mathcal{U} is a “two-point homogeneous” Riemannian manifold as defined by Helgason [1] p. 335.) In each such sphere we choose standard spherical coordinates θ and $\phi, 0 \leq \theta < 2\pi, -\pi/2 \leq \phi \leq \pi/2$, constant on each geodesic emanating from p . The functions (ϱ, θ, ϕ) make up a regular local coordinate system except where ϱ or θ is zero or $\phi = \pm \pi/2$.

Now let $r = (x^2 + y^2 + z^2)^{1/2}$. Then ϱ is a function of r in \mathcal{U} , and $f(\varrho(r)) = r, 0 < r < \infty$, since the induced Riemann metric on a sphere $\varrho = \text{constant}$ can also be induced from a 3-plane $t = \text{constant}$, which has the standard Euclidean structure.

To find the function $\varrho(r)$, and hence its inverse f , it suffices to find the Riemann structure induced on the hyperbola $H: t^2 - x^2 = 1, y = z = 0$. Here $r = |x|$. x is a coor-

dinate function on H , and the tangent vector d/dx is $(1, 0, 0, x/\sqrt{1+x^2})$ at the point $(x, 0, 0, \sqrt{1+x^2})$ of H . This vector has "magnitude" $1/\sqrt{1+x^2}$ for the Riemann metric induced on H . Thus a unit vector $d/d\rho$ for the Riemann metric is $\sqrt{1+x^2} d/dx$ for $x > 0$, so that $dx/d\rho = \sqrt{1+x^2}$, $\rho = \log(x + \sqrt{1+x^2})$ (since $\rho = 0$ when $x = 0$), $|x| = \sinh \rho$ on H , and $r = \sinh \rho$ throughout \mathcal{U} .

Now let g be the induced Riemann form on \mathcal{U} :

$$g_{ij} = g(\partial/\partial x_i, \partial/\partial x_j), \quad \text{where } x_1 = \rho, x_2 = \theta, x_3 = \phi.$$

Then $g_{ij} = 0$ for $i \neq j$, $g_{11} = 1$, $g_{22} = \sinh^2 \rho$, and $g_{33} = \sinh^2 \rho \sec^2 \phi$ (except at the exceptional points $\theta = 0$ etc.). Now the "Laplace-Beltrami" operator Δ on \mathcal{U} is (see Helgason [1] p. 386)

$$\frac{1}{\sqrt{\bar{g}}} \sum_k \frac{\partial}{\partial x_k} \sum_i g^{ik} \sqrt{\bar{g}} \frac{\partial}{\partial x_i}$$

(\bar{g} = determinant of (g_{ij}) , g^{ik} = inverse matrix)

$$= \partial^2/\partial \rho^2 + 2 \coth \rho \partial/\partial \rho + \operatorname{csch}^2 \rho \partial^2/\partial \phi^2 - \tan \phi \operatorname{csch}^2 \rho \partial/\partial \phi + \operatorname{csch}^2 \rho \sec^2 \phi \partial^2/\partial \theta^2$$

(wherever ρ, θ, ϕ form a coordinate system).

8. Semigroups of measures on \mathcal{U}

At the end of §6 we had obtained a collection of nonnegative, K_p -invariant measures $P_t(p, \cdot)$ on $B(\mathcal{U})$ for $p \in \mathcal{U}$, $t \geq 0$, of total mass at most 1. For a fixed p , we call finite, K_p -invariant, nonnegative measures on \mathcal{U} "radial". The "convolution" of two radial measures μ and ν on \mathcal{U} is given by

$$(\mu * \nu)(A) = \int_{\mathcal{U}} \nu(A_x) \mu(dx)$$

for any $A \in B(\mathcal{U})$, where A_x is a set $T_x(A)$ and T_x is any element of \mathcal{L} taking x into p . By K_p -invariance, it is irrelevant which such element is chosen, and we can take T_x continuous in x to insure measurability.

Since $P_t(U, B) = P_t(L(U), L(B))$ for all $L \in \mathcal{L}$, $B \in B(\mathcal{U})$, we have

$$P_s * P_t = P_{s+t} \quad \text{for all } s, t \geq 0.$$

Thus the P_t form a semigroup under convolution. Now each radial measure on \mathcal{U} has a "Fourier transform", defined in terms of the eigenfunctions of Δ which are functions of ρ alone ("spherical functions": see Helgason [1] Chap. X §3 p. 398). As given by Tutubalin, Karpelevich and Shur [1] (hereafter called "TKS") the non-constant spherical functions on \mathcal{U} , normalized to the value 1 at p , are

$$F_\lambda(\rho) = \sin \lambda \rho / \lambda \sinh \rho,$$

where λ is any complex number and $\Delta F_\lambda = (-1 - \lambda^2) F_\lambda$. For λ or ρ equal to 0, $F_\lambda(\rho)$ is defined by continuity. These functions are bounded for $|\operatorname{Im} \lambda| \leq 1$. The Fourier transform of a radial measure μ is given by

$$\hat{\mu}(\lambda) = \int_{\mathcal{U}} F_\lambda(x) d\mu(x),$$

defined at least for $|Im \lambda| \leq 1$. For radial measures convolution is commutative, and

$$(\mu * \nu)^\wedge(\lambda) = \hat{\mu}(\lambda) \hat{\nu}(\lambda)$$

(see TKS [1]).

