Ergod. Th. & Dynam. Sys. (2003), **23**, 293–305
DOI: 10.1017/S0143385702000925
© *2003 Cambridge University Press*
Printed in the United Kingdom

Group actions and singular martingales

JOSEPH ROSENBLATT†, DANIEL STROOCK‡ and MICHAEL TAYLOR§

† *Mathematics Department, University of Illinois at Urbana-Champaign, IL 43210, USA*
(e-mail: jrsnbltt@math.uiuc.edu)

‡ *Mathematics Department, Massachusetts Institute of Technology, MA 02139, USA*
(e-mail: dws@math.mit.edu)

§ *Mathematics Department, University of North Carolina at Chapel Hill, NC 27599, USA*
(e-mail: met@math.unc.edu)

(*Received* 12 *December* 2000 *and accepted in revised form* 10 *November* 2001)

Abstract. We investigate properties of a class of martingales formed by picking a measurable set A in a compact group G, taking random rotates of A, and taking measures of the resulting intersections, suitably normalized. Such martingales are shown to yield measures on G^∞ that are singular with respect to product Haar measure, and their ergodic decompositions with respect to the shift map are analyzed. Connections are made with De Finetti theory.

1. *Introduction*

Let G be a compact group, with Haar measure m, normalized to have total mass 1. Let $\mathcal{Z} = G^\infty$, with product measure μ. Fix $A \subset G$, measurable, with $m(A) > 0$. For $x = (x_1, x_2, x_3, \dots) \in \mathcal{Z}$, let

$$f_N(x) = \frac{m\big(\bigcap_{i=1}^N x_i A\big)}{m(A)^N}. \tag{1.1}$$

The numerator provides a measure of the extent to which the N-fold translates of A tend to disperse. The denominator is a normalizing factor, uniquely chosen so that the following important property holds.

PROPOSITION 1.1. *The sequence* (f_N) *is a martingale, satisfying* $\int_{\mathcal{Z}} f_N \, d\mu = 1$.

To be precise, let $\mathcal{C}_N$ consist of all Borel functions on $\mathcal{Z}$ of the form $g(x_1, \dots, x_N)$, and let $\mathcal{B}_N$ be the σ-algebra generated by $\mathcal{C}_N$. We assert that $\mathbb{E}(f_{N+1}|\mathcal{B}_N) = f_N$, for each N, or equivalently

$$\int_{\mathcal{Z}} f_{N+1} g \, d\mu = \int_{\mathcal{Z}} f_N g \, d\mu, \quad \text{for all bounded } g \in \mathcal{C}_N. \tag{1.2}$$

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

The proof of (1.2) is a straightforward consequence of the identity

$$\int_G m(yA \cap B)\, dm(y) = m(A)m(B), \tag{1.3}$$

for measurable subsets A and B of G, itself a simple consequence of Fubini's Theorem and the bi-invariance of Haar measure on a compact group. Indeed, to prove (1.2) for $g = g(x_1, \ldots, x_N)$, write the left-hand side of (1.2) as

$$\int \frac{m\left(x_{N+1}A \cap \bigcap_{i=1}^{N} x_i A\right)}{m(A)^{N+1}} g(x_1, \ldots, x_N)\, dm(x_1) \ldots dm(x_{N+1}).$$

If we integrate first with respect to x_{N+1} and use (1.3), we see that this is equal to the right-hand side of (1.2).

Thus (f_N) is a martingale, consisting of non-negative functions. Clearly $f_1 \equiv 1$, so $\int_{\mathcal{Z}} f_N\, d\mu = 1$ for each N. Hence the martingale convergence theorem of Fatou guarantees the existence of a pointwise limit f:

$$f_N \longrightarrow f, \quad \mu\text{-a.e. on } \mathcal{Z}. \tag{1.4}$$

One of us (JR) defined this class of martingales, as an offshoot of work in [**R1**] (cf. also [**R2**]), and noticed that f must be a constant $c \in [0, 1]$, and also speculated that in fact $c = 0$ whenever $m(A) < 1$. One goal of this paper is to present a proof of this result.

The following restatement of this result is useful. Given a martingale (f_N) of non-negative $\mathcal{B}_N$-measurable functions satisfying $\int f_N\, d\mu = 1$, there may not be a final element in $L^1(\mathcal{Z}, d\mu)$, but there will be a unique probability measure ν on $\mathcal{Z}$ such that

$$\int_{\mathcal{Z}} g f_N\, d\mu = \int_{\mathcal{Z}} g\, d\nu, \quad \forall g \in \mathcal{C}_N. \tag{1.5}$$

Furthermore, we have the Lebesgue–Radon–Nikodym decomposition

$$\nu = f\mu + \nu_s, \quad f \in L^1(\mathcal{Z}, d\mu), \quad \nu_s \perp \mu, \tag{1.6}$$

with f as in (1.4). This result is a refinement of Fatou's martingale convergence theorem. It was established in [**AJ1, AJ2**]; cf. also the discussion in [**D**, pp. 631–632]. In view of this, the result that $f = c = 0$ is equivalent to the result that $\nu = \nu_s$, i.e. $\nu \perp \mu$.

In §2 we will prove that $f = c$, μ-a.e., for some $c \in [0, 1]$, and we give one proof that $c = 0$. In §3 we present a formula for the ergodic decomposition, under the shift map, of the measure ν, which is very useful for revealing the structure of this measure. In particular, it leads to a second proof that $\nu \perp \mu$. Furthermore, in concert with results on uniqueness of ergodic decompositions, it yields strong results on the extent to which the martingale (1.1) determines the set A.

