A CRITERION FOR WEAK COMPACTNESS OF MEASURES ON PRODUCT SPACES WITH APPLICATIONS

By

Jun Kawabe*

(Received October 3, 1994)

Abstract. We show that a subset of τ-smooth measures on a product of regular spaces is relatively compact with respect to the weak topology of measures if and only if the sets of its marginals on the factor spaces are relatively compact. This criterion has a lot of applications.

1. Introduction and notation

Since the notion of τ-smoothness was introduced by LeCam [13], it has been known that the object best suited for the study of weak convergence of measures on regular spaces is the space of not Radon but τ-smooth measures (see Topsøe [18]). In this paper, we give a weak compactness criterion for τ-smooth measures on a product of regular spaces. In fact, we show that a subset of τ-smooth measures on a product of regular spaces is relatively compact with respect to the weak topology of measures if and only if the sets of its marginals on the factor spaces are relatively compact. This criterion has a lot of applications. We first apply it to prove the celebrated Strassen's theorem for not Radon but τ-smooth probability measures on a product of two regular spaces. Since Strassen [16] first gave a necessary and sufficient condition for the existence of probability measures with given marginals, this result has been extended by many authors in more general settings under the assumption that the measures are Radon (see section 3). This is indeed a convenient assumption, but is not necessary to extend Strassen's theorem for regular spaces. We only assume the τ-smoothness of measures. However it is not easy to see

this, and indeed we need the above compactness criterion. Next we apply it to establish the weak convergence of product measures and compound measures on a product of completely regular spaces.

In the rest of this section, we prepare basic results and notation. Let X be a topological space (which is always assumed to be Hausdorff) and $\mathscr{B}(X)$ be the σ-algebra of all Borel subsets of X. By $\mathscr{M}^+(X)$ we denote by the set of all non-negative, totally finite measures defined on $\mathscr{B}(X)$. We endow $\mathscr{M}^+(X)$ with the weak topology in the sense of Topsøe [18], that is, the weakest topology for which all functionals $\mu \mapsto \int_X f(x)\mu(dx)$ are upper semicontinuous (u.s.c.) whenever f is a bounded u.s.c. real-valued function. In this topology, a net $\{\mu_\alpha\}$ converges to μ (we write $\mu_\alpha \overset{w}{\to} \mu$) if and only if $\lim_\alpha \mu_\alpha(X)=\mu(X)$ and $\lim \sup_\alpha \mu_\alpha(F) \leq \mu(F)$ for all closed subsets F of X. (Equivalent conditions are given in Topsøe's Portmanteau theorem; see [18; Theorem 8.1].)

A family $\mathcal{A}$ of subsets of X is said to be *filtering downwards* (resp. *upwards*) to a subset A_0 of X, and we write $\mathcal{A} \downarrow A_0$ (resp. $\mathcal{A} \uparrow A_0$) if for any A_1, $A_2 \in \mathcal{A}$ we can find $A_3 \in \mathcal{A}$ such that $A_3 \subset A_1 \cap A_2$ (resp. $A_3 \supset A_1 \cup A_2$) and $A_0 = \cap_{A \in \mathcal{A}} A$ (resp. $A_0 = \cup_{A \in \mathcal{A}} A$). In this paper, the following concept of regularity for Borel measures is useful. We say that $\mu \in \mathscr{M}^+(X)$ is τ-*smooth* if $\mu(F_0)=\inf_{F \in \mathcal{F}}\mu(F)$ holds for any family $\mathcal{F}$ of closed subsets of X which is filtering downwards to F_0, and this is equivalent to the condition that for any family $\mathcal{G}$ of open subsets of X with $\mathcal{G} \uparrow G_0$, we have $\mu(G_0)=\sup_{G \in \mathcal{G}}\mu(G)$. Let us denote by $\mathscr{M}_\tau^+(X)$ the set of all τ-smooth measures in $\mathscr{M}^+(X)$. For further necessary definitions and results, we refer the reader to Kelley [12], Schwartz [14], Topsøe [18] and Vakhania *et al.* [20].

2. A criterion for weak compactness

Given two sets X and Y, π_X and π_Y denotes the projections $X \times Y \to X$ and $X \times Y \to Y$, respectively. For a measure $\gamma \in \mathscr{M}^+(X \times Y)$, we define its *marginals* $\pi_X(\gamma)$ and $\pi_Y(\gamma)$ by $\pi_X(\gamma)(A)=\gamma(\pi_X^{-1}(A))$ and $\pi_Y(\gamma)(B)=\gamma(\pi_Y^{-1}(B))$ for all $A \in \mathscr{B}(X)$ and $B \in \mathscr{B}(Y)$. The following theorem is a main result of this paper.

