A note on Lorentz-invariant Markov processes

By R. M. DUDLEY

1. Introduction

This note will remove some unnecessary regularity assumptions from my first paper on the subject [1]. There, beyond the conditions of spatial and temporal homogeneity, Lorentz-invariance, the Markov property, and the requirement that speeds be less than that of light, there were additional assumptions: continuity in probability ((E), p. 248) and existence of derivatives with probability one at each time (first half of (C), p. 248). Both these assumptions will be shown to follow from the others. This makes easier the proof of the converse part of the main theorem of [1], Theorem 8.2, as will be indicated below.

It was assumed in [1] that speeds remain less than (resp. equal to) 1, taken as the speed of light, if the initial speed is less than (resp. equal to) 1. This follows the physical distinction between particles of positive and zero rest mass. In section 3 we show that this assumption is unnecessary for initial speed 1, but that there are processes with initial speeds less than 1, satisfying all our other assumptions, which jump to speed 1 at later times.

2. Removing regularity assumptions

We shall use freely the notation and terminology of [1, sections 2 and 4]. Also, for any function f from the real line to three-space X let

$$\hat{f}(t) = \langle f(t), t \rangle \in M.$$

As in [1], we consider families of measures $\{P_x^v\}$ where for each x in X and v in V, P_x^v is a starting probability at x on $(\mathcal{A}, I, X, B(X))$. We retain assumptions (A) (spatial homogeneity), (B) (initial velocity v), and (D) (Lorentz-invariance) of [1, section 4], and a weakened form of the Markov property (C) as follows.

Each f in $\mathcal{A}$ is differentiable almost everywhere (Lebesgue) and $f'(t) \in V$ whenever it is defined. By [1, Lemma 4.1]

$$\{\langle f, t \rangle : f'(t) \text{ exists}\}$$

is product measurable for $B^0(\mathcal{A}) \times B(R^+)$. Let

$$T(v) = \{t > 0 : P_x^v(f'(t) \text{ exists}) = 1\}.$$

(Clearly $T(v)$ is independent of x.) We know that the complement of $T(v)$ has measure 0 in R^+ for each v in V. Now rather than assuming $T(v) = (0, \infty)$ we simply assume the stationary Markov property wherever it makes sense:

(C') For any v in V, x in X, t in $T(v)$, and A in $B^t(\mathcal{A})$, $P^v_x(f \in A \mid B^0_t(\mathcal{A})) = P^{f'(t)}_{f(t)}(\theta_t A)$ almost everywhere for P^v_x.

Theorem 1. *Let $\{P^v_x\}$ on $\mathcal{A}$ satisfy conditions (A), (B), (C'), and (D). Then they also satisfy (C) and (E) [1, section 4], thus define a process given by an infinitely divisible radial probability on Lobachevsky space [1, Theorem 8.2], and conversely.*

Proof. We first establish continuity in probability at 0 through $T(v)$ (a weakened form of (E)). We put a natural Lorentz-invariant metric on V. Let U be the transformation

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

Then U transforms V one-to-one onto the hyperbolic space $\mathcal{U}$ (three-velocity to four-velocity). On $\mathcal{U}$, there is a natural Lorentz-invariant Riemannian metric ϱ [1, section 7]. Let e be the metric on V defined by

$$e(v, w) = \varrho(U(v), U(w)).$$

For any L in $\mathcal{L}$, $V_L = U^{-1}LU$. Hence e is V_L-invariant. For any v in V, let

$$L_v\langle x, t \rangle = \langle x + vt, t + v \cdot x \rangle / (1 - |v|^2)^{1/2}.$$

Then $L_v \in \mathcal{L}$ and if $L = L_v$, $V_L(0) = v$.

Lemma 1. *For any v and w in V, (a) $\sinh e(v, w) = |v - w| / ((1 - v \cdot w)^2 - |v - w|^2)^{1/2}$. (b) $e(v, w) \geqslant |v - w|$.*

Proof. Let $L = L_{-v}$. Then

$$e(v, w) = e(0, V_L(w)) = e(0, (w - v)/(1 - v \cdot w))$$

$$= \varrho(U(0), \langle (w - v)/(1 - v \cdot w), 1 \rangle / (1 - |v - w|^2/(1 - v \cdot w)^2)^{1/2})$$

and (a) follows [1, section 7].