We consider various more general settings for the martingale construction (1.1). To begin, we mention an apparently more general setting for this construction. Namely, suppose $X = G/K$ is a compact homogeneous space. Then X has a unique G-invariant probability measure, which we also denote m, and for measurable $A \subset X$, with $m(A) > 0$, we can define the sequence f_N on $\mathcal{Z} = G^\infty$ by (1.1). This can be reduced to the

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

previous case, since we have the same sequence of functions on $\mathcal{Z}$ if A is replaced by $A' = \pi^{-1}(A) \subset G$, where $\pi : G \to G/K$ is the standard projection. Thus this does not actually give an extended theory, but it is of conceptual interest, for example when $G = SO(n)$ and $X = S^{n-1}$, the unit sphere in $\mathbb{R}^n$.

The martingale (1.1) is generated by a single set A. One can use a sequence of sets $A_j \subset G$, each of positive measure. Then an argument parallel to the proof of Proposition 1.1 shows that

$$f_N(x) = m(A_1)^{-1} \cdots m(A_N)^{-1} m\left(\bigcap_{i=1}^{N} x_i A_i\right) \tag{1.7}$$

is also a martingale. In §4 we study some examples of such martingales, which can have interesting non-constant limits.

In §§5–6 we extend the setting of §§1–3 in another way, in which the invariance of the probability measures of interest under the action of the group $\mathcal{S}$ of finite permutations of $\mathbb{Z}^+$ plays the central role. Contact is made with De Finetti's theory of exchangeable processes.

2. *Hands-on proof that $m(A) < 1 \Rightarrow f = 0$ almost surely*

The first part of our task in this section is to show the following proposition.

PROPOSITION 2.1. *There exists $c \in [0, 1]$ such that*

$$f = c, \quad \mu\text{-a.e.} \tag{2.1}$$

Proof. Let $\mathcal{S}$ be the group of automorphisms of $\mathcal{Z} = G^\infty$ consisting of finite permutations of the coordinates. More precisely, $\mathcal{S} = \cup \mathcal{S}_N$, where an element of $\mathcal{S}_N$ has the form

$$(x_1, x_2, \dots) \mapsto (x_{\pi(1)}, x_{\pi(2)}, \dots),$$

where π permutes $\{1, \dots, N\}$ and $\pi(j) = j$ for $j > N$. Clearly each f_N in (1.1) is invariant under $\mathcal{S}_N$, so f in (1.4) is invariant under $\mathcal{S}$. Thus (2.1) is a consequence of the following. □

PROPOSITION 2.2. *The group $\mathcal{S}$ acts ergodically on $(\mathcal{Z}, \mu)$.*

This result is well known; for example, it is a consequence of the Hewitt–Savage 0–1 law. A simple direct proof can be obtained by the following route. Say $f \in L^1(\mathcal{Z}, d\mu)$ is invariant under $\mathcal{S}$. In particular, with

$$f^N(x) = f(x_2, x_3, \dots, x_N, x_1, x_{N+1}, \dots),$$

we have $f = f^N$ a.e. for each N. It is then straightforward to show that $f = f^b$ a.e., where $f^b(x) = f(x_2, x_3, \dots)$. In other words, f is invariant under the shift map $\Sigma : \mathcal{Z} \to \mathcal{Z}$, given by $\Sigma(x) = (x_2, x_3, \dots)$. Thus Proposition 2.2 also follows from the well-known result that the shift map Σ acts ergodically on $(\mathcal{Z}, \mu)$.

Our next goal is to show that $c = 0$ whenever $m(A) < 1$. We mention that it is relatively easy to prove that if $G \setminus A$ has non-empty interior, then $c = 0$. In fact, in this case one

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

easily verifies that there is a set $\mathcal{X} \subset \mathcal{Z} = G^{\infty}$ with $\mu(\mathcal{X}) = 1$ such that for all $x \in \mathcal{X}$ we have $f_N(x) = 0$ for all sufficiently large N. Since our main result encompasses this case, we leave the details to the interested reader.

If $G \setminus A$ has empty interior but positive measure, the situation is more difficult to analyze. The following result helps.

PROPOSITION 2.3. *We have*

$$c = \inf_{N,x} f_N(x). \tag{2.2}$$

Proof. From (1.6) and (2.1) we have

$$f_N \mu|_{\mathcal{B}_N} = \nu|_{\mathcal{B}_N} = c\mu|_{\mathcal{B}_N} + \nu_s|_{\mathcal{B}_N}. \tag{2.3}$$

Hence, for all N,

$$f_N \geq c, \quad \mu\text{-a.e. on } \mathcal{Z}. \tag{2.4}$$

Note that, whenever $m(A) > 0$, f_N in (1.1) is continuous. Thus (2.4) together with $f_N \to c$ gives (2.2). □

We can be more precise about (2.2). Suppose

$$c_N = \min_x f_N(x) = f_N(a_1, \ldots, a_N). \tag{2.5}$$

(Here the components a_j may depend on N, but we suppress this.) Then from

$$\int_G f_{N+1}(a_1, \ldots, a_N, y)\, dm(y) = c_N, \quad f_{N+1}(a_1, \ldots, a_N, a_1) = \frac{c_N}{m(A)},$$

we deduce that (if $m(A) < 1$)

$$c_{N+1} \leq \min_y f_{N+1}(a_1, \ldots, a_N, y) < c_N. \tag{2.6}$$

Hence

$$c_N \searrow c. \tag{2.7}$$

We are now ready to prove our first main result.

