ISRAEL JOURNAL OF MATHEMATICS 78 (1992), 281–307

# A "PLANAR" REPRESENTATION FOR GENERALIZED TRANSITION KERNELS

BY

DOROTHY MAHARAM

*Department of Mathematics, Northeastern University*
*Boston, MA 02115 USA*

ABSTRACT

In [9], Mauldin, Preiss and von Weizsäcker have given a theorem representing transition kernels (atomless and between standard Borel spaces) by a planar model. Here, motivated by measure-theoretic as well as probabilistic considerations, we generalize by allowing the parametrizing space $X$ to be arbitrary, with an arbitrary $\sigma$-field of "Borel" subsets, and allowing the corresponding measures to have atoms. (We also, for convenience rather than generality, allow arbitrary finite measures rather than probability ones.) The transition kernel is replaced by a substantially equivalent one from $X$ to $X \times I$ that is "sectioned", hence completely orthogonal. This is shown to be isomorphic to a model in which the image space consists of 3 specifically defined subsets of $X \times \mathbb{R}$: an ordinate set (in which vertical sections have Lebesgue measure), an "atomic" set contained in $X \times (-\mathbb{N})$, and a "singular" set with null sections. The method incidentally produces and exploits a "reverse" transition kernel from $X$ to $X \times I$. Some further extensions are briefly discussed; in particular, allowing "uniformly $\sigma$-finite" measures (in the "standard" case) leads to a generalization that includes the planar representation theorem of Rokhlin [10] and the author [5]; cf. also [7, 2].

## 1. Introduction

1.1 Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be "Borel structures" in the sense of Mackey [4] (or "measurable spaces" in the sense of Halmos); that is, $\mathcal{A}$ is a given $\sigma$-field of subsets of the set $X$, and $\mathcal{B}$ is a $\sigma$-field of subsets of $Y$. We often refer to the members of $\mathcal{A}$ and $\mathcal{B}$ as "Borel sets", though in general there is no topology

Received March 26, 1991

involved. A "transition kernel" from $(X, \mathcal{B})$ to $(Y, \mathcal{B})$ (in the sense of [9]) is a map $\mu$ assigning to each $x \in X$ a probability measure $\mu_x$ on $(Y, \mathcal{B})$ in such a way that, for all $B \in \mathcal{B}$, the function $x \to \mu_x(B)$ is $\mathcal{A}$-measurable. In [9, p. 974] a representation theorem for transition kernels was proved under the assumptions that (a) the spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are uncountable and "standard" (in effect, isomorphic to the unit interval $I$ with its usual Borel sets), (b) the measures $\mu_x$ are non-atomic. Our object here is to consider the more general situation in which $(X, \mathcal{A})$ is an *arbitrary* Borel structure and the measures $\mu_x$ may have atoms. (But we shall still require $(Y, \mathcal{B})$ to be standard.) In fact, we allow the $\mu_x$'s to be arbitrary finite (non-negative, $\sigma$-additive) measures on $\mathcal{B}$, a situation later extended (9.4 below) to allow certain infinite measures (the "uniformly $\sigma$-finite" ones).

1.2 We start, then, with arbitrary Borel structures $(X, \mathcal{A})$, $(Y, \mathcal{B})$ and a given transition kernel $\mu$ assigning to each $x \in X$ a finite (non-negative, $\sigma$-additive) measure $\mu_x$ on $(Y, \mathcal{B})$. As a first step, we replace $\mu$ by another, substantially equivalent, transition kernel $\gamma$ incorporating additional structure (much as in [9]), by defining, for each $x \in X$ and $L \in \mathcal{A} \otimes \mathcal{B}$ (the $\sigma$-field of subsets of $X \times Y$ $\sigma$-generated by the "rectangles" $A \times B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$),

$$\gamma_x(L) = \mu_x(L_x) \qquad \text{where } L_x = \{y \in Y : (x, y) \in L\}$$

(it is easy to check that $L_x \in \mathcal{B}$). We have at once:

PROPOSITION: *$\gamma$ is a transition kernel from $(X, \mathcal{A})$ to $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ satisfying the condition*

$$(S) \qquad \text{for all } x \in X \text{ and } L \in \mathcal{A} \otimes \mathcal{B}, \quad \gamma_x(X \times L_x) = \gamma_x(L).$$

Conversely, a transition kernel $\gamma$ from $(X, \mathcal{A})$ to $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ that satisfies condition (S) always arises in this way from a unique transition kernel $\mu$ from $(X, \mathcal{A})$ to $(Y, \mathcal{B})$. (Define $\mu_x(K) = \gamma_x(X \times K)$ for each $x \in X$ and $K \in \mathcal{B}$.)

Because (S) implies that $\gamma_x(L)$ depends only on the section $L_x$, we shall say that a transition kernel satisfying (S) is "sectioned". It is then possible, and often convenient, to think of $\gamma_x$ as a measure on the "slice" $\{x\} \times Y$ rather than on $Y$ (taking $\gamma_x(L \cap (\{x\} \times Y))$ to be $\gamma_x(L)$), so that $\gamma$ can be regarded as a sort of "direct sum" of the slice-measures $\gamma_x, x \in X$, and $\gamma$ becomes "completely orthogonal" in the sense of [9]. Note, however, that here the slice $\{x\} \times Y$ need not be $\mathcal{A} \otimes \mathcal{B}$-measurable, because $\mathcal{A}$ need not contain singletons.

1.3 One can generalize further, allowing $X \times Y$ to be replaced by a "Borel" subset (say) $W \in \mathcal{A} \otimes \mathcal{B}$, so that $\gamma$ becomes a transition kernel from $(X, \mathcal{A})$ to $(W, \mathcal{A} \times \mathcal{B}|W)$ (where $\mathcal{A} \otimes \mathcal{B}|W$ means $\{W \cap L : L \in \mathcal{A} \otimes \mathcal{B}\}$ or equivalently $\{K \in \mathcal{A} \otimes \mathcal{B} : K \subset W\}$). It is still required to be "sectioned"; that is, for all $x \in X$ and "Borel" $L \subset W$,

$$\gamma_x(W \cap (X \times L_x)) = \gamma_x(L).$$

Under suitable conditions on $W, X$ and $Y$, our main result can be extended to this more general situation; we return to this in §9 below.

1.4 But from now on we shall specialize by requiring $(Y, \mathcal{B})$ to be an uncountable standard Borel structure – say the real line $\mathbb{R}$ or unit interval $I = [0, 1]$, with $\mathcal{B}$ the family of (genuine) Borel sets. The main theorem of this paper (8.1) asserts that every sectioned transition kernel from $(X, \mathcal{A})$ to $(X \times I, \mathcal{A} \otimes \mathcal{B})$ is isomorphic, in a measure-preserving and first-coordinate-preserving way, to a "model" sectioned transition kernel $\rho$ from $(X, \mathcal{A})$ to $(M, \mathcal{A} \otimes \mathcal{B}|M)$, of the type described in 1.3 above, but with $M$ ($\subset X \times \mathbb{R}$) having a simple and specific structure. (We refer to 8.1 for a detailed description of $M$; a rough description is in 1.5 below.)

1.5 The fact that we are allowing arbitrary finite (positive, $\sigma$-additive) measures, rather than only probability measures, is not in itself significant, for we can always reduce to the "probability" case by dividing each $\gamma_x$ by $\gamma_x(X \times Y)$ when $\gamma_x(X \times Y) \neq 0$, thus arranging that each slice has total measure either 0 or 1. But the extra generality here is convenient in §7, where the "continuous part" $\gamma^c$ of $\gamma$, and $\gamma$ itself, will not usually have the same total slice-measures. However, for the preliminary considerations in §§2–7, we shall make the simplifying and easily-removed assumption that each $\gamma_x$ has total measure 1.

On the other hand, the fact that we are allowing both $\gamma_x$-atoms and an arbitrary $\sigma$-field $\mathcal{A}$, does cause considerable complications, especially in connection with measurability. (If $(X, \mathcal{A})$ is a standard space, allowing $\gamma_x$-atoms does not present much difficulty.) Accordingly §§2, 3 are concerned (respectively) with elementary properties of $\mathcal{A}$-measurable and $\mathcal{A} \otimes \mathcal{B}$-measurable functions. In dealing with the latter, we use the technique of [6] (based on an idea of Ursell [11]) exploiting functions $f(x, y)$ that are measurable in $x$ and monotone in $y$. We then study the one- and two-variable distribution functions of the measures (§4). One curious consequence here (which will be useful in §6) is the existence of a "reverse" transition kernel $\gamma^r$, obtained from the pseudoinverse of the distribution function for $\gamma$. The atoms of the measures $\gamma_x$ are studied in §5; the main result here (5.4) is that these form a countable union of graphs of measurable functions. The $\mathcal{A}$-measurability of the union of the "intervals of discontinuity", and of the union of the "intervals of constancy", of the distribution function then follows in §6. The continuous part $\gamma^c$ of $\gamma$ (obtained by subtracting off the atomic part) is again a transition kernel (§7), which enables us to define a "proper" (i.e., first-coordinate-preserving) isomorphism of "most" of the non-atomic part onto an ordinate set $\mathcal{O}$. This leads to the sought-for "model" representation (§8), which gives a proper Borel isometry of $X \times I$ onto a model set $M \subset X \times \mathbb{R}$ consisting of 3 parts: the ordinate set $\mathcal{O}$, a sequence of "linear" sets corresponding to the atoms, and a null "singular set" (analogous to the "garbage set" in [2] and [7]). In the first instance, the displacing of the atoms produces a sequence of lacunae in the ordinate set, but these are later filled by a special device (8.5), for which I am indebted to A.H. Stone. Finally, in §9 we consider some extensions in which the transition kernel is further generalized – the image space $X \times I$ replaced by a "Borel" subset of $X \times \mathbb{R}$ (9.1 – 9.3) or the measures $\gamma_x$ allowed to be infinite (but uniformly $\sigma$-finite in a suitable sense) in 9.4 – 9.6. We conclude by remarking that the result in 9.6 (which concerns the case in which $(X, \mathcal{A})$ is a standard Borel structure) contains the planar decomposition theorem of Rokhlin [10] and the author [5]; cf. also [7,9].