A radial measure μ on \mathcal{U} such that for all $n = 1, 2, \dots$, μ is the n -fold convolution $(\mu_n)^n$ of a radial measure μ_n with itself, is called "infinitely divisible". Clearly the measures P_t have this property. Such measures have been characterized by Tutubalin [1] and Gangolli [1]: μ is infinitely divisible of total mass $\beta > 0$ if and only if

$$(*) \quad \hat{\mu}(\lambda) = \beta \exp \left((-1 - \lambda^2) \alpha - \int_{\mathcal{U}} (1 - F_\lambda(u)) dL(u) \right)$$

where $\alpha \geq 0$ and L is a nonnegative, K_p -invariant measure on \mathcal{U} satisfying

$$\int_0^\infty \frac{\varrho(u)^2}{1 + \varrho(u)^2} dL(u) < \infty$$

(so that the integral of $1 - F_\lambda$ is absolutely convergent). $\alpha = \alpha(\mu)$ and $L = L(\mu)$ are uniquely determined. Conversely, for any such measure L and numbers α and $\beta \geq 0$, there is a radial, infinitely divisible measure μ satisfying (*).

Tutubalin [1] proved the representation formula (*) under a definition of infinite divisibility requiring that for every $\varepsilon > 0$, μ is a convolution of possibly distinct measures each concentrated within Riemann distance ε of p , except for mass ε . An argument in his proof, partly reproduced below, shows that this is equivalent to the definition given above. A proof under our definition was also given by R. Gangolli [1] for general symmetric spaces of non-compact type.

We say a radial measure μ divides another one, ν , if $\mu * \varrho = \nu$ for some radial ϱ . Clearly this implies

$$\alpha(\mu) + \alpha(\varrho) = \alpha(\nu) \quad \text{and} \quad L(\mu) + L(\varrho) = L(\nu),$$

so that μ divides ν if and only if $\alpha(\mu) \leq \alpha(\nu)$ and $L(\mu) \leq L(\nu)$.

Now given a convolution semigroup P_t , we have $L(P_{rt}) = rL(P_t)$ and $\alpha(P_{rt}) = r\alpha(P_t)$ for all rational $r > 0$ by uniqueness, and hence for all $r > 0$ by the ordering just mentioned. Thus the semigroup is completely determined by $L(P_1)$ and $\alpha(P_1)$.

$P_t(p, \cdot)$ converges weak* to the unit mass concentrated at p as $t \downarrow 0$, since \hat{P}_t converges to 1 pointwise and each $|F_\lambda|$ for λ real is bounded below 1 outside any neighborhood of p . Note that the original continuity assumption (E) was used only to obtain a Markov process with proper time as parameter; for the latter, the continuity properties follow from the invariance conditions, and the first part of assumption (C) is taken care of by Lemma 6.3.

We now show that ζ is actually infinite:

Lemma 8.1. *For any $z \in M$ and $U \in \mathcal{U}$, $\prod_{z, U}(\zeta = +\infty) = 1$.*

Proof. The Markov process

$$\{(\prod_{z, U} P^{-1}), P(\mathcal{V}), \eta, \mathcal{U}, B(\mathcal{U})\}$$

is right continuous and Fellerian, hence strongly Markovian. Given $f \in P(\mathcal{V})$, let $t_0(f) = 0$. Given $t_0(f), \dots, t_n(f)$, let $t_{n+1}(f)$ be the least t , if any, such that the Riemann distance from

$$f(t_0(f) + \dots + t_n(f)) \quad \text{to} \quad f(t_0(f) + \dots + t_n(f) + t)$$

is greater than or equal to 1 (if such a t exists, $t_{n+1}(f)$ is well-defined by right continuity). Clearly $t_1(f) + \dots + t_n(f)$ is a stopping time for each n .

If $\eta(f)$ is finite, then all the $t_n(f)$ are defined and their sum converges. In any case, each t_n is defined with $\prod_{z, U} P^{-1}$ probability one. The random variables t_n for $n \geq 1$ are independent by the strong Markov property and have the same probability distribution on the nonnegative real axis. Each is almost surely strictly positive by right continuity. Thus the probability of their sum converging is zero, q.e.d.

It follows that under our assumptions the measures $P_t(p, \cdot)$ are all probability measures. We summarize our results as follows:

Theorem 8.2. *There is a natural 1 - 1 correspondence between sets $\{P_x^v\}$ of measures on \mathcal{A} satisfying (A)-(E) of § 4, and infinitely divisible radial probability measures P_1 on \mathcal{U} , and hence with pairs (L, α) as described earlier in this section. Each P_x^v gives outer measure 1 to the set of functions $f \in \mathcal{A}$ having left and right derivatives at all points (except left derivatives at 0), equal except at countably many points, with right derivatives continuous from the right and left derivatives from the left.*

9. Further discussion of radial semigroups on \mathcal{U}

In this section we discuss further facts which were not needed for the proof of Theorem 8.2, but may well be of interest.