THEOREM 2.4. *If $m(A) < 1$, then $c = 0$; hence $\nu \perp \mu$.*

Proof. It follows from (2.6) that $c_2 < 1$. We will show that

$$c_{2N} \leq c_2^N, \tag{2.8}$$

which will suffice. To prove this, consider

$$B = yA \cap A. \tag{2.9}$$

Pick $y \in G$ such that

$$\frac{m(B)}{m(A)^2} = c_2. \tag{2.10}$$

Now consider

$$f_N^B(x) = \frac{m\left(\bigcap_{i=1}^N x_i B\right)}{m(B)^N}. \tag{2.11}$$

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

We have

$$f_N^B(x) = \frac{m\left(\bigcap_{i=1}^N x_i A \cap \bigcap_{i=1}^N x_i y A\right)}{m(A)^{2N}} \left[\frac{m(A)^2}{m(B)}\right]^N$$
$$= f_{2N}(x_1, \ldots, x_N, x_1 y, \ldots, x_N y) \cdot c_2^{-N}. \tag{2.12}$$

By the analysis above, with A replaced by B, we can pick $x \in \mathcal{Z}$ (depending on N) such that $f_N^B(x) \leq 1$. For such a choice we have

$$f_{2N}(x_1, \ldots, x_N, x_1 y, \ldots, x_N y) \leq c_2^N, \tag{2.13}$$

which gives (2.8). □

3. *Ergodic decomposition of* ν

We give a formula for ν as an integral of simpler measures. To begin, given $A \subset G$ with $m(A) > 0$, define the following probability measures on G:

$$m_y^A = m(A)^{-1} \mathbf{1}_{yA^{-1}} m, \tag{3.1}$$

where $\mathbf{1}_A$ denotes the indicator function of the set A (equal to 1 on A and 0 off A). Then define the probability measures μ_y^A on $\mathcal{Z}$ to be the associated product measures:

$$\mu_y^A = (m_y^A)^{\mathbb{Z}^+}. \tag{3.2}$$

PROPOSITION 3.1. *If (f_N) is given by (1.1), then the measure ν on $\mathcal{Z}$ in (1.5) satisfies the identity*

$$\nu = \int_G \mu_y^A \, dm(y). \tag{3.3}$$

Proof. If $g = g(x_1, \ldots, x_N)$, then

$$\int_{\mathcal{Z}} g \, d\mu_y^A = m(A)^{-N} \int g(x_1, \ldots, x_N) \mathbf{1}_{yA^{-1}}(x_1) \cdots \mathbf{1}_{yA^{-1}}(x_N) \, dm(x_1) \cdots dm(x_N). \tag{3.4}$$

Using the fact that $\mathbf{1}_{yA^{-1}}(x) = \mathbf{1}_{xA}(y)$, we then have

$$\int_G \int_{\mathcal{Z}} g \, d\mu_y^A \, dm(y) = m(A)^{-N} \int g(x_1, \ldots, x_N) m(\cap x_j A) \, dm(x_1) \cdots dm(x_N)$$
$$= \int g(x) f_N(x) \, d\mu$$
$$= \int g \, d\nu, \tag{3.5}$$

giving (3.3). □

Part of the significance of formula (3.3) arises from the fact that each measure μ_y^A is invariant and ergodic under the shift $\Sigma : \mathcal{Z} \to \mathcal{Z}$. Furthermore, $\mu_{y_1}^A \perp \mu_{y_2}^A$ unless $y_1 A^{-1} = y_2 A^{-1}$, i.e. $A = A(y_2^{-1} y_1)$ (modulo sets of measure zero). This yields the following.

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

COROLLARY 3.2. *Let $H_A \subset G$ be the subgroup of y such that $Ay = A$, modulo sets of measure zero, and set $X_A = G/H_A$, a compact homogeneous space with invariant measure which we also denote by m. Then*

$$\nu = \int_{X_A} \mu_y^A \, dm(y) \tag{3.6}$$

gives the ergodic decomposition of ν with respect to the shift on $\mathcal{Z}$.

It remains only to note that the various measures μ_y^A in the decomposition (3.6) are mutually disjoint. This is a consequence of the fact that

$$m(A \triangle B) \neq 0 \Longrightarrow \mu_y^A \perp \mu_y^B. \tag{3.7}$$

To see this, note that the measure μ_y^A is supported on $\mathcal{S}_{A,y} = yA^{-1} \times yA^{-1} \times \cdots \subset \mathcal{Z}$. Now $m(A \triangle B) \neq 0 \Rightarrow m_y^B(y^{-1}A) = \alpha_{AB} < 1$, and we have $\mu_y^B(\mathcal{S}_{A,y}) \leq \alpha_{AB}^K$ for all $K \in \mathbb{Z}^+$, hence $\mu_y^B(\mathcal{S}_{A,y}) = 0$.

We can also establish the following uniqueness result, showing that the martingale (f_N), or equivalently the measure ν, uniquely determines the set $A \subset G$, up to a right translation (and modulo null sets). At this point it is convenient to denote f_N in (1.1) by f_N^A, to record explicitly its dependence on A. Similarly we denote the measure ν in (1.5), and in (3.3) and (3.6), by ν^A.

PROPOSITION 3.3. *If $A, B \subset G$ and $f_N^A \equiv f_N^B$ for all N, then there exists $y \in G$ such that $B = Ay$, modulo null sets.*

Proof. The identity $f_N^B \equiv f_N^A$ for all N is equivalent to $\nu^B = \nu^A$. We have the ergodic decompositions

$$\nu^B = \int_{X_B} \mu_y^B \, dm(y), \quad \nu^A = \int_{X_A} \mu_y^A \, dm(y). \tag{3.8}$$

Suppose for the moment that G is metrizable. Then so is $\mathcal{Z}$, and we can appeal to the uniqueness of ergodic decompositions of a probability measure on $\mathcal{Z}$ invariant under Σ (cf. [**Ph**, pp. 82–83]) to say that if $\nu^A = \nu^B$ then there exist $y_1, y_2 \in G$ such that $\mu_{y_1}^B = \mu_{y_2}^A$. In view of (3.7), this yields (modulo null sets) $B = Ay$, with $y = y_2^{-1} y_1$. □

To remove the hypothesis that G is metrizable, we can use the following result, whose proof can be found in [**HR**].