To prove (b), note that $v \cdot w \geqslant -|v - w|^2/4$. Letting $x = |v - w|$, $y = e(v, w)$, this yields

$$\sinh y \geqslant x/((1 + x^2/4)^2 - x^2)^{1/2} \geqslant x/(1 - x^2/4) = 4x/(4 - x^2),$$

$$y \geqslant \arg \sinh (4x/(4 - x^2)) = \ln (4x/(4 - x^2) + (1 + 16x^2/(4 - x^2)^2)^{1/2})$$

$$= \ln ((4x + (x^4 + 8x^2 + 16)^{1/2})/(4 - x^2)) = \ln ((2 + x)/(2 - x)) = f(x).$$

Now $f(0) = 0$ and $f'(x) \geqslant 1$ for $0 \leqslant x < 2$, so $f(x) \geqslant x$ and $y \geqslant x$, q.e.d.

Lemma 2. *For any $\varepsilon > 0$ there is a $\delta > 0$ such that for all v in V, $0 \leqslant t \leqslant \delta$ and t in $T(v)$,*

$$P_0^v(f : e(f'(t), v) > \varepsilon) < \varepsilon.$$

Proof. We may assume $\varepsilon \leqslant 1$. Take $\alpha > 0$ such that $|v| \leqslant 5\alpha$ implies $e(0, v) \leqslant \varepsilon$. Thus $\alpha < 1/5$. Since

$$P_0^0(f'(0+) = 0) = 1,$$

there is a $\beta > 0$ such that

$$P_0^0(|f(t)| > \alpha t \quad \text{for some} \quad t \leqslant 2\beta) < \varepsilon/2.$$

(The set of all f in $\mathcal{A}$ satisfying the given condition is in $B^0(\mathcal{A})$ since it suffices to consider rational t.) Let

$$K = \{\langle x, t\rangle : |x| \leqslant \alpha t \leqslant 2\alpha\beta\}, \quad D = \{\langle x, 2\beta\rangle \in K\}.$$

The idea of the rest of the proof is that a large derivative for small t would cause departure from the cone K too soon.

We have for any w in V

$$1 - \varepsilon/2 \leqslant P_0^0(f : \hat{f}(t) \text{ first leaves } K \text{ through } D)$$

$$\leqslant P_0^w(g : \hat{g}(s) \text{ first leaves } L_w(K) \text{ through } L_w(D))$$

$$\leqslant P_0^w(g : \hat{g}(s) \in L_w(K) \quad \text{for all} \quad s \leqslant \beta),$$

since if $L = L_w$ and $\langle x, t\rangle \in D$, $L_{(2)}\langle x, t\rangle \geqslant \beta$.

Now we let $\delta = \beta/2$. Given v in V, let $L = L_v$, $t \in T(v)$, and $t \leqslant \delta$. Then

$$P_0^v(e(g'(t), v) \geqslant \varepsilon) = P_0^0(e(f'(\sigma), 0) \geqslant \varepsilon),$$

where $\sigma = \sigma(f, t) = (L^{-1})_{(2)}(L^* f)\hat{}(t)$, i.e.

$L\hat{f}(\sigma) = (L^* f)\hat{}(t)$, $\sigma + v \cdot \hat{f}(\sigma) = t(1 - |v|^2)^{1/2}$. Thus if $\hat{f}(\sigma) \in K$, $\sigma \leqslant 2t \leqslant \beta$. Now

$$1 - \varepsilon/2 \leqslant \int_{\hat{g}(t) \in L(K)} P_{\hat{g}(t)}^{g'(t)}(h : \langle h(s), s + t\rangle \text{ first leaves } L(K) \text{ through } L(D)) \, dP_0^v(g)$$

$$= \int_{\hat{g}(t) \in L(K)} P_0^{g'(t)}(h : \hat{h}(s) + \hat{g}(t) \text{ first leaves } L(K) \text{ through } L(D)) \, dP_0^v(g)$$

$$= \int_{\hat{f}(\sigma) \in K} P_0^{f'(\sigma)}(j : \hat{j}(r) + \hat{f}(\sigma) \text{ first leaves } K \text{ through } D) \, dP_0^0(f)$$

$$\leqslant \int_{\hat{f}(\sigma) \in K} P_0^{f'(\sigma)}(j : \hat{j}(\beta) + \hat{f}(\sigma) \in K) \, dP_0^0(f).$$