Suppose μ is an infinitely divisible radial probability measure on \mathcal{U} , and $\{\prod_{z, \sigma}\}$ the corresponding Markov process on \mathcal{V} . It follows from results of Gangolli [3] that if $n(M, f, t)$ for $M, t \geq 0$ and $f = \langle \phi, \psi \rangle \in \mathcal{V}$ is the number of values of $\tau < t$ such that the Riemann distance from $\phi'(\tau -)$ to $\phi'(\tau +)$ is at least M , then

$$\int n(M, f, t) d\prod_{z, \sigma}(f) = tL(\mu)(U: \varrho(U) \geq M).$$

Thus if $L = 0, \alpha > 0, \psi$ is almost surely continuous. These processes ("Brownian motions") will be considered in § 10.

On the other hand, we have so-called "Poisson" processes defined as follows: let μ_σ be normalized K_p -invariant "surface area" measure on the sphere $\varrho = \sigma$ in \mathcal{U} , for any $\sigma > 0$. Then the "Poisson measure" $\pi(\sigma, c)$ is defined by

$$\pi(\sigma, c) = e^{-c} \sum_{n=0}^{\infty} (c\mu_\sigma)^n / n!,$$

where the powers represent convolutions. It is easy to verify that $P_t = \pi(\sigma, ct)$ gives a convolution semigroup. The associated process in \mathcal{U} will remain at one point for some time, then jump to one of the points at Riemann distance σ , with uniform probability distribution over the sphere of such points. The probability of such a jump during an interval $t_0 < t < t_0 + \varepsilon$ is $1 - e^{-\varepsilon c}$ (see Loève [1] p. 548), which is asymptotic to $c\varepsilon$ as $\varepsilon \downarrow 0$.

A Poisson process seems very natural as a velocity process, where a particle undergoes collisions at various times which cause discontinuities in its velocity, constant between collisions.

If I is the identity operator on bounded continuous functions on \mathcal{U} , then

$$\lim_{t \rightarrow 0} (P_t - I)/t$$

will be, in general, an unbounded operator defined on a subspace, called the “generator” of the semigroup P_t . Transferring our semigroups from \mathcal{U} to \mathcal{L} in a rather obvious way, we obtain information on their generators from Hunt [1] (see also Wehn [1]).

Since any radial measure μ on \mathcal{U} is an average

$$\mu = \int \mu_\sigma d\nu(\sigma)$$

for some finite measure ν on R^+ , we can explicitly calculate any convolution of radial measures if we can find $\mu_a \times \mu_b$ for $a, b > 0$. In doing this, we can assume $b \leq a$. Suppose given $c \geq 0$; let us find

$$f(a, b, c) = (\mu_a \times \mu_b)(\varrho \leq c).$$

Clearly

$$\begin{aligned} f(a, b, c) &= 0 & \text{if } c < a - b, \\ &= 1 & \text{if } c > a + b. \end{aligned}$$

For any $u \in \mathcal{U}$ and $t > 0$ let $S(u, t)$ be the set of points of \mathcal{U} at Riemann distance t from u . Let

$$q = (\sinh a, 0, 0, \cosh a);$$

then $\varrho(q) = a$. Also $L_a \in \mathcal{L}$ takes $p = (0, 0, 0, 1)$ into q , where

$$L_a(x, y, z, t) = (x \cosh a + t \sinh a, y, z, t \cosh a + x \sinh a).$$

$S(p, b)$ is the set of $(X, y, z, T) \in \mathcal{U}$ with $T = \cosh b$ and

$$X^2 + y^2 + z^2 = \sinh^2 b.$$

Thus $S(q, b)$ is the set of points

$$(X \cosh a + \cosh b \sinh a, y, z, \cosh b \cosh a + X \sinh a)$$

with $X^2 + y^2 + z^2 = \sinh^2 b$. Then $S(p, c) \cap S(q, b)$ is the set of (x, y, z, t) with $t = \cosh c$, $x = X \cosh a + \cosh b \sinh a$ where $X = (\cosh c - \cosh b \cosh a) / \sinh a$, and $y^2 + z^2 = \sinh^2 b - X^2$. Applying L_a^{-1} we obtain a circle of radius $\sinh^2 b - X^2$ in $S(p, b)$, a sphere of radius $\sinh b$, in the given coordinate system for M_0 . We now want to find the areas of the two zones demarcated by this circle.

For c small enough so that $\cosh c \leq \cosh b \cosh a$, the area of the zone which concerns us is

$$2\pi \sinh b \left(\sinh b - \frac{\cosh b \cosh a - \cosh c}{\sinh a} \right).$$

For $\cosh c > \cosh b \cosh a$ we want the area of the larger of the two zones marked off by the circle, which is given by the same formula. The total area of the sphere being $4\pi \sinh^2 b$, we obtain

$$f(a, b, c) = \frac{1}{2} \left[1 + \frac{\cosh c - \cosh b \cosh a}{\sinh b \sinh a} \right],$$

$$df(a, b, c)/dc = \sinh c/2 \sinh a \sinh b.$$

Proposition 9.1. For any $a, b, \geq 0$,

$$\mu_a * \mu_b = \int_{|a-b|}^{a+b} \frac{\sinh x}{2 \sinh a \sinh b} \mu_x dx.$$

We now turn to proving “transience” or “non-recurrence” of our invariant processes in \mathcal{U} . Proposition 9.1 will not be used.

