FRANK B. KNIGHT

On the Ray topology

Séminaire de probabilités (Strasbourg), tome 18 (1984), p. 56-69

<http://www.numdam.org/item?id=SPS_1984__18__56_0>

On the Ray Topology

F. B. Knight

0. Introduction: In a short note [2, (1965)] the author introduced a method of topologizing a general measurable space $(E,\mathcal{E})$ by using a Markov resolvent family on the space. This method was specialized by H. Kunita and T. Watanabe [4, (1967)] who connected it in a certain way with a given locally compact Hausdorff topology on E. A similar route was followed by J. B. Walsh and P. A. Meyer [6, (1971)], who applied the method to study left limits of processes satisfying certain "hypotheses droites". This work was extended and given a definitive form in the well-known booklet of R. K. Getoor [1]. Since that time, the same method has been used in several papers, often in connection with a reversal of time (Walsh, Glover, and others).

It seems to the present author that, despite its success in achieving results, this connection of the method of [2] with a given topology on E has not been very well justified. It often appears to us as an ad hoc device to obtain a connection between the new Ray topology and the original topology, in such a way that E will be measurable in its R.-K. compactification $\bar{E}$. The object of the present work is to make a deeper investigation of this device, in terms of the author's general construction called the "prediction process" [3, Essay I], thus obtaining a new characterization of the Ray topology.

The characterization which we will obtain is not limited to right processes. In fact, one can show that it applies to the prediction process itself. Hence it automatically applies to any Markov process which identifies (up to equivalence) with its own prediction process by the mapping $x \leftrightarrow P^x$.

However, this is an unfamiliar setting, and would require more abstraction without, perhaps, adding much to the ideas involved. Since we are concerned here with basic ideas rather than with new results, for brevity we only study the case of Borel right processes. It should be stated also that we do not know whether our characterization applies to non-Borel right processes.

The present work was first developed during the author's visit to the University of Strasbourg during October - December, 1982. He is greatly indebted to Professor P. A. Meyer for that visit.

1. The Prediction Process of a Borel Right Processes. Let $(E,\mathcal{E})$ be a Lusin space with its Borel σ-field (E is homeomorphic to a Borel subset of a compact metric space $\hat{E}$). For our definition of a Borel right process, we will follow [1, §9].

Definition 1.1. Let P_t, $t \geq 0$, be a Markov semigroup on $(E,\mathcal{E})$, where $P_t f(x)$ is $\mathcal{E}$-measurable in x for $f \in b(\mathcal{E})$. Let $(\Omega,\mathcal{F}_t^o)$ be the space of all right-continuous paths $w(s) : R^+ \to E$ with the coordinate σ-fields $(\mathcal{F}^o = \bigvee_{t\geq 0} \mathcal{F}_t^o)$, and identify $X_t = w(t)$. Then X_t is a Borel right process with semigroup P_t if

a) For each probability μ on $(E,\mathcal{E})$, there is a (unique) P^μ on $(\Omega,\mathcal{F}^o)$ such that X_t is a Markov process with μ as initial distribution ($\mu(A) = P^\mu\{X_0 \in A\}$, $A \in \mathcal{E}$) and P_t as transition function, and

b) Whenever f is α-excessive $(\alpha > 0)$ for the resolvent R_λ of P_t, and μ is as in a), then $f(X_t)$ is P^μ-a.s. right-continuous in $t \geq 0$.

From the standpoint of the present paper, hypothesis a) is largely superfluous. It is needed only to the extent that it guarantees $P^x\{X_0 = x\} = 1$ and provides a sequence g_n of bounded, α-excessive functions separating points of E. Indeed, if f_n are uniformly dense in $C^+(\hat{E})$, since $\lim_{\alpha\to\infty} \alpha R_\alpha f_n = f_n$ and the range of R_α on $b^+(\mathcal{E})$ is free of α, one sees that for each fixed α, $g_n = R_\alpha f_n$ is such a sequence. Once we have such

g_n $(\in \mathcal{E})$ it is easy to see by Lusin's Theorem that $\mathcal{E}$ is also the Borel σ-field of E with respect to the coarsest topology on E making all g_n continuous, and E is a Lusin space in this topology. Then b) and Theorem (9.4), (i), of [1] imply that there is a (unique) Borel right process with semigroup P_t when $(E, \mathcal{E})$ is considered in the topology generated by (g_n). From that point on, we can discard the original $(\Omega, \mathcal{F}^o)$ and assume that X_t is the latter Borel right process. Here it should be noted that only a) involves the given topology of E, since the definition of excessive function is purely measure-theoretic.