LEMMA 3.4. *Given $A, B \subset G$ measurable, of positive measure, there exists a compact normal subgroup $K \subset G$ such that G/K is metrizable, and $A_0 = A$, $B_0 = B$, modulo null sets, such that $K \subset A_0 \cap B_0$.*

Given this, we work with $A' = A_0/K$, $B' = B_0/K \subset G/K$. Clearly $f_N^A \equiv f_N^B \Rightarrow f_N^{A'} \equiv f_N^{B'}$.

The following result complements Proposition 3.3 and also strengthens Theorem 2.4.

THEOREM 3.5. *If $A, B \subset G$ have positive measure, then*

$$\nu^A \neq \nu^B \Longrightarrow \nu^A \perp \nu^B. \tag{3.9}$$

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

Proof. As above, we can assume without loss of generality that G is metrizable. Suppose we have two probability measures ν^1 and ν^2 on $\mathcal{Z}$, each invariant under Σ, with ergodic decompositions

$$\nu^j = \int_{X_j} \mu_y^j \, dm_j(y). \tag{3.10}$$

Take the Lebesgue–Radon–Nikodym decomposition

$$\nu^2 = h\nu^1 + \nu_s^2, \quad \nu_s^2 \perp \nu^1. \tag{3.11}$$

Then h and ν_s^2 are invariant under Σ; cf. [**Ph**, p. 78]. Take also the ergodic decomposition of $\nu^3 = \nu_s^2$, as in (3.10). Then

$$\nu^2 = \int_{X_1} h\mu_y^1 \, dm_1(y) + \int_{X_3} \mu_y^3 \, dm_3(y) \tag{3.12}$$

is an ergodic decomposition of ν^2, and we have $h = h(y)$ for m_1-a.e. $y \in X_1$. Suppose we know that

$$\mu_{y_1}^1 \perp \mu_{y_2}^2, \quad \forall y_1 \in X_1, y_2 \in X_2. \tag{3.13}$$

Then by uniqueness of the ergodic decomposition of ν^2 and comparison of (3.12) with the $j = 2$ case of (3.10), we have $h \equiv 0$, and hence $\nu^2 = \nu_s^2$. Interchanging the roles of ν^1 and ν^2, we see that when (3.13) holds we must have $\nu^1 \perp \nu^2$. By (3.7) this argument applies to $\nu^1 = \nu^A$ and $\nu^2 = \nu^B$ when $\nu^A \neq \nu^B$, so Theorem 3.5 is proven. □

Note how Theorem 3.5 contains Theorem 2.4. If $m(A) < 1$, then $\nu^A \neq \nu^G$, so $\nu^A \perp \nu^G$, but $\nu^G = \mu$.

Remark 3.6. While we have emphasized that (3.6) is an ergodic decomposition with respect to the shift map, it is also an ergodic decomposition with respect to the action of the group $\mathcal{S}$, defined at the beginning of §1. In §5 we will consider an extension of our set-up, in which the group $\mathcal{S}$ plays the major role.

Remark 3.7. In a follow-up paper we will examine the question of to what extent does the knowledge of f_N^A, not for all N, but for a finite number of Ns, determine A.

4. *Martingales arising from varying sequences of sets*

This section is devoted to a class of martingales that differ from those of the form (1.1) in that we consider $\cap x_i A_i$, with different sets A_i. We look at some examples with interesting non-constant limits.

We take $G = \mathbb{T} = \mathbb{R}/\mathbb{Z}$, which we represent as $[0, 1]$ with 0 and 1 identified. We consider the following closed sets $E_j \subset \mathbb{T}$, each of measure $1/2$. To form E_j, take the standard partition of $[0, 1]$ into dyadic intervals of length 2^{-j}, and let E_j be the union of the odd-numbered (closed) dyadic intervals in the collection. One verifies (e.g., by induction on N) that, for each $x_j \in \mathbb{T}$,

$$m(x_1E_1 \cap \cdots \cap x_NE_N) = 2^{-N}. \tag{4.1}$$

Now, given $x = (x_1, x_2, \dots) \in \mathcal{Z} = \mathbb{T}^\infty$, form the following sequence of measures on $\mathbb{T}$:

$$\lambda_x^N(A) = 2^N m(x_1E_1 \cap \cdots \cap x_NE_N \cap A). \tag{4.2}$$

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

Note that $\lambda_x^N(\mathbb{T}) = 1$, so each λ_x^N is a probability measure on $\mathbb{T}$. Given $\varphi \in C(\mathbb{T})$, set

$$\begin{aligned} f_N^\varphi(x) &= \int_{\mathbb{T}} \varphi(y)\, d\lambda_x^N(y) \\ &= 2^N \int_{\mathbb{T}} \mathbf{1}_{E_1}(x_1^{-1}y) \cdots \mathbf{1}_{E_N}(x_N^{-1}y)\varphi(y)\, dm(y). \end{aligned} \tag{4.3}$$

An argument similar to one in §1 shows that, for each $\varphi \in C(\mathbb{T})$, the sequence (f_N^φ) is a martingale. Note that, for each x and N, the map $\varphi \mapsto f_N^\varphi(x)$ is linear. We have

$$f_N^1(x) = 1, \tag{4.4}$$

$$|f_N^\varphi(x)| \le \|\varphi\|_{\sup}, \tag{4.5}$$

$$\int_{\mathcal{Z}} f_N^\varphi\, d\mu = \int_{\mathbb{T}} \varphi\, dm. \tag{4.6}$$