Now let $w = w(f) = f'(\sigma)$. If $\hat{f}(\sigma) \in K$, $|w| > 5\alpha$, and $\hat{j}(\beta) \in L_w(K)$, then $|j(\beta) - w\beta| \leqslant \alpha(\beta - w \cdot \hat{j}(\beta)) \leqslant 2\alpha\beta$,

$$|\hat{j}(\beta)| > 3\alpha\beta, |\hat{j}(\beta) + \hat{f}(\sigma)| > 3\alpha\beta - \alpha\beta = 2\alpha\beta, \hat{j}(\beta) + \hat{f}(\sigma) \notin K.$$

Thus $\hat{f}(\sigma) \in K$ and $|w| > 5\alpha$ imply

$$P_0^w(j : \hat{j}(\beta) + \hat{f}(\sigma) \in K) \leqslant P_0^w(j : \hat{j}(\beta) \notin L_w(K)) \leqslant \varepsilon/2 \leqslant 1/2.$$

Thus

$$P_0^0(e(f'(\sigma), 0) > \varepsilon) \leqslant P_0^0(|f'(\sigma)| > 5\alpha) \leqslant \varepsilon, \text{ q.e.d.}$$

Next, let $v \in V$ and let W be a countable dense subset of $T(v) \cup \{0\}$ with $0 \in W$. Given $\varepsilon > 0$ and $\delta > 0$ let

$$\alpha(\varepsilon, \delta) = \sup \left[P_0^w(e(f'(t), w) \geqslant \varepsilon : w \in V, t \in T(w), 0 \leqslant t \leqslant \delta \right].$$

For each positive integer k and $a \geqslant 0$ let
$A(k, a, \varepsilon, \delta) = \{ f \in \mathcal{A} : \text{ for some } s_i \text{ in } W \text{ with } a \leqslant s_1 < s_2 < \ldots < s_{2k} \leqslant a + \delta, e(f'(s_{2i-1}),$
$f'(s_{2i})) \geqslant 4\varepsilon \text{ for all } i = l, \ldots, k \}$.

Lemma 3. *For any* $\varepsilon > 0, \delta > 0$, *positive integer* $k, a \geqslant 0, x \in X$, *and* $v \in V, P_x^v(A(k, a, \varepsilon, \delta)) \leqslant [2\alpha(\varepsilon, \delta)]^k$.

The statement and proof of Lemma 3 are essentially those of Lemma 6.4 of Dynkin [2]. We make only the following remarks: application of the strong Markov property is permissible since W is countable, as mentioned at the end of section 5.8 of [2]. We need only the Markov property at each t in W. We cannot formally apply the statement of [2, Lemma 6.4] since we have the Markov property only on a set of times $T(v)$, depending on the initial v. Our assumptions imply that for any $s \leqslant t$ in $T(v), t - s \in T(f'(s))$ for P_x^v-almost all f in $\mathcal{A}$.

Lemma 2 implies that for every $\varepsilon > 0, \lim_{\delta \to 0} \alpha(\varepsilon, \delta) = 0$. Thus for any $M > 0$, we can take k large enough so that $\alpha(\varepsilon, M/k) < 1/2$ and obtain from Lemma 3

$$\lim_{n \to \infty} P_x^v(A(n, 0, \varepsilon, M)) \leqslant \lim_{n \to \infty} kP_x^v(A([(n-k)/k], 0, \varepsilon, M/k)) = 0,$$

where $[x]$ is the largest integer $\leqslant x$.