This replacement of X_t has an immediate dividend. Since $g_n(X_t)$ is an α-super-martingale for each P^μ, it follows from right-continuity that there also exist for all $t > 0$ the left limits $\lim_{s \to t-} g_n(X_s)$, P^μ-a.s. Therefore, X_t has left limits in the topology generated by (g_n), P^μ-a.s. In defining our Borel right process we can therefore replace $(\Omega, \mathcal{F}^o_t)$ by the space $(\hat{\Omega}, \mathcal{F}^o_t)$ of all right-continuous with left limits paths $w(s)$: $R^+ \to E$, with the new topology of E and the new coordinate σ-fields. This $\hat{\Omega}$ is well-known to be itself a Lusin space, with Borel σ-field $\mathcal{F}^o$ (unlike the original Ω, which is only the complement of an analytic set). Since we are concerned with topologies on E, it will be convenient to assume this choice for X_t in setting up the prediction process. But it must be kept in mind that our object is to understand the connection between the originally given topology of E and a certain other topology which depends both on the given topology and on P_t (it will be defined immediately below). Thus the original E-topology cannot be dispensed with entirely.

We now review briefly how the Ray topology and the R.-K. compactification are defined, following [1, §10]. One first constructs

Definition 1.2. (a). The least convex cone, denoted by C_∞, containing $\{R_\lambda f;\ \lambda > 0,\ f \in C^+(\hat{E})\}$, and closed under the two operations

i) application of R_λ, any $\lambda > 0$, and

ii) formation of minima $f_1 \wedge f_2$.

This is done inductively, starting with

$$C_1 = \{R_{\lambda_1} f_1 + \cdots + R_{\lambda_n} f_n;\ \forall\, n,\ \lambda_k > 0,\ f_k \in C^+(\hat{E})\},$$

$$C_2 = \{f_1 \wedge f_2 \wedge \ldots \wedge f_n; \forall n,\ f_k \in C_1\},$$

$$C_{2n+1} = C_{2n} + \{R_{\lambda_1} f_1 + \cdots + R_{\lambda_n} f_n; \forall n,\ \lambda_k > 0,\ f_k \in C_{2n}\},$$

$$C_{2(n+1)} = \{f_1 \wedge f_2 \wedge \ldots \wedge f_n;\ \forall n, f_k \in C_{2n+1}\}.$$

Then we have $C_\infty = \bigcup_n C_n$. Next,

<u>Definition 1.2. (b)</u>. The Ray topology of E is the coarsest topology making all elements of C_∞ continuous.

It should be emphasized that the given topology of E is used only in defining C_1, in the sense that $R_\lambda f$ must be Ray-continuous on E for $f \in C^+(\hat{E})$. Obviously, the topology on E induced by C_1 is at least as fine as the topology used in the above construction of $\hat{\Omega}$.

A basic lemma [2] now asserts that there is a countable dense subset of C_∞ in the uniform norm on E. It follows that the topology of E induced by C_∞ is metrizable by a bounded metric $d(x,y) = \sum_n c_n |d_n(x) - d_n(y)|$, where $\{d_n(x)\}$ is such a countable dense set and $\sum_n c_n \max |d_n| < \infty$, $c_n > 0$.

<u>Definition 1.3</u>. The compactification $\overline{E}$ of E for this metric does not depend on choice of the c_n or d_n. It is called the R.-K. compactification of E, relative to $\hat{E}$.