Hence we can apply Fatou's martingale convergence theorem to deduce the existence of a pointwise limit:

$$f_N^\varphi(x) \longrightarrow f^\varphi(x), \quad \forall x \in \mathcal{Z} \setminus \mathcal{R}, \tag{4.7}$$

where $\mu(\mathcal{R}) = 0$. *A priori*, $\mathcal{R}$ could depend on φ. In fact we claim $\mathcal{R}$ can be taken independent of φ. We can see this as follows. Let $V \subset C(\mathbb{T})$ be a countable dense set. Assume $1 \in V$ and V is a linear space over the field $\mathbb{Q}$ of rational numbers. We can pick $\mathcal{R} \subset \mathcal{Z}$ such that $\mu(\mathcal{R}) = 0$ and (4.7) holds for all $\varphi \in V$. Then we claim that (4.7) holds for all $\varphi \in C(\mathbb{T})$. This follows readily from (4.5) and the linearity of $\varphi \mapsto f_N^\varphi(x)$.

One way to interpret the preceding is to say that, for each $x \in \mathcal{Z} \setminus \mathcal{R}$, the sequence $\{\lambda_x^N\}$ converges weakly (i.e. when tested against continuous functions) to a probability measure λ_x. That is, we can write

$$\begin{aligned} f^\varphi(x) = \int_{\mathbb{T}} \varphi\, d\lambda_x \quad &\text{for all } x \in \mathcal{Z} \setminus \mathcal{R} \text{ and } \varphi \in C(\mathbb{T}), \\ &\text{where } \lambda_x^N \longrightarrow \lambda_x \text{ weakly.} \end{aligned} \tag{4.8}$$

Moreover, by the Lebesgue Dominated Convergence Theorem, we see, from (4.6), that

$$\int_{\mathcal{Z}} \int_{\mathbb{T}} \varphi(y)\, d\lambda_x(y)\, d\mu(x) = \int_{\mathcal{Z}} f^\varphi\, d\mu = \int_{\mathbb{T}} \varphi\, dm \quad \text{for all } \varphi \in C(\mathbb{T}). \tag{4.9}$$

Notice that, for each $x \in \mathcal{Z} \setminus \mathcal{R}$,

$$\operatorname{supp} \lambda_x \subset \mathcal{E}_x = \bigcap_{j \ge 1} x_j E_j, \tag{4.10}$$

and that, for μ-almost every x, $\mathcal{E}_x$ is a rather intriguing, Cantor-like subset of $\mathbb{T}$ having Lebesgue measure zero. It would be interesting to know whether $\operatorname{supp}(\lambda_x) = \mathcal{E}_x$ for $x \in \mathcal{Z} \setminus \mathcal{R}$.

Remark. It should be clear that we can use the prescription in (4.3) to associate a martingale f_N^φ equally well with any bounded measurable φ on $\mathbb{T}$ and that exactly the same line of reasoning leads to the μ-almost convergence of $\{f_N^\varphi\}$ to a limit f^φ. On the other hand, because the convergence of λ_x^N to λ_x is weak, it would be naive to think that (4.8) continues

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

to hold in general. Indeed, choose any $p \in \mathcal{Z} \setminus \mathcal{R}$, and take $\varphi_p = \mathbf{1}_{\mathcal{E}_p}$. Because $\mathcal{E}_p$ has m-measure 0, $f_N^{\varphi_p} \equiv 0$. On the other hand, $\lambda_p(\mathcal{E}_p) = 1$. Hence, (4.8) fails in this case. Be that as it may, standard facts about weak convergence of probability measures guarantee that (4.8) holds for all $x \in \mathcal{Z} \setminus \mathcal{R}$ with the property that φ is continuous λ_x-almost everywhere. In particular, if φ is a bounded, Borel measurable function on $\mathbb{T}$ which is continuous m-almost everywhere, then there will exist a μ-null set $\mathcal{R}_\varphi \supseteq \mathcal{R}$ such that

$$f^\varphi(x) = \int \varphi \, d\lambda_x \quad \text{for all } x \in \mathcal{Z} \setminus \mathcal{R}_\varphi. \tag{4.11}$$

In view of the preceding, this comes down to showing that φ is continuous λ_x-almost everywhere for μ-almost every $x \in \mathcal{Z} \setminus \mathcal{R}$. To see this, note that because $\int \varphi \, dm = \iint \varphi(y)\lambda_x(dy)\mu(dx)$ for all $\varphi \in C(\mathbb{T})$, the equality remains true for all bounded measurable φ. In particular, this proves that $m(\Gamma) = 0$ if and only if $\lambda_x(\Gamma) = 0$ for μ-almost every $x \in \mathcal{Z} \setminus \mathcal{R}$.

5. *More general setting: exchangeable measures*

In this section we will give a substantial extension of the setting of §§1–3, to a setting that removes all group theoretic and topological considerations.

We set some notation. Let $(E, \mathcal{B})$ be a measurable space, and set $\Omega = E^\infty$ and $\mathcal{F} = \mathcal{B}^\infty$. For each $N \in \mathbb{Z}^+$, define

$$\omega = (x_1, \ldots, x_N, \ldots) \in \Omega \longmapsto \Pi_N \omega \equiv (x_1, \ldots, x_N) \in E^N, \tag{5.1}$$

and take $\mathcal{F}_N = \Pi_N^{-1}(\mathcal{B}^N)$. Next, let $\mathcal{S} = \cup \mathcal{S}_N$ be the group acting on Ω by permuting finitely many of the coordinates x_j, as in §2. Finally, let $\mathbf{M}_1(\Omega)$ denote the space of all probability measures on Ω, and let $\widehat{\mathbf{M}}_1(\Omega)$ and $\widehat{\mathcal{F}}$ denote those elements of $\mathbf{M}_1(\Omega)$ and $\mathcal{F}$, respectively, that are $\mathcal{S}$-invariant. These probability measures were studied by De Finetti, in his theory of exchangeable processes. The following result is a modern version, due to [**HS**].