## 2. $\mathcal{A}$-measurable functions

2.1 As in §1, let $(X, \mathcal{A})$ be a Borel structure. As usual, a function $f : X \to \mathbb{R}$ is said to be "$\mathcal{A}$-measurable", or "measurable" for short, if the sets $f^{-1}(-\infty, t)$ (or equivalently, $f^{-1}(-\infty, t])$ $(t \in \mathbb{R})$, and hence $f^{-1}(B)$ for Borel $B \subset \mathbb{R}$, are all in $\mathcal{A}$. As usual, the measurable functions form a real linear space: if $f, g$ are measurable then so are $fg$ and (if $g(x)$ is never 0) $f/g$; and the limit of a pointwise convergent sequence of measurable functions is measurable. Somewhat less trivially, we have:

2.2 Proposition: *Let $f : D \to [0, \infty) = \mathbb{R}^+$ be a non-negative real-valued function defined on a subset $D$ of $X$. Then the following are equivalent:*

(1) *$f$ is $\mathcal{A}$-measurable,*
(2) *the "upper ordinate set" $\{(x, t) : x \in D \text{ and } 0 \le t \le f(x)\}$ of $f$ is $\mathcal{A} \otimes \mathcal{B}$-measurable,*

(3) *the "lower ordinate set"* $\{(x,t) : x \in D$ *and* $0 \le t < f(x)\}$ *of* $f$ *is* $\mathcal{A} \otimes \mathcal{B}$*-measurable,*

(4) $D \in \mathcal{A}$*, the graph* $\Gamma(f)$ $(= \{(x, f(x)) : x \in D\})$ *is* $\mathcal{A} \otimes \mathcal{B}$*-measurable, and the projection map* $p$ *of* $\Gamma(f)$ *onto* $D$ *(where* $p(x, f(x)) = x$*) is a Borel isomorphism.*

*Remark:* In (1), the meaning is that $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$, the family of Borel sets of $\mathbb{R}$. In (4) it is understood that the Borel structure on $\Gamma(f)$ is $\mathcal{A} \otimes \mathcal{B}|\Gamma(f)$, and the Borel structure on $D$ is $\mathcal{A}|D$ (that is, $\{H \cap D : H \in \mathcal{A}\}$).

*Proof:* (1) $\Rightarrow$ (2): Assume (1), and define (for $n = 1, 2, \ldots$ and $k = 0, 1, \ldots$)

$$Q_{nk} = \{(x,t) : x \in D, \quad (k-1)/n < f(x) \le k/n, \quad 0 \le t \le k/n\},$$
$$Q_n = \bigcup_k Q_{nk}, \text{ and } Q = \bigcap_n Q_n.$$

Clearly $Q \in \mathcal{A} \otimes \mathcal{B}$; and it is easy to verify that $Q$ is the upper ordinate set of $f$. (2) $\Rightarrow$ (1): Assuming (2), let $t_0 \ge 0$ be given. The intersection of the upper ordinate set with $\{(x,t) : x \in D, t = t_0\}$, namely $\{(x,t) : x \in D, t = t_0$ and $f(x) \ge t_0\}$, is in $\mathcal{A} \otimes \mathcal{B}$; hence the set $\{x \in D : f(x) \ge t_0\}$ is in $\mathcal{A}$, as required. The proof that (1) and (3) are equivalent is similar.

To see that (1), (2) and (3) imply (4), we get from (1) that $D(= f^{-1}[0, \infty)) \in \mathcal{A}$ and from (2) and (3) that the graph $\Gamma(f)$ (=(upper ordinate set) $\setminus$ (lower ordinate set)) $\in \mathcal{A} \otimes \mathcal{B}$. The map $p$ is a bijection of $\Gamma(f)$ onto $D$, its inverse being $i \times f$ (where $i$ is the identity map on $D$); we must show that both maps are measurable here. If $C \subset D$ and $C \in \mathcal{A}$, then $(i \times f)(C) = \Gamma(f) \cap (C \times \mathbb{R}^+) \in \mathcal{A} \otimes \mathcal{B}$ so $p$ is measurable. All that remains is to show that if $L \subset \Gamma(f)$ and $L \in \mathcal{A} \otimes \mathcal{B}$ then $p(L) \in \mathcal{A}$. Consider the family $\mathcal{M}$ of all $J \in \mathcal{A} \otimes \mathcal{B}$ with the property $p(J \cap \Gamma(f)) \in \mathcal{A}$. It is not hard to see that $\mathcal{M}$ is a monotone class containing the (finitely additive) field generated by the "rectangles" $H \times K$ $(H \in \mathcal{A}, K \in \mathcal{B})$, so that $\mathcal{M}$ coincides with $\mathcal{A} \otimes \mathcal{B}$. In particular, if $L \subset \Gamma(f)$ and $L \in \mathcal{A} \otimes \mathcal{B} = \mathcal{M}$, then $p(L) \in \mathcal{A}$.

Finally, (4) implies (1): Assuming (4), let $t_0 \ge 0$ be given. Then

$$\{x \in D : f(x) \ge t_o\} = p(\Gamma(f) \cap (X \times [t_o, \infty)) \in \mathcal{A}.$$

2.3 We conclude this section with a result that will be needed in §8. Let $\mathcal{L}$ be the set of all $\mathcal{A}$-measurable functions $f : X \to \mathbb{R}^+$, where (as before) $\mathbb{R}^+$ denotes

$[0,\infty)$ with its usual Borel structure. The topology $\mathcal{C}$ of "close approximation" on $\mathcal{L}$ (cf. [1]) is defined as follows. A neighborhood base at $f \in \mathcal{L}$ consists of the sets

$$C(f,\varepsilon) = \{g \in \mathcal{L} : |f(x) - g(x)| < \varepsilon(x) \text{ for all } x \in X\}$$

where $\varepsilon$ is an everywhere positive $\mathcal{A}$-measurable function. We write $\mathcal{L}(\Phi)$ for the subspace

$$\{f \in \mathcal{L} : 0 \le f(x) \le \Phi(x) \text{ for all } x \in X\}$$

in the induced topology, where $\Phi$ is an everywhere positive $\mathcal{A}$-measurable function on $X$.

PROPOSITION: *The space $\mathcal{L}$, in the topology of close approximation, is a Baire space (that is, the intersection of a countable family of dense open sets is dense).*

The straightforward proof is exactly the same as for the special case $X = \mathbb{R}$, $\mathcal{A} =$ Lebesgue measurable sets, given in [1].

COROLLARY: *$\mathcal{L}(\Phi)$ is also a Baire space.*

For $\mathcal{L}(\Phi)$ is the closure of the open subspace $\{f \in \mathcal{L} : 0 \le f(x) < \Phi(x)\}$ of the Baire space $\mathcal{L}$.

## 3. $\mathcal{A} \otimes \mathcal{B}$-measurable functions

3.1 As in [6], our technique for controlling measurability of real-valued functions on $X \times I$ (or $X \times \mathbb{R}$) is one due originally to Ursell [11]; we require $\mathcal{A}$-measurability with respect to the first variable and monotoneity with respect to the second. We first recall some notations and results from [6, §2] (in which the present $\mathcal{A}$ was denoted by $\mathcal{B}(X)$).

Given an increasing (i.e., non-decreasing) $f : I \to I$, we make the unorthodox conventions $f_-(0) = 0$, $f_+(1) = 1$, but otherwise

$$f_-(t) = f(t-)) = \lim_{\delta \to 0+} (t - \delta) \quad \text{and} \quad f_+(t) = f(t+) = \lim_{\delta \to 0+} f(t + \delta),$$

as usual. Thus $f_-$ and $f_+$ are also increasing functions from $I$ to $I$; $f_-$ is continuous on the left, and $f_+$ on the right. Two increasing functions $f$ and $f'$ from $I$ to $I$ are "equivalent" if $\{t \in I : f(t) \ne f'(t)\}$ is (at most) countable, or (what comes to the same thing) $f$ and $f'$ agree on a dense subset of $I$ (and hence at their points of continuity). The equivalence class of $f$ consists precisely of all

$f' : I \to I$ such that $f_-(t) \le f'(t) \le f_+(t)$ for all $t \in I$. And for all $f'$ in this equivalence class, $f'_- = f_-$ and $f'_+ = f_+$.

A map $g : I \to I$ is a "pseudoinverse" of $f$ provided $f_+(g(s)) \ge s \ge f_-(g(s))$ for all $s \in I$, or equivalently if (for all $s, t \in I$) $t > g(s) \Rightarrow f(t) \ge s$ and $t < g(s) \Rightarrow f(t) \le s$. Then $g$ is also pseudoinverse to every (increasing) function equivalent to $f$. Moreover, $g$ must also be increasing, and $f$ is a pseudoinverse of $g$. The pseudoinverses $g$ of $f$ form a single equivalence class; we denote its largest member (the common value of $g_+$) by $f^{\leftarrow}$ (the "upper pseudoinverse" of $f$) and its smallest member by $f_{\leftarrow}$ (the "lower pseudoinverse" of $f$).

We define the "jump set" $J(f)$ to be the set $\{t \in I : f_-(t) < f_+(t)\}$. The "set of values of constancy of $f$" is $K(f) = \{p \in I : f^{-1}(p)$ has more than one point$\}$. Both $J(f)$ and $K(f)$ are countable (at most), and depend only on the equivalence class of $f$. For each $p \in K(f)$, $f^{-1}(p)$ is a nondegenerate interval ("interval of constancy") from $f_{\leftarrow}(p)$ to $f^{\leftarrow}(p)$; the endpoints themselves may or may not be included. The "constancy set" $C(f)$ was defined in [6, p.6] to be $f^{-1}(K(f)) = \bigcup\{f^{-1}(p) : p \in K(f)\}$; but it will be convenient here to change this definition, redefining

$$C(f) = \bigcup\{(f_{\leftarrow}(p), f^{\leftarrow}(p)] : p \in K(f)\}.$$

This differs from the previous definition only at endpoints in $J(f)$; so $C(f) \cup J(f)$ has the same meaning as in [6], and hence so also does the "good set" $G(f) = I\backslash(J(f) \cup C(f))$. Incidentally, the redefined $C(f)$ now depends only on the equivalence class of $f$, unlike the previously defined one.

If $g$ is any pseudoinverse of $f$, we have $K(g) = J(f)$ (and hence $J(g) = K(f)$); and the restriction $f|G(f)$ is a homeomorphism of $G(f)$ onto $G(g)$, the (genuine) inverse map being $g|G(g)$.