Let us recall the reason for introducing this construction. We can define a new family $\overline{R}_\lambda$ on $C(\overline{E})$, using the facts that if $\overline{f} \in C(\overline{E})$, and if f denotes $(\overline{f}|E)$, then f is in the uniform closure of $C_\infty - C_\infty$, and moreover clearly $R_\lambda(C_\infty - C_\infty) \subset C_\infty - C_\infty$. Thus we can define $\overline{R}_\lambda \overline{f} = \overline{R_\lambda f}$, $\overline{f} \in C(\overline{E})$, and clearly $\overline{R}_\lambda$ maps $C(\overline{E}) \to C(\overline{E})$. Now it is shown that $\overline{R}_\lambda$ is a Ray resolvent on $(\overline{E}, \overline{\mathcal{E}})$, so by a theorem of Ray [5] there is a unique right-continuous Markov semigroup $\overline{P}_t$ on $C(\overline{E})$ having $\overline{R}_\lambda$ as its resolvent. Moreover, it is shown that for $x \in E$ we have $\overline{P}_t \overline{f}(x) = P_t f(x)$, $\overline{f} \in C(\overline{E})$. It follows that $E \in \overline{\mathcal{E}}$ and that, for initial distribution μ concentrated on E, the Ray process with semigroup $\overline{P}_t$ is indistinguishable from X_t restricted to the subset of $\hat{\Omega}$ of paths which are r.c.l.l. ("right-continuous, with left limits for $t > 0$") in the Ray topology, with the trace of P^{μ} on this subset. Indeed, since C_{∞} contains only λ-excessive functions, X_t is already r.c.l.l. in the Ray topology except on a P^{μ}-null set for each μ. Hence one can consider X_t as a Ray process in the Ray topology. We shall not attempt to analyze the advantages of a Ray process over a right process, but merely remark again that experience has shown the above construction to be quite powerful.

The emphasis here is on the role of the given topology of E, and the above description unfortunately does little to make this clear. Indeed, in so far as the above conclusions are concerned, in defining C_1 one might just as well use any other convex cone $\{R_{\lambda_1} f_1 + \cdots + R_{\lambda_n} f_n\}$ with $f_k \in b^+(\mathcal{E})$ which separates points and has a countable uniformly dense subset. In particular, any other Lusin topology of $(E, \mathcal{E})$ would provide such a cone. To elucidate the role of the topology of E, it is useful to take a more general point of view based on the prediction process of $(\hat{\Omega}, \mathcal{F}^o_t)$. At least initially, this is a purely measure-theoretic construction. We begin with <u>any</u> probability P on $(\hat{\Omega}, \mathcal{F}^o)$. Letting $(M, \mathcal{M})$ denote the Lusin space of all such P ($\mathcal{M}$ is generated by $\{P(S);\ S \in \mathcal{F}^o\}$), we can construct two $(M, \mathcal{M})$-valued processes $Z^P_{t\pm} = Z^P_{t\pm}(S, w)$, unique up to a P-null set, such that

a) Z^P_{t+} (resp. Z^P_{t-}) is an $\mathcal{F}^o_{t+}$-optional (resp. $\mathcal{F}^P_{t+}$-previsible) process, and

b) for each t, $P(\theta_t^{-1} S | \mathcal{F}^o_{t+}) = Z^P_t(S)$, $S \in \mathcal{F}^o$,

(resp. $P(\theta_t^{-1} S | \mathcal{F}^P_{t-}) = Z^P_{t-}(S)$, $S \in \mathcal{F}^o$, $t > 0$) in the usual sense of conditional probability. It is shown [3, Essay I] that all the processes Z^P_{t+}, $P \in M$, are homogeneous strong Markov processes relative to $\mathcal{F}^o_{t+}$, all the Z^P_{t-} are homogeneous moderately Markov processes relative to $\mathcal{F}^P_{t-}$, and all the $Z^P_{t\pm}$

have a single Borel transition function $q(t,z,A)$ on $(M,\mathcal{M})$. Further, for $\mathcal{F}^o_{t+}$-optional $T < \infty$ one has $P(\theta_T^{-1}S|\mathcal{F}^o_{T+}) = Z^P_{T+}(S)$, $S \in \mathcal{F}^o$, and for $\mathcal{F}^P_{t+}$-previsible $0 < T < \infty$ one has $P(\theta_T^{-1}S|\mathcal{F}^P_{T-}) = Z^P_{T-}(S)$, $S \in \mathcal{F}^o$.