Let $G = \bigcup_{m=1}^\infty \bigcap_{n=1}^\infty A(n, 0, 1/m, m)$. Then $P_x^v(G) = 0$. For $f \notin G$, the limits

$$w(t) = \lim_{s \downarrow t, \, s \in W} f'(s), \, t \geqslant 0, \text{ and}$$

$$w(t^-) = \lim_{s \uparrow t, \, s \in W} f'(s), \, t > 0$$

always exist, w is continuous from the right, and for all $t > 0$

$$\lim_{s \uparrow t} w(s) = w(t^-).$$

For each fixed t in $T(v), w(s) \to f'(t)$ in P_x^v-probability as $s \downarrow t$ through W, by the Markov property at t and Lemma 2. Thus

$$P_x^v(w(t) = f'(t)) = 1.$$

Now since the maps $\langle f, t \rangle \to f'(t)$ and $\langle f, t \rangle \to w(t)$ are both measurable [1, Lemmas 4.1 and 2.1], we have $P_x^v(f'(t) = w(t)$ for (Lebesgue) almost all $t \geqslant 0) = 1$. Since f is Lipschitzian, $f' = w$ almost everywhere implies that f is the indefinite integral of w. Then by the left and right limit properties of w,

$$P_x^v(f'(t^+) = w(t) \quad \text{for all } t \geqslant 0)$$

$$= P_x^v(f'(t^-) = w(t-) \quad \text{for all } t > 0) = 1.$$

If $t \notin T(v)$, we take $s_n \uparrow t$ and $t_n \downarrow t, s_n, t_n \in T(v)$. For any $\varepsilon > 0$, $\lim_{n \to \infty} P_x^v (e(f'(s_n),$ $f'(t_n)) > \varepsilon) = 0$, by Lemma 2 and the Markov property at s_n. Thus

$$P_x^v(f'(t^+) \neq f'(t^-)) = 0,$$

and in fact $T(v) = (0, \infty)$, i.e. assumption (C) holds. We have

$$P_x^v(\lim_{t \downarrow 0} |f(t) - x| = 0) = 1$$

since the functions in $\mathcal{A}$ are Lipschitzian. Thus assumption (E) holds by Lemmas 1(b) and 2. We have now obtained the full assumptions (A) through (E) of [1], and can draw all the inferences of [1, sections 4–8]. (In particular, we have by [1, Lemma 5.3] that for any $t \geq 0$ and A in $B(X \times V)$,

$$\langle x, v \rangle \to P_x^v(\langle f(t), f'(t) \rangle \in A),$$

is $B(X \times V)$-measurable. Such a condition is made part of the definition of Markov process by Dynkin [2], but not by all authors in the field. It is not needed for the earlier results in this paper (even for Lemma 3), since the weaker measurability implied by (C′) suffices.)

To finish the proof of Theorem 1 we now sketch a proof of the converse part of Theorem 8.2 of [1], which was not proved there. Let P_1 be an infinitely divisible, radial probability measure on the Lobachevsky space $\mathcal{U}$. Then P_1 can be imbedded in a convolution semigroup $\{P_\tau, \tau \geq 0\}$ of such measures, as is obvious from the "Lévy–Khinchin" representation formula [1, section 8]. Such a semigroup is always weak-star continuous, hence defines a strong Markov process $\{x_\tau, \tau \geq 0\}$ having sample functions continuous from the right with left limits everywhere (see the remarks before [1, Lemma 8.1]). We take an "indefinite integral" of the process, obtaining paths in M. Then we can invert the arguments of [1, section 6], making a "random change of time" from proper time τ to co-ordinate time t. We thus obtain a process which satisfies (A), (B), (C′), and (D), as desired.

3. Mixing speeds less than 1 and equal to 1

Let $\mathcal{J}$ be the set of functions from R^+ to X satisfying

$$|f(s) - f(t)| \leq |s - t| \quad \text{for all } s, t \geq 0.$$

Each f in $\mathcal{J}$ is differentiable almost everywhere and $|f'(t)| \leq 1$ whenever $f'(t)$ is defined.

