ANNALES DE L'INSTITUT FOURIER

DAVID ALBERT EDWARDS

On the existence of probability measures with given marginals

Annales de l'institut Fourier, tome 28, nº 4 (1978), p. 53-78

<http://www.numdam.org/item?id=AIF_1978__28_4_53_0>

Ann. Inst. Fourier, Grenoble
28, 4 (1978), 53-78.

ON THE EXISTENCE OF PROBABILITY MEASURES WITH GIVEN MARGINALS

by D. A. EDWARDS

1. Introduction.

Let μ, ν be probability measures defined respectively on spaces X, Y and suppose that K is a given set of probability measures on the product $X \times Y$. Then it is in general a non-trivial question whether there exists a measure in K whose projections onto the factor spaces X, Y are μ and ν. A number of problems of this type have been investigated by Strassen in his important paper [10] (see also the works cited by Strassen there, especially the papers of Kellerer). The main aim of the present paper is to investigate a class of such problems concerning ordered topological spaces (we also obtain some results in the same spirit for infinite products). Various aspects of our subject have been studied by Nachbin [7], Strassen [10], Preston [8] and Hommel [5], and I comment below on some of their work. (Of recent work on non-topological ordered measure spaces (not considered here), it seems proper to mention [8], [4] and [6].)

It has been found necessary to develop here some general theory for compact and for completely regular (unordered) spaces (see § 3 and § 5). This material has been kept to the minimum needed for present purposes; for a broader perspective the reader is urged to consult [10].

For valuable comments on parts of this paper I must thank Dr. C. J. Preston and Pr G. Choquet. I must also thank M. J. Saint Raymond for an example used in § 7.

2. Notation.

Let X be a completely regular space. Then we denote by $\mathscr{C}(X)$ (resp. $\mathscr{C}^b(X)$) the space of all continuous real (resp. continuous and bounded) functions on X. If, moreover, X is partially ordered then we denote by $\mathscr{C}^{\uparrow}(X)$ (resp. $\mathscr{C}^{\downarrow}(X)$) the set of all increasing (resp. decreasing) *bounded* continuous real functions on X. The σ-algebra of Borel subsets of X will be denoted by $\mathscr{B}(X)$, the set of all *bounded* Radon measures on X by $\mathscr{M}(X)$. The positive cone in $\mathscr{M}(X)$ is denoted by $\mathscr{M}_+(X)$; $\mathscr{P}(X)$ is the set of all probability Radon measures on X. Given $f \in \mathscr{C}^b(X)$ and $\mu \in \mathscr{M}(X)$ we shall usually write $\mu(f)$ as a shorthand for $\int_X f\,d\mu$. The support of a measure $\mu \in \mathscr{M}_+(X)$ is denoted by $\operatorname{supp}\mu$. By a *compactification* X^ξ of X we mean a compact Hausdorff space X^ξ together with a function $\xi : X \to X^\xi$ that maps X homeomorphically onto a dense subspace $\xi(X)$ of X^ξ. In these circumstances we shall write

$$\mathscr{C}^\xi(X) = \{f \circ \xi : f \in \mathscr{C}(X^\xi)\}.$$

For example, if X^β is the Čech-Stone compactification of X then $\mathscr{C}^\beta(X) = \mathscr{C}^b(X)$.

Given a topological product $X \times Y$ of two completely regular spaces we normally denote by π_X, π_Y the natural projections of it onto the factor spaces X, Y. We also use the same symbols to denote the induced maps of $\mathscr{M}(X \times Y)$ onto $\mathscr{M}(X)$ and $\mathscr{M}(Y)$. But when dealing with a product written in the form $\prod_{\alpha \in A} X_\alpha$ it will be more convenient to denote the α^{th} natural projection by π_α.

Given a convex set K we shall denote by K_e the set of all extreme points of K.

Other notations will be explained as they arise.

3. Products of compact spaces.

We treat first the case of compact spaces, not merely because of its attractive simplicity, but also because it is the foundation of our approach to non-compact spaces.

We consider first the product of two spaces.

THEOREM 3.1. — *Let* X , Y *be compact Hausdorff spaces, let* $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, *and let* K *be a non-empty vaguely compact convex subset of* $\mathscr{P}(X \times Y)$. *Then the following statements are equivalent:*

(i) *there exists a measure* $\theta \in K$ *such that*

$$\pi_X\theta = \mu , \qquad \pi_Y\theta = \nu ;$$

(ii) *for all* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *we have*

$$\mu(u) + \nu(v) \leqslant \max \{\lambda(u \circ \pi_X + v \circ \pi_Y) : \lambda \in K\} ;$$

(iii) *whenever* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *are such that*

$$\lambda(u \circ \pi_X + v \circ \pi_Y) \geqslant 0 \qquad (\lambda \in K) ,$$

we have $\mu(u) + \nu(v) \geqslant 0$.

This result is based on Theorem 7 of Strassen [10] who, however, worked with Polish spaces (and also introduced some additional complications concerning moments). We shall use Theorem 3.1 later to prove *inter alia* a mild generalization of part of Strassen's result. The proof of Theorem 3.1 is a simple application of the Hahn-Banach theorem.

Replacing u, v by $-u, -v$ respectively, we see that statement (ii) is equivalent to

(ii′) *for all* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *we have*

$$\min \{\lambda(u \circ \pi_X + v \circ \pi_Y) : \lambda \in K\} \leqslant \mu(u) + \nu(v) .$$

It is obvious that (ii′) implies (iii). On the other hand if (iii) is true and if $u \in \mathscr{C}(X)$, $v \in \mathscr{C}(Y)$ and

$$\alpha \equiv \min \{\lambda(u \circ \pi_X + v \circ \pi_Y) : \lambda \in K\}$$

then, on writing $w(y) \equiv v(y) - \alpha$, we obtain

$$\lambda(u \circ \pi_X + w \circ \pi_Y) \geqslant 0$$

for all $\lambda \in K$, and hence $\mu(u) + \nu(w) \geqslant 0$, from which statement (ii′) follows at once. Thus (ii) and (iii) are equivalent.

Next, we note that (i) implies (iii) trivially. To complete the proof we show that statement (ii) implies (i). For this we write, for all $f \in \mathscr{C}(X \times Y)$,

$$p(f) = \max \{\lambda(f) : \lambda \in K\} .$$

Then p is a real sublinear functional on $\mathscr{C}(X \times Y)$. Consider the linear subspace

$$L = \{u \circ \pi_X + v \circ \pi_Y : u \in \mathscr{C}(X), v \in \mathscr{C}(Y)\}$$

of $\mathscr{C}(X \times Y)$ and define $\theta : L \to \mathbf{R}$ via

$$\theta(u \circ \pi_X + v \circ \pi_Y) = \mu(u) + \nu(v).$$

It is easy to see that θ is well defined and linear and that $\theta(1) = 1$. By (ii) we have $\theta(f) \leqslant p(f)$ for all $f \in L$. By the Hahn-Banach theorem θ can be extended to a linear functional (call it θ again) on $\mathscr{C}(X \times Y)$ such that $\theta(f) \leqslant p(f)$ for all $f \in \mathscr{C}(X \times Y)$. If $f \leqslant 0$ then $\theta(f) \leqslant p(f) \leqslant 0$. Consequently θ is positive and $\theta(1) = 1$, whence $\theta \in \mathscr{P}(X \times Y)$. It is obvious that $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$. Finally we note that because

$$K = \cap \{P_h : h \in \mathscr{C}(X \times Y)\},$$

where

$$P_h = \{\rho \in \mathscr{P}(X \times Y) : \rho(h) \leqslant p(h)\},$$

we have $\theta \in K$. This completes the proof.

A special case of some interest is that in which K is of the form

$$\{\theta \in \mathscr{P}(X \times Y) : \theta_1 \leqslant \theta \leqslant \theta_2\}$$

where θ_1, θ_2 are given elements of $\mathscr{M}(X \times Y)$.

In some of the applications of Theorem 3.1 the following observation is useful.