We emphasize again that this involves neither any topology on E, nor any Markov property of P. We can define a single pair $Z_{t\pm}$ of coordinate processes, on two copies of the canonical space of $\mathcal{B}^+|\mathcal{M}$-measurable paths, in such a way that Z_{t+} is a strong Markov process with state space $(M,\mathcal{M})$ and transition function q relative to the coordinate σ-fields $\mathcal{F}^Z_{t+}$, while Z_{t-} is a moderately Markov process with state space $(M,\mathcal{M})$ and transition function q, relative to the coordinate σ-fields $\mathcal{F}^Z_t$ (since the coordinate Z_{t-} is not $\mathcal{F}^Z_{t-}$-measurable, we cannot use $\mathcal{F}^Z_{t-}$ here; nevertheless, for each initial value P, Z_{t-} is in the P-completion of $\mathcal{F}^Z_{t-}$). However, this separation of Z_{t+} and Z_{t-} appears unnatural from the standpoint of Markov processes. It is natural to demand that $(M,\mathcal{M})$ be topologized in such a way that Z^P_{t+} is r.c.l.l. for each P (where r.c.l.l. now denotes right-continuous with left limits, P-a.s.) and that the left limit process then have (for each initial value P) the same probability law as the corresponding Z^P_{t-}.

There are, of course, many different Lusin topologies on $(M,\mathcal{M})$ which satisfy the above requirement. The object here is to introduce one which corresponds to the given topology of E, in such a way that we can understand the Ray topology of E as a consequence of a topology on $(M,\mathcal{M})$. The connection of E and M is to be given by the natural mapping of

<u>Definition 1.4</u>. For a given Borel right process X_t, let $\varphi(x) : E \to M$, denote the mapping $\varphi(x) = P^x$, when we define $\hat{\Omega}$ as above.

<u>Remark</u>. We use a fixed choice of $\hat{E} \supset E$, a fixed $\alpha > 0$, and a fixed sequence $g_n = R_\alpha f_n$, $\{f_n\}$ dense in $C^+(\hat{E})$. Obviously this does not depend on the particular selection of f_n, but it may depend on α as well as $\hat{E}$. This dependence is of no real importance below, and will be suppressed.

We observe next that, for any topology on M which renders Z^P_{t+} right-continuous with left limits Z^P_{t-}, P-a.s. for $P \in M$, it will follow that X_t is also right-continuous in the topology of E induced by φ from this topology of M. Indeed, we have easily

Theorem 1.5. For any initial distribution μ on $(E,\mathcal{E})$, one has

$$P^\mu\{\varphi(X_t) = Z^{P^\mu}_t \text{ for all } t \geq 0\} = 1.$$

Proof. It will suffice to indicate the proof . Both sides of $\varphi(X_t) = Z^{P^\mu}_t$ are $\mathcal{F}^o_{t+}$-optional, hence by the optional section theorem it suffices to show that, for $\mathcal{F}^o_{t+}$-optional $T < \infty$, $P^\mu\{\varphi(X_T) = Z^{P^\mu}_T\} = 1$. But for each $S \in \mathcal{F}^o$, the strong Markov property of X_t gives $P^{X_T}(S) = P^\mu(\theta^{-1}_T S|\mathcal{F}^o_{T+}) = Z^{P^\mu}_T(S)$, P^μ-a.s., and since $\mathcal{F}^o$ is countably generated this completes the proof.

2. Characterization of the Ray topology. Simple as it is, Theorem 1.5 contains the key idea behind our approach to the Ray topology. Namely, we will regard X_t as synonymous with P^{X_t}, in such a way that any topology on $(M,\mathcal{M})$ induces a corresponding topology for X_t on $(E,\mathcal{E})$. The Ray topology tends, in some sense, to make points x and y near when the measures P^x and P^y are near, so to make this precise we have to obtain the appropriate topology of measures. A topology of measures on $(\hat{\Omega},\mathcal{F}^o)$, however, generally presupposes some topology on $(\hat{\Omega},\mathcal{F}^o)$. Consequently, we must attend to this first.