THEOREM 5.1. *The set $\widehat{\mathbf{M}}_1(\Omega)$ is convex, and its extreme elements are precisely those $P \in \widehat{\mathbf{M}}_1(\Omega)$ for which $P(\Gamma) \in \{0, 1\}$ whenever $\Gamma \in \widehat{\mathcal{F}}$. Alternatively, $P \in \widehat{\mathbf{M}}_1(\Omega)$ is extreme if and only if $P = \mu^{\mathbb{Z}^+}$ where $\mu = (\Pi_1)_* P$.*

The first statement here is a special case of the fact that ergodic probability measures for a group action generally form the extreme elements of the set of invariant probability measures. The second statement follows by De Finetti's classification of probability measures that are invariant and ergodic under the action of $\mathcal{S}$. Using Theorem 5.1, we establish the following result.

THEOREM 5.2. *Let P be an extreme element of $\widehat{\mathbf{M}}_1(\Omega)$ and Q a second element of $\widehat{\mathbf{M}}_1(\Omega)$. If $Q \restriction \mathcal{F}_N \ll P \restriction \mathcal{F}_N$ for each $N \in \mathbb{Z}^+$, then either $P \perp Q$ or $P \ll Q$. In fact, if $P \not\perp Q$, then there exists a $G \in \widehat{\mathcal{F}}$ such that $Q(G) > 0$ and $P = Q(\cdot|G)$.*

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

Proof. We can choose $\{X_N : N \in \mathbb{Z}^+\} \subset L^1(\Omega, P)$ so that, for each $N \in \mathbb{Z}^+$, the following properties hold:

(a) X_N is $\mathcal{F}_N$-measurable and $[0, \infty)$-valued;

(b) $X_N = X_N \circ \sigma$ for all $\sigma \in \mathcal{S}_N$;

(c) X_N is a Radon–Nikodym derivative of $Q \restriction \mathcal{F}_N$ with respect to $P \restriction \mathcal{F}_N$.

Set $X = \liminf_{N\to\infty} X_N$. If Q_a is the absolutely continuous part in the Lebesque–Radon–Nikodym decomposition of Q relative to P, then X is the Radon–Nikodym derivative of Q_a with respect to P. At the same time, because X is $\widehat{\mathcal{F}}$-measurable, $P(X = c) = 1$ for some $c \in [0, \infty)$. Hence, $Q_a = cP$, from which the dichotomy $P \perp Q$ or $P \ll Q$ is immediate. Finally, if $c > 0$ and therefore $P \ll Q$, then $Q \restriction \mathcal{F}_N$ is equivalent to $P \restriction \mathcal{F}_N$ for each $N \in \mathbb{Z}^+$; and therefore $cP(\Gamma) = Q_a(\Gamma) = Q(\Gamma \cap G)$, $\Gamma \in \mathcal{F}$, where $G \equiv \{X \in (0, \infty)\}$. But clearly this means that $0 < c = Q(G)$ and so $P = Q(\cdot|G)$. □

We can use Theorem 5.2 to extend the scope of the study in §§1–3 as follows. Let μ be a probability measure on $(E, \mathcal{B})$, $(F, \mathcal{C}, \nu)$ a second probability space, and $\Phi : E \times F \longrightarrow [0, \infty)$ a $\mathcal{B} \times \mathcal{C}$-measurable function with the properties that

$$\int_E \Phi(x, y)\, d\mu(x) = 1 \quad \text{for every } y \in F$$

and

$$\mu(\{x : \Phi(x, y) = 1\}) < 1 \quad \text{for } \nu\text{-almost every } y \in F. \tag{5.2}$$

Also, for each $y \in F$, let μ_y^Φ be the probability measure on $(E, \mathcal{B})$ given by $\mu_y^\Phi = \Phi(x, y)\mu$. Finally, set

$$P = \mu^{\mathbb{Z}^+}, \quad Q_y^\Phi = (\mu_y^\Phi)^{\mathbb{Z}^+} \quad \text{and} \quad Q = \int_F Q_y^\Phi\, d\nu(y). \tag{5.3}$$

COROLLARY 5.3. *Assume that $\mathcal{B}$ is countably generated, and refer to the preceding paragraph. Then $Q \perp P$. Equivalently, if*

$$f_N^\Phi(\omega) = \int_F \Phi(x_1, y) \cdots \Phi(x_N, y)\, d\nu(y) \quad \textit{when } \omega = (x_1, \ldots, x_N, \ldots) \in \Omega, \tag{5.4}$$

then $f_N^\Phi(\omega) \longrightarrow 0$ as $N \to \infty$ for P-almost every $\omega \in \Omega$.