Addendum to Theorem 3.1. — *In Theorem* 3.1 *conditions* (ii) *and* (iii) *are respectively equivalent to :*

(ii)$_e$ *for all* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *we have*

$$\mu(u) + \nu(v) \leqslant \sup \{\lambda(u \circ \pi_X + v \circ \pi_Y) : \lambda \in K_e\};$$

(iii)$_e$ *whenever* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *are such that*

$$\lambda(u \circ \pi_X + v \circ \pi_Y) \geqslant 0 \qquad (\lambda \in K_e),$$

we have $\mu(u) + \nu(v) \geqslant 0$.

This follows from the Krein-Milman theorem. By way of application we have the following result.

Corollary 3.2. — *Let* X, Y, μ, ν *be as in Theorem* 3.1 *and let* R *be a non-empty closed subset of* X $\times$ Y. *Then the following statements are equivalent:*

(i) *there exists a measure* $\theta \in \mathscr{P}(X \times Y)$, *with* supp $\theta \subseteq R$, *such that* $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$;

(ii) *for all* $u \in \mathscr{C}(X)$ *and* $v \in \mathscr{C}(Y)$ *we have*

$$\mu(u) + \nu(v) \leqslant \max\{u(x) + v(y) : (x, y) \in R\} ;$$

(iii) *if* $u \in \mathscr{C}(X)$, $v \in \mathscr{C}(Y)$ *and* $u(x) + v(y) \geqslant 0$ *for all* $(x, y) \in R$ *then* $\mu(u) + \nu(v) \geqslant 0$.

To see this we need only take

$$K = \{\theta \in \mathscr{P}(X \times Y) : \text{supp}\,\theta \subseteq R\}$$

and apply Theorem 3.1 (with Addendum).

In many applications it is important to have criteria for the existence of θ expressed in terms of sets rather than functions. For example (compare [10]) in connection with Corollary 3.2 we have the following result.

Proposition 3.3. — *Let* X, Y *be compact Hausdorff spaces, let* $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$ *and let* R *be a closed non-empty subset of* X $\times$ Y. *Then the following statements are equivalent:*

(i) *there exists a measure* $\theta \in \mathscr{P}(X \times Y)$, *with* supp $\theta \subseteq R$, *such that* $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$;

(ii) *for all* $E \in \mathscr{B}(X)$ *and* $F \in \mathscr{B}(Y)$ *such that*

$$\left(E \times \complement F\right) \cap R = \varphi \qquad \textit{we have} \qquad \mu(E) \leqslant \nu(F) ;$$

(iii) *for all open* $U \subseteq X$ *and open* $V \subseteq Y$ *such that*

$$\left(U \times \complement V\right) \cap R = \varphi \qquad \textit{we have} \qquad \mu(U) \leqslant \nu(V) .$$

The implication (i) $\Longrightarrow$ (ii) requires only an easy calculation; (ii) $\Longrightarrow$ (iii) is evident. To prove that (iii) $\Longrightarrow$ (i) let $u \in \mathscr{C}(X)$, $v \in \mathscr{C}(Y)$ be such that $u(x) \leqslant v(y)$ for all $(x, y) \in R$ and let

$$U_t = \{x \in X : u(x) > t\} , \qquad V_t = \{y \in Y : v(y) > t\}$$

for all real t. Then $\left(U_t \times \complement V_t\right) \cap R = \varnothing$ and so, if (iii) is assumed, we have $\mu(U_t) \leqslant \nu(V_t)$. Adding a positive constant to u, v if necessary, we can suppose that they are positive. We now have

$$\mu(u) = \int_0^\infty \mu(U_t)\, dt \leqslant \int_0^\infty \nu(V_t)\, dt \leqslant \nu(v) .$$

It follows that statement (iii) of Corollary 3.2 has been verified, and the existence of θ is therefore proved.

For much more information about criteria involving sets rather than functions see [10].

The foregoing results, with the exception of Proposition 3.3, extend at once to finite products of compact spaces. Moreover, one can go to infinite products in the following way.

THEOREM 3.4. — *Let* $(X_\alpha)_{\alpha\in A}$ *be a non-empty family of compact Hausdorff spaces and let* X *be the topological product* $\prod_{\alpha\in A} X_\alpha$. *Suppose that* $(\mu_\alpha)_{\alpha\in A}$ *is an element of* $\prod_{\alpha\in A} \mathcal{P}(X_\alpha)$ *and that* K *is a non-empty vaguely compact subset of* $\mathcal{P}(X)$. *Then the following statements are equivalent.*

(i) *there exists a measure* $\theta \in K$ *such that* $\pi_\alpha\theta = \mu_\alpha$ *for all* $\alpha \in A$;

(ii) *whenever* $n \geqslant 1$, $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in A^n$ *and* $u_{\alpha_r} \in \mathcal{C}(X_{\alpha_r})$ *for* $r = 1, 2, \ldots, n$, *we have*

$$\sum_{r=1}^n \mu_{\alpha_r}(u_{\alpha_r}) \leqslant \max \left\{\lambda\left(\sum_{r=1}^n u_{\alpha_r} \circ \pi_{\alpha_r}\right) : \lambda \in K\right\} ;$$

(iii) *whenever* $n \geqslant 1$, $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in A^n$ *and* $u_{\alpha_r} \in \mathcal{C}(X_{\alpha_r})$ *for* $r = 1, 2, \ldots, n$, *and*

$$\lambda\left(\sum_{r=1}^n u_{\alpha_r} \circ \pi_{\alpha_r}\right) \geqslant 0 \qquad (\lambda \in K) ,$$

we have $\sum_{r=1}^n \mu_{\alpha_r}(u_{\alpha_r}) > 0$.

The proof is a mild complication of that for Theorem 3.1 and may be left to the reader, who will also see that Corollary 3.2 may be extended in the same fashion.

4. Compact ordered spaces.

By a compact ordered space we shall understand, following Nachbin [7], a partially ordered compact Hausdorff space X the graph of whose partial ordering,

$$R = \{(x, y) \in X^2 : x \leqslant y\}$$

is a closed subset of the topological product X^2. Throughout this section X will be such a space. We recall that, by a theorem of Nachbin [7]. the space $\mathscr{C}^\uparrow(X)$ is *order determining* for X in the sense that, for all $(x, y) \in X^2$, we have $x \leqslant y$ if and only if $g(x) \leqslant g(y)$ for all $g \in \mathscr{C}^\uparrow(X)$. Thus if $x \nleqslant y$ then there exists an $h \in \mathscr{C}^\uparrow(X)$ such that $h(x) > h(y)$. It follows that $\mathscr{C}^\uparrow(X)$ separates the points of X and hence, by the Kakutani-Stone theorem, that the vector difference $\mathscr{C}^\uparrow(X) - \mathscr{C}^\uparrow(X)$ is a dense sublattice of $\mathscr{C}(X)$.

Given $\mu, \nu \in \mathscr{M}_+(X)$, we shall write $\mu \prec \nu$ whenever $\mu(g) \leqslant \nu(g)$ for all $g \in \mathscr{C}^\uparrow(X)$. Since $\pm 1 \in \mathscr{C}^\uparrow(X)$ this condition implies that $\mu(1) = \nu(1)$. Since $\mathscr{C}^\uparrow(X)$ generates a dense subspace of $\mathscr{C}(X)$ the relation $\prec$ is a partial ordering for $\mathscr{M}_+(X)$.

The first theorem of this section is as follows.

THEOREM 4.1. — *Suppose that* $\mu, \nu \in \mathscr{P}(X)$ *and that* $\mu \prec \nu$. *Then there exists a measure* $\theta \in \mathscr{P}(X^2)$ *such that*

(i) $\operatorname{supp} \theta \subseteq R$,

(ii) *the first and second marginals of* θ *are* μ, ν *respectively.*

A result in this spirit, but for Polish spaces, was given by Strassen [10] who, however, defined his ordering of measures in terms of sets rather than functions. (We shall deduce Strassen's result from Theorem 3.1 in § 7.) For finite X, Theorem 4.1 is an immediate consequence of Strassen's result; and both the finite case and Strassen's theorem have been shown by Preston (in [8] and an unpublished communication) to follow from the min-cut max-flow theorem of programming theory. Theorem 4.1 was announced in [3].