We note, for later use, two identities not in [6]. (As always in the present section, $f$ is an increasing function from $I$ to $I$, and $t \in I$.)

$$\begin{aligned} f[0,t] =& ([0, f(t+)]\backslash \bigcup\{[f(d-), f(d+]: d \in J(f), 0 \le d \le t\}) \\ & \cup \bigcup\{K(f) \cap \{f(d-), f(d+)\} : d \in J(f), 0 \le d < t\} \\ & \cup \{f(d) : d \in J(f), 0 \le d \le t\}, \end{aligned} \tag{1}$$

$$\begin{aligned} f[0,t) =& ([0, f(t-))\backslash \bigcup\{[f(d-), f(d+)] : d \in J(f), 0 \le d < t\}) \\ & \cup \bigcup\{\{f(d)\} \cup (K(f) \cap \{f(d-), f(d+)\}) : d \in J(f), 0 \le d < t\}. \end{aligned} \tag{2}$$

The verifications, though tedious, are straightforward and are omitted. As immediate consequences, we have

$$f[0,t] =([0,f(t+)]\setminus\bigcup\{[f(d-),f(d+)] : d\in J(f), 0\le d\le t\}) \cup \text{ countable set}, \tag{3}$$

$$f[0,t) =([0,f(t-))\setminus\bigcup\{[f(d-),f(d+)] : d\in J(f), 0\le d< t\}) \cup \text{ countable set}\ . \tag{4}$$

3.2 Now suppose $F : X\times I \to I$ is such that, for each $x\in X$, the map $F_x : I\to I$, where $F_x(t) = F(x,t)$, is increasing. As in [6, p.10] we write $F^t(x)$ for $F(x,t), F_{\leftarrow}(x,t)$ for $(F_x)_{\leftarrow}(t)$ and $F^{\leftarrow}(x,t)$ for $(F_x)^{\leftarrow}(t)$; similarly $F_-(x,t)$ denotes $(F_x)_-(t) = F_x(t-)$, and so on.

Two such functions $F$, $F'$ (from $X\times I$ to $I$, with $F_x$ and $F'_x$ increasing) are said to be "equivalent" provided $F_x$ and $F'_x$ are equivalent for each $x\in X$. It is easy to see that $F$ and $F'$ are equivalent if and only if (for all $x$ and $t$) $F_-(x,t)\le F'(x,t)\le F_+(x,t)$, so that $F_-$ and $F_+$ are the least and greatest members of the equivalence class of $F$; and that $F'$ and $F$ are equivalent if and only if $F' = F_-$ and $F'_+ = F_+$. Further, $F_{\leftarrow}$ and $F^{\leftarrow}$are equivalent; and if $F$ and $F'$ are equivalent we have $F'_{\leftarrow} = F_{\leftarrow}$ and $(F')^{\leftarrow} = F^{\leftarrow}$. Note that the first variable ($x$) plays a purely passive role in these considerations; it will be relevant mainly for questions of measurability.

We define $\tilde F : X\times I\to X\times I$ by $\tilde F(x,t) = (x, F(x,t))$ and have [6, p.8] that $\tilde F$ is measurable if $F$ is (of course the Borel structures involved are $(X\times I, \mathcal{A}\otimes\mathcal{B})$ and $(I,\mathcal{B})$, where $\mathcal{B}$ is the family of (genuine) Borel subsets of $I$). The key property is given by the following theorem from [6, p.8]:

THEOREM: *Suppose $F : X\times I\to I$ satisfies*

(i) *for each $t\in I$, the map $F^t : X\to I$ is measurable,*

(ii) *for each $x\in X$, the map $F_x : I\to I$ is increasing,*

(iii) *either (a) $F_x$ is continuous on the left, for all $x\in X$, or (b) is continuous on the right, for all $x\in X$.*

*Then $F$ is measurable, and consequently $\tilde F$ is measurable.*

We remark that if $F$ satisfies (i) and (ii) then so do $F_-$ and $F_+$, and they further satisfy (iii) and are therefore measurable. (In fact, $F_-$ is continuous on the left and $F_+$ is continuous on the right.)

One further result from [6, p. 10] will be useful:

(1) If $F$ satisfies conditions (i) and (ii) above, then $F^{\leftarrow}$ and $F_{\leftarrow}$ satisfy (i), (ii) and (iii), and are therefore measurable.

## 4. Distribution functions

4.1 Given a Borel probability measure $\nu$ on $I$, its distribution function $f$ (defined by: $f(t) = \nu[0,t), t \in I$) is an increasing function from $I$ to $I$ satisfying (for all $t \in I$), in the notation of §3,

$$\nu[0,t) = f(t) = f(t-) = f_-(t) \qquad \text{and} \qquad \nu[0,t] = f(t+) = f_+(t).$$

As is well known, the converse also holds. More precisely:

PROPOSITION: *Given an increasing function $f$ from $I$ to $I$, there is a unique Borel probability measure $\nu$ on $I$ such that $\nu[0,t) = f_-(t)$ and $\nu[0,t] = f_+(t)$ $(t \in I)$.*

This is implicit, for example, in [3, p.53]. We sketch a direct proof, since it is short. Let $\lambda$ denote Lebesgue measure on $I$. Using the notation of §3, define (for each Borel $H \subset I$)

$$\nu(H) = \lambda(f(G \cap H)) + \Sigma\{f(d+) - f(d-) : d \in J(f) \cap H\},$$

where $G = I \backslash (C(f) \cup J(f))$. (Note that $G \cap H$ is Borel, because $G$ is and $f|G$ is a homeomorphism.) Because $f(J(f) \cup C(f))$ is countable, $\lambda(f([0,t) \cap G)) = \lambda(f[0,t))$; so, from 3.1(4) and the definition of $\nu$, $\nu[0,t) = f(t-)$. Similarly $\nu[0,t] = f(t+)$ from 3.1(3). Since a finite Borel measure is determined by its values on intervals, $\nu$ is unique.

We remark that each $d \in J(f)$ will be a $\nu$-atom, of weight $f(d+) - f(d-)$; and $\nu$ is zero on $C(f) \backslash J(f)$. And if $f'$ is another increasing function from $I$ to $I$, it defines the same measure $\nu$ (as in the Proposition) if, and only if, $f'$ and $f$ are equivalent (in the sense of §3).

4.2 For an arbitrary finite Borel measure $\nu$ on $I$, we denote the non-atomic or "continuous" part of $\nu$ by $\nu^c$; that is, if $J$ denotes the (countable) set of atoms (singletons of positive measure) of $\nu$, we have for each $H \in \mathcal{B} = \mathcal{B}(I)$,

$$\nu^c(H) = \nu(H) - \Sigma\{\nu\{d\} : d \in J \cap H\}. \tag{1}$$

Thus $\nu^c$ is a Borel measure on $I$, though in general not a probability measure (even if $\nu$ is).

Correspondingly, if $f$ is an increasing function from $I$ to $I$, the "continuous part" of $f, f^c$, is defined by

$$f^c(t) = f(t) - f^s(t) \qquad (t \in I), \tag{2}$$

where $f^s$, the "saltus part", is given by

$$f^s(t) = \Sigma\{f(d+) - f(d-) : d \in J(f), d \leq t\} \qquad (t \in I). \tag{3}$$

Both $f^c$ and $f^s$ are increasing functions from $I$ to $I$, vanishing at 0; and $f^c$ is continuous. We also have, for all $t \in I$,

$$\begin{aligned} f^c(t) &= f_+(t) - \Sigma\{f_+(d) - f_-(d) : d \in J(f), d \leq t\} \\ &= f_-(t) - \Sigma\{f_+(d) - f_-(d) : d \in J(f), d < t\} \end{aligned} \tag{4}$$

whence

$$\text{if } f' \text{ is equivalent to} f, \text{ then } (f')^c = f^c. \tag{5}$$

In particular,

$$(f_-)^c = f^c = (f_+)^c. \tag{6}$$

Of course, we also have (from continuity)

$$(f^c)_- = f^c = (f^c)_+. \tag{7}$$

Since $f^c$ is continous and increasing, $f^c[0,t] = [0, f^c(t)]$, so that

$$\lambda(f^c[0,t)) = f^c(t) = \lambda(f^c[0,t]) \qquad (t \in I). \tag{8}$$

4.3 The next proposition asserts that if $f$ is the distribution function of $\gamma$ (as in 4.1) then $f^c$ is the distribution function of $\gamma^c$. More precisely:

PROPOSITION: *If $\nu$ is a finite Borel measure on $I$, and $f$ is a function (necessarily increasing) on $I$ such that (for all $t \in I$) $\nu[0,t) = f_-(t)$ and $\nu[0,t] = f_+(t)$, then $\nu^c[0,t) = f^c(t) = \nu^c[0,t]$ for all $t \in I$.*

*Proof:* First note that, for each $t \in I$,

$$\nu\{t\} = \nu[0,t] - \nu[0,t) = f_+(t) - f_-(t).$$

Hence, from 4.2(1),

$$\begin{aligned}\nu^c[0,t) &= \nu[0,t) - \Sigma\{f_+(d) - f_-(d) : d \in J(f), d < t\}\\ &= f^c(t) \text{ by 4.2(4).}\end{aligned}$$

Similarly $\nu^c[0,t] = f^c(t)$. ∎

4.4 Until further notice, $\gamma$ will denote a sectioned transition kernel from $(X, \mathcal{A})$ to $(X \times I, \mathcal{A} \otimes \mathcal{B})$ (where $\mathcal{B}$ is the family $\mathcal{B}(I)$ of Borel sets of $I$), as in 1.4, and such that each $\gamma_x(I) = 1$ ($x \in X$). We define the "distribution function" $F : X \times I \to I$ of $\gamma$ by setting $F(x,t) = \gamma_x[0,t)$ ($x \in X, t \in I$); thus $F_-(x,t) = F(x,t) = \gamma_x[0,t)$ and $F_+(x,t) = \gamma_x[0,t]$; $F_+$ and $F_-$ are equivalent in the sense of §3, and are respectively the largest and smallest members of their equivalence class. To simplify printing, we sometimes write $\underline{F}$ for $F_-$ and $\bar{F}$ for $F_+$; thus $\underline{F}_x(t) = F_-(x,t) = \gamma_x[0,t)$, and so on. (As usual, the parameter $x$ is relevant only for considerations of measurability.) Though $F_-$ is the same as $F$ here, we often write $F_-$ for symmetry of notation.