Theorem 1. *Let X and Y be regular spaces and let $\{\gamma_\alpha\}$ be a net in $\mathscr{M}^+(X \times Y)$ with $\lim \sup_\alpha \gamma_\alpha(X \times Y)<\infty$. If $\pi_X(\gamma_\alpha) \overset{w}{\to} \mu \in \mathscr{M}_\tau^+(X)$ and $\pi_Y(\gamma_\alpha) \overset{w}{\to} \nu \in \mathscr{M}_\tau^+(Y)$ then every subnet of $\{\gamma_\alpha\}$ has a further subnet converging weakly to a measure $\gamma \in \mathscr{M}_\tau^+(X \times Y)$ such that $\pi_X(\gamma)=\mu$ and $\pi_Y(\gamma)=\nu$.*

To prove Theorem 1 we need the following

Lemma 1. *Let X be a regular space. Assume that a family $\mathcal{G}$ of subsets of X satisfies the following two conditions:*

(1) $\mathcal{G}$ *is closed under finite unions.*

(2) $\mathcal{G}$ *contains an open basis for the topology of* X.

Then for any open subset G *of* X, *we can find a subfamily* $\mathcal{H}$ *of* $\mathcal{G}$ *such that* $\mathcal{H} \uparrow G$.

Proof. Put $\mathcal{H} = \{H \in \mathcal{G} : H \subset \bar{H} \subset G\}$. Then, using assumption (1), it is easy to see that $\mathcal{H}$ is filtering upwards. Hence we have only to show that the equality $\bigcup_{H \in \mathcal{H}} H = G$ holds. Fix $x \in G$. Since X is regular, there exists an open subset U of X such that $x \in U \subset \bar{U} \subset G$. On the other hand, we can find $H \in \mathcal{G}$ such that $x \in H \subset U$ by assumption (2). Consequently we have $x \in H \in \mathcal{H}$, and this implies $\bigcup_{H \in \mathcal{H}} H \supset G$. The reverse inclusion is obvious. $\square$

Proof of Theorem 1. We first show that every subnet of $\{\gamma_\alpha\}$ contains a further subnet converging weakly to a measure $\gamma \in \mathcal{M}_\tau^+(X \times Y)$. To do this by Theorem 6 of [17], we have only to show that the condition of "τ-smoothness"

$$\inf_{F \in \mathcal{F}} \limsup_\alpha \gamma_\alpha(F) = 0$$

holds for every family $\mathcal{F}$ of closed subsets of $X \times Y$ with $\mathcal{F} \downarrow \emptyset$. (Here we remark that in Theorem 6 of [17] we need not assume that a net (μ_α) is in $\mathcal{M}_+(X ; \tau)$—it is enough to consider a net in $\mathcal{M}_+(X)$; then the conditions of the theorem are the necessary and sufficient conditions that every subnet of (μ_α) contains a further subnet converging to a measure in $\mathcal{M}_+(X ; \tau)$).

Fix $\varepsilon > 0$ and let $\mathcal{F}$ be an arbitrary family of closed subsets of $X \times Y$ with $\mathcal{F} \downarrow \emptyset$. Put $\mathcal{E}_X = \{\pi_X(G)^c : G \text{ is open and } G^c \supset F \text{ for some } F \in \mathcal{F}\}$ and $\mathcal{E}_Y = \{\pi_Y(G)^c : G \text{ is open and } G^c \supset F \text{ for some } F \in \mathcal{F}\}$. Since the projections π_X and π_Y are open mappings, $\mathcal{E}_X$ and $\mathcal{E}_Y$ are families of closed subsets of X and Y, respectively. Moreover we can show that $\mathcal{E}_X \downarrow \emptyset$ and $\mathcal{E}_Y \downarrow \emptyset$ as in the proof of Lemma 1. Since μ and ν are τ-smooth, we have

$$\inf_{E \in \mathcal{E}_X} \mu(E) = 0 \quad \text{and} \quad \inf_{E \in \mathcal{E}_Y} \nu(E) = 0,$$

which imply that there exists an open subset G_ε of $X \times Y$ with $G_\varepsilon^c \supset F_\varepsilon$ for some $F_\varepsilon \in \mathcal{F}$ such that

$$\mu(\pi_X(G_\varepsilon)) > 1 - \frac{\varepsilon}{3} \quad \text{and} \quad \nu(\pi_Y(G_\varepsilon)) > 1 - \frac{\varepsilon}{3}.$$