Our main object here is to give a proof of Theorem 4.1 that will generalize to higher products, but it may be of interest to give first a very short proof, even though it does not so generalize.

By a theorem of Nachbin (see p. 46 of [7]) and one of Priestley (Theorem 3 of [9]), compact ordered spaces have the following property : *Let* X *be a compact ordered space and let* $u, v \in \mathscr{C}(X)$ *be functions such that, for all* $(x, y) \in R$, $u(x) \leqslant v(y)$. *Then there exists a function* $w \in \mathscr{C}^{\uparrow}(X)$ *such that, for all* $x \in X$,

$$u(x) \leqslant w(x) \leqslant v(x) .$$

Admit this, and let u, v, w be as in the statement. Then

$$\mu(u) \leqslant \mu(w) \leqslant \nu(w) \leqslant \nu(v) .$$

We have thus proved that, whenever $u, v \in \mathscr{C}(X)$ with

$$-u(x) + v(y) \geqslant 0$$

for all $(x, y) \in R$ it follows that

$$\mu(-u) + \nu(v) \geqslant 0 .$$

Hence, by Corollary 3.2, a measure $\theta \in \mathscr{P}(X^2)$ with the desired properties (i) and (ii) exists.

The above proof will not generalize to higher products, and we shall therefore give another one which will. This second proof is an adaptation of some standard arguments of Choquet theory. (In fact the argument is a special case of parts of [2] and [11], but it is simpler for present purposes to argue directly)

When $f \in \mathscr{C}(X)$ and $x \in X$ we shall write

$$\hat{f}(x) = \inf \{g(x) : g \in \mathscr{C}^{\downarrow}(X), g \geqslant f\} .$$

This definition provides an upper semicontinuous decreasing function $\hat{f}: X \to \mathbf{R}$ that majorizes f but is bounded above by $\max_X f$.

Proposition 4.2. — *Let* $\mu \in \mathscr{M}_+(X)$ *and let* ν *be a linear functional on* $\mathscr{C}(X)$. *Then* $\nu(f) \leqslant \mu(\hat{f})$ *for all* $f \in \mathscr{C}(X)$ *if and only if* $\nu \in \mathscr{M}_+(X)$ *with* $\nu \succ \mu$.

For the reader's convenience we give the (short) proof.

Suppose that $\nu \in \mathscr{M}_+(X)$ and that $\nu \succ \mu$, and let $g \in \mathscr{C}^{\downarrow}(X)$. Then $\mu(g) \geqslant \nu(g)$. Since $\{g \in \mathscr{C}^{\downarrow}(X) : g \geqslant f\}$ is downward filtering with pointwise limit $\hat{f}$, it follows that $\mu(\hat{f}) \geqslant \nu(\hat{f}) \geqslant \nu(f)$.

Conversely, let ν be a linear functional on $\mathscr{C}(X)$ that satisfies $\nu(f) \leqslant \mu(\hat{f})$ for all $f \in \mathscr{C}(X)$. Then $\nu(f) \leqslant 0$ whenever $f \leqslant 0$, and hence $\nu \in \mathscr{M}_+(X)$. And if $g \in \mathscr{C}^{\downarrow}(X)$ then $\hat{g} = g$ and we have $\nu(g) \leqslant \mu(\hat{g}) = \mu(g)$, whence $\nu \succ \mu$.

Proposition 4.3. — *Let μ, μ_1, μ_2, ν belong to $\mathscr{M}_+(X)$ and suppose that $\mu \prec \nu$, $\mu = \mu_1 + \mu_2$. Then there exist measures $\nu_1, \nu_2 \in \mathscr{M}_+(X)$ such that $\mu_1 \prec \nu_1$, $\mu_2 \prec \nu_2$ and $\nu = \nu_1 + \nu_2$.*

If $f \in \mathscr{C}(X)$ then, by Proposition 4.2,

$$\nu(f) \leqslant \mu(\hat{f}) = \mu_1(\hat{f}) + \mu_2(\hat{f}).$$

If we write $p(f) = \mu_1(\hat{f})$, $q(f) = \nu(f) - \mu_2(\hat{f})$ then p is sublinear on $\mathscr{C}(X)$, q is superlinear, and $q(f) \leqslant p(f)$ for all $f \in \mathscr{C}(X)$. By a well-known sharpening of the Hahn-Banach theorem, there is a linear functional ν_1 on $\mathscr{C}(X)$ such that

$$q(f) \leqslant \nu_1(f) \leqslant p(f) \qquad (f \in \mathscr{C}(X)).$$

On taking $\nu_2 = \nu - \nu_1$ and using Proposition 4.2 we see that ν_1 and ν_2 have the desired properties.

Now let

$$S = \{(\mu, \nu) : \mu, \nu \in \mathscr{P}(X), \mu \prec \nu\}.$$

Then S is a compact convex subset of $\mathscr{P}(X)^2$, and we proceed to characterize the extreme points of S.

Theorem 4.4. — $S_e = \{(\varepsilon_x, \varepsilon_y) : x, y \in X, x \leqslant y\}$.

Let $(\mu, \nu) \in S$ and suppose that μ has a non-trivial convex decomposition

$$\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$$

with $\mu_1, \mu_2 \in \mathscr{P}(X)$. By Proposition 4.3, there exist measures $\nu_1, \nu_2 \in \mathscr{P}(X)$ such that $\nu_1 \succ \mu_1$, $\nu_2 \succ \mu_2$ and

$$\nu = \alpha\nu_1 + (1 - \alpha)\nu_2.$$

But now

$$(\mu, \nu) = \alpha(\mu_1, \nu_1) + (1 - \alpha)(\mu_2, \nu_2),$$

which shows that (μ, ν) is not extreme. It follows that if $(\mu , \nu) \in S_e$ then $\mu = \varepsilon_x$ for some $x \in X$. Reversing the order in X we see that we must likewise also then have $\nu = \varepsilon_y$ for some y in X. But since $\mathscr{C}^{\uparrow}(X)$ is order determining, we have $\varepsilon_x \prec \varepsilon_y$ if and only if $x \leqslant y$. The result is therefore clear.

We can now prove Theorem 4.1. For this it is enough by Corollary 3.2 to show that, for all $u , v \in \mathscr{C}(X)$,

$$(4.1) \quad \mu(u) + \nu(v) \leqslant \max \{u(x) + v(y) : (x , y) \in R\} ,$$

where R is the graph of the order in X. Since the map

$$S \ni (\mu , \nu) \longmapsto \mu(u) + \nu(v)$$

is affine and continuous on S it is enough, by the Krein-Milman theorem, to check that the inequality (4.1) is true for all $(\mu , \nu) \in S_e$. By Theorem 4.5 this means that we must show that, for all $(\xi , \eta) \in R$,

$$u(\xi) + v(\eta) \leqslant \max \{u(x) + v(y) : (x , y) \in R\} ,$$

which is obvious.

Proposition 4.5. — *Suppose that* $\mu , \nu \in \mathscr{P}(X)$ *and that* $\mu \prec \nu$. *Let* $f , g \in \mathscr{C}(X)$ *be such that* $f(x) \leqslant g(y)$ *whenever* $(x , y) \in R$. *Then* $\mu(f) \leqslant \nu(g)$.

This result was demonstrated during the first proof of Theorem 4.1. It also follows directly from Theorem 4.1 and Corollary 3.2.

The measure θ of Theorem 4.1 is in general not unique, as one can see by considering for example the matrices

$$\begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where the rows are to be numbered from the bottom, the columns from the left. Uniqueness does occur in the special case $\mu = \nu$.

PROPOSITION 4.6. — *If in Theorem* 4.1 *it is assumed that* $\mu = \nu$ *then* θ *is uniquely determined and, moreover,* $\operatorname{supp} \theta \subseteq \Delta$, where Δ *is the diagonal in the space* X^2.