From these definitions we have immediately:

1 For each $t \in I$, the maps $\underline{F}^t$ and $\bar{F}^t$ (from $(X, \mathcal{A})$ to $(I, \mathcal{B})$) are measurable.

2 For each $x \in X$, the maps $\underline{F}_x$ and $\bar{F}_x$ (from $I$ to $I$) are increasing.

3 $\underline{F}_x$ is continuous on the left, and $\bar{F}_x$ is continuous on the right; also $\underline{F}_x(0) = 0$ and $\bar{F}_x(1) = 1$ (in harmony with the conventions in 3.1).

Theorem 3.2 now gives:

COROLLARY: *The four functions $\underline{F}, \bar{F}, \underline{F}^{\sim}$ and $\bar{F}^{\sim}$ are all $\mathcal{A}$-measurable.*

4.5 As in the one-variable case, the "distribution function" determines the transition kernel. More precisely:

PROPOSITION: *Suppose $F : X \times I \to I$ satisfies*

(i) *for each $t \in I$, the map $F^t : X \to I$ is $\mathcal{A}$-measurable,*

(ii) *for each* $x \in X$*, the map* $F_x : I \to I$ *is increasing.*
*Let* $\mu_x$ *be the Borel probability measure on* $I$ *determined by* $F_x$ $(x \in X)$*, as in Proposition 4.1, and define* $\gamma'$ *by*

$$\gamma'_x(L) = \mu_x(L_x) \qquad (L \in \mathcal{A} \otimes \mathcal{B}).$$

*Then* $\gamma'$ *is a sectioned transition kernel from* $(X, \mathcal{A})$ *to* $(X \times I, \mathcal{A} \otimes \mathcal{B})$*, with each* $\gamma'_x(I) = 1$*; and the distribution function* $F'$ *of* $\gamma'$ *is equivalent to* $F$*.*

*Proof:* In view of the considerations of 1.2, it is enough to prove that $\mu$ is a transition kernel from $(X, \mathcal{A})$ to $(I, \mathcal{B})$; and the only thing requiring proof here is that, for each $H \in \mathcal{B}$ the function $x \mapsto \mu_x(H)$ is $\mathcal{A}$-measurable. Let $\mathcal{H}$ denote the family of all $H \in \mathcal{B}$ for which this is true. It is easy to see that $\mathcal{H}$ contains all intervals in $I$, and thence that $\mathcal{H}$ contains the (finitely additive) field that they generate; and that $\mathcal{H} = \mathcal{B}$ then follows via the monotone class theorem. ∎

*Remarks:* (a) The difference between $\mu_x$ and $\gamma'_x$ here is really only a matter of notation, since $\gamma'_x$ can be thought of as $\mu_x$ transferred to the slice $\{x\} \times I$.

(b) Functions equivalent to $F$ determine the same transition kernel.

(c) If $F$ satisfies conditions (i) and (ii) above, then so do $F_-$ and $F_+$; and moreover each $\underline{F}_x$ is continuous on the left, and each $\bar{F}_x$ on the right (cf. 4.4(3)). ∎

4.6 Let $F : X \times I \to I$ satisfy conditions (i) and (ii) above (4.5). Then the upper and lower pseudoinverse functions, $F^{\leftarrow}$ and $F_{\leftarrow}$, also satisfy these conditions (see the proof of Lemma 3.5 of [6, p.11]), and they are equivalent. Hence, by Proposition 4.5, they both determine the same transition kernel (from $(X, \mathcal{A})$ to $(X \times I, \mathcal{A} \otimes \mathcal{B})$), which is in a sense reversed from that determined by $F$. In particular, if $F$ is (equivalent to) the distribution function of a transition kernel $\gamma$ (from $(X, \mathcal{A})$ to $(X \times I, \mathcal{A} \otimes \mathcal{B})$), $F_{\leftarrow}$ and $F^{\leftarrow}$ are equivalent to the distribution function of a kernel which we denote by $\gamma^r$, the "reverse" of $\gamma$. Since the pseudoinverses of $F_{\leftarrow}$ and $F^{\leftarrow}$ are equivalent to $F$, the reverse of $\gamma^r$ is $\gamma$. We shall make use of $\gamma^r$ in §6 below. It would be interesting to know the probabilistic significance of $\gamma^r$, if any.

## 5. The atomic set

5.1 Let $\gamma$ be a sectioned transition kernel (as in subsection 4.4) from $(X, \mathcal{A})$ to $(X \times I, \mathcal{A} \otimes \mathcal{B})$, each $\gamma_x$ being a probability measure. Let $F$ be the distribution function of $\gamma$ (cf. 4.4), so that, for each $x \in X$, $F_x$ is the distribution function of $\gamma_x$. Recall that $J(F_x)$ denotes the "jump set" of $F_x$ (3.1), and depends only on the equivalence class of $F_x$. We define $J_\gamma$, or $J$ for short, by

$$J = J_\gamma = \bigcup \{\{x\} \times J(F_x) : x \in X\} \subset X \times I.$$

Since $J(F_x)$ is the set of all atoms of $\gamma_x$, the "atomic set" $J$ of $\gamma$ satisfies

$$J = \{(x,t) : x \in X, t \in I, \gamma_x\{t\} > 0\}. \tag{1}$$

Our object in §5 is to show (Theorem 5.4 below) that $J$ is a countable union of graphs of $\mathcal{A}$-measurable functions. First we require some elementary results about $\mathcal{A}$-measurability.

For each $c > 0$, define

$$J(c) = \{(x,t) \in X \times I : \gamma_x\{t\} \leq c\}. \tag{2}$$

From 4.4, Corollary, we have that $F_-(= F)$ and $F_+$ are $\mathcal{A} \otimes \mathcal{B}$-measurable functions; hence so is their difference, the function $(x,t) \longmapsto \gamma_x\{t\}$. In particular, the sets $J$ and $J(c)$ are in $\mathcal{A} \otimes \mathcal{B}$. More generally, for each $B \in \mathcal{B}$, the set $B^*$ defined by

$$B^* = \{(x,t) \in X \times I : \gamma_x\{t\} \in B\} \tag{3}$$

will be in $\mathcal{A} \otimes \mathcal{B}$.

5.2 Let $L$ be an $\mathcal{A} \otimes \mathcal{B}$-measurable subset of $J$, fixed for the present. Note that its projection $\pi(L) = \{x \in X : L_x \neq \emptyset\}$ also equals $\{x \in X : \gamma_x(L) > 0\}$, and therefore

$$\pi(L) \in \mathcal{A}. \tag{1}$$

Since $\gamma_x(I)$ is finite, we have that, for each $c > 0$, the set $\{t \in L_x : \gamma_x\{t\} \geq c\}$ is finite. Accordingly we define, for each $x \in X$, the "weight"

$$w_L(x) = \max\{\gamma_x\{t\} : t \in L_x\}, \tag{2}$$

with the convention that $\max \emptyset = 0$. Note that

(3) $$\text{the function } w_L : X \to I \text{ is } \mathcal{A}\text{-measurable},$$

because, for all $t > 0$,

$$\{x \in X : w_L(x) \geq t\} = \{x \in X : \gamma_x(L \cap J(t)) > 0\} \in \mathcal{A}.$$

Let $L^\wedge$ denote $\{(x,t) \in L : \gamma_x\{t\} = w_L(x)\}$. We claim

(4) $$L^\wedge \in \mathcal{A} \otimes \mathcal{B}, \quad \pi(L^\wedge) = \pi(L), \quad \text{and} \quad w_{L^\wedge} = w_L.$$

For the function $(x,t) \longmapsto w_L(x)$ is $\mathcal{A} \otimes \mathcal{B}$-measurable, from (3); also, as noted in 5.1, so is $(x,t) \longmapsto \gamma_x\{t\}$, so that the set $\{(x,t) \in X \times I : w_L(x) - \gamma_x\{t\} = 0\}$ is in $\mathcal{A} \otimes \mathcal{B}$. And $L^\wedge$ is the intersection of this set with $L$. That $\pi(L^\wedge) \subset \pi(L)$ is trivial; the reverse inclusion is straightforward. That $w_{L^\wedge} = w_L$ is immediate from the definitions.

5.3 Now let $M$ be an $\mathcal{A} \otimes \mathcal{B}$-measurable subset of $J$ with the further property that, for each $x \in X$, the section $M_x$ is finite. Define the "height"of $M$ at $x$ by

(1) $$h_M(x) = \max M_x,$$

again with the convention $\max \emptyset = 0$ (so that, in particular, $h_M(x) = 0$ if $x \notin \pi(M)$). Then

(2) $$h_M : X \to I \text{ is } \mathcal{A}\text{-measurable},$$

because, for each $t > 0$, $\{x \in X : h_M(x) \geq t\} = \pi(M \cap (X \times [t,1])) \in \mathcal{A}$ in view of 5.2(1).

In particular, taking $M = L^\wedge$, we obtain from (2) and 5.2(4) that

(3) $$h_{L^\wedge} \text{ is } \mathcal{A}\text{-measurable and is zero outside } \pi(L).$$

Write $h_{L^\wedge}|\pi(L)$ as $g_L$, for short, and let $\Gamma_L$ denote its graph $\{(x,t) : x \in L, t = g_L(x)\}$. As the graph of an $\mathcal{A}$-measurable function, $\Gamma_L \in \mathcal{A} \otimes \mathcal{B}$ (Proposition 2.2); and we have

(4) $$\Gamma_L \subset L^\wedge \subset L \subset J.$$

We observe

(5) if $(x,t) \in L \backslash \Gamma_L$, then $\gamma_x\{t\} \le w_{L^\wedge}(x)$, with equality only if $t < g_L(x)$.

For (since $t \in L_x$) $\gamma_x\{t\} \le w_L(x) = w_{L^\wedge}(x)$ by 5.2(4). If there is equality, then $(x,t) \in L^\wedge$ and therefore $t \le h_{L^\wedge}(x) = g_L(x)$. But $t \ne g_L(x)$ (else $(x,t) \in \Gamma_L$), so $t < g_L(x)$.