Definition 2.1. The prediction topology on $(\hat{\Omega},\mathcal{F}^o)$, relative to the given topology of E, is the topology generated by the functions

$\int_0^t f(X_s)ds$, $f \in C(\hat{E})$, $0 < t$.

Remark. Although this depends formally on choice of $\hat{E}$, this dependence is really nil. To see this, set $\mu(t,B) = \int_0^t I_B(X_s)ds$, $B \in \mathcal{E}$, $t > 0$, as function of $w \in \hat{\Omega}$. Thus, $\mu(t,B)$ is the sojourn kernel of w.

Now $\int_0^t f(X_s)ds = \int_{\hat{\Omega}} f(b)\mu(t,db)$, from which we see that the prediction topology of $\hat{\Omega}$ is simply the topology of weak convergence of sojourn measures on $(E,\mathcal{E})$ for each rational r (and hence, by uniform continuity, for each t). It is well-known (see [1, Proposition (14.7)]) that this is defined independently of $\hat{E}$, and indeed implies convergence of $\int f(b)\mu(r,db)$ for all bounded continuous functions f on E. This is, in our approach, the reason that the Ray topology does not depend on $\hat{E}$. On the other hand, since $C(\hat{E})$ is separable, it is easy to see from Lusin's Theorem that $\hat{\Omega}$ is a Lusin space in the prediction topology, with Borel field ${}^{r}\mathcal{F}^{o}$.

Definition 2.2. The prediction topology of $(M,\mathcal{M})$ is the topology of vague (weak-*) convergence of measures with respect to the prediciton topology of $(\hat{\Omega},{}^{r}\mathcal{F}^{o})$.

As noted above, this is well-defined independently of any compactification of M. It suffices for $z_n \to z \in M$ that $E^{z_n}f \to E^{z}f$ for $f \in C(\bar{\hat{\Omega}})$, for any metric compactification $\bar{\hat{\Omega}} \supset \hat{\Omega}$. It is also clear that the Borel field of M in this (Lusin) topology is again $\mathcal{M}$. Indeed, for any $f \in b({}^{r}\mathcal{F}^{o})$, $E^{z}f$ is Borel measurable. Then it is also clear that the topology induced on E by φ from the prediction topology of $(M,\mathcal{M})$ has Borel field $\mathcal{E}$.

This brings us to an identification of the above induced topology.

Theorem 2.3. The topology induced by φ from the prediction topology of M is the topology generated by the cone C_1 of Definition 1.2, or again, by $\{R_\lambda f,\ \lambda > 0,\ f$ bounded and continuous on $E\}$. Thus, it is coarser than the Ray topology, but finer than the topology used in defining $\hat{\Omega}$. In particular, since X_t is Ray-r.c.l.l., X_t is a Borel right process in the topology induced by φ, and $Z_t^{\varphi(x)}$ is r.c.l.l. for $x \in E$.

Note. It can be shown that Z_t^P is r.c.l.l. (in the prediction topology) for any $P \in M$, but we omit the argument.

Proof. The equivalence of the two topologies in the first sentence follows immediately from the topological fact already cited [1, (14.7)], so we need only prove the first assertion. Let $f \in C^+(\hat{E})$. Then for $\lambda > 0$ we have, uniformly on $\hat{\Omega}$,

$$\int_0^\infty e^{-\lambda s} f(X_s)ds = \lim_{N\to\infty} \lambda \int_0^N e^{-\lambda t}\left(\int_0^t f(X_s)ds\right)dt$$

$$= \lim_{N\to\infty} \lambda/N \sum_{k=1}^{N^2} e^{-(\lambda\, k/N)}\left(\int_0^{k/N} f(X_s)ds\right).$$

Thus the left side is a bounded continuous function on $\hat{\Omega}$, and hence $E^z \int_0^\infty e^{-\lambda s} f(X_s)ds$ is bounded and continuous on M. Consequently, in the topology generated by φ, $R_\lambda f(x)$ is continuous, and hence so are all the functions in C_1.