Theorem 9.2. For any invariant Markov process $\{\prod_{z,U}\}$ as described in § 6, with $\prod_{z,U}$ not concentrated in one function, and any compact subset K of \mathcal{U} ,

$$\prod_{z,U} \{ \langle \phi, \psi \rangle : \psi(\tau) \notin K \text{ for all sufficiently large } \tau \} = 1.$$

We first prove the following

Lemma 9.3. For any radial probability measure μ on \mathcal{U} not concentrated at p , $0 < \hat{\mu}(0) < 1$ and for any $R \geq 0$ there is a $K > 0$ such that

$$\mu^n(\varrho \leq R) \leq K[\hat{\mu}(0)]^n.$$

Proof. Since μ is not concentrated at p ,

$$0 < \hat{\mu}(0) = \int_{\mathcal{U}} \frac{\varrho}{\sinh \varrho} d\mu < 1.$$

Then

$$\int \frac{\varrho}{\sinh \varrho} d\mu^n = (\mu^n)^\wedge(0) = [\hat{\mu}(0)]^n.$$

Thus given $R \geq 0$,

$$\mu^n(\varrho \leq R) < \frac{\sinh R}{R} [\hat{\mu}(0)]^n, \text{ q.e.d.}$$

Now to prove Theorem 9.2, suppose for some z, U , and compact $K \subset \mathcal{U}$,

$$P_K = \prod_{z,U} \{ \langle \phi, \psi \rangle \in \mathcal{V} : \psi(t) \in K \text{ for } t \text{ arbitrarily large} \} > 0.$$

For $k = 1, 2, \dots, f = \langle \phi, \psi \rangle \in \mathcal{V}$, let $n_k(f)$ be the k th integer n such that $\psi(t) \in K$ for some t in the interval $[n, n + 1)$, or $+\infty$ if there is no such n ; let $t_k(f)$ be the least such t , or $+\infty$ if $n_k = +\infty$. Then each t_k is a stopping time.

Let d be the Riemann distance in \mathcal{U} . There is an $R > 0$ such that

$$\prod_{z,U} \{ \langle \phi, \psi \rangle : d(\psi(t), \psi(0)) \leq R, 0 \leq t \leq 1 \} = \varepsilon > 0.$$

ε is clearly independent of z and U . Let A_k be the set of all $f \in \mathcal{V}$ such that $n_k(f)$ and $t_k(f)$ are finite, and let C_k be the set of $\langle \phi, \psi \rangle \in A_k$ such that

$$d(\psi(t_k), \psi(t_k + t)) \leq R, 0 \leq t \leq 1.$$

Then by the the strong Markov property of $\{\prod_{z,v}\}$

$$\prod_{z,v}(C_k | B_{t_k}) = \varepsilon$$

almost everywhere on A_k . Thus

$$\prod_{z,v}(C_k) = \varepsilon \prod_{z,v}(A_k) \geq \varepsilon P_K > 0 \quad \text{for } k = 1, 2, \dots$$

Thus $\prod_{z,v}\{f : f \in C_k \text{ for infinitely many } k\} \geq \varepsilon P_K$.

Hence if K_R is the set of $U \in \mathcal{U}$ such that $d(U, V) \leq R$ for some $V \in K$,

$$\prod_{z,v}\{\langle \phi, \psi \rangle : \psi(n_k + 1) \in K_R \text{ for infinitely many } k\} > 0.$$

This contradicts Lemma 9.3, so the proof of Theorem 9.2 is complete.

10. Diffusion processes in \mathcal{U} and $M \times \mathcal{U}$

A convolution semigroup P_t in a symmetric space, e.g. in \mathcal{U} , is called a *diffusion* semigroup if

$$\lim_{t \downarrow 0} P_t(\varrho \geq t)/t = 0,$$

where ϱ is the Riemannian distance from the fixed point used in defining convolution. The theory of such semigroups tells us that the generator A is defined at least on all functions which, together with their first- and second-order partial derivatives, are bounded and uniformly continuous on \mathcal{U} , and is equal on these functions to a second-order differential operator (Yosida [1]).

On \mathcal{U} , or other "two-point homogeneous" symmetric spaces G/K , G -invariance then implies that $A = m\Delta + b$ where Δ is the Laplace-Beltrami operator. By definition of A it is zero on constants and $Af \leq 0$ where f has a relative maximum, so that $b = 0$ and $m \geq 0$.

Thus we have the "Brownian motion" semigroup with parameter m as defined by Yosida [1]. On \mathcal{U} , the P_t for this semigroup are given explicitly by Tutubalin [1] as

$$P_t = (4\pi mt)^{-3/2} \frac{\varrho}{\sinh \varrho} \exp\left(-mt - \frac{\varrho^2}{4mt}\right) \cdot N,$$

where N is the \mathcal{L} -invariant measure on \mathcal{U} given by $4\pi(\sinh^2 \varrho) d\varrho d\Omega$ and Ω is normalized orthogonally invariant surface area measure on a sphere ($d\Omega = \cos \phi d\theta d\phi/4\pi$). In this case the measures P_t^2 are concentrated in the set of functions f having continuous first derivatives $f'(t)$, since the P_t are a diffusion semigroup (Yosida [1], Dynkin [1] Theorem 6.5).