Proof. In that f_N^Φ is a Radon–Nikodym derivative of $Q \restriction \mathcal{F}_N$ with respect to $P \restriction \mathcal{F}_N$, it suffices to prove that $Q \perp P$. To this end, assume the contrary and apply Theorem 5.2 to find a $G \in \widehat{\mathcal{F}}$ such that $Q(G) > 0$ and $P = Q(\cdot|G)$. Because $Q_y^\Phi(G) \in \{0, 1\}$ for each $y \in F$,

$$\nu(S) = Q(G) \quad \text{and} \quad P = \frac{1}{\nu(S)} \int_S Q_y^\Phi\, d\nu(y), \tag{5.5}$$

where

$$S = \{y \in F : Q_y^\Phi(G) = 1\}. \tag{5.6}$$

Hence, for any $\mathcal{C}$-measurable $C \subset S$ of strictly positive ν-measure, the fact that P is extreme in $\widehat{\mathbf{M}}_1(\Omega)$ leads to the conclusion that

$$P = \frac{1}{\nu(C)} \int_C Q_y^\Phi\, d\nu(y). \tag{5.7}$$

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

Equivalently, for any $\Gamma \in \mathcal{F}$, $P(\Gamma) = Q_y^\Phi(\Gamma)$ for ν-almost every $y \in S$. Since $\mathcal{B}$ and therefore $\mathcal{F}$ are countably generated, this means that $P = Q_y^\Phi$ for ν-almost every $y \in S$. On the other hand, for ν-almost every $y \in F$, $(\Pi_1)_* P = \mu \neq \mu_y^\Phi = (\Pi_1)_* Q_y^\Phi$, and so it cannot be that $\nu(S) > 0$. Equivalently, it must be that $Q \perp P$. □

Remark 5.4. Corollary 5.3 contains Theorem 2.4 as a special case. Indeed, all that one has to do is take $E = G = F$, $\mu = m = \nu$, and $\Phi(x, y) = \mathbf{1}_A(y^{-1}x)$. In fact, one can easily make considerable extensions of the setting described in §§1–3. For example, let $(E, \mathcal{B})$ be a separable (i.e. $\mathcal{B}$ is countably generated) measurable space on which a measurable space $(F, \mathcal{C})$ acts measurably. That is, there is a measurable map $(x, y) \in E \times F \longmapsto yx \in E$. Furthermore, assume that E admits a probability measure μ with the property that $\mu \ll y_*\mu$ for each $y \in F$. Then, because $(E, \mathcal{B})$ is separable, we can find a $\mathcal{B} \times \mathcal{C}$-measurable $(x, y) \in E \times F \longmapsto R_y(x) \in [0, \infty)$ such that, for each $y \in F$, R_y is a Radon–Nikodym derivative of μ with respect to $y_*\mu$. Finally, let $\varphi : E \longrightarrow [0, \infty)$ be a $\mathcal{B}$-measurable function with μ-integral 1, and let ν be a probability measure on $(F, \mathcal{C})$ with the property that

$$\int_F (\varphi(yx)R_y(x))^N \, d\nu(y) < \infty \quad \text{for all } N \in \mathbb{Z}^+ \text{ and } x \in E. \tag{5.8}$$

We can then take $\Phi(x, y) = \varphi(yx)R_y(x)$ in Corollary 5.3.

6. *Ergodic decomposition of exchangeable measures*

In this section we continue in the setting described in §5, only now we will be assuming that E is a compact metric space, F is a Polish space (i.e. a separable topological space that admits a complete metric), and that $\mathcal{B}$ and $\mathcal{C}$ are the Borel fields over E and F, respectively. Define $\Omega = E^\infty$, $\mathcal{F}$, $\widehat{\mathcal{F}}$, etc., as in §6. Clearly, Ω is again a compact metric space. Moreover, when we consider it under what probabilists call the weak topology and analysts call the weak* topology (i.e. the topology in which convergence corresponds to convergence of integrals against continuous functions), $\mathbf{M}_1(\Omega)$ is itself a compact metric space. Hence, since the maps $\sigma \in \mathcal{S}$ are continuous, $\widehat{\mathbf{M}}_1(\Omega)$ is also compact. Furthermore, if $\mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$ denotes the set of extreme elements of $\widehat{\mathbf{M}}_1(\Omega)$, then $\mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$ too is compact. To see this, use the facts (cf. Theorem 5.1) that $\mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$ can be identified with $\{\mu^{\mathbb{Z}^+} : \mu \in \mathbf{M}_1(E)\}$, $\mathbf{M}_1(E)$ is compact, and $\mu \mapsto \mu^{\mathbb{Z}^+}$ is continuous.

THEOREM 6.1. *Let ρ and ρ' be a pair of probability measures on $\mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$. Then*

$$\begin{aligned} \rho = \rho' &\iff \int P \, d\rho(P) = \int P \, d\rho'(P), \\ \rho \perp \rho' &\iff \int P \, d\rho(P) \perp \int P \, d\rho'(P). \end{aligned} \tag{6.1}$$

Proof. The proof of the non-trivial part of the first line of (6.1) is an immediate consequence of the uniqueness assertion in the highly non-trivial Choquet Representation Theorem (cf. [**Cho**, Theorem 31.3] or [**Ph**, p. 83]), since $\widehat{\mathbf{M}}_1(\Omega)$ is a simplex.

Turning to the second line of (6.1), set $Q = \int P \, d\rho(P)$ and $Q' = \int P \, d\rho'(P)$. If $Q \perp Q'$, then there is a $B \in \mathcal{F}$ such that $Q(B) = 0 = Q'(B\complement)$. In fact, by