Now we set

$$\mathcal{G} = \left\{ \bigcup_{i=1}^{n} (U_i \times V_i) : \right.$$

the U_i's are open subsets of X and the V_i's are open subsets of $Y \Big\}$,

then $\mathcal{G}$ satisfies conditions (1) and (2) of Lemma 1. Consequently we can find a subfamily $\mathcal{H}$ of $\mathcal{G}$ such that $\mathcal{H} \uparrow G_\varepsilon$, and hence we have $\pi_X(\mathcal{H}) \uparrow \pi_X(G_\varepsilon)$ and $\pi_Y(\mathcal{H}) \uparrow \pi_Y(G_\varepsilon)$. Then, noting that π_X and π_Y are open mappings, by the τ-smoothness of μ and ν, we have

$$\sup_{H \in \mathcal{H}} \mu(\pi_X(H)) = \mu(\pi_X(G_\varepsilon)) > 1 - \frac{\varepsilon}{3}$$

and

$$\sup_{H \in \mathcal{H}} \mu(\pi_Y(H)) = \nu(\pi_Y(G_\varepsilon)) > 1 - \frac{\varepsilon}{3},$$

which imply that there exists $H_\varepsilon \in \mathcal{H}$ with $H_\varepsilon \subset G_\varepsilon$ such that

$$\mu(\pi_X(H_\varepsilon)) > 1 - \frac{\varepsilon}{2} \quad \text{and} \quad \nu(\pi_Y(H_\varepsilon)) > 1 - \frac{\varepsilon}{2}.$$

Since $H_\varepsilon \in \mathcal{G}$, it can be expressed by the form $H_\varepsilon = \bigcup_{i=1}^n (U_i \times V_i)$, where the U_i's are non-empty open subsets of X and the V_i's are non-empty open subsets of Y, and thus $\pi_X(H_\varepsilon) = \bigcup_{i=1}^n U_i$ and $\pi_Y(H_\varepsilon) = \bigcup_{i=1}^n V_i$. If we notice that $F_\varepsilon \subset G_\varepsilon^c \subset H_\varepsilon^c$ and the equality

$$H_\varepsilon^c = \left\{ \left(\bigcap_{i=1}^n U_i^c \right) \times Y \right\} \cup \left\{ X \times \left(\bigcap_{i=1}^n V_i^c \right) \right\}$$

holds, putting $K = \bigcap_{i=1}^n U_i^c = \pi_X(H_\varepsilon)^c$ and $L = \bigcap_{i=1}^n V_i^c = \pi_Y(H_\varepsilon)^c$, then by the assumption we have

$$\limsup_\alpha \gamma_\alpha(F_\varepsilon) \leqq \limsup_\alpha \gamma_\alpha(H_\varepsilon^c)$$

$$= \limsup_\alpha \gamma_\alpha((K \times Y) \cup (X \times L))$$

$$\leqq \limsup_\alpha \gamma_\alpha(K \times Y) + \limsup_\alpha \gamma_\alpha(X \times L)$$

$$= \limsup_\alpha \pi_X(\gamma_\alpha)(K) + \limsup_\alpha \pi_Y(\gamma_\alpha)(L)$$

$$\leqq \mu(K) + \nu(L)$$

$$= \mu(\pi_X(H_\varepsilon)^c) + \nu(\pi_Y(H_\varepsilon)^c)$$

$$\leqq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and this implies that the condition of "τ-smoothness" holds.

It remains to prove that γ has marginals μ and ν, that is, $\pi_X(\gamma) = \mu$ and $\pi_Y(\gamma) = \nu$. Let us assume that $\gamma_\alpha \xrightarrow{w} \gamma \in \mathcal{M}_\tau^+(X \times Y)$, $\pi_X(\gamma_\alpha) \xrightarrow{w} \mu \in \mathcal{M}_\tau^+(X)$ and $\pi_Y(\gamma_\alpha) \xrightarrow{w} \nu \in \mathcal{M}_\tau^+(Y)$. Then by the continuity of the projections π_X and π_Y, we have $\pi_X(\gamma_\alpha) \xrightarrow{w} \pi_X(\gamma)$ and $\pi_Y(\gamma_\alpha) \xrightarrow{w} \pi_Y(\gamma)$, and it is obvious that $\pi_X(\gamma) \in \mathcal{M}_\tau^+(X)$ and $\pi_Y(\gamma) \in \mathcal{M}_\tau^+(Y)$ by the τ-smoothness of γ. Consequently from P19 of [18; page XIV], it follows that $\pi_X(\gamma) = \mu$ and $\pi_Y(\gamma) = \nu$, and the proof is complete. $\square$

Remark 1. In the case μ and ν being Radon measures, Theorem 1 is known (see Hoffmann-Jørgensen [8], Lemma I.5.1). In this case, the limiting measure on the product space is also Radon.

A net $\{x_\alpha\}$ in a topological space X is said to be *compact* if every subnet of $\{x_\alpha\}$ has a further subnet which converges, and a subset A of X is said to be *net-compact* if every net in A has a convergent subnet. It is known that if X is regular, then a subset A of X is net-compact if and only if it is relatively compact (see Bourbaki [2]; chap. I, §9, ex. 22 and 23). By Theorem 1, we have the following compactness criterion for τ-smooth measures on a product of regular spaces.