5.4 Now write $J_1 = J, \Gamma_1 = \Gamma_J =$ graph of $g_1$ where $g_1 = g_J = h_{J^\wedge}|\pi(J)$, and define recursively, for $n = 2,3,\ldots, J_n = J_{n-1} \backslash \Gamma_{n-1}, \Gamma_n = \Gamma_{J_n} =$ graph of $g_n$ where

$$g_n = g_{J_n} = h_{(J_n)^\wedge}|\pi(J_n).$$

The results in 5.1, 5.2 and 5.3 show that all these sets and functions are $\mathcal{A}$-measurable. We now prove:

THEOREM: *The graphs $\Gamma_1, \Gamma_2, \ldots,$ are pairwise disjoint; their union is $J$; and their domains of definition decrease as $n$ increases.*

The last assertion (useful later) is trivial, because $J_1 \supset J_2 \supset \ldots$. From 5.3(4), $\Gamma_n \subset (J_n)^\wedge \subset J_n$, which is disjoint from $\Gamma_1 \cup \cdots \cup \Gamma_{n-1}$, so the graphs $\Gamma_1, \Gamma_2, \ldots,$ are pairwise disjoint. All that remains is to prove $\bigcup_n \Gamma_n = J$.

One inclusion here is trivial; for the other, fixing $x \in \pi(J)$, we show that $\{x\} \times J_x \subset \bigcup_n \Gamma_n$.

Keeping $x$ fixed throughout, define (for each $v \in J_x$) the "rank" of $v$ by

$$\begin{aligned} r(v) = 1 + {} & \operatorname{card}\{s \in J_x : \gamma_x\{s\} > \gamma_x\{v\}\} \\ & + \operatorname{card}\{t \in J_x : \gamma_x\{t\} = \gamma_x\{v\} \text{ and } t > v\} \end{aligned}$$

$= 1+$ number of points preceding $v$ in the "lexicographic" linear ordering of $J_x$ specified by:

$$s \succ t \Leftrightarrow \text{ either } \gamma_x\{s\} > \gamma_x\{t\}, \text{ or } \gamma_x\{s\} = \gamma_x\{t\} \text{ and } s > t.$$

If we list the members of $J_x$ in this order, as (say) $s_1 \succ s_2 \succ \ldots$, each $t \in J_x$ gets listed (because $\gamma_x\{t\} > 0$ and $t$ has only finitely many predecessors); and the rank of $s_r$ is $r$. Now we show, by induction over $r$, that $(x, s_r) \in \Gamma_r$. For $r = 1$ this follows from the fact that $s_1$ will be the highest of the points of greatest weight of $J_x$. And if it holds for all $r < k$, we note that the $k-1$ predecessors $s_1, s_2, \ldots, s_{k-1}$

of $s_k$ are in $(\Gamma_1 \cup \cdots \cup \Gamma_{k-1})_x$, so $(x, s_k) \in J \backslash (\Gamma_1 \cup \cdots \cup \Gamma_{k-1}) = J_k$ and $s_k$ will be the highest of the points of greatest weight of $J_k$ — that is, $s_k = h_{(J_k)^\wedge}(x)$— showing $(x, s_k) \in \Gamma_k$. Thus $(x, s_r) \in \Gamma_r$ for $r = 1, 2, \ldots$, and $\{x\} \times J_x \subset \bigcup_n \Gamma_n$, as required.

## 6. The intervals of discontinuity and of constancy

6.1 With $\gamma$ and $F$ as in 4.4 and §5, we say that an interval of the form $(\gamma_x[0,t),\ \gamma_x[0,t]]$, where $\{t\}$ is an atom of $\gamma_x$ — in other words, an interval of the form $(F_x(t-), F_x(t+)]$ where $(x,t) \in J_\gamma$ — is an "interval of discontinuity" over $x$; and we denote the corresponding union (over $x \in X$) by $D_\gamma$, or $D$ for short. More precisely,

$$D = D_\gamma = \bigcup \{\{x\} \times (\gamma_x[0,t), \gamma_x[0,t]] : (x,t) \in J_\gamma\}.$$

PROPOSITION: $D \in \mathcal{A} \otimes \mathcal{B}$.

*Proof:* Since $J_\gamma = \bigcup_n \Gamma_n$ (Theorem 5.4), we can write $D = \bigcup_n D_n$ where $D_n = \bigcup\{\{x\} \times (\gamma_x[0,t), \gamma_x[0,t]] : (x,t) \in \Gamma_n\}$; and it will suffice to prove that $D_n \in \mathcal{A} \otimes \mathcal{B}$ for each $n$. For the rest of the argument, $n$ is fixed; and to save notation we omit the suffix $n$. Thus we are concerned with $(x,t) \in \Gamma$ where $\Gamma$ is the graph of the measurable function $g$ (abbreviating $g_n$) as in 5.4. By Proposition 2.2, the upper and lower ordinate sets of $g$, namely

$$\{(x,y) : x \in \pi(\Gamma), 0 \le y \le g(x)\} \text{ and } \{(x,y) : x \in \pi(\Gamma), 0 \le y < g(x)\},$$

are in $\mathcal{A} \otimes \mathcal{B}$. Because $\gamma$ is a transition kernel it follows that the functions defined by $x \longmapsto \gamma_x[0, g(x)]$ and by $x \longmapsto \gamma_x[0, g(x))$ for $x \in \pi(\Gamma)$ (and zero on $X \backslash \pi(\Gamma)$) are $\mathcal{A}$-measurable. Hence the upper (and lower) ordinate sets of these last two functions are in $\mathcal{A} \otimes \mathcal{B}$; and on intersecting them with $(X \backslash \pi(\Gamma)) \times I$ we have that (in particular) the sets

$$\mathcal{O}_1 = \{(x,s) : x \in \pi(\Gamma), 0 \le s \le \gamma_x[0, g(x)]\}$$

and

$$\mathcal{O}_2 = \{(x,s) : x \in \pi(\Gamma), 0 \le s \le \gamma_x[0, g(x))\}$$

are in $\mathcal{A} \otimes \mathcal{B}$. But $\mathcal{O}_1 \backslash \mathcal{O}_2 = D_n$, proving $D_n \in \mathcal{A} \otimes \mathcal{B}$ as required. ∎

6.2 REMARK: *A similar argument will prove that the 3 other sets produced by replacing the intervals $(\gamma_x[0,t),\ \gamma_x[0,t]]$ in $D$ throughout by their closures, or by their interiors, or by $[\gamma_x[0,t),\ \gamma_x[0,t])$ respectively, are also in $\mathcal{A}\otimes\mathcal{B}$.*

6.3 As in 3.1, the set of "values of constancy" of $F_x$ is by definition $K(F_x) = \{p \in I : (F_x)^{-1}(p)$ has more than one point$\}$. We define the set of values of constancy of $F$, denoted by $K(F)$, to be $\bigcup\{K(F_x) : x \in X\}$. It is easy to see that $K(F)$ is not altered if $F$ is replaced by an equivalent function (e.g., $F^+$), and accordingly we sometimes write $K_\gamma$ instead of $K(F)$.

Now consider any measurable pseudoinverse of $F$ (3.2), say $F^{\leftarrow}$ for definiteness. As in 4.6, $F^{\leftarrow}$ is (equivalent to) the distribution function of the "reverse" transition kernel $\gamma^r$. From [6, p.7] we have that $K(F_x) = J((F^{\leftarrow})_x)$ and $J(F_x) = K((F^{\leftarrow})_x)$ — that is, the jumps of the one function are exactly the constancy-values of the other. Hence we have

$$J_\gamma = K(F^{\leftarrow}) = K_{\gamma^r} \text{ and } K_\gamma = J_{\gamma^r}. \tag{1}$$

*Remark:* Hence, by 5.4 applied to $\gamma^r$, $K_\gamma$ too is a countable union of graphs of $\mathcal{A}$-measurable functions. ∎

6.4 We have (re)defined the "constancy set" $C(F_x)$ to be

$$\bigcup\{((F_x)_{\leftarrow}(p), (F_x)^{\leftarrow}(p)]; p \in K(F_x)\}$$

(cf. 3.1), and we now define the "constancy set" $C(F)$, or $C_\gamma$, or $C$ for short, to be

$$\bigcup\{\{x\}\times C(F_x) : x \in X\} = \bigcup\{\{x\}\times((F_x)_{\leftarrow}(p), (F_x)^{\leftarrow}(p)] : (x,p) \in K_\gamma\}.$$

From [6, 2.(13)] and 6.3 above, we can write this as

$$\bigcup\{\{x\}\times((\gamma^r)_x[0,p), (\gamma^r)_x[0,p]] : (x,p) \in J_{\gamma^r}\};$$

thus $C_\gamma = D_{\gamma^r}$, and 6.1 gives:

PROPOSITION: $C_\gamma \in \mathcal{A}\otimes\mathcal{B}$.

Of course, it also follows that $D_\gamma = C_{\gamma^r}$.

From the Remark (6.2) we see that if the intervals of constancy were all redefined so as to include (for all of them) both endpoints, or neither, or the left endpoint but not the right, the 3 resulting redefined constancy sets would also be in $\mathcal{A}\times\mathcal{B}$. However, the normalization adopted here (including the right endpoint but not the left) is convenient later on (in §7).

## 7. The continuous part

7.1 With $\gamma$ and $F$ as in §§5,6, we define the "continuous part" $\gamma^c$ of $\gamma$ by

$$\gamma^c(x, L) = \gamma_x^c(L) = \gamma_x(L \backslash J_\gamma) \quad (x \in X, L \in \mathcal{A} \otimes \mathcal{B})$$

and the "continuous part" $F^c$ of $F$ by

$$F_x^c = (F_x)^c \ \ (x \in X),$$

the continuous part of the distribution function $F_x$ of $\gamma_x$. Note that, from 4.2(5), equivalent $F$'s have the same continuous part.

We write $X^c = \{x \in X : \gamma_x^c(X \times I) > 0\}$; thus $X^c \in \mathcal{A}$ and for each $x \in X \backslash X^c$ the probability measure $\gamma_x$ is purely atomic. We use $(X^c, \mathcal{A})$ as an abbreviation for $(X^c, \mathcal{A}|X^c)$, and similarly for $(X^c \times I, \mathcal{A} \otimes \mathcal{B})$.