Now, the space $M \times \mathcal{U}$ is a homogeneous space under the action of the "Poincaré group" $\tilde{\mathcal{L}}$ generated by \mathcal{L} (acting on M and \mathcal{U} together) and by translations of M which leave \mathcal{U} pointwise fixed. The subgroup of $\tilde{\mathcal{L}}$ leaving the point $(0, p)$ of $M \times \mathcal{U}$ fixed remains equal to K_p , so that $M \times \mathcal{U}$ can be regarded as the space $\tilde{\mathcal{L}}/K_p$ of right cosets of K_p in $\tilde{\mathcal{L}}$.

One can define a convolution for finite, (left) K_p -invariant measures on $M \times \mathcal{U}$, as before, by

$$(\mu * \nu)(A) = \int_{M \times \mathcal{U}} \mu(T_z(A)) \nu(dz), \quad A \in B(M \times \mathcal{U}),$$

where $T_z \in \mathcal{L}$ takes z into $(0, p)$ and μ and ν are invariant under K_p .

Our conditions (I) and (II) of § 6 can be regarded as conditions of $\bar{\mathcal{L}}$ -invariance of transition probabilities

$$Q_t((x, p); B) = \prod_{x,p}(f : f(t) \in B),$$

$t \geq 0, x \in M, p \in \mathcal{U}, B \in B(M \times \mathcal{U})$. Our assumptions imply, letting

$$Q_t = R_t((0, p); \cdot),$$

that $Q_s * Q_t = Q_{s+t}$ on $M \times \mathcal{U}$. If P_t is a diffusion semigroup on \mathcal{U} with generator $m\Delta$, the generator of Q_t is the differential operator whose value at the point (x, U) of $M \times \mathcal{U}$ is $-U + m\Delta$, where $m\Delta$ acts on \mathcal{U} and U acts as a first-order differential operator on M at x .

G. Schay [1] also studied $\bar{\mathcal{L}}$ -invariant diffusion processes in $M \times M_0$, and arrived at essentially the same generator or "diffusion equation" (Schay [1] Theorem 4 p. 39, Equation 3.33). Our use of the proper time permits a considerable simplification of the result.

11. Non-existence results

In § 8 above we proved the existence of a class of Lorentz-invariant processes with speeds less than 1 in $M \times V$. In this section we show that such processes with speeds equal to 1 (in $M \times S^2$, where S^2 is the sphere $|x| = 1$ in R^3) are deterministic (trivial), and that this remains essentially true if we allow states to be specified by momenta rather than velocities only. We also prove the non-existence of invariant processes in M itself. The results in this section do not require the Markov property or continuity in probability.

On a trajectory with speed almost everywhere equal to 1 the proper time τ is a constant and hence not suitable as a parameter.

Let \mathcal{D} be the set of all functions f from R^+ to X such that for $0 \leq s \leq t$,

$$|f(t) - f(s)| \leq t - s,$$

and if $f'(t)$ is defined, $|f'(t)| = 1$.

Now suppose given starting probabilities P_x^v on the stationary, measurable path space $(\mathcal{D}, I, X, B(X))$, satisfying (A), (B) and (D) of § 4, with V replaced by S^2 .

A collection $\{P_x^v\}$ will be called *deterministic* if each P_x^v gives mass 1 to a set containing only one function.

For any $f \in \mathcal{D}$ with f' non-constant there is a unique $t = t_0(f) > 0$ such that

$$t^2 - f(t)^2 = 1$$

(i.e., $\langle f(t), t \rangle \in \mathcal{U}$). Clearly $t_0(\cdot)$ is $B^0(\mathcal{D})$ -measurable.

Suppose $\{P_x^v\}$ is non-deterministic. Then by (A) and (D) no P_x^v gives mass 1 to one function. Let $x=0$ and fix $v \in S^2$. Then there exist Lorentz transformations $L_n, n = 1, 2, \dots$, such that $V_{L_n}(v) = v$ for all n and for any compact subset C of \mathcal{U} , $L_n(C)$ and C are disjoint for n large enough (we take L_n as Lorentz transformations defined by relative velocities v_n parallel to v with $|v_n| \uparrow 1$). For $A \in \mathcal{B}(\mathcal{U})$, let

$$P(A) = P_x^v\{f : \langle f(t_0(f)), t_0(f) \rangle \in A\}.$$

Then P is a finite, non-zero Borel measure on \mathcal{U} , invariant under all the L_n , which is impossible. We have proved

Theorem 11.1. *If $(\{P_x^v\}, \mathcal{D})$ satisfies (A), (B) and (D) of §4, then each P_x^v is concentrated in the function $f \in \mathcal{D}$ with $f(0) = x$ and $f'(t) = v$ for all $t \geq 0$. Conversely, such P_x^v satisfy (A) – (E).*

We can try to avoid the paucity of processes in Theorem 11.1 by introducing different possible “states” of a particle moving with speed 1 in a given direction. One possibility is to use an “energy” analogous to the energy of a photon or other particle of zero rest mass. We take such an energy as defined by a function

$$E : (\mathcal{C}, s, f) \rightarrow E(\mathcal{C}, s, f),$$

where $f \in \mathcal{D}, s \geq 0$, and \mathcal{C} is an arbitrary Lorentz coordinate system $\langle x(\cdot), t(\cdot) \rangle$, where $x(\cdot)$ and $t(\cdot)$ take M onto X and R respectively. Then the trajectory defined by f in the original coordinate system of M is defined in \mathcal{C} by another function $f_c \in \mathcal{D}$ with

$$x(\langle f(s), s \rangle) = f_c(t(\langle f(s), s \rangle)), s \geq 0.$$

We assume that $E(\mathcal{C}, s, f)$ is defined if and only if $f_c(s)$ is defined.