Corollary 1. *Let X and Y be regular spaces.*

(1) *Let $\{\gamma_\alpha\}$ be a net in $\mathcal{M}_\tau^+(X \times Y)$ with $\lim \sup_\alpha \gamma_\alpha(X \times Y) < \infty$. Then $\{\gamma_\alpha\}$ is compact in $\mathcal{M}_\tau^+(X \times Y)$ if and only if the nets $\{\pi_X(\gamma_\alpha)\}$ and $\{\pi_Y(\gamma_\alpha)\}$ are compact in $\mathcal{M}_\tau^+(X)$ and $\mathcal{M}_\tau^+(Y)$, respectively.*

(2) *Let P be a subset of $\mathcal{M}_\tau^+(X \times Y)$ with $\sup_{\gamma \in P} \gamma(X \times Y) < \infty$. Then P is relatively compact in $\mathcal{M}_\tau^+(X \times Y)$ if and only if the sets $\pi_X(P)$ and $\pi_Y(P)$ are relatively compact in $\mathcal{M}_\tau^+(X)$ and $\mathcal{M}_\tau^+(Y)$, respectively.*

Proof. (1) It easily follows from Theorem 1. (2) Since the weak topologies on $\mathcal{M}_\tau^+(X)$, $\mathcal{M}_\tau^+(Y)$ and $\mathcal{M}_\tau^+(X \times Y)$ are regular (see [18; Theorem 11.2]), we can replace "relatively compact" by "net-compact" in the statement (2). Consequently, (2) also follows from Theorem 1. □

Remark 2. We say that a subset P of $\mathcal{M}^+(X)$ is *uniformly tight* if for each $\varepsilon > 0$, there exists a compact subset K_ε of X such that $\mu(X - K_\varepsilon) < \varepsilon$ for all $\mu \in P$. Then the following criterion for uniform tightness is well-known and easily verified: A subset P of $\mathcal{M}^+(X \times Y)$ is uniformly tight if and only if the marginals $\pi_X(P)$ and $\pi_Y(P)$ are uniformly tight. Moreover every uniformly tight set is relatively compact with respect to the weak topology of measures (see [18; Theorem 9.1]). However we know that even for a subset of measures on Suslin spaces, the relative compactness does not imply the uniform tightness in general (see Example I.6.4 of Fernique [6]). Therefore Corollary 1 cannot be inferred directly from the above criterion for uniform tightness.

Remark 3. For a topological space X, denote by $\mathcal{M}_t^+(X)$ the set of all Radon measures in $\mathcal{M}^+(X)$. Using Lemma I.5.1 of [8] (see also Remark 1), it can be easily shown that a Radon version of Corollary 1 also remains valid for arbitrary topological spaces X and Y if we replace $\mathcal{M}_\tau^+(X)$, $\mathcal{M}_\tau^+(Y)$ and $\mathcal{M}_\tau^+(X \times Y)$ by $\mathcal{M}_t^+(X)$, $\mathcal{M}_t^+(Y)$ and $\mathcal{M}_t^+(X \times Y)$ respectively. However, τ-smooth measures are not Radon in general (see Varadarajan [21; page 200] and remark that the

measure constructed there is not only σ-smooth but also τ-smooth), and thus Corollary 1 is not contained in the above stated Radon version of Corollary 1.

3. Some applications

(I) *Strassen's theorem for τ-smooth probability measures.* In a celebrated paper, Strassen [16] gave a necessary and sufficient condition for the existence of probability measures with given marginals. Although his theorem has been extended by many authors in more general settings (cf. Hoffmann-Jørgensen [8], Edwards [5], Kellerer [11], Hansel and Troallic [7], Skala [15] and so on), they all treat only Radon measures. The following theorem states that Strassen's theorem still holds if we replace Polish spaces by regular spaces and restrict the possible solutions to a weakly closed, convex set of τ-*smooth* probability measures. Since τ-smooth measures on a regular space are not Radon in general (see Remark 3), the following Theorem 2 is not contained in any previous result cited above. For a regular space X, let us denote by $\mathcal{P}(X)$ (resp. $\mathcal{P}_\tau(X)$) the set of all probability (resp. τ-smooth probability) measures on $\mathcal{B}(X)$.

Theorem 2. *Let X and Y be regular spaces and let Λ be a non-empty closed convex subset of $\mathcal{P}_\tau(X \times Y)$. In order that there exists $\lambda \in \Lambda$ with given marginals $\mu \in \mathcal{P}_\tau(X)$ and $\nu \in \mathcal{P}_\tau(Y)$, that is, $\pi_X(\lambda) = \mu$ and $\pi_Y(\lambda) = \nu$, it is necessary and sufficient that*

$$\int_X f d\mu + \int_Y g d\nu \leq \sup\left\{\int_{X \times Y} (f \oplus g) d\gamma : \gamma \in \Lambda\right\}$$

for all bounded, Borel measurable real-valued functions f on X and g on Y. Here $(f \oplus g)(x, y) \equiv f(x) + g(y)$ for all $(x, y) \in X \times Y$.