PROPOSITION: *$\gamma^c$ is a sectioned transition kernel from $(X^c, \mathcal{A})$ to $(X^c \times I, \mathcal{A} \otimes \mathcal{B})$; and its distribution function is $F^c|X^c \times I$.*

*Remarks:* (a) In general $\gamma_x^c(X^c \times I)$ will be less than 1: but the previous theory continues to apply, essentially unchanged; to justify it (as in 1.5) we need only imagine dividing $\gamma_x^c$ by $\gamma_x^c(X^c \times I)$, for each $x \in X^c$.

(b) Having shown that $\gamma^c$ is sectioned, we shall often use the convenient "abuse of notation" $\gamma_x^c(H)$ for $\gamma_x^c(X^c \times H)$, $H \in \mathcal{B}$.

*Proof of the Proposition:* Since $J_\gamma \in \mathcal{A} \otimes \mathcal{B}$, each $\gamma_x^c$ is a countably additive measure on $\mathcal{B}$; and since the function

$$x \longmapsto \gamma_x^c(L) = \gamma_x(L \backslash J_\gamma)$$

is $\mathcal{A}$-measurable (for fixed $L \in \mathcal{A} \otimes \mathcal{B}$), $\gamma_c$ is a transition kernel. Since $\gamma_x^c(L) = \gamma_x(X^c \times (L \backslash J_\gamma)x)$, which depends only on $x, \gamma^c$ is sectioned. From Proposition 4.3 and the facts that $\gamma_x[0, t) = (F_x)_-(t)$ and $\gamma_x[0, t] = (F_x)_+(t)$, we see that $F^c$ is (equivalent to) the distribution function of $\gamma^c$. ∎

7.2 COROLLARIES:

(1) *$F_x^c[0, t] = [0, F^c(x, t)]$ and hence*

$$\lambda(F_x^c[0, t]) = F^c(x, t) = \gamma_x^c[0, t] \qquad (x \in X^c, t \in I).$$

*(From 4.2(8).)*

(2) $F^c$, *as a map from* $(X^c \times I, \mathcal{A} \otimes \mathcal{B})$ *to* $(I, \mathcal{B})$, *is measurable. (From the remark following Theorem 3.2, since here* $F^c_- = F^c_+ = F^c$ *because* $\gamma^c_x$ *has no atoms.)*

7.3 Recalling (3.1) that $C(F^c_x)$ denotes the (revised) "constancy set" of $F^c_x$, we now show:

PROPOSITION: *For each* $x \in X^c$, $(F^c_x)_{\leftarrow}$ *is a bijection of* $[0, F^c_x(1)]$ *onto* $I \backslash C(F^c_x)$, *with inverse map* $F^c_x | (I \backslash C(F^c_x))$.

Since $x$ will be fixed throughout the proof, we simplify the notation by writing $f$ for $F^c_x$, $\nu$ for the corresponding measure $\gamma^c_x$, and $C$ for $C(F^c_x)$. Thus $f$ is a continuous increasing function from $I$ onto $[0, f(1)]$.

Enumerate $K(f)$ as $\{p_1, p_2, \ldots\}$, as in 3.1, and let $f_{-1}(p_n) = [a_n, b_n]$ $(n = 1, 2, \ldots)$. Thus $C = \bigcup\{(a_n, b_n] : n = 1, 2, \ldots\}$ where we note for future use that $\nu(a_n, b_n] = f(b_n) - f(a_n) = 0$ for each $n$, so that $\nu(C) = 0$. On $I \backslash C$, $f$ is *strictly* increasing, hence 1-1. Also $f(I \backslash C) = f(I)$, as can be seen as follows. Given $p \in f(I)$, there is some $s \in I$ such that $f(s) = p$. If $p \notin K(f)$ then $x \in I \backslash C$ as required; but if $p \in K(f)$, say $p = p_n$, then $a_n \in I \backslash C$ and $f(a_n) = p$. Thus in either case $p \in f(I \backslash C)$. The same reasoning shows that there is only *one* $s \in I \backslash C$ with $f(s) = p$, namely $s = f^{-1}(p)$ if $p \notin K(f)$ and $s = a_n$ if $p = p_n$; and in either case we have $s = f_{\leftarrow}(p)$.

7.4 Now let $x$ vary (in $X^c$). Write $\Phi$ for the function $x \longmapsto F^c_x(1) = \gamma^c_x[0, 1]$, and $\mathcal{O}$ for its upper ordinate set $\{(x, t) : x \in X^c, 0 \leq t \leq \Phi(x) = F^c_x(1)\}$. Since $\gamma^c$ is a transition kernel, $\Phi$ is $\mathcal{A}$-measurable and therefore (by 2.2) $\mathcal{O} \in \mathcal{A} \otimes \mathcal{B}$.

Define $E = (X^c \times I) \backslash C(F^c)$ (so that $E_x = I \backslash C(F^c_x), x \in X^c$). Both $E$ and $\mathcal{O}$ are given the Borel structure they inherit from $\mathcal{A} \otimes \mathcal{B}$.

As in 3.2 (and [6, p.8]) we have from 7.2(2) that $(F^c)^{\sim}$ is measurable, where $(F^c)^{\sim}(x, t) = (x, F^c(x, t))$. Now we show:

PROPOSITION: $(F^c)^{\sim} | E$ *is a Borel isomorphism of* $E$ *onto* $\mathcal{O}$, *with inverse* $((F^c)_{\leftarrow})^{\sim}$; *and further, for all* $x \in X^c$ *and* $t \in I$, $((F^c)^{\sim} | E)(\{x\} \times ([0, t] \cap E_x)) = \{x\} \times F^c(\{x\} \times [0, t])$.

*Proof:* For each $x \in X^c, (F^c)^{\sim} | E_x$ takes $\{x\} \times E_x$, in a 1-1 way, onto $\{x\} \times [0, \Phi(x)]$, by Proposition 7.3, the inverse map being $((F^c_x)_{\leftarrow})^{\sim} = ((F^c_{\leftarrow})^{\sim})_x$. This shows that $(F^c)^{\sim} | E$ is a "proper" (i.e., first-coordinate-preserving) bijection of

$E$ onto $\mathcal{O}$, with inverse $((F^c)_{\leftarrow})\check{}$. Both maps are $\mathcal{A}\otimes\mathcal{B}$-measurable. All that remains to be proved is that (for all $x \in X^c$ and $t \in I$),

$$F_x^c([0,t]\cap E_x) = F_x^c[0,t];$$

that is, in the abbreviated notation used in 7.3,

$$f([0,t]\backslash C) = f[0,t] \qquad (= [0, f(t)]).$$

And this follows by reasoning similar to that in 7.3. ∎

7.5 COROLLARY: *$(F^c)\check{}\,|E$ is a proper Borel isometry, taking the measure $\gamma_x^c$ on $E_x$ to Lebesgue measure $\lambda$ on $\mathcal{O}_x$.*

That is, for each $H \in \mathcal{B}$ and $x \in X^c$,

$$\gamma_x^c((X^c \times H)\backslash C(F_x^c)) = \lambda(F_x^c(H_x\backslash C(F_x^c))).$$

This is true when $H$ has the form $[0,t], t \in I$, by 7.2(1); and the general case follows routinely.

## 8. The model representation

8.1 In this section we return to the more general situation (1.4) in which the measures $\gamma_x$, though still required to be finite, do not necessarily have total measure 1. Our object is to define a "proper" (first-coordinate-preserving) Borel isometry taking the sectioned transition kernel $\gamma$ (from $(X,\mathcal{A})$ to $(X\times I, \mathcal{A}\otimes\mathcal{B})$) to the model representation promised in 1.4. This will be done in two main steps, the first of which produces an approximation to the desired model, which the second step adjusts to give the final model. We emphasize that, throughout the construction, all the isomorphisms used are to preserve first coordinates — that is, to be "proper".

Write $X^0 = \{x \in X : \gamma_x(I) = 0\}$, $X^a = \{x \in X : \gamma_x(I) > 0 = \gamma_x^c(I)\}$, and (as before) $X^c = \{x \in X : \gamma_x^c(I) > 0\}$. (Here again we are taking advantage of the fact that $\gamma$ is sectioned, writing $\gamma_x^c(B)$ for $\gamma_x^c(X\times B), B \in \mathcal{B}$.) Of course $X^a$ consists of the $x$'s for which $\gamma_x$ is purely atomic and nonzero; and $X^0, X^a, X^c$ partition $X$ into three $\mathcal{A}$-measurable sets.

The model $M$ will be a subset of $X \times \mathbb{R}$ consisting of three pairwise disjoint $\mathcal{A}\otimes\mathcal{B}(\mathbb{R})$-measurable sets, as follows.

(1) A "singular set" $S \subset (X^0 \cup X^a) \times K^0$, where $K^0$ is a Cantor set contained in the interval (-1,0) and of Lebesgue measure 0.

(2) The set $\bigcup_n (A_n \times \{-n\}) = T$, say, where (cf. 5.4) the sets $A_n$ $(n = 1, 2, \dots)$ are the domains of definition of the $\mathcal{A}$-measurable functions $g_n$ whose graphs $\Gamma(g_n) = \Gamma_n$ make up the "atomic set" $J$ of $\gamma$ — or rather of the restriction of $\gamma$ to $X \backslash X^0 = X^a \cup X^c$. (In §5 the $\gamma_x$'s were assumed to be probability measures, but the usual "renormalization" process, replacing $\gamma_x$ by $(1/\gamma_x(I))\gamma_x$, shows that the finiteness of the $\gamma_x$'s suffices.) We recall that the graphs $\Gamma_n$ are pairwise disjoint, and that each $A_n \in \mathcal{A}$ and $A_1 \supset A_2 \supset \cdots$.

(3) The set $\mathcal{O}$, where (as in 7.4) $\mathcal{O}$ is the upper ordinate set of the function $\Phi$ (where $\Phi(x) = \gamma_x^c(I), x \in X^c$).

The model is $M = S \cup T \cup \mathcal{O}$, equipped with the "Borel" field $\mathcal{A} \otimes \mathcal{B}(\mathbb{R})|M$. The corresponding (sectioned) measure $\rho_x$ $(x \in X)$ on $M$ is defined to consist of Lebesgue measure $\lambda$ on the "slice" $\mathcal{O}_x \cup S_x$ of $\mathcal{O} \cup S$, together with a purely atomic measure on $T$, each atom $(x, -n)$ (where $x \in A_n$) having weight $\gamma_x\{g_n(x)\}$. Of course, $\rho_x$ will be zero on $S_x$.