Conversely, suppose that a topology on E makes these functions continuous, and let us show it must be finer than the topology generated by φ. Let f_n, $1 \leq n$, be uniformly dense in $C^+(\hat{E})$, and let $\bar{\hat{\Omega}}$ be the compactification of $\hat{\Omega}$ with respect to the family $\int_0^r f_n(X_s)ds$, $0 < r$ rational, $1 \leq n$. Thus $\bar{\hat{\Omega}}$ is a compact metric space with $\hat{\Omega}$ as Borel subset. From the previous analysis, each term $\int_0^\infty e^{-\lambda s} f(X_s)ds$ extends to a continuous function on $\bar{\hat{\Omega}}$, for which we retain the same notation. It is clear that $\int_0^\infty e^{-\lambda s} f(X_s)ds = \lambda\int_0^\infty e^{-\lambda t}\left(\int_0^t f(X_s)ds\right)dt$ remains valid on $\bar{\hat{\Omega}}$, in terms of the extension of $\int_0^t f(X_s)ds$ to $\bar{\hat{\Omega}}$ (which remains uniformly continuous in t, uniformly on $\bar{\hat{\Omega}}$). Consequently, by the inversion theorem for Laplace transforms, we see that the terms $\int_0^\infty e^{-\lambda s} f_n(X_s)ds$ separate points of $\bar{\hat{\Omega}}$ along with $\left\{\int_0^r f_n(X_s)ds\right\}$, and it follows by the Stone-Weierstrass Theorem that finite linear combinations of terms of the form $\prod_{k=1}^m \left(\int_0^\infty e^{-\lambda_k s} f_{n_k}(X_s)ds\right)$ are uniformly dense on $\bar{\hat{\Omega}}$. Hence it suffices to show that its expetations is continuous in E. We proceed by introduction on m, the case $m = 1$ being true by hypothesis. Writing f_k for f_{n_k}, we can write it as a sum of m! similar terms obtained by permuting the functions $e^{-\lambda_k s} f_k(x)$ and integrating over $\{s_1 < s_2 < \cdots < s_m\}$. Denoting the first term by T_m, we have

$$E^xT_m = E^x\left[\int_0^\infty e^{-\lambda_1 s_1} f_1(X_{s_1}) \int_{s_1}^\infty e^{-\lambda_2 s_2} f_2(X_{s_2}) \cdots \int_{s_{m-1}}^\infty e^{-\lambda_m s_m} f_m(X_{s_m})\,ds_m \cdots ds_1\right]$$

$$= E^x\left[\int_0^\infty e^{-\lambda_1 s} f_1(X_{s_1}) E^{X_{s_1}}\left(\int_0^\infty e^{-\lambda_2(s_1+s_2)} f_2(X_{s_2}) \int_{s_2}^\infty e^{-\lambda_3(s_1+s_3)} f_3(X_{s_3}) \cdots \int_{s_{m-1}}^\infty e^{-\lambda_m(s_1+s_m)} f_m(X_{s_m})\,ds_m \cdots ds_2\right) ds_1\right]$$

$$= mE^x\left[\int_0^\infty e^{-(\lambda_1+\cdots+\lambda_m)s_1} f_1(X_{s_1}) E^{X_{s_1}}(T_{m-1})\, ds_1\right.$$

The last expectation has the form E^xT_{m-1}, so by induction it is bounded and continuous on E, along with $f_1(x)$. Hence by the original hypothesis (since it implies continuity of $R_\lambda f$ for all bounded, continuous f on E), E^xT_m is also bounded and continuous. This completes the induction step. The remaining assertions of Theorem 2.3 follow directly (but we must except a negligible set in asserting that X_t is a right process in the new topology; this can be discarded from $\hat{\Omega}$ if desired).