Theorem 2. *Suppose $\{P_x^v\}$ for x in X and v in S^2 (i.e. $v \in X, |v| = 1$) is a family of probability measures on $\mathcal{J}$ satisfying assumptions* (A), (B), *and* (D). *Then each P_x^v is concentrated in the one function $f(t) = x + vt, t \geq 0$.*

Proof. The proof of [1, Theorem 11.1] is sufficient since it only uses the fact that initial speeds are 1, not that speeds remain equal to 1.

Next, we describe informally a class of invariant Markov processes for which transitions *from* speeds less than 1 *to* speeds equal to 1 *do* occur.

Suppose $\{P_x^v\}$ satisfy conditions (A), (B), (C'), and (D) on $\mathcal{A}$, as in section 2. We construct a new process on $\mathcal{J}$ as follows. Let τ_0 be a random variable with a distribution given, for some $k > 0$, by

$$P(\tau_0 > c) = e^{-kc} \quad \text{for all} \quad c \geqslant 0.$$

We assume τ_0 is independent of the P_x^v. Now let a particle move according to P_x^v until proper time τ_0 has passed, say at time t. From then on, let it move in a straight line with velocity w, where w is uniformly distributed over the sphere $\{w : |w| = 1\}$ in a co-ordinate system where the particle has zero velocity at time t^- (and w is independent of τ_0 and other events before time t). This defines a set $\{R_x^v\}$ of probability measures on $\mathcal{J}$ (using the usual deterministic ones for $|v| = 1$) which have all the properties (A) through (E) of the P_x^v without being concentrated in $\mathcal{A}$ for $|v| < 1$.

4. Diffusion processes

Here we make some additions and corrections to [1, section 10]. First, the definition of diffusion semigroup given there is too strong. It should read: for every $\varepsilon > 0$,

$$\lim_{t \downarrow 0} P_t(\varrho \geqslant \varepsilon)/t = 0.$$

The phase space $M \times \mathcal{U}$ is acted on transitively by the Poincaré group (inhomogeneous Lorentz group) $\overline{\mathcal{L}}$. I thank C. C. Moore for the following construction of $\overline{\mathcal{L}}$-invariant Riemannian metrics. Take a tangent space V to $M \times \mathcal{U}$ at one point, say $(0, p)$. Then V is a seven-dimensional real vector space $R^3 \times R \times R^3$. It suffices to find the positive definite quadratic froms B on V such that

$$B(\langle x, t, y \rangle) = B(\langle A(x), t, A(y) \rangle)$$

for all orthogonal transformations A of R^3 with determinant 1 (the latter restriction is also needed in the first paragraph of [1, section 7]). B satisfies the above conditions if and only if

$$B(\langle x, t, y \rangle) = a_1 t^2 + a_2 |x|^2 + 2a_3 x \cdot y + a_4 |y|^2,$$

where
$$a_1 > 0, a_2 > 0, a_3^2 < a_2 a_4.$$

The definition of diffusion process depends only superficially on the metric. For a homogeneous Riemannian manifold, as here, it clearly depends only on the original topology of the manifold. On a general differentiable manifold, Dynkin [3] defines a diffusion process as one whose generator is a second-order differential operator.

In any case, the diffusion processes in $M \times \mathcal{U}$ which are also Lorentz-invariant Markov processes as in Theorem 1 above are those defined by "Brownian motions" in $\mathcal{U}$ [1, section 10].

5. Note

In [1, section 5], use was made of the relations between weak-star convergence of probability measures and convergence in the dual space of the bounded Lipschitzian functions. Details of these relations are given in a paper of mine, "Convergence of Baire measures", to appear in *Studia Mathematica*.

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

REFERENCES

1. DUDLEY, R. M., Lorentz-invariant Markov processes in relativistic phase space, *Arkiv för Matematik 6*, 241–268 (1965).
2. DYNKIN, E. B., Theory of Markov Processes (Moscow, 1959; English translation, London, Pergamon, 1960).
3. DYNKIN, E. B., Markov Processes (Moscow, 1963; English translation, Berlin, Springer, 1965).