Suppose if possible that $(x_0, y_0) \in \operatorname{supp} \theta$ with $x_0 < y_0$. Then $y_0 \not\preccurlyeq x_0$ and so we can find a function $f \in \mathscr{C}^{\uparrow}(X)$ such that $f(x_0) < f(y_0)$. Then

$$0 = \mu(f) - \mu(f) = \iint_{X^2} (f(y) - f(x))\theta\,(dx, dy) > 0,$$

so we have a contradiction, and it follows that $\operatorname{supp} \theta \subseteq \Delta$. It follows easily now that θ can be identified with the measure on Δ that is the image of μ under the map $X \ni x \longmapsto (x, x)$.

In Proposition 4.6, all the mass is carried by Δ. In the next result none of it is.

PROPOSITION 4.7. — *Let* $\mu, \nu \in \mathscr{P}(X)$ *with* $\mu \prec \nu$ *and suppose that* μ, ν *are mutually singular. Then any measure* $\theta \in \mathscr{P}(X^2)$ *that satisfies conditions* (i) *and* (ii) *of Theorem* 4.1 *must also satisfy the equation* $\theta(\Delta) = 0$.

To see this let θ_1 be the measure on X^2 defined by

$$\theta_1(E) = \theta(E \cap \Delta) \qquad (E \in \mathscr{B}(X^2)).$$

Then $0 \leqslant \theta_1 \leqslant \theta$. The two projections of θ_1 are equal, and their common value, ρ say, satisfies $0 \leqslant \rho \leqslant \mu$, $0 \leqslant \rho \leqslant \nu$. Since μ, ν are mutually singular, these inequalities imply that $\rho = 0$, whence $\theta_1 = 0$. Hence $\theta(\Delta) = 0$, as asserted.

COROLLARY 4.8 (Nachbin). — *Let* $\sigma \in \mathscr{M}(X)$ *be such that* $\sigma(f) \geqslant 0$ *for all* $f \in \mathscr{C}^{\uparrow}(X)$. *Then there exists a measure* $\theta \in \mathscr{M}_+(X^2)$ *such that*

(i) $\operatorname{supp} \theta \subseteq R$ *but* $\theta(\Delta) = 0$;

(ii) $\|\theta\| = 1/2\,\|\sigma\|$;

(iii) $\sigma(f) = \iint_{R \setminus \Delta} (f(y) - f(x))\theta\,(dx, dy)$ *for all* $f \in \mathscr{C}(X)$.

To see this consider the Jordan decomposition $\sigma = \sigma^+ - \sigma^-$ of σ. Then $\sigma^+, \sigma^- \in \mathscr{M}_+(X)$, $\sigma^- \prec \sigma^+$ and, without loss of generality, we can suppose that $\sigma^+, \sigma^- \in \mathscr{P}(X)$. By Theorem 4.1 there is a measure $\theta \in \mathscr{P}(X^2)$ supported by R and having σ^- , σ^+ as its first and second projections. By Proposition 4.7 this θ satisfies $\theta(\Delta) = 0$. Moreover since θ , σ^- , σ^+ are probabilities,

$$\|\theta\| = 1/2\,(\|\sigma^+\| + \|\sigma^-\|) = 1/2\,\|\sigma\| .$$

Property (iii) is evident. This completes the proof.

Nachbin's own proof of Corollary 4.8 was rather different from this one (see pp. 120-121 of [7]).

It should be remarked that Theorem 4.1 and Corollary 4.8 are in fact equivalent, as we shall now show by adapting a construction of Hommel [5]. To derive Theorem 4.1 from Nachbin's result let Y be a homeomorphic copy of X and let Z be the topological sum of X and Y . Partially order Z by taking (with a convenient abuse of notation) the set

$$\Gamma \equiv \{(x , y) : x \in \mathrm{X} , y \in \mathrm{Y} , x \leqslant y\} \cup \{(z , z) : z \in \mathrm{Z}\}$$

in Z^2 as the graph of the order relation. Let $\mu \in \mathscr{P}(\mathrm{X})$, $\nu \in \mathscr{P}(\mathrm{Y})$ with (abuse of notation) $\mu \prec \nu$. Interpret μ , ν as measures on Z and take $\sigma = \nu - \mu$ in Corollary 4.8. Then the measure θ provided by that corollary lives on

$$\Gamma \setminus \{(z , z) : z \in \mathrm{Z}\} = \{(x , y) \in \mathrm{X} \times \mathrm{Y} : x \leqslant y\}$$

and its marginals are μ , ν (again interpreted as measures on Z). Hence we get the conclusions of Theorem 4.1.

The route to Theorem 4.1 via Nachbin's theorem is somewhat similar to the proof given above that was based on Priestley's theorem. Like that argument, it has no natural generalization to higher products.

It can be shown that if μ , $\nu \in \mathscr{P}(\mathrm{X})$ then $\mu \prec \nu$ if and only if $\mu(\mathrm{E}) \leqslant \nu(\mathrm{E})$ for every increasing $\mathrm{E} \in \mathscr{B}(\mathrm{X})$. We postpone discussion of this type of criterion until § 7, and turn now instead to the question of higher products.

Suppose that A is a totally ordered set and let us, for the rest of this section, use R to denote the set of all those $(x_\alpha)_{\alpha \in \mathrm{A}}$ in X^A for which the map $\alpha \to x_\alpha$ is an increasing function from A into X . Let us also write now S for the set of all those $(\mu_\alpha)_{\alpha \in \mathrm{A}}$ in $\mathscr{P}(\mathrm{X})^\mathrm{A}$ for which the map $\alpha \to \mu_\alpha$ is increasing with respect to the partial order $\prec$ in $\mathscr{P}(\mathrm{X})$.

THEOREM 4.9. — *Let* $(\mu_\alpha)_{\alpha\in A} \in S$. *Then there exists a measure* $\theta \in \mathscr{P}(X^A)$ *such that*

(i) $\operatorname{supp}\theta \subseteq R$,

(ii) $\pi_\alpha\theta = \mu_\alpha \ (\alpha \in A)$.

We begin by considering the special case where $A = \{1, 2, \ldots, n\}$ with the usual order. Points of X^A, $\mathscr{P}(X)^A$ are then denoted by n-tuples. The crucial step in this case is to prove that

$$(4.2) \qquad S_e = \{(\varepsilon_{x_1}, \varepsilon_{x_2}, \ldots, \varepsilon_{x_n}) : (x_1, x_2, \ldots, x_n) \in R\}.$$

Suppose, for this, that $(\mu_1, \mu_2, \ldots, \mu_n) \in S$ and that one of these measures, μ_r say, has a non-trivial convex decomposition

$$\mu_r = t\sigma_r + (1-t)\rho_r,$$

where $0 < t < 1$ and $\sigma_r, \rho_r \in \mathscr{P}(X)$. Working backwards and forwards from μ_r with the aid of Proposition 4.3 (as in the proof of Theorem 4.4) we obtain elements

$$(\sigma_1, \sigma_2, \ldots, \sigma_n), \qquad (\rho_1, \rho_2, \ldots, \rho_n)$$

of S such that

$$\mu_j = t\sigma_j + (1-t)\rho_j \qquad (j = 1, 2, \ldots, n).$$

Since $\sigma_r \neq \rho_r$ this shows that $(\mu_1, \mu_2, \ldots, \mu_n)$ is not an extreme point of S. The equation (4.2) is therefore now clear. We can now apply Theorem 3.4 (see the remark at the end of § 3) with

$$K = \{\theta \in \mathscr{P}(X^A) : \operatorname{supp}\theta \subseteq R\}$$

to conclude the proof for this special case (see the final steps for Theorem 4.1).

To prove the general case it is evidently enough, by Theorem 3.4, to show that whenever $n \geq 1$, $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in A^n$ and $u_{\alpha_r} \in \mathscr{C}(X)$ for $r = 1, 2, \ldots, n$ we have

$$\sum_{r=1}^{n} \mu_{\alpha_r}(u_{\alpha_r}) \leq \max\left\{\sum_{r=1}^{n} u_{\alpha_r}(x_{\alpha_r}) : (x_\alpha)_{\alpha\in A} \in R\right\}.$$

But this inequality is a consequence of our argument for the case of finite A. The proof of Theorem 4.9 is therefore complete.

5. Completely regular spaces.

Here we consider Radon probability measures on completely regular spaces.