Now we are ready to turn to the Ray topology. We first observe that, by Theorems 1.5 a) and 2.3, there is a Borel right process with transition function $q(t,z,A)$ on $\varphi(E)$, namely the image of X_t under φ, which has the same law as $Z_t^{P^\mu}$ for each initial distribution $\varphi(\mu)$. Consequently, we can apply Theorem 2.3 to this new right process. In fact, since $Z_t^{P^\mu}$ is already r.c.l.l., we can take the new $\hat{\Omega}(= \hat{\Omega}_M)$ to consist of the r.c.l.l. paths with values in M (for the prediction topology of M, relative to the given topology of E). We will label the new mapping φ as φ_2 and the previous φ as φ_1 for the sake of clarity. We next show

Lemma 2.4. The topology on E generated by $\varphi_2\varphi_1$ is finer than that generated by φ_1, but coarser than the Ray topology.

Proof. We prove the second assertion first because it is easier. Indeed, we know from Theorem 2.3 that the topology generated by $\varphi_2\varphi_1$ on E is coarser than the Ray topology relative to the topology on E generated by φ_1. But since the topology generated by φ_1 is coarser than the Ray topology of E relative to its original topology, it is clear that the Ray topology relative to the topology generated by φ_1 is coarser than the Ray topology relative to the Ray topology itself. Since the resolvent R_λ already maps $C(\overline{E}) \to C(\overline{E})$ for the Ray topology, however, it is clear that the second Ray topology (relative to the Ray topology itself) is coarser than the first Ray topology. Combining these three (not necessarily strict) inequalities yields the second assertion.

As to the first, by Theorem 2.3 the topology induced by φ_2 on $\varphi_1(E)$ is generated by the functions $R^Z_\lambda g$ for g bounded and continuous in the prediction topology of $\varphi_1(E)$, where R^Z_λ denotes the resolvent of $q(t,z,A)$. Suppose, in particular, that $g(z) = E^z h$ where h is bounded and continuous on $\overline{\hat{\Omega}}$ (as in the proof of Theorem 2.3). Then we have

$$\begin{aligned} R^Z_\lambda g(z) &= E^z \int_0^\infty e^{-\lambda s} E^{Z^z_s} h \, ds \\ &= E^z \int_0^\infty e^{-\lambda s} E^z(h\circ\theta_s | \mathcal{F}^o_{s+}) ds \\ &= \int_0^\infty e^{-\lambda s} E^z(h\circ\theta_s) ds. \end{aligned}$$

We observe now that the generators of the topology of $\overline{\hat{\Omega}}$, viz. $\left\{\int_0^r f_n(X_s)ds\right\}$, are all uniformly continuous under translation by θ_t, uniformly on $\hat{\Omega}$, since $\theta_t\left(\int_0^r f_n(X_s)ds\right) = \int_t^{r+t} f_n(X_s)ds$ and f_n is bounded. By the Stone-Weierstrass Thecrem, functions of the form $g_m\left(\int_0^{r_1} f_{n_1}(X_s)ds, \ldots, \int_0^{r_m} f_{n_m}(X_s)ds\right)$, $g_m(x_1,\ldots,x_m)$ continuous, $1 \le m$, are uniformly dense in $\overline{\hat{\Omega}}$, and obviously they have the same continuity under translation (g_m being uniformly continuous on compact sets). Hence all continuous h on $\overline{\hat{\Omega}}$ share this property, so that $E^z h \circ \theta_s$ is uniformly continuous in s. Then it is easy to see from the continuity theorem for Laplace Transforms that, since $R^Z_\lambda(E^z h)$ is continuous in z for each λ, $E^z(h \circ \theta_s)$ must be continuous for each s. In particular, $E^z h$ is continuous. Since these generate the prediction topology, the proof of Lemma 2.4 is complete.

We can view this Lemma in a slightly different way as follows. If we give E the topology induced by φ_1, then φ_1 is a homeomorphism of $E \to \varphi_1(E)$. For each $x \in E$, $Z_t^{\varphi_1(x)}$ $(= \varphi_1(X_t)$ for all t, P^x-a.s.) is the image of X_t under this homeomorphism by Theorem 1.5 a). Thus Z_t on $\varphi_1(E)$ is just the homeomorphic image of X_t on E with the φ_1-induced topology. It follows that in defining φ_2 we can just as well begin with X_t in the φ_1-induced topology. The topology generated by $\varphi_2\varphi_1$ on E is thus the same as the topology generated on E by φ alone when it is considered as a map into M with the prediction topology relative to the φ_1-induced topology of E.