We require that

$$\langle E(\mathcal{C}, t, f) f_c(t), E(\mathcal{C}, t, f) \rangle$$

transform as an “energy-momentum vector” under Lorentz transformations L (in \mathcal{L}) of \mathcal{C} (see Møller [1] p. 72). This is simply the natural action of \mathcal{L} on M_0 , the tangent space of M , through its isomorphism with M . The space of possible energy-momentum vectors is the open half-cone Q of all points $\langle \xi, E \rangle, \xi \in R^3, E > 0$, with $|\xi|^2 = E^2$. Q is acted on transitively by \mathcal{L} .

The subgroup \mathcal{K} of \mathcal{L} leaving the point $(1, 0, 0, 1)$ fixed contains all transformations K_{bc} with matrices

$$\begin{pmatrix} 1 - a & b & c & a \\ -b & 1 & 0 & b \\ -c & 0 & 1 & c \\ -a & b & c & 1 + a \end{pmatrix}$$

where $a = (b^2 + c^2)/2$ and b and c are arbitrary real numbers. Thus (in its relative topology from the general linear group) \mathcal{K} is not compact.

For $k \geq 0$ let \mathcal{U}_k be the set of $\langle x_1, x_2, x_3, t \rangle \in \mathcal{U} \subset M$ with $x_2 \geq 0, x_3 \geq 0$, and $t \leq k$. In $\mathcal{U}_k, t - x_1 > 0$ (since $t > 0$ and $t^2 > x_1^2$) so by compactness there is an $\varepsilon > 0$ such that $t + x_1 \geq \varepsilon$ for all $\langle x, t \rangle \in \mathcal{U}_k$. Now

$$K_{bc}\langle x, t \rangle = \langle , -ax_1 + bx_2 + cx_3 + (1+a)t \rangle.$$

Thus for $\langle x, t \rangle \in \mathcal{U}_k$ and $b, c \geq 0$ the time component of $K_{bc}\langle x, t \rangle$ is at least ae . Hence for b and c sufficiently large, $K_{bc}(\mathcal{U}_k)$ is disjoint from \mathcal{U}_k , but, on the other hand, it is included in a larger set \mathcal{U}_k . Thus we can find an infinite sequence of disjoint sets in \mathcal{U} taken into \mathcal{U}_k by transformations in \mathcal{K} . Thus for the initial energy-momentum $(1, 0, 0, 1)$ the probability that the trajectory passes through \mathcal{U}_k is zero. We can apply the same argument to the cases where $y \geq 0, z \geq 0$ in the definition of \mathcal{U}_k is replaced by $y \geq 0, z \leq 0$ or $y \leq 0, z \geq 0$ or $y \leq 0, z \leq 0$ (letting b and c have the same signs as y and z respectively). Thus there is probability zero of passing through \mathcal{U} .

Let \mathcal{M} be the set of all functions $f = \langle g, h \rangle$ from R^+ to $M \times Q$ such that $g \in \mathcal{D}$ and $g'(t)$, where defined, is proportional to $h(t)$ (where the constant may vary with t). We have proved

Theorem 11.2. *Suppose starting probabilities $\{P_{z, \langle \xi, E \rangle}\}$ on $(\mathcal{M}, I, M \times Q, B(M \times Q))$ are homogeneous in z and Lorentz-invariant in the sense that for any $L \in \mathcal{L}$, defining a map $L \times L$ of $M \times Q$ onto itself and hence of \mathcal{M} onto itself,*

$$P_{z, \langle \xi, E \rangle} \circ (L \times L)^{-1} = P_{L(z), L\langle \xi, E \rangle}.$$

Then for each $z \in M$ and $\langle \xi, E \rangle \in Q$, $P_{z, \langle \xi, E \rangle}$ is concentrated in the set of functions $f = \langle g, h \rangle$ where g defines a straight half-line in M and $h(t)$ is proportional to $\langle \xi, E \rangle$ for almost all t (Lebesgue measure), possibly with a varying proportionality factor.

Thus the energy-momentum approach does not yield any essentially non-trivial processes either. We do not consider here the possibility of allowing still more information, e.g. a "polarization", in the definition of states.

Random processes with speed 1 have been considered by Rudberg [1] in one or two space dimensions. His approach is different from ours; to help clarify the situation we now consider the cases of one and two space dimensions.

For the "velocity" approach (as in Theorem 11.1) in any number of space dimensions, the same arguments as above yield the conclusion that invariant processes do not exist. This seems to be reflected in Rudberg's condition (p. 12, above Equ. (19), and p. 28, A) that there is a distinguished time-axis direction for which the probabilities of scattering in all directions are equal.

For the "momentum" approach, apparently not treated by Rudberg, the situation is as follows: in two or more space dimensions, the situation for invariant processes is essentially the same; in the argument in two dimensions $(1, 0, 0, 1)$ and K_{bc} can be replaced respectively by $(1, 0, 1)$ and

$$K_b = \begin{pmatrix} 1-a & b & a \\ -b & 1 & b \\ -a & b & 1+a \end{pmatrix}, a = b^2/2.$$

In one space dimension, the argument which proved Theorem 11.2 does not apply, since the group \mathcal{K} reduces to the identity. The question of the truth of Theorem 11.2 in one space dimension will be left open here.