We first recall some facts about Radon measures. Let X be a completely regular space and let X^{ξ} be a compactification of X (see § 2). Then the map ξ, as a map of Borel measures (see § 2), maps $\mathcal{M}(X)$ injectively into $\mathcal{M}(X^{\xi})$, and maps probabilities to probabilities. If $\mu_1 \in \mathcal{M}_+(X^{\xi})$ then in order that there exist a measure $\mu \in \mathcal{M}_+(X)$ such that $\xi\mu = \mu_1$ it is necessary and sufficient that for each $\varepsilon > 0$ we can find a compact subset C of X such that $\mu_1(X^{\xi}\setminus\xi(C)) < \varepsilon$. On these matters see [1], especially Exposé n° 6.

In our first theorem we consider two completely regular spaces X, Y and their product $Z = X \times Y$. We take compactifications X^{ξ} of X and Y^{η} of Y and let $\zeta = \xi \times \eta$, $Z^{\zeta} = X^{\xi} \times Y^{\eta}$.

Proposition 5.1. — *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, $\theta_1 \in \mathcal{P}(Z^{\zeta})$ and suppose that the projections of θ_1 onto the factor spaces of $X^{\xi} \times Y^{\eta}$ are $\xi\mu$, $\eta\nu$. Then there is a unique measure $\theta \in \mathcal{P}(Z)$ such that $\zeta\theta = \theta_1$. Moreover, $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$.*

Let $\varepsilon > 0$ and choose compact sets $X_\varepsilon \subseteq X$ and $Y_\varepsilon \subseteq Y$ such that

$$\mu(X\setminus X_\varepsilon) < \frac{\varepsilon}{2}, \qquad \nu(Y\setminus Y_\varepsilon) < \frac{\varepsilon}{2}.$$

Then

$$\theta_1(\xi X_\varepsilon \times Y^{\eta}) = (\xi\mu)(\xi X_\varepsilon) = \mu(X_\varepsilon) > 1 - \frac{\varepsilon}{2}.$$

Similarly

$$\theta_1(X^{\xi} \times \eta Y_\varepsilon) > 1 - \frac{\varepsilon}{2}.$$

Hence

$$\theta_1(Z^{\zeta}\setminus\xi Z_\varepsilon) < \varepsilon .$$

where $Z_\varepsilon = X_\varepsilon \times Y_\varepsilon$. By our opening remarks this shows that there is a unique $\theta \in \mathcal{P}(Z)$ such that $\zeta\theta = \theta_1$.

Now let $E \in \mathcal{B}(X)$. Then there exists an $E_1 \in \mathcal{B}(X^{\xi})$

such that $E = \xi^{-1}(E_1)$. Then

$$\theta(E \times Y) = \theta(\zeta^{-1}(E_1 \times Y^{\eta})) = \theta_1(E_1 \times Y^{\eta}) = (\xi\mu)(E_1) = \mu(E) ,$$

and so $\pi_X\theta = \mu$. Similarly $\pi_Y\theta = \nu$.

Continuing with the same notation, we can now prove:

THEOREM 5.2. — *Let* $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ *and let* K *be a non-empty* $\sigma(\mathcal{M}(Z), \mathcal{C}^{\zeta}(Z))$ — *closed convex subset of* $\mathcal{P}(X \times Y)$. *Then the following statements are equivalent:*

(i) *there exists a measure* $\theta \in K$ *such that*

$$\pi_X\theta = \mu , \pi_Y\theta = \nu ;$$

(ii) *for all* $u \in \mathcal{C}^{\xi}(X)$, $v \in \mathcal{C}^{\eta}(Y)$ *we have*

$$\mu(u) + \nu(v) \leqslant \sup \{\lambda(u \circ \pi_X + v \circ \pi_Y) : \lambda \in K\} ;$$

(iii) *whenever* $u \in \mathcal{C}^{\xi}(X)$, $v \in \mathcal{C}^{\eta}(Y)$ *and*

$$\lambda(u \circ \pi_X + v \circ \pi_Y) \geqslant 0 \qquad (\lambda \in K)$$

we have $\mu(u) + \nu(v) \geqslant 0$.

The implications (i) $\Longrightarrow$ (ii) $\Longleftrightarrow$ (iii) are much the same as for Theorem 3.1. It will accordingly suffice for us to give the proof that (ii) $\Longrightarrow$ (i) .

Let us denote by $\tilde{K}$ the $\sigma(\mathcal{M}(Z^{\zeta}), \mathcal{C}(Z^{\zeta}))$ — closure in $\mathcal{M}(Z)$ of the convex set $\zeta(K)$. Then $\tilde{K}$ is a non-empty compact convex subset of $\mathcal{P}(Z^{\zeta})$ and it is easy to see that

$$\tilde{K} \cap \zeta(\mathcal{P}(Z)) = \zeta(K) . \tag{5.1}$$

Now assume that statement (ii) of Theorem 5.2 is true. Then, for all $u \in \mathcal{C}(Y^{\xi})$ and all $v \in \mathcal{C}(Y^{\eta})$ we have

$$(\xi\mu)(u) + (\eta\nu)(v) \leqslant \max \{\lambda(u \circ \pi_{X^{\xi}} + v \circ \pi_{Y^{\eta}}) : \lambda \in \tilde{K}\} .$$

By Theorem 3.1 there exists a measure $\theta_2 \in \tilde{K}$ such that $\pi_{X^{\xi}}\theta = \xi\mu$, $\pi_{Y^{\eta}}\theta = \eta\nu$. Proposition 5.1 now supplies a measure $\theta \in \mathcal{P}(Z)$ such that $\zeta\theta = \theta_1$, $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$.

COROLLARY 5.3. — *Let* X , Y , μ , ν , ξ , η *be as in Theorem* 5.2 *and let* R *be a non-empty closed subset of* $X \times Y$. *Then the following statements are equivalent:*

(i) *there exists a measure* $\theta \in \mathcal{P}(X \times Y)$, *with* $\operatorname{supp}\theta \subseteq R$, *such that* $\pi_X\theta = \mu$, $\pi_Y\theta = \nu$;

(ii) *for all* $u \in \mathscr{C}^{\xi}(X)$ *and* $v \in \mathscr{C}^{\eta}(Y)$ *we have*

$$\mu(u) + \nu(v) \leqslant \sup \{u(x) + v(y) : (x, y) \in R\} ;$$

(iii) *if* $u \in \mathscr{C}^{\xi}(X)$, $v \in \mathscr{C}^{\eta}(Y)$ *and* $u(x) + v(y) \geqslant 0$ *for all* $(x, y) \in R$ *then* $\mu(u) + \nu(v) \geqslant 0$.

To prove this one applies Theorem 5.2 with

$$K = \{\theta \in \mathscr{P}(X \times Y) : \operatorname{supp} \theta \subseteq R\},$$

noting that this choice makes the supremum in that theorem have the same value as in Corollary 5.3. The proof that this K is $\sigma(\mathscr{M}(Z), \mathscr{C}^{\zeta}(Z))$ — closed may be left to the reader.

Proposition 5.4. — *If in Proposition* 3.3 *the conditions are relaxed to allow* X, Y *to be completely regular spaces, the conclusions remain true.*

The proof is as before, except that one uses Corollary 5.3 in place of Corollary 3.2.

In order to treat infinite products we introduce some notation. We let $(X_n)_{n \geqslant 1}$ be an infinite sequence of completely regular spaces, with topological product $X = \prod_{n=1}^{\infty} X_n$. For each n we suppose given a compactification $X_n^{\xi_n}$ of X_n. We denote by ξ the product map $\prod_{n=1}^{\infty} \xi_n$, so that X^{ξ} can be identified with the topological product $\prod_{n=1}^{\infty} X_n^{\xi_n}$. The natural projections $X \to X_n$ and $X^{\xi} \to X_n^{\xi_n}$ will be denoted by π_n and $\tilde{\omega}_n$ respectively.

Proposition 5.5. — *Suppose that* $\mu_n \in \mathscr{P}(X_n)$ *for all* n, *that* $\theta_1 \in \mathscr{P}(X^{\xi})$ *and that* $\tilde{\omega}_n \theta_1 = \xi_n \mu_n$ *for all* n. *Then there is a unique measure* $\theta \in \mathscr{P}(X)$ *such that* $\xi\theta = \theta_1$. *Moreover* $\pi_n \theta = \mu_n$ *for all* n.