In the above form, it is clear that Lemma 2.4 applies inductively (with the Lemma as case n = 2) to yield for all n > 1 a Lusin topology on E generated by φ from the prediction topology of $\varphi(E)$ for the $(n-1)^{st}$-topology of E, in such a way that

a) the n^{th} topology is finer than the $(n-1)^{st}$, and

b) the n^{th} topology is coarser than the Ray topology, and

c) X_t is a right process in the n^{th} topology.

(It should be remarked, however, that this does not apply for n = 1: the topology generated by φ_1 need not be comparable to the original topology of E.)

Denoting the n^{th} topology in this sequence by $\mathcal{T}_n$, we have at last

Theorem 2.5. The Ray topology is given by $\lim_{n\to\infty} \mathcal{T}_n$. In other words, it is the coarsest topology finer than each $\mathcal{T}_n$.

Proof. If $d_n(x,y)$ for each n is a metric bounded by 1 on E and generating $\mathcal{T}_n$, then $\lim_{n\to\infty} \mathcal{T}_n$ is generated by the metric $d(x,y) = \sum_{n=1}^{\infty} 2^{-n} d_n(x,y)$. Of course, the Borel field with respect to this metric is again $\mathcal{E}$ since each $d_n(x,y)$ is continuous (hence measurable), and $(E,\mathcal{E})$ is a Lusin space in the limit topology. Since the Ray topology is finer than each $\mathcal{T}_n$, it is finer than the limit. Conversely, by Theorem 2.3 the limit topology is finer than the topology generated by the cone C_1, and this is obviously the same as generated by C_2 (Definition 1.2 a)). Assuming for induction that it is finer than that generated by C_{2n+1}, it follows by Theorem 2.3 that it is finer than that generated by C_{2n+2} together with $R_\lambda(C_{2n+2})$, which is the topology generated by C_{2n+3}. Hence by induction it is finer than that generated by C_∞, which is the Ray topology. This completes the proof.

Final Remarks. This characterization of the topology does not suffice to characterize the R.-K. compactification, for which we need the particular cone C_∞. Nevertheless, it seems to suggest that the compactification procedure is a device to obtain topological identification of X_t with its prediction process P^{X_t} on a compact state space. For example, if E is already compact and we start with a Ray process whose resolvent separates points, instead of a right process, then the same operations are meaningful but yield simply the given topology of E. Indeed, we have $C_1 \subset C(E)$ in that case, whence easily $C_\infty \in C(E)$. Conversely, since C_1 separates points, it is easy to see by the Stone-Weierstrass Theorem that it generates a topology as fine as the given one. Moreover, by [7, Theorem 2] one can achieve separation of points by a preliminary identification of equivalence classes, hence this does not seem to be a serious restriction. Finally, another characterization of the Ray topology, in purely analytical terms, is given in [1, (15.3)] together with a comparison of different compactifications.

References

1. R. K. Getoor, Markov Processes: Ray Processes and Right Processes, Springer Lecture Notes #440 (1975).

2. F. B. Knight, Note or regularization of Markov processes, Illinois J. Math. 9, 548-552 (1965).

3. F. B. Knight, Essays on the Prediction Process, Institute of Math. Statistics Lecture Notes #1 (1981).

4. H. Kúnita and T. Watanabe, Some theorems concerning resolvents over locally compact spaces, Fifth Berkeley Symposium on Probability and Statistics, Vol. V, Univ. of Calif. Press, J. Neyman, Ed.., 131-161 (1967).

5. D. B. Ray, Resolvents, transition functions, and strongly Markovian processes, Ann. Math. 70(1959), 43-72.

6. J. B. Walsh et P. A. Meyer, Quelques application des resolvantes de Ray, Inventiones Math. 14, 143-166 (1971).

7. J. B. Walsh, Transition functions of Markov processes. Sem. de Prob. VI, Springer Lecture Notes #258, 215-233 (1972).