Proof. By the argument in the proof of Theorem 1 of [15], we have only to show that a net $\{\gamma_\alpha\}$ in $\mathcal{P}(X \times Y)$ satisfying the condition that $\pi_X(\gamma_\alpha) \xrightarrow{w} \mu$ and $\pi_Y(\gamma_\alpha) \xrightarrow{w} \nu$ is compact, but this immediately follows from Theorem 1. $\square$

(II) *Weak convergence of product probability measures.* Let X and Y be topological spaces and let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}_\tau(Y)$. Then by the τ-smoothness of ν, for each $B \in \mathcal{B}(X \times Y)$, the function $x \in X \mapsto \nu(B_x)$ is Borel measurable (see Lemma 4.1 of [20] or Example 1-(1) of Kawabe [9]), and hence we can define a product measure $\mu \times \nu$ as a Borel measure on $X \times Y$ by

$$(\mu \times \nu)(B) = \int_X \nu(B_x) \mu(dx) \qquad \text{for all} \quad B \in \mathcal{B}(X \times Y).$$

Here for a subset B of $X \times Y$ and $x \in X$, B_x denotes the section determined by x, that is, $B_x = \{y \in Y : (x, y) \in B\}$. Thus it makes sense that we consider the weak convergence of product measures, and we have the following

Theorem 3. *Let X and Y be completely regular spaces. Let $\{\mu_a\}$ be a net in $\mathscr{P}(X)$ and $\{\nu_a\}$ be a net in $\mathscr{P}_\tau(X)$. If $\mu_a \xrightarrow{w} \mu \in \mathscr{P}_\tau(X)$ and $\nu_a \xrightarrow{w} \nu \in \mathscr{P}_\tau(Y)$, then $\mu_a \times \nu_a \xrightarrow{w} \mu \times \nu$.*

In order to prove Theorem 3, we shall make use of the following result which establishes the relation between τ-smooth measures and their characteristic functions. For a topological space X, denote by $C(X)$ (resp. $C_b(X)$) the set of all continuous (resp. bounded continuous) real-valued functions defined on X. For $\mu \in \mathscr{P}(X)$, we define its characteristic function by

$$\hat{\mu}(f) = \int_X e^{if(x)} \mu(dx), \qquad f \in C(X).$$

Lemma 2 (cf. Vakhania *et al.* [20], Theorem IV.2.2 and IV.3.1). *Let X be a completely regular space. Assume that a linear subspace Γ of $C(X)$ generates the topology of X, that is, the topology of X coincides with the weakest topology for which all the functions in Γ are continuous. Then the following statements hold:*

(1) *Let $\mu, \nu \in \mathscr{P}_\tau(X)$. If $\hat{\mu}(f) = \hat{\nu}(f)$ for all $f \in \Gamma$, then $\mu = \nu$ on $\mathscr{B}(X)$.*

(2) *Let $\{\mu_a\}$ be a net in $\mathscr{P}(X)$. If every subnet of $\{\mu_a\}$ has a further subnet converging weakly to a τ-smooth probability measure on X, and for each $f \in \Gamma$, $\chi(f) = \lim_a \hat{\mu}_a(f)$ exists, then $\{\mu_a\}$ converges weakly to a measure $\mu \in \mathscr{P}_\tau(X)$ with $\hat{\mu} = \chi$.*

Proof. This lemma can be proved by Lemma 5 of [10] and a standard argument (see [20], Theorem IV.3.1). □

Proof of Theorem 3. Put $\gamma_a = \mu_a \times \nu_a$. Then it is clear that $\pi_X(\gamma_a) \xrightarrow{w} \mu$ and $\pi_Y(\gamma_a) \xrightarrow{w} \nu$, and hence by Theorem 1, $\{\gamma_a\}$ has a further subnet converging weakly to a τ-smooth probability measure on $X \times Y$. It is easily verified that for each $f \in C_b(X)$ and $g \in C_b(Y)$, $\lim_a \hat{\gamma}_a(f \oplus g) = \lim_a \hat{\mu}_a(f)\hat{\nu}_a(g) = \hat{\mu}(f)\hat{\nu}(g) = (\mu \times \nu)^\wedge(f \oplus g)$. Since $\Gamma \equiv C_b(X) \oplus C_b(Y)$ is a linear subspace of $C(X \times Y)$ generating the topology of $X \times Y$, by Lemma 2 we have $\mu_a \times \nu_a \xrightarrow{w} \mu \times \nu$, and the proof is complete. □

Remark 4. Theorem 3 was known in the case when X and Y are separable metric spaces (see, e.g., Billingsley [1; Theorem 3.2]), and has been extended by Vakhania *et al.* [20; Proposition I.4.1] to completely regular spaces. But their technique is that the weak convergence $\mu_a \xrightarrow{w} \mu$ can be proved by showing that $\mu_a(A) \to \mu(A)$ for some special class of sets A, and is different from ours.