THEOREM: *There is a proper Borel isomorphism $\xi$ of $X \times I$ onto $M$, taking each $\gamma_x$ to the measure $\rho_x$ $(x \in X)$.*

As remarked above, we first prove this for an "approximate model", in which $\mathcal{O}$ is replaced by the set (say) $U = \mathcal{O} \backslash J''$, where $J''$ is the image of $J' = J \cap E$ under the isomorphism (cf. 7.4) $(F^c)^{\tilde{}}|E$ of $E$ onto $\mathcal{O}$. (Since each $J''_x$ is countable, the measures $\rho_x$ are not affected by this replacement.) Later (8.5) we shall restore $\mathcal{O}$ by means of a further proper isomorphism of $U$.

We define $\xi$ on $J$ as follows: We have $J = \bigcup_n \Gamma_n$ where (Proposition 2.2) $\Gamma_n$ is "Borel" isomorphic to $A_n$ by projection. Hence we define $\xi$ on $\Gamma_n$ by $\xi(x, g_n(x)) = (x, -n)$ (for all $x \in A_n, n = 1, 2, \dots$).

Next consider $\xi$ on $(((X^0 \cup X^a) \times I) \cup C(F^c)) \backslash J$. Fix a Borel isomorphism $\theta$ of $I$ onto $K^0$, and define $\xi(x, t) = (x, \theta(t))$ $(x \in X, t \in I)$; the resulting image set $(\subset X \times K^0)$ is defined to be $S$. Here, of course, $\xi$ is measure-preserving, because both $\gamma_x$ and $\rho_x$ are 0.

Finally, on the remaining part $E \backslash J$ (where, as in 7.4, $E = (X^c \times I) \backslash C(F^c))$ we take $\xi$ to be the restriction of the (proper) Borel isomorphism $(F^c)^{\tilde{}}$ (as in 7.5), mapping $E \backslash J$ onto $\mathcal{O} \backslash J'' = U$. Thus here $\rho_x|U_x$ provides the continuous (dispersed) part, and $\rho_x|T_x$ the atomic part, of $\gamma_x$.

8.2 To replace $U$ by $\mathcal{O}$ we have to "fill in" the gaps in $\mathcal{O}$ left by the removal of $J''$; and the first step is to obtain an improved representation of the set $J' = J \cap E$. The reasoning in §5 applies virtually unchanged when $J$ is replaced by $J'$, and leads to the following result (cf. Theorem 5.4): $J'$ is the union of a sequence (possibly finite) of pairwise disjoint graphs $\Gamma'_n = \Gamma(g'_n)$ of $\mathcal{A}$-measurable functions $g'_n$ $(n = 1, 2, \ldots)$; and the domain of definition of $g'_n$ is an $\mathcal{A}$-measurable set $A'_n$ such that $X^c \supset A'_1 \supset A'_2 \supset \cdots$.

Now the proper isomorphism $(F^c)\check{}|E$ (of $E$ onto $\mathcal{O}$) takes each $\Gamma'_n = \Gamma(g'_n)$ to the graph $\Gamma''_n = \Gamma(g''_n)$ of the $\mathcal{A}$-measurable function $g''_n$, where $g''_n(x) = (F^c_x \circ g'_n)(x)$. Since this isomorphism is proper, the domain of $g''_n$ is still $A'_n$; and the graphs $\Gamma''_n$ are pairwise disjoint since the $\Gamma'_n$ are. The graphs $\Gamma''_n$ are all contained in $\mathcal{O}$, and are precisely the graphs whose removal from $\mathcal{O}$ has left the set $U$ $(= (F^c)\check{}(E \backslash J'))$. The remainder of the proof, for which I am indebted to A.H. Stone, is devoted to reinserting these graphs, by means of a suitable proper Borel isometry.

8.3 We recall (7.4) that $\mathcal{O}$ is the upper ordinate set of the $\mathcal{A}$-measurable everywhere positive function $\Phi : X^c \to \mathbb{R}$, where $\Phi(x) = F^c_x(1) = \gamma^c_x[0,1]$. As in 2.3, consider the space $\mathcal{L}(\Phi)$ of all $\mathcal{A}$-measurable functions $f : X^c \to \mathbb{R}$ such that $0 \le f(x) \le \Phi(x)$ for all $x \in X^c$. The construction depends on the following lemma:

LEMMA: *Given a countable family of $\mathcal{A}$-measurable functions $\phi_n : B_n \to \mathbb{R}$, where each $B_n \subset X^c$, such that $0 \le \phi_n(x) \le \Phi(x)$ for all $x \in B_n$, there exists an $\mathcal{A}$-measurable $\psi \in \mathcal{L}(\Phi)$ whose graph $\Gamma(\psi)$ is disjoint from all the graphs $\Gamma(\phi_n)$.*

*(Remark:* From 2.2, we have that each $B_n \in \mathcal{A}$.)

*Proof:* As noted in 2.3, $\mathcal{L}(\Phi)$ with the topology of close approximation is a Baire space. Put

$$H_n = \{f \in \mathcal{L}(\Phi) : \Gamma(f) \cap \Gamma(\phi_n) = \phi\}, \qquad n = 1, 2, \ldots.$$

It is not hard to see that $H_n$ is both open and dense in $\mathcal{L}(\Phi)$. Hence $\bigcap_n H_n$ is nonempty, and we just pick $\psi \in \bigcap_n H_n$. ∎

8.4 COROLLARY: *Under the hypotheses of the Lemma, there is an infinite sequence of $\mathcal{A}$-measurable functions $\psi_1, \psi_2, \ldots,$ on $X^c$ such that (for all $n \in \mathbb{N}$ and*

*$x \in X^c$) $0 \le \psi_n(x) \le \Phi(x)$ and the graphs $\Gamma(\psi_n)$ are disjoint from each other and from all the graphs $\Gamma(\phi_n)$.*

We get $\psi_n$ by applying the Lemma inductively to the functions $\psi_1, \psi_2, \ldots, \psi_{n-1}, \phi_1, \phi_2, \phi_3, \ldots$.

8.5 We apply Corollary 8.4 to the graphs $\Gamma''_n = \Gamma(g''_n)$ of 8.2, obtaining $\mathcal{A}$-measurable functions $f_n : X^c \to \mathbb{R}$ such that $0 \le f_n(x) \le \Phi(x)$ for all $x \in X^c$ and *all* the graphs $\Gamma(f_n), \Gamma(g''_m)(m, n = 1, 2, \ldots)$ are pairwise disjoint. By construction, all of them are subsets of $\mathcal{O}$.

Put $B_0 = X^c \backslash A'_1, B_1 = A'_1 \backslash A'_2, B_2 = A'_2 \backslash A'_3, \ldots$, and $B_\infty = \bigcap_n A'_n$; these sets partition $X^c$ into $\mathcal{A}$-measurable sets. For each $k \in \mathbb{N}$ we have that $g''_1, g''_2, \ldots, g''_k$ are defined on $B_k$, while the domains of definition of $g''_n$ for $n > k$ are disjoint from $B_k$. Now map $\mathcal{O} \backslash \bigcup_n \Gamma''_n$ $(= U)$ onto $\mathcal{O}$ by the proper "Borel" isomorphism $\eta$ defined as follows. The map $\eta$ is to be the identity except on the union of the graphs $\Gamma(f_n), n = 1, 2, \ldots$. For each finite $k$, and each $x \in B_k$, $\eta$ maps $f_1(x)$ to $g''_1(x), f_2(x)$ to $g''_2(x), \ldots, f_k(x)$ to $g''_k(x), f_{k+1}(x)$ to $f_1(x), \ldots, f_{k+n}(x)$ to $f_n(x), \ldots$. And, for $x \in B_\infty$, $\eta$ maps $f_{2n-1}(x)$ to $g''_n(x)$ and $f_{2n}(x)$ to $f_n(x)$ $(n = 1, 2, \ldots)$.

The combined isomorphism $\eta \circ \xi$ sends $X \times I$ to the model $M$, and $\gamma$ to the transition kernel $\rho$, as required.

## 9. Some extensions

9.1 Suppose $\gamma$ is, more generally, a sectioned transition kernel from $(X, \mathcal{A})$ to $(W, \mathcal{A} \otimes \mathcal{B}|W)$, where $W$ is a "Borel" subset of $X \times \mathbb{R}$, and $\mathcal{B}$ is the family $\mathcal{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$; cf. 1.3. We wish to extend Theorem 8.1 to cover this more general situation.

The first step is to extend $\gamma$ to a sectioned transition kernel $\hat{\gamma}$ from $(X, \mathcal{A})$ to $(X \times \mathbb{R}, \mathcal{A} \otimes \mathcal{B})$, simply by defining $\hat{\gamma}$ to be 0 everywhere outside $W$ (and to agree with $\gamma$ on $W$). By Theorem 8.1, $\hat{\gamma}$ is isomorphic, by a proper isometry $\hat{\xi}$, to a model sectioned transition kernel, $\hat{\rho}$ say, from $(X, \mathcal{A})$ to $(\hat{M}, \mathcal{A} \otimes \mathcal{B}|\hat{M})$ where $\hat{M} = \hat{S} \cup \hat{T} \cup \hat{\mathcal{O}}$ as in 8.1. We consider 3 cases:

CASE I: Suppose $W$ is "sectionally closed" in the sense that all its sections $W_x$ $(x \in X)$ are closed subsets of $\mathbb{R}$. In this case, since $\hat{\gamma}_x$ is zero on each complementary interval of $(\{x\} \times \mathbb{R}) \backslash W_x$, the construction of the model automatically ensures that $\hat{\xi}$ maps each such interval into the singular set $\hat{S}$. Thus the restriction $\xi$ of $\hat{\xi}$ to $W$ provides a proper isometric isomorphism of the original kernel $\gamma$ onto a model $M = S \cup T \cup \mathcal{O}$ in which $T = \hat{T}$ and $\mathcal{O} = \hat{\mathcal{O}}$ but $S \subset \hat{S}$; that is, $M$ differs from $\hat{M}$ only in having a smaller singular set. Thus Theorem 8.1 applies unchanged in this case.