To see this choose $\varepsilon > 0$ and then a sequence $(\varepsilon_n)_{n \geqslant 1}$ of positive reals such that $\sum_{n=1}^{n} \varepsilon_n < \varepsilon$. For each n we can find a compact subset Y_n of X_n such that $\mu_n(X_n \setminus Y_n) < \varepsilon_n$.

Now let

$$E_n = \left(\prod_{j=1}^{n} \xi_j(Y_j)\right) \times \left(\prod_{k=n+1}^{\infty} X_k^{\xi_k}\right).$$

Then a simple calculation (see the proof of Proposition 5.1 for the idea) shows that

$$\theta_1(E_n) > 1 - (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n) .$$

Letting $n \to \infty$ we obtain

$$\theta_1\left(\prod_{j=1}^{\infty} \xi_j(Y_j)\right) \geqslant 1 - \prod_{k=1}^{\infty} \varepsilon_k > 1 - \varepsilon .$$

By the opening remarks of this section this implies that there is a unique $\theta \in \mathscr{P}(X)$ such that $\xi\theta = \theta_1$. The rest of the proof is similar to that for Proposition 5.1.

THEOREM 5.6. — *Let* $(\mu_n) \in \prod_{n=1}^{\infty} \mathscr{P}(X_n)$ *and let* K *be a non-empty* $\sigma(\mathscr{M}(X), \mathscr{C}^{\xi}(X))$ *— closed convex subset of* $\mathscr{P}(X)$. *Then the following statements are equivalent:*

(i) *there exists a* $\theta \in K$ *such that* $\pi_n\theta = \mu_n$ *for* $n = 1, 2, \ldots$;

(ii) *whenever* $m \geqslant 1$ *and* $u_r \in \mathscr{C}^{\xi_r}(X_r)$ *for* $r = 1, 2, \ldots, m$ *we have*

$$\sum_{r=1}^{m} \mu_r(u_r) \leqslant \sup \left\{ \lambda\left(\sum_{s=1}^{m} u_s \circ \pi_s\right) : \lambda \in K \right\} ;$$

(iii) *whenever* $m \geqslant 1$ *and* $u_r \in \mathscr{C}^{\xi_r}(X_r)$ *for* $r = 1, 2, \ldots, m$ *and, for all* $\lambda \in K$,

$$\lambda\left(\sum_{r=1}^{m} u_r \circ \pi_r\right) \geqslant 0$$

it follows that $\sum_{r=1}^{m} \mu_r(u_r) \geqslant 0$.

The proof is the (now) obvious extension of that for Theorem 5.2. One appeals to Theorem 3.4 and to Proposition 5.6 instead of (respectively) Theorem 3.1 and Proposition 5.1.

If in Theorem 5.2 we take X^{ξ} and Y^{η} to be the Čech-Stone compactifications then we obtain an extension to completely regular spaces of a special case of Strassen's Theorem 7 [10]. We have dealt here with more general compactifications because of the needs of § 6 below.

6. Completely regular ordered spaces.

Our aim here is to extend some of the results of § 4 to a reasonable class of non-compact ordered spaces. For this purpose Nachbin's [7] *completely regular ordered spaces* (CRO-spaces) will serve. He defines such a space to be a topological space X on which a partial order (with closed graph R) is defined such that:

(i) for each $a \in X$ and each neighbourhood G of a we can find continuous functions $f, g : X \to [0, 1]$ such that f is increasing, g is decreasing, $f(a) = g(a) = 1$, and $\min(f, g)$ is identically zero in G;

(ii) if $x, y \in X$ then $x \leqslant y$ if and only if $f(x) \leqslant f(y)$, for all $\mathscr{C}^{\uparrow}(X)$.

Such spaces are completely regular in the usual sense. A compact ordered space is always a CRO-space; so is every subspace of a CRO-space. A Euclidean space $\mathbf{R}^k$, with the ordering induced by its positive cone, is a CRO-space. More generally, let E be a topological abelian group with a partial order such that (i) the set $E_+ = \{x \in E : x \geqslant 0\}$ is closed, (ii) there is a neighbourhood base $\mathscr{U}$ at the origin such that if $0 \leqslant x \leqslant y \in U \in \mathscr{U}$ then $x \in U$. In these circumstances E is a CRO-space; this example includes many ordered topological vector spaces. (On all this see Chapter II of [7].)

Nachbin [7] (see Hommel [5] for a more convenient account) has shown that if X is a CRO-space then there exists a compactification X^γ of X such that:

(*a*) X^γ is a compact ordered space;

(*b*) if $x, y \in X$ then $x \leqslant y$ in X if and only if $\gamma(x) \leqslant \gamma(y)$ in X^γ;

(*c*) for every $f \in \mathscr{C}^{\uparrow}(X)$ there exists a function $f_1 \in \mathscr{C}^{\uparrow}(X^\gamma)$ such that $f_1 \circ \gamma = f$.

This compactification is unique up to isomorphisms of ordered topological spaces, and will be called the *Nachbin compactification* of X.

If X is a CRO-space and $\mu, \nu \in \mathscr{M}_+(X)$ we shall write $\mu \prec \nu$ if $\mu(f) \leqslant \nu(f)$ for all $f \in \mathscr{C}^{\uparrow}(X)$. This is evidently the case if and only if $\gamma(\mu) < \gamma(\nu)$ in $\mathcal{M}_+(X^\Upsilon)$. Consequently the relation $\prec$ is a partial order in $\mathcal{M}_+(X)$.

THEOREM 6.1. — *Let* X *be a* CRO-*space, and let* $\mu, \nu \in \mathcal{P}(X)$ *with* $\mu \prec \nu$. *Then there exists a measure* $\theta \in \mathcal{P}(X^2)$ *such that*

(i) $\theta(R) = 1$;

(ii) *the first and second marginals of* θ *are* μ *and* ν.

Since $\gamma(\mu) \prec \gamma(\nu)$ there exists, by Theorem 4.1, a measure $\theta_1 \in \mathcal{P}(X^\Upsilon \times X^\Upsilon)$, whose first and second marginals are $\gamma(\mu)$ and $\gamma(\nu)$, and which satisfies $\theta_1(\tilde{R}) = 1$, where $\tilde{R}$ is the graph in $X^\Upsilon \times X^\Upsilon$ of the order relation on X^Υ. By Proposition 5.1, there is a unique measure $\theta \in \mathcal{P}(X^2)$ such that $\theta_1 = (\gamma \times \gamma)(\theta)$; and this measure θ has μ, ν as its first and second marginals. It remains only for us to prove that $\theta(R) = 1$. By condition (*b*) in the definition of X^Υ we have

$$\tilde{R} \cap (\gamma(X) \times \gamma(X)) = (\gamma \times \gamma)(R).$$

Since $\theta_1 = (\gamma \times \gamma)(\theta)$ this implies that

$$\theta(R) = \theta((\gamma \times \gamma)^{-1}\tilde{R}) = \theta_1(\tilde{R}) = 1.$$

COROLLARY 6.2. — *Let* X, μ, ν *be as in Theorem* 6.1 *and let* $f, g \in \mathcal{C}^b(X)$ *be such that* $f(x) \leq g(y)$ *for all* $(x, y) \in R$. *Then* $\mu(f) \leq \nu(g)$.

To see this, consider $\int_R (g \circ \pi_2 - f \circ \pi_1)\, d\theta$, where θ is as in Theorem 6.1.

Corollary 6.2 generalizes Proposition 4.5. Other results of § 4 which carry over, *mutatis mutandis*, to the present situation are Proposition 4.6 (uniqueness of θ when $\mu = \nu$) and Proposition 4.7 ($\theta(\Delta) = 0$ if $\mu \prec \nu$ with μ, ν mutually singular). From this last remark it follows that we have the following generalization of Nachbin's theorem.