(III) *Weak convergence of compound probability measures.* Let X be a topological space and Y be a completely regular space. A (Borel) *transition*

probability λ on $X \times Y$ is defined to be a mapping from X into $\mathscr{P}(Y)$ which satisfies the condition that the function $x \in X \mapsto \lambda(x, B)$ is Borel measurable for every $B \in \mathscr{B}(Y)$. We say that a transition probability λ is *τ-smooth* if the probability measure $\lambda_x \equiv \lambda(x, \cdot)$ is τ-smooth for each $x \in X$, that is, it is a mapping from X into $\mathscr{P}_\tau(Y)$. We also say that λ is *continuous* if the mapping $x \in X \mapsto \lambda_x \in \mathscr{P}(Y)$ is continuous. According to Proposition 1-(1) of [9], for any $\mu \in \mathscr{P}(X)$ and any continuous τ-smooth transition probability λ on $X \times Y$, we can define a Borel probability measure $\mu \circ \lambda$ on $X \times Y$, which is called the *compound probability measure* of μ and λ, by

$$\mu \circ \lambda(D) = \int_X \lambda(x, D_x) \mu(dx) \qquad \text{for all} \quad D \in \mathscr{B}(X \times Y).$$

Moreover if μ is τ-smooth, then $\mu \circ \lambda$ is also τ-smooth (see Proposition 2 of [9]).

In information theory, a transition probability λ on $X \times Y$ is called an information channel with input space X and output space Y, and for a probability measure μ on X (which is called the input source), the compound probability measure $\mu \circ \lambda$ plays an important role (see, e.g., Umegaki [19]).

On the other hand, compound probability measures can be viewed as a generalization of convolution measures. In fact, if $X = Y$ is a topological group and λ is a transition probability given by $\lambda(x, B) = \nu(Bx^{-1})$ for all $x \in X$ and all Borel subsets B of Y, where ν is a probability measure on X, then the marginal $\pi_Y(\mu \circ \lambda)$ is the convolution measure $\mu * \nu$. The weak convergence of convolution measures has been looked into in great detail by Csiszár [3, 4].

Let X be a topological space and Y be a uniform space with its uniformity $\mathscr{V}$. Denote by $C(X, Y)$ the set of all continuous mappings from X into Y. We say that a subset H of $C(X, Y)$ is *equicontinuous at* $x \in X$ if for each $V \in \mathscr{V}$, there exists a neighborhood U of x such that $(f(x), f(u)) \in V$ for all $u \in U$ and all $f \in H$. H is *equicontinuous on* X if it is equicontinuous at every $x \in X$. In the following, we need a variant of the notion of equicontinuity. We say that a subset H of $C(X, Y)$ is *equicontinuous on a set* A of X if the set of all restrictions of functions of H to A is equicontinuous on A. By Proposition 1 of [9], $C(X, \mathscr{P}_\tau(Y))$ coincides with the set of all continuous τ-smooth transition probabilities on $X \times Y$. Moreover since Y is completely regular, the weak topology on $\mathscr{P}_\tau(Y)$ is also completely regular, and hence it is uniformizable (see [18], Theorem 11.2). Thus we can introduce the notion of equicontinuity for a set of transition probabilities. Then it is not too hard to prove that a subset Q of $C(X, \mathscr{P}_\tau(Y))$ is equicontinuous on every compact subset of X if and only if for each $h \in C_b(Y)$, the set of the functions

$$x \in X \longmapsto \int_Y h(y) \lambda(x, dy), \qquad \lambda \in Q,$$

is equicontinuous on every compact subset of X. Some examples of equicon-

tinuous sets of transition probabilities are given in [9, 10]. Now we state our result about the weak convergence of compound probability measures.

Theorem 4. *Let* X *and* Y *be completely regular spaces. Assume that a net* $\{\lambda_\alpha\}$ *in* $C(X, \mathscr{P}_\tau(Y))$ *satisfies*

(a) $\{\lambda_\alpha\}$ *is equicontinuous on every compact subset of* X,

(b) *there exists* $\lambda \in C(X, \mathscr{P}_\tau(Y))$ *such that* $\lambda_\alpha(x, \cdot) \xrightarrow{w} \lambda(x, \cdot)$ *for every* $x \in X$. *Then for any uniformly tight net* $\{\mu_\alpha\}$ *in* $\mathscr{P}(X)$ *converging weakly to* $\mu \in \mathscr{P}_\tau(X)$, *we have* $\mu_\alpha \circ \lambda_\alpha \xrightarrow{w} \mu \circ \lambda$.