Now we turn to processes in M itself. Let \mathcal{J} be the class of all functions f from R^+ to M with

$$|f(s) - f(t)| \leq |s - t|, s, t \geq 0.$$

Suppose given starting probabilities $\{P_z\}$ on

$$(\mathcal{J}, I, M, B(M))$$

which satisfy (A) and (B) below:

(A) For $t_i \geq 0$, $A_i \in B(M)$, $z \in M$,

$$P_0\{f: f(t_i) \in A_i, i = 1, \dots, n\} = P_z\{f: f(t_i) \in A_i + z, i = 1, \dots, n\}$$

(B) For any $L \in \mathcal{L}$ and $A \in B^0(\mathcal{J})$,

$$P_0(A) = P_0(L(A)).$$

As before, for those $f \in \mathcal{J}$ with $\langle f(t), t \rangle \in \mathcal{U}$ for some t we let $t_0(f)$ be the unique such t . Assuming P_0 gives positive probability to the set of such f ,

$$P(A) = P_0(f: \langle f(t_0(f)), t_0(f) \rangle \in A)$$

is a finite, \mathcal{L} -invariant Borel measure on \mathcal{U} , which is impossible. Thus P_0 is concentrated in the set of $f \in \mathcal{J}$ with $f(0) = 0$ and f' equal to a constant $v \in S_2$. (This result was obtained under additional hypotheses by G. Schay [1] Theorem 1 pp. 17–18.) Since orthogonal transformations of X are in \mathcal{L} , the distribution of v is an orthogonally invariant probability measure, hence the standard surface area measure divided by 4π .

Let \mathcal{D} be the set of half-lines

$$t \rightarrow (at, bt, ct, t), t \geq 0$$

on which $a^2 + b^2 + c^2 = 1$ and $a, b, c \geq 0$, so that $P_0(\mathcal{D}) = \frac{1}{8}$. If $w > 0$, the Lorentz transformation X_w :

$$\langle x, y, z, t \rangle \rightarrow \left\langle \frac{x + wt}{\sqrt{1 - w^2}}, y, z, \frac{t + wx}{\sqrt{1 - w^2}} \right\rangle$$

takes the half-line defined by $v = (a, b, c)$ into the half-line of all points

$$\left\langle \frac{a + w}{1 + wa} t', b \frac{\sqrt{1 - w^2}}{1 + wa} t', c \frac{\sqrt{1 - w^2}}{1 + wa} t', t' \right\rangle, t' \geq 0.$$

Thus \mathcal{D} is taken into itself. The second components of the velocities associated with elements of $X_w(\mathcal{D})$ are at most equal to $\sqrt{1 - w^2}$. Thus an open subset of S^2 , the (a, b, c) with a and c sufficiently small and positive, has measure zero for the distribution of v in S^2 , a contradiction. Thus we have

Theorem 11.3. *No starting probabilities $\{P_z\}$ satisfying (A) and (B) (of this section) exist, i.e. there are no Lorentz-invariant processes in space-time M .*

12. Unsolved problems

A first set of problems is the explicit calculation of transition probabilities for the processes of Theorem 8.2, such as the following: the distribution of velocity in \mathcal{U} at proper time τ for nondiffusion processes; the distribution of position at proper

time τ for all processes; and the distribution of position and velocity at coordinate time t for all processes. The latter would be of special interest for diffusion processes; to simplify it, one can first seek only the distribution of velocity at time t . For a diffusion process with parameter m this leads to a parabolic partial differential equation in \mathcal{U} :

$$\cosh \varrho \partial f / \partial t = m \Delta f$$

(cf. Schay [1] equation 3.60 for the case of one space dimension).

Also of interest are the “relativistic Maxwell(-Boltzmann)” distributions originally defined by Jüttner [1] (see also Synge [1] equation (118) p. 36). These are radial measures on \mathcal{U} of the form

$$M_{\alpha\beta} = \alpha \exp(-\beta \cosh \varrho) \cosh \varrho \cdot N$$

where N is the \mathcal{L} -invariant measure

$$4\pi \sinh^2 \varrho d\varrho d\Omega$$

and $\alpha, \beta > 0$. (Note that if \mathcal{U} is projected into the spacelike hyperplane $t=0$ perpendicular to $p = (0, 0, 0, 1) \in \mathcal{U}$, $\cosh \varrho dN$ goes into Lebesgue measure.)

Given $\beta > 0$, $M_{\alpha\beta}$ is a probability measure if and only if

$$\alpha = 1/4\pi \int_0^\infty \exp(-\beta \cosh \varrho) \cosh \varrho \sinh^2 \varrho d\varrho = \beta/4\pi K_2(\beta),$$

where K_2 is a Bessel function (see e.g. Synge [1] §14). Let M_β be $M_{\alpha\beta}$ for this value of α .

Schay [1] asserts that M_β for fixed β is a “steady-state” solution of a diffusion equation with a term representing “internal friction.” Thus it may be irrelevant to ask whether M_β is infinitely divisible, etc., but it seems that Schay’s result should be followed up.