THEOREM 6.3. — *Let* X *be a* CRO-*space and let* $\sigma \in \mathcal{M}(X)$ *be such that* $\sigma(f) \geq 0$ *for all* $f \in \mathcal{C}^\uparrow(X)$. *Then there exists a measure* $\theta \in \mathcal{M}_+(X^2)$ *such that*

(i) $\operatorname{supp} \theta \subseteq R$ *but* $\theta(\Delta) = 0$;

(ii) $\|\theta\| = 1/2\,\|\sigma\|$;

(iii) $\sigma(f) = \iint_{R\setminus\Delta} (f(y) - f(x))\theta\,(dx, dy)$ *for all* $f \in \mathscr{C}^b(X)$.

The proof is the obvious extension of that for Corollary 4.8.

Hommel [5] has investigated rather fully the possibility of extending Nachbin's theorem to locally compact ordered spaces (see Theorem 4.2.5 and Corollary 4.2.8 in [5]). Theorem 6.3 shows that in part (1) of Hommel's Corollary 4.2.8. we can drop the local compactness condition, provided that we continue to insist that all measures be Radon.

We can extend Theorem 6.1 to a power X^A of X, for any countable totally ordered set A. We obtain the following theorem, in which R denotes the set of all $(x_\alpha)_{\alpha\in A} \in X^A$ such that $\alpha \to x_\alpha$ is an increasing map.

THEOREM 6.4. — *Let* X *be a* CRO-*space, let* A *be a totally ordered countable set, and let* $\alpha \to \mu_\alpha$ *be an increasing map of* A *into* $\mathscr{P}(X^A)$ *such that*

(i) supp $\theta \subseteq R$;

(ii) $\pi_\alpha\theta = \mu_\alpha$ *for all* $\alpha \in A$.

To prove this, let Y be the Nachbin compactification $X\gamma$ of X and consider the map $\alpha \longmapsto \gamma\mu_\alpha$ of A into $\mathscr{P}(Y)$. Since this map is increasing, Theorem 4.9 provides a measure $\theta_1 \in \mathscr{P}(Y^A)$ such that

$$\text{supp}\,\theta_1 \subseteq \tilde{R}, \qquad \tilde{\omega}_\alpha\theta_1 = \gamma\mu_\alpha \quad (\alpha \in A),$$

where $\tilde{R}$ is the set of all $(x_\alpha)_{\alpha\in A}$ in Y^A such that $\alpha \to x_\alpha$ is an increasing map, and $\tilde{\omega}_\alpha$ is the α th projection of Y^A onto Y. Since A is countable there exists, by Proposition 5.5, a measure $\theta \in \mathscr{P}(X^A)$ such that $\theta_1 = \gamma\theta$ and $\pi_\alpha\theta = \mu_\alpha$ for all α. It remains to prove that $\theta(R) = 1$, but this follows (see the proof of Theorem 6.1) from the relation

$$\tilde{R} \cap \gamma(X)^A = (\gamma^A)(R).$$

7. Further remarks on ordered spaces.

By an ordered completely regular space we shall mean a completely regular space with a partial order whose graph is closed. Such a space need not be a CRO-space (see p. 88 of [5]). The following result is very close to part of Theorem 4.2.4 in [5], and we shall merely sketch the proof.

THEOREM 7.1. — *Let* X *be an ordered completely regular space and let* μ , $\nu \in \mathscr{P}(X)$. *Then the following conditions are equivalent :*

(i) $\mu(f) \leqslant \nu(f)$ *for all increasing bounded upper semicontinuous* $f: X \to \mathbf{R}$;

(ii) $\mu(F) \leqslant \nu(F)$ *for all increasing closed* $F \subseteq X$;

(iii) $\mu(G) \leqslant \nu(G)$ *for all increasing open* $G \subseteq X$;

(iv) $\mu(f) \leqslant \nu(f)$ *for all increasing bounded lower semicontinuous* $f: X \to \mathbf{R}$.

If X *is in fact a* CRO-*space then these conditions are equivalent to :*

(v) $\mu(f) \leqslant \nu(f)$ *for all* $f \in \mathscr{C}^{\uparrow}(X)$.

To show that (i) $\Longrightarrow$ (ii) it suffices to take $f = 1_F$ in (i), with F a closed increasing subset of X. To prove that (ii) $\Longrightarrow$ (i) it suffices to consider the case of increasing upper semicontinuous $f: X \to [0, 1]$. For such, take

$$F(t) = \{x : f(x) \geqslant t\} \quad (t \in \mathbf{R})$$

and let $f_n = \frac{1}{n} \sum_{r=0}^{n-1} 1_{F\left(\frac{r}{n}\right)}$. Then, by (ii), $\mu(f_n) \leqslant \nu(f_n)$. But $0 \leqslant f_n - f \leqslant \frac{1}{n}$, and hence we see, letting $n \to \infty$, that $\mu(f) \leqslant \nu(f)$. We have thus proved that (i), (ii) are equivalent. The proof that (iii) $\Longleftrightarrow$ (iv) is similar, and we omit it.

To prove that (ii) $\Longrightarrow$ (iii) take G as in (iii), let $\varepsilon < 0$, and choose a compact set K such that

$$\mu(G) < \mu(K) + \varepsilon , \qquad \nu(G) < \nu(K) + \varepsilon .$$

Because K is compact the smallest increasing subset $i(K)$ of X that contains K is closed (see p. 44 of [7]). By condition (ii) we therefore have

$$\mu(G) < \mu(K) + \varepsilon \leqslant \mu(i(K)) + \varepsilon \leqslant \nu(i(K)) + \varepsilon \leqslant \nu(G) + \varepsilon .$$

Hence we obtain (iii). The proof that (iii) $\Longrightarrow$ (ii) is similar (consider $\complement$ G and $\complement$ F).

Now suppose that X is a CRO-space. Then (i) $\Longrightarrow$ (v). To complete the proof of Theorem 7.1 it will suffice to show that (v) $\Longrightarrow$ (*i*). To see this, consider

$$\iint_{R} (f(y) - f(x))\theta\,(dx\,,\,dy)\,,$$

where θ is as in Theorem 6.1 and f is bounded increasing and upper semicontinuous.

If X is an ordered completely regular space which is not a CRO-space then, as we shall see below, it can happen that the space $\mathscr{C}^{\uparrow}(X)$ does not determine the order in X. In that case we can find $a\,,\,b$ in X such that $(a\,,\,b) \notin R$ yet $f(a) \leqslant f(b)$ for all $f \in \mathscr{C}^{\uparrow}(X)$. Consequently $\varepsilon_a \prec \varepsilon_b$ in the sense of § 6, but there is no measure θ on R having $\varepsilon_a\,,\,\varepsilon_b$ as its first and second marginals.

For an example of an ordered completely regular space X for which $\mathscr{C}^{\uparrow}(X)$ does not determine the order I am indebted to M. J. Saint-Raymond. Here is his example. Let

$$S = \left\{\frac{1}{n} :\; n = 1\,,\,2\,,\,3\,,\,\ldots\right\} \cup \{0\}\,,$$
$$J = [0\,,\,1]\setminus Q\,,$$

and let $\{q_n\}$ be an enumeration of the set $Q \cap [0\,,\,1]$. Let X be the following subspace of $\mathbf{R}^2$:

$$X = (J \times \{0\}) \cup \left\{\left(q_n\,,\frac{1}{mn}\right) :\; m\,,\,n = 1\,,\,2\,,\,\ldots\right\} \cup (\{2^k : k = 1\,,\,2\,,\,\ldots\} \times S)\,.$$

One can show that X is a G_δ, and hence is a Polish space, and hence completely regular. We partially order X by postulating that

$$(x\,,\,0) \leqslant (y\,,\,0) \;\text{ if }\; (x\,,\,0) \in X\,,\;\; (y\,,\,0) \in X \;\text{ and }\; x \leqslant y\,;$$
$$\left(2^k\,,\frac{1}{n}\right) \leqslant \left(q_k\,,\frac{1}{kn}\right) \quad \text{for} \quad k\,,\,n = 1\,,\,2\,,\,3\,,\,\ldots\,.$$

The graph of this partial order is closed. However, one can easily show that every f in $\mathscr{C}^{\uparrow}(X)$ is constant on $X \cap (\mathbf{R} \times \{0\})$.