In order to prove Theorem 4, we need the following

Lemma 3 ([9; Lemma 3]). *Let* X *be a completely regular space and let* $\{\mu_\alpha\}$ *be a net in* $\mathscr{P}(X)$ *which is uniformly tight. Assume that a net* $\{\varphi_\alpha\}$ *in* $C_b(X)$ *satisfies*

(a) $\{\varphi_\alpha\}$ *is uniformly bounded, and*

(b) $\{\varphi_\alpha\}$ *is equicontinuous on every compact subset of* X.

If $\mu_\alpha \xrightarrow{w} \mu \in \mathscr{P}_\tau(X)$, *and if* $\varphi \in C_b(X)$ *and* $\varphi_\alpha(x) \to \varphi(x)$ *for each* $x \in X$, *then we have*

$$\lim_\alpha \int_X \varphi_\alpha(x) \mu_\alpha(dx) = \int_X \varphi(x) \mu(dx).$$

Proof of Theorem 4. We first show that $\pi_Y(\mu_\alpha \circ \lambda_\alpha) \xrightarrow{w} \pi_Y(\mu \circ \lambda)$. Fix $g \in C_b(Y)$ and define bounded continuous functions φ_α and φ on X by

$$\varphi_\alpha(x) = \int_Y g(y) \lambda_\alpha(x, dx) \quad \text{and} \quad \varphi(x) = \int_Y g(y) \lambda(x, dx).$$

Then it is clear that $\{\varphi_\alpha\}$ satisfies condition (a) of Lemma 3, while by assumption (a), $\{\varphi_\alpha\}$ satisfies condition (b) of Lemma 3. On the other hand, from assumption (b) it follows that $\varphi_\alpha(x) \to \varphi(x)$ for each $x \in X$ and $\varphi \in C_b(X)$. Therefore by Lemma 3, we have

$$\lim_\alpha \int_X \varphi_\alpha(x) \mu_\alpha(dx) = \int_X \varphi(x) \mu(dx),$$

which implies $\pi_Y(\mu_\alpha \circ \lambda_\alpha) \xrightarrow{w} \pi_Y(\mu \circ \lambda)$. Thus we have $\pi_X(\mu_\alpha \circ \lambda_\alpha) = \mu_\alpha \xrightarrow{w} \mu = \pi_X(\mu \circ \lambda)$ $\in \mathscr{P}_\tau(X)$ and $\pi_Y(\mu_\alpha \circ \lambda_\alpha) \xrightarrow{w} \pi_Y(\mu \circ \lambda) \in \mathscr{P}_\tau(Y)$, and hence by Theorem 1, we can find a subnet $\{\mu_{\alpha'} \circ \lambda_{\alpha'}\}$ of $\{\mu_\alpha \circ \lambda_\alpha\}$ converging weakly to a measure $\gamma \in \mathscr{P}_\tau(X \times Y)$.

Now for each $\gamma \in \mathscr{P}(X \times Y)$, we define its characteristic function by

$$\hat{\gamma}(f) = \int_{X \times Y} e^{if(x, y)} \gamma(dx, dx), \qquad f \in C(X \times Y).$$

Then by Lemma 2 and the fact obtained above, in order to prove that $\mu_\alpha \circ \lambda_\alpha$ $\xrightarrow{w} \mu \circ \lambda$, it is sufficient to show that for each $f \in C_b(X)$ and $g \in C_b(Y)$, we have

$(\mu_\alpha \circ \lambda_\alpha)^\wedge(f \oplus g) \to (\mu \circ \lambda)^\wedge(f \oplus g)$, since $\Gamma \equiv C_b(X) \oplus C_b(Y)$ is a linear subspace of $C(X \times Y)$ generating the completely regular topology of $X \times Y$.

Fix $f \in C_b(X)$ and $g \in C_b(Y)$, and put

$$\phi_\alpha(x) = e^{if(x)} \int_Y e^{ig(y)} \lambda_\alpha(x, dy) \quad \text{and} \quad \phi(x) = e^{if(x)} \int_Y e^{ig(y)} \lambda(x, dy).$$

Then by assumptions (a) and (b) of Theorem 4, it is easily verified that ϕ_α and ϕ satisfy conditions of Lemma 3, and hence we have

$$\lim_\alpha \int_X \phi_\alpha(x) \mu_\alpha(dx) = \int_X \phi(x) \mu(dx).$$

This implies that $\lim_\alpha(\mu_\alpha \circ \lambda_\alpha)^\wedge(f \oplus g) = (\mu \circ \lambda)^\wedge(f \oplus g)$, and the proof is complete.

□

Remark 5. Theorem 4 shows that the following conditions of Theorem 1 of [9] are superfluous: (1) X is a k-space, and (2) $\{\lambda_\alpha(x, \cdot)\}$ is uniformly tight for each $x \in X$. In [9], an application of Theorem 3 to Gaussian transition probabilities is also discussed.