9.2 CASE II: Even if $W$ is not "sectionally closed", the preceding method can be applied. The restriction $\xi$ of $\hat{\xi}$ to $W$ provides a proper isometric isomorphism of $\gamma$ onto a model $M = S \cup T \cup \mathcal{O}'$, where $S \subset \hat{S}, T = \hat{T}$ and $\mathcal{O}'$ is of the form $\mathcal{O}\backslash Z$, where $\mathcal{O}$ is an ordinate set as in Theorem 8.1 and where a "Borel" set $Z$ has been removed from it. By a further proper Borel isometry we can compress $\mathcal{O}\backslash Z$ into an "almost ordinate set" in the sense of [7]; that is, we may assume each $Z_x$ has zero Lebesgue measure. This still produces a "model representation" for $\gamma$, but of a less tidy nature than in Theorem 8.1.

9.3 CASE III: But in Case II, if $(X, \mathcal{A})$ is itself a standard Borel structure, one can use Mauldin's Borel parametrization theorem [8] (as in [2]) to eliminate the gaps in $\mathcal{O}$ caused by the removal of $Z$, and we end with Theorem 8.1 applying unchanged.

9.4 INFINITE MEASURES. We return to our usual situation, except that it is now more natural to replace $I$ by the equivalent standard space $\mathbb{R}^+ = [0, \infty)$. Thus $\gamma$ is a sectioned transition kernel from $(X, \mathcal{A})$ to $(X \times \mathbb{R}^+, \mathcal{A} \otimes \mathcal{B}(\mathbb{R}^+))$. Now the measures $\gamma_x$ so far have always been assumed to be finite; can this restriction be removed? It seems natural to require them to be $\sigma$-finite, but we shall need something stronger. We say that $\gamma$ is "uniformly $\sigma$-finite" provided $X \times \mathbb{R}^+$ can be expressed as $\bigcup\{H_n : n = 1, 2, \ldots\}$ where $H_n \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R}^+)$ and $\gamma_x(H_n) < \infty$ for all $x \in X$ and $n \in \mathbb{N}$. Then Theorem 8.1 applies to each $H_n$; however, for the present purpose, it is convenient to modify the "model representation" of Theorem 8.1 slightly, as follows. The "ordinate set" $\mathcal{O}$ in that model was defined to be the "upper ordinate set" $\{(x, t) : x \in X, 0 \leq t \leq \Phi(x)\}$ of the function $\Phi$. It will here be more convenient to replace $\mathcal{O}$ by the "lower ordinate set" $\{(x, t) : x \in X^c, 0 \leq t < \Phi(x)\}$. This can be done by a proper Borel isometry, similar to that in 8.5 but simpler (map each $(x, \Phi(x)/k)$ to $(x, \Phi(x)/(k+1)), k = 1, 2, \ldots$). In what follows, $\mathcal{O}$ will denote this lower ordinate set.

Again we consider 3 cases.

CASE I: Suppose $\gamma_x(X \times [0,t]) < \infty$ for all $t \in \mathbb{R}^+$ and $x \in X$. Then we may (and do) take $H_n = X \times [0,n]$ ($n \in \mathbb{N}$). We shall show:

THEOREM: *In Case I, there is a proper isometric Borel isomorphism $\xi$ taking the sectioned, uniformly $\sigma$-finite transition kernel $\gamma$ to a "model" sectioned transition kernel $\rho$ from $(X, \mathcal{A})$ to $(M, \mathcal{A} \otimes \mathcal{B}(\mathbb{R})|M)$, where $M$ is the disjoint union of three $\mathcal{A} \otimes \mathcal{B}(\mathbb{R})$-measurable sets, $S \cup T \cup \mathcal{O}$, such that*

(1) *the singular set $S$ is a subset of $X \times K^0$, where $K^0$ is a null Cantor subset of (-1,0),*
(2) *the atomic set $T$ is the union of a sequence of $\mathcal{A}$-measurable sets contained respectively in distinct sets of the form $X \times \{-n\}$ for certain values of $n \in \mathbb{N}$,*
(3) *the ordinate set $\mathcal{O}$ is the lower ordinate set of an $\mathcal{A}$-measurable non-negative extended-real-valued function $\Phi$.*

*The measure $\rho_x$ (thought of as located on the $x$-section) is zero on $X_x$, purely atomic on $T_x$, and coincides with Lebesgue measure $\lambda$ on $\mathcal{O}_x$.*

*Sketch of proof:* Apply Theorem 8.1 to each $H_n$, obtaining a proper isometric isomorphism $\xi_n$ onto a "model" $M_n = S_n \cup T_n \cup \mathcal{O}_n$, but in which we now make the following modifications. We first fix a sequence of pairwise disjoint Cantor sets $K_n^0, n = 1, 2, \ldots$, in the null Cantor set $K^0$ (cf. 8.1), and arrange that $\xi_n$ places $S_n \backslash S_{n-1}$ in $X \times K_n^0$. The subset $T_n \backslash T_{n-1}$ of the atomic set is placed on the "lines" $X \times \{-(p_n)^k\}, k = 1, 2, \ldots$, where $p_n$ is the $n$th prime. And, as arranged above, $\mathcal{O}_n$ is the lower ordinate set of the corresponding function $\Phi_n$ (where $\Phi_n(x) = \gamma_x^c[0,n]$). The construction proving Theorem 8.1 automatically ensures that $\xi_n$ extends $\xi_{n-1}$ so far as $\mathcal{O}_{n-1}$ is concerned, and we easily arrange that $\xi_n$ extends $\xi_{n-1}$ elsewhere. Thus the maps $\xi_n$ combine to give an isomorphism $\xi$ as required. ∎

9.5 CASE II. Here we suppose merely that $\gamma$ is uniformly $\sigma$-finite as defined in 9.4, so that the sets $H_n$ are merely "Borel" sets with union $X \times \mathbb{R}^+$. We may of course suppose them pairwise disjoint. As pointed out in 9.2, Theorem 8.1 applies to each $H_n$ except that $\mathcal{O}_n$ will here be merely an "almost ordinate set". We deduce: *Theorem 9.4 still applies, except that the ordinate set $\mathcal{O}$ is to be replaced by an "almost ordinate set".*

*Sketch of proof:* Again Theorem 8.1 produces a proper isometric isomorphism $\xi_n$ taking the restriction of $\gamma$ (from $(X, \mathcal{A})$ to $(H_n, \mathcal{A} \otimes \mathcal{B}|H_n)$) to $\rho_n$, from $(X, \mathcal{A})$

to $(M_n, \mathcal{A} \otimes \mathcal{B}|M_n)$, where $M_n = S_n \cup T_n \cup \mathcal{O}_n$. Much as in Case I (9.4) we arrange that $S_n \subset X \times K_n^0$ and $T_n \subset X \times \{-(p_n)^k, k = 1, 2, \ldots\}$. And we raise each "almost ordinate set" $\mathcal{O}_n$ vertically so that it fits just above $\mathcal{O}_{n-1}$. More precisely, $\mathcal{O}_n$ was by definition (almost) the lower ordinate set of the positive finite function $\Phi_n$; now $\Phi_n$ can here be regarded as defined on all $X$, with value 0 outside its original domain of definition, and we arrange that $\mathcal{O}_n$ is translated vertically by $\Phi_1 + \Phi_2 + \cdots + \Phi_{n-1}$. Then $\mathcal{O}_1 \cup \mathcal{O}_2 \cup \cdots \cup \mathcal{O}_n$ will be (almost) the lower ordinate set of $\Phi_1 + \Phi_2 + \cdots + \Phi_n$; and finally $\mathcal{O} = \bigcup\{\mathcal{O}_n : n \in \mathbb{N}\}$ is (except for an omitted subset of section-measure 0 for all $x$) the lower ordinate set of $\sum_{n=1}^{\infty} \Phi_n$. ∎

9.6 Case III. As in 9.3, if $(X, \mathcal{A})$ is a standard Borel structure, we can fill the gaps in $\mathcal{O}$ by means of Mauldin's theorem, so that Theorem 9.4 then applies even in Case II.

9.7 Planar decomposition Finally we remark that (as with the theorem of 9) the theorem in 9.6 contains the planar decomposition theorem of Rokhlin [10] and the author [5] (see also [7] and [2]). For a suitable strict disintegration of the given standard ($\sigma$-finite) measure space provides a transition kernel $\gamma$ in which $(X, \mathcal{A})$ is a standard Borel structure with a ($\sigma$-finite) measure $\mu$. In the resulting planar model for $\gamma$, the atoms of the vertical factor have been taken care of, and we have only to adjust $X$ so as to separate out the $\mu$-atoms.

## References


[1] E. K. van Douwen and A. H. Stone, *The topology of close approximation*, Topology and Appl. **35** (1990), 261–275.

[2] S. Graf and R. D. Mauldin, *A classification of disintegrations of measures*, in *Measurable Dynamics*, Contemporary Math. 94 (Am. Math. Soc. 1989), pp. 147–158.

[3] K. Krickelberg, *Probability Theory*, Addison-Wesley, New York, 1963.

[4] G. W. Mackey, *Borel structures in groups and their duals*, Trans. Am. Math. Soc. **85** (1957), 134–165.

[5] D. Maharam, *Decomposition of measure algebras and spaces*, Trans. Am. Math. Soc. **69** (1950), 142–160.

[6] D. Maharam, *On smoothing compact measure spaces by multiplication*, Trans. Am. Math. Soc. **204** (1975), 1–39.

[7] D. Maharam, *On the planar representation of a measurable subfield,* Lecture Notes in Math. 1089, Springer-Verlag, Berlin, 1984, pp. 47–57.

[8] R. D. Mauldin, *Borel parametrizations,* Trans. Am. Math. Soc. **250** (1979), 223–234.

[9] R. D. Mauldin, D. Preiss and H. von Weizäcker, *Orthogonal transition kernels,* Ann. Probab. **11** (1983), 970–988.

[10] V. A. Rokhlin, *On the fundamental ideas in measure theory,* Mat. Sb. **25** (1949) = Am. Math. Soc. Transl. **71** (1952).

[11] H. D. Ursell, *Some methods of proving measurability,* Fund. Math. **32** (1939), 311–330.