Accordingly $\mathscr{C}^{\uparrow}(X)$ does not determine the order, and the pathology indicated above (failure of conclusion of Theorem 6.1) occurs here.

To obtain a positive result for an ordered completely regular space X that is not CRO let us consider the following conditions for X :

O_1 : If G is an open subset of X then so is $i(G)$;
O_2 : If F is a closed subset of X then so is $i(F)$.

THEOREM 7.2. — *Let* X *be an ordered completely regular space which satisfies either of the conditions* O_1 , O_2 . *Let* μ , $\nu \in \mathscr{P}(X)$ *be such that the (mutually equivalent) conditions* (i) — (iv) *of theorem* 7.1 *are satisfied. Then there exists a measure* $\theta \in \mathscr{P}(X^2)$ *such that*

(i) $\theta(R) = 1$;

(ii) *the first and second marginals of* θ *are* μ *and* ν .

Assume that O_1 is true and let $u , v \in \mathscr{C}^b(X)$ be such that $u(x) \leqslant v(y)$ for all $(x , y) \in R$. For all real t let $G_t = \{x : u(x) > t\}$. Note that G_t , and hence $i(G_t)$, decreases as t increases. Hence, for all $x \in X$,

$$\inf \{t : x \notin i(G_t)\} = \sup \{t : x \in i(G_t)\} .$$

Let $w(x)$ be the (finite) common value of these two expressions. I claim that w is a lower semicontinuous increasing function such that $u \leqslant w \leqslant v$. Admit this for the moment. Then

$$\mu(u) \leqslant \mu(w) \leqslant \nu(w) \leqslant \nu(v) .$$

Consequently, by Corollary 5.3 a measure θ satisfying conditions (i) and (ii) exists. It remains for us to check that w has the properties asserted.

Suppose that $y \in X$, $s \in \mathbf{R}$ with $w(y) > s$. Choose t such that $w(y) > t > s$. Then

$$y \in i(G_t) \subseteq \{x : w(x) > s\} .$$

Since, by O_1 , $i(G_t)$ is open, this shows that w is lower semicontinuous.

Next suppose that x , t are such that $u(x) > t$. Then $x \in G_t \subseteq i(G_t)$ and hence $w(x) \geqslant t$. This shows that $u \leqslant w$.

Suppose now that y, t satisfy $w(y) > t$. Then $y \in i(G_t)$, and so $y \geqslant x$ for some $x \in G_t$. Then $t < u(x) \leqslant v(y)$. This proves that $w \leqslant v$.

The fact that w is increasing is clear because $t \longmapsto i(G_t)$ is a decreasing map. This concludes the proof when O_1 is true.

When O_2 is true a similar proof applies. For this case one takes

$$F_t = \{x : u(x) \geqslant t\}$$

and now defines $w(x)$ as the common value of the two sides of

$$\inf \{t : x \notin i(F_t)\} = \sup \{t : x \in i(F_t)\}.$$

One then shows that $u \leqslant w \leqslant v$ and that w is increasing and upper semicontinuous. The rest of the proof is as before. We leave the details of this case to the reader.

Conditions O_1 and O_2 are in general not satisfied. For example if $X = [0, 1]$ and

$$R = \Delta \cup \left\{\left(\frac{1}{4}, \frac{3}{4}\right)\right\}$$

then O_2 is true, but O_1 is false. On the other hand, if $X = \mathbf{R}^2$, with the partial order induced by its positive cone (i.e. positive quadrant) then O_1 is true but O_2 false. (Indeed, O_1 is always true for a Banach space with the partial order induced by a closed cone).

If we assume neither O_1 nor O_2, but take X to be Polish then a theorem of Strassen [10] provides a positive conclusion:

THEOREM 7.3 (Strassen). — *Let* X *be an ordered Polish space, let* $\mu, \nu \in \mathcal{P}(X)$, *and suppose that* $\mu(E) \leqslant \nu(E)$ *whenever* E *is an increasing analytic subset of* X. *Then there exist a measure* $\theta \in \mathcal{P}(X^2)$ *such that*

(i) $\theta(R) = 1$;

(ii) *the first and second marginals of* θ *are* μ *and* ν.

Here is a proof of Theorem 7.3 somewhat different from Strassen's. Let $u, v \in \mathcal{C}^b(X)$ be such that $u(x) \leqslant v(y)$

for all $(x, y) \in R$. Let $w : X \to \mathbf{R}$ be defined by

$$w(x) = \sup \{u(y) : y \leqslant x\} \quad (x \in X).$$

I claim that w is increasing and universally measurable and that $u \leqslant w \leqslant v$.

To prove that w is increasing and that $u \leqslant w \leqslant v$ is easy, so we omit this part. For the measurability of w observe that

$$\{y : w(y) > \alpha\} = \{y : \exists x \text{ such that } x \leqslant y, u(x) > \alpha\}.$$

Thus $\{y : w(y) > \alpha\}$ is a projection of the Borel set

$$R \cap (\{x : u(x) > \alpha\} \times X)$$

and hence is analytic. Hence w is universally measurable.

If now we can show that $\mu(w) \leqslant \nu(w)$ then the proof can be concluded as for Theorem 7.2.

To prove that $\mu(w) \leqslant \nu(w)$ we can suppose that $0 \leqslant w \leqslant 1$. Write

$$W(t) = \{x : w(x) \geqslant t\} \quad (t \in \mathbf{R})$$

and let $w_n = \frac{1}{n} \sum_{r=0}^{n-1} 1_{W\left(\frac{r}{n}\right)}$. Then $\mu(w_n) \leqslant \nu(w_n)$ and $0 \leqslant w_n - w \leqslant \frac{1}{n}$. Letting $n \to \infty$ we get $\mu(w) \leqslant \nu(w)$ as desired.

Note added in proof, 2 *August* 1978. Since this paper was submitted new material on the existence of probability measures with prescribed marginals has appeared in J. Hoffmann-Jørgensen's lectures « Probability in Banach spaces » in *Ecole d'Eté de Probabilités de Saint-Flour VI*-1976, Springer-Verlag, Berlin, 1977.

There is some overlap between our two accounts; but, as regards the strictly new material, the only significant overlap lies in the fact that Hoffmann-Jørgensen's Theorem I. 5.12 contains the present Theorem 7.2. CRO-spaces ar enot dealt with by Hoffmann-Jørgensen.

BIBLIOGRAPHIE

[1] A. Badrikian, Séminaire sur les fonctions aléatoires linéaires et les mesures cylindriques, Springer-Verlag, Berlin, 1970.

[2] D. A. Edwards, Choquet boundary theory for certain spaces of lower semicontinuous functions, in *Function algebras*, Scott Foresman and Co., Chicago 1966.

[3] D. A. Edwards, Measures on product spaces and the Holley-Preston inequalities, *Bull. Lond. Math. Soc.*, 8 (1976), 7.
[4] D. A. Edwards, On the Holley-Preston inequalities, to appear in *Proc. Roy. Soc. of Edinburgh*, Section A (Mathematics).
[5] G. Hommel, Increasing Radon measures on locally compact ordered spaces, *Rendiconti di Matematica*, 9 (1976), 85-117.
[6] J. H. B. Kemperman, On the FKG-inequality for measures on a partially ordered space (to appear).
[7] L. Nachbin, Topology and order, van Nostrand, Princeton, 1965.
[8] C. J. Preston, A generalization of the FKG inequalities, *Commun. Math. Phys.*, 36 (1974), 233-241.
[9] H. A. Priestley, Separation theorems for semi-continuous functions on normally ordered topological spaces, *J. Lond. Math. Soc.*, 3 (1971), 371-377.
[10] V. Strassen, The existence of probability measures with given marginals, *Ann. Math. Statist.*, 36 (1965), 432-439.
[11] G. F. Vincent-Smith, Filtering properties of wedges of affine functions, *Journ. Lond. Math. Soc.*, 8 (1974), 621-629.

Manuscrit reçu le 21 septembre 1977
Proposé par G. Choquet.

D. A. Edwards,
Mathematical Institute
24-29 St. Giles
Oxford, 0X1 3LB
Grande-Bretagne.