Thirdly, it would be interesting to move from the “special relativity” assumptions of this paper to the case of general relativity, in which M is no longer a vector space but a 4-dimensional manifold with a Lorentz quadratic form on its tangent spaces. The proper time is still available, but spatial homogeneity and Lorentz-invariance require reformulation. Instead of the product $M \times \mathcal{U}$ one has a subset $\mathcal{U}(M)$ of the “tangent bundle” $T(M)$: $\mathcal{U}(M)$ consists of all “forward” timelike vectors of unit magnitude at all points of M . (We assume it is possible to choose a “forward” direction at all points in a continuous way.)

In particular, there should be diffusion processes in $\mathcal{U}(M)$, generated by the differential operators which, roughly speaking, have the form $-U + m\Delta$ at a point (x, U) of $\mathcal{U}(M)$, where $m > 0$, $x \in M$, U belongs to the tangent space at x , and Δ is the Laplace-Beltrami operator in a hyperboloid in this tangent space. U may be regarded as a first-order partial derivation, or tangent vector, to $\mathcal{U}(M)$ by way of the pseudo-Riemannian “affine connection” (or “parallel displacement”: see Helgason [1] Chapter 1 §§4-6). Diffusions in sufficient generality to cover this case have been considered by Nelson [1] and Gangolli [2], but their results are not as complete as might be desired for our purposes. For example, it apparently is not known whether the semigroup generated by the operator mentioned above on bounded measurable functions actually takes bounded continuous functions into bounded continuous functions. This may require supplementary, but physically reasonable, hypotheses on M .

Finally, given a "Brownian motion" diffusion in $\mathcal{U}(M)$, even where M is a vector space, one might let the parameter m in the generator approach infinity for use in defining a sort of "Feynman integral" as in Ito's approach [1] to the non-relativistic Feynman integral. Of course, there is also the problem of finding a suitable replacement for the classical Lagrangian function.

ACKNOWLEDGEMENT

This research was partially supported by National Science Foundation grant GP-2 and ONR contract 222 (60).

Department of Mathematics, University of California, Berkeley, Cal., U.S.A.

REFERENCES

- DYNKIN, E. B., *Osnovaniya Teorii Markovskikh Protsessov* (Moscow, Fizmatgiz, 1959). English translation; *Theory of Markov Processes* (New York, Pergamon, 1960). There are also French and German translations.
- *Markovskie Protsessi* (Moscow, Gostekhizdat, 1963).
- FEYNMAN, R. P., "Space-time approach to nonrelativistic quantum mechanics", *Reviews of Modern Physics* 20, 368-387 (1948).
- GELFAND, I. M., and F. A. BEREZIN. "Some remarks on the theory of spherical functions" (in Russian), *Trudy Moskovskovo Matematicheskovo Obshchestva* 5, 311-351 (1956).
- GANGOLLI, R., "Isotropic infinitely divisible measures on symmetric spaces", *Acta Mathematica* (Uppsala) 111, 213-246 (1964).
- "On the construction of certain diffusions on a differentiable manifold", *Zeitschrift für Wahrscheinlichkeitstheorie* 2, 406-419 (1964).
- "Sample functions of certain differential processes on symmetric spaces" (preprint).
- HUNT, G. A., "Semi-groups of measures on Lie groups", *Transactions of the American Mathematical Society* 81, 264-293 (1956).
- ITO, K., "Wiener integral and Feynman integral", *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley and Los Angeles, University of California Press, 227-238 (1961).
- JÜTTNER, F., "Das Maxwell'sche Gesetz der Geschwindigkeitsverteilung in der Relativtheorie", *Annalen der Physik* 34, 856-882 (1911).
- LOÈVE, M., *Probability Theory* (second edition) (Princeton, Van Nostrand, 1960).
- NELSON, E., "An existence theorem for second-order parabolic equations", *Transactions of the American Mathematical Society* 88, 414-429 (1958).
- RAO, R. RANGA, "Relations between weak and uniform convergence of measures with applications", *Annals of Mathematical Statistics* 33, 659-680 (1962).
- RUDBERG, H., *On the Theory of Relativistic Diffusion* (Uppsala, Almqvist and Wiksells, 1957).
- SCHAY, G., "The equations of diffusion in the special theory of relativity" (physics Ph. D. thesis Princeton University, 1961; available on Xerox through University Microfilms, Ann Arbor, Michigan).
- SYNGE, J. L., *The Relativistic Gas* (North-Holland, Amsterdam, 1957).
- TUTUBALIN, V. N., "On the limit behavior of convolutions of measures in the Lobachevsky plane and space" (in Russian), *Teoriya Veroyatnosti i eyo Primeneniya* 7, 197-204 (1962). English translation in *Theory of Probability and its Applications* 7, 189-196.
- TUTUBALIN, V. N., KARPELEVICH, F. I., and SHUR, M. G., "Limit theorems for convolutions of distributions in the Lobachevsky plane and space", *Teoriya Veroyatnosti i eyo Primeneniya* 4, 432-436 (1959). English translation: *Theory of Probability and its Applications* 4, 399-402.
- WEHN, D., "Probabilities on Lie groups", *Proceedings of the National Academy of Science of the U.S.A.* 48, 791-795 (1962).
- YOSIDA, K., "On Brownian motion in a homogeneous Riemannian space", *Pacific Journal of Mathematics* 2, 263-270 (1952).

Tryckt den 15 december 1965

Uppsala 1965. Almqvist & Wiksells Boktryckeri AB