Let G be a topological group and let $\mu \in \mathscr{P}(G)$ and $\nu \in \mathscr{P}_\tau(G)$. By [3] (or Example 1-(2) of [9]), we can define a *convolution* $\mu * \nu$ by $\mu * \nu = \pi_Y(\mu \circ \lambda)$, where $\lambda(x, B) = \nu(Bx^{-1})$ for all $x \in G$ and all $B \in \mathscr{B}(G)$. Then we have Corollary to Theorem 1 of [3]:

Corollary 2. *Let G be a topological group, and let $\{\mu_\alpha\}$ be a net in $\mathscr{P}(G)$ and $\{\nu_\alpha\}$ be a net in $\mathscr{P}_\tau(G)$. Assume that $\{\mu_\alpha\}$ is uniformly tight. If $\mu_\alpha \overset{w}{\to} \mu \in \mathscr{P}_\tau(G)$ and $\nu_\alpha \overset{w}{\to} \nu \in \mathscr{P}_\tau(G)$, then $\mu_\alpha * \nu_\alpha \overset{w}{\to} \mu * \nu$.*

Proof. Put $\lambda_\alpha(x, B) = \nu_\alpha(Bx^{-1})$ and $\lambda(x, B) = \nu(Bx^{-1})$ for all $x \in G$ and all $B \in \mathscr{B}(G)$. Then $\{\lambda_\alpha\}$ is equicontinuous on every compact subset of G by Example 1-(2) of [10], and it is clear that $\lambda_\alpha(x, \cdot) \overset{w}{\to} \lambda(x, \cdot)$ for each $x \in G$. Therefore by Theorem 4, we have $\mu_\alpha \circ \lambda_\alpha \overset{w}{\to} \mu \circ \lambda$, which implies $\mu_\alpha * \nu_\alpha \overset{w}{\to} \mu * \nu$. □

References

[1] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[2] N. Bourbaki, *Eléments de mathématique*, Livre III, *Topologie générale*, 4th ed., Paris, 1965.

[3] I. Csiszár, Some problems concerning measures on topological spaces and convolutions of measures on topological groups, *Les Probabilités sur les Structures Algébriques*, Clermont-Ferrand, Colloq. Internat. CNRS, Paris (1969), 75-96.

[4] ——————, On the weak* continuity of convolution in a convolution algebra over an arbitrary topological group, *Studia Sci. Math. Hungar.*, **6** (1971), 27-40.

[5] D. A. Edwards, On the existence of probability measures with given marginals, *Ann. Inst. Fourier (Grenoble)*, **28** (1978), 53-78.

[6] X. Fernique, Processus linéaires, processus généralisés, *Ann. Inst. Fourier (Grenoble)*, **17** (1967), 1-92.

[7] G. Hansel and J. P. Troallic, Sur le probléme des marges, *Probab. Th. Rel. Fields*, **71** (1986), 357-366.

[8] J. Hoffmann-Jørgensen, *Probability in Banach spaces*, Lecture Notes in Math. **598**, Springer-Verlag, New York, 1977.

[9] J. Kawabe, Convergence of compound probability measures on topological spaces, *Collog. Math.*, **67** (1994), 161-176.

[10] ——————, Weak convergence of compound probability measures (submitted for publication).

[11] H. G. Kellerer, Duality theorems for marginal problems, *Z. Wahrsch. Verw. Gebiete*, **67** (1984), 399-432.

[12] J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955.

[13] L. LeCam, Convergence in distribution of stochastic processes, *Univ. California Publ. Statist.*, **2** (1957), 207-236.

[14] L. Schwartz, *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press, 1973.

[15] H. J. Skala, The existence of probability measures with given marginals, *Ann. Probab.*, **21** (1993), 136-142.

[16] V. Strassen, The existence of probability measures with given marginals, *Ann. Math. Statist.*, **36** (1965), 423-439.

[17] F. Topsøe, Compactness in spaces of measures, *Studia Math.*, **36** (1970), 195-212.

[18] ——————, *Topology and Measure*, Lecture Notes in Math. **133**, Springer-Verlag, New York, 1970.

[19] H. Umegaki, Representations and extremal properties of averaging operators and their applications to information channels, *J. Math. Anal. Appl.*, **25** (1969), 41-73.

[20] N. N. Vakhania, V. I. Tarieladze and S. A. Chobanyan, *Probability Distributions on Banach Spaces*, D. Reidel Publishing Company, 1987.

[21] V. S. Varadarajan, Measures on topological spaces, *Amer. Math. Soc. Transl. Ser.* II, **48** (1965), 161-228.

Department of Mathematics
Faculty of Engineering
Shinshu University
Wakasato, Nagano 380
Japan
e-mail: jkawabe@gipwc. shinshu-u.ac.jp