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

replacing B by $\bigcup_{\sigma\in\mathcal{S}}\sigma B$ if necessary, we may and will assume that $B \in \widehat{\mathcal{F}}$. But, because $P(B) \in \{0,1\}$ for each $P \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$, this means that $\rho(\Gamma) = 0 = \rho'(\Gamma\complement)$ when $\Gamma = \{P \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega) : P(B) = 1\}$. To go in the other direction, assume that $Q \not\perp Q'$, and let Q_{a} denote the absolutely continuous part of Q with respect to Q'. Then we can find a $G \in \mathcal{F}$ such that $Q_{\mathrm{a}}(\Gamma) = Q(\Gamma \cap G)$ for all $\Gamma \in \mathcal{F}$. Moreover, we again may and will assume that $G \in \widehat{\mathcal{F}}$. Similarly, we may and will take a version of the Radon–Nikodym derivative $h = dQ_{\mathrm{a}}/dQ'$ to be non-negative and $\widehat{\mathcal{F}}$-measurable. Hence, if $S = \{P \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega) : P(G) = 1\}$ and $H(P) = \mathbb{E}^P[h]$ for $P \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$, then, for all $\Gamma \in \mathcal{F}$,

$$\begin{aligned}\int \mathbf{1}_S(P)P(\Gamma)\,d\rho(P) &= \int P(\Gamma\cap G)\,d\rho(P) = Q_{\mathrm{a}}(\Gamma)\\ &= \int_\Gamma h(\omega)\,dQ'(\omega) = \int \mathbb{E}^P[\mathbf{1}_\Gamma h]\,d\rho'(P) = \int P(\Gamma)H(P)\,d\rho'(P). \qquad (6.2)\end{aligned}$$

Hence, by the first line of (6.1), $\mathbf{1}_S\,d\rho = H\,d\rho'$, which, because $\rho(S) = Q(G) = Q_{\mathrm{a}}(\Omega) > 0$, means that $\rho \not\perp \rho'$. □

From Theorem 6.1 we will obtain a result that extends Proposition 3.3 and Theorem 3.5. In the following application, μ is a probability measure on $(E,\mathcal{B})$, Φ and Φ' are non-negative, $\mathcal{B}\times\mathcal{C}$-measurable maps on $E\times F$ with the properties that

$$\int \Phi(x,y)\,d\mu(x) = 1 = \int \Phi'(x,y)\,d\mu(x) \quad \text{for all } y \in F, \qquad (6.3)$$

and (cf. the paragraph preceding Corollary 5.3) $y \mapsto \mu_y^\Phi$ and $y \mapsto \mu_y^{\Phi'}$ are continuous maps from F into $\mathbf{M}_1(E)$. Furthermore, ν is a probability measure on $(F,\mathcal{C})$, and $Q = \int P_y^\Phi\,d\nu(y)$ and $Q' = \int P_y^{\Phi'}\,d\nu(y)$ where $P_y^\Phi = (\mu_y^\Phi)^{\mathbb{Z}^+}$ and $P_y^{\Phi'} = (\mu_y^{\Phi'})^{\mathbb{Z}^+}$.

COROLLARY 6.2. *$Q \perp Q'$ if and only if there is a $S \in \mathcal{C}$ of full ν-measure such that*

$$\{\mu^\Phi : y \in S\} \cap \{\mu^{\Phi'} : y \in S\} = \emptyset. \qquad (6.4)$$

Also, $Q = Q'$ if and only if there exists an $S \in \mathcal{C}$ of full ν-measure such that

$$\{\mu_y^\Phi : y \in S\} \subset \{\mu_y^{\Phi'} : y \in F\} \quad \textit{and} \quad \{\mu_y^{\Phi'} : y \in S\} \subset \{\mu_y^\Phi : y \in F\}. \qquad (6.5)$$

In fact, if $Q = Q'$ and the support $\mathrm{supp}(\nu)$ *of ν is compact in F, then*

$$\{\mu_y^\Phi : y \in \mathrm{supp}(\nu)\} = \{\mu_y^{\Phi'} : y \in \mathrm{supp}(\nu)\}. \qquad (6.6)$$

Proof. In both cases, the sufficiency is clear. To prove the necessity, let ρ and ρ' be the distributions under ν of $y \in F \longmapsto P_y^\Phi \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$ and $y \in F \longmapsto P_y^{\Phi'} \in \mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$, respectively.

First, suppose that $Q \perp Q'$. Then, by the second line of (6.1), $\rho \perp \rho'$. Choose a Borel set $B \subseteq \mathbf{E}\widehat{\mathbf{M}}_1(\Omega)$ so that $\rho(B) = 0 = \rho'(B\complement)$, and set $S = \{y \in F : P_y^\Phi \notin B \;\&\; P_y^{\Phi'} \in B\}$. Then $S \in \mathcal{C}$, $\nu(S) = 1$, and $\{P_y^\Phi : y \in S\} \cap \{P_y^{\Phi'} : y \in S\} = \emptyset$.

Now assume that $Q = Q'$. Then, by the first line of (6.1), $\rho = \rho'$. Next, using Ulam's Lemma, choose a σ-compact subset $\widetilde{F}$ of F with $\nu(\widetilde{F}) = 1$.

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press

Then, $A \equiv \{P_y^{\Phi} : y \in \widetilde{F}\} \cap \{P_y^{\Phi'} : y \in \widetilde{F}\}$ is a Borel set, which, because $\rho = \rho'$, has ρ measure 1. Hence, $S \equiv \{y \in \widetilde{F} : P_y^{\Phi} \in A\}$ is a Borel set with ν-measure 1, and clearly S has the required property.

Finally, suppose that $Q = Q'$ and that $\mathrm{supp}(\nu)$ is compact. Without loss of generality, we may and will assume $F = \mathrm{supp}(\nu)$. Notice that both $\{\mu_y^{\Phi} : y \in F\}$ and $\{\mu_y^{\Phi'} : y \in F\}$ are compact in $\mathbf{M}_1(E)$. Hence, if S is chosen as in the preceding paragraph, then its closure $\overline{S}$ will have all the required properties. In particular, because $\nu(\overline{S}) = 1$ and $\mathrm{supp}(\nu) = F$, it follows that $\overline{S} = F$. □

Acknowledgements. Conversations with H. Furstenberg and B. Weiss have been stimulating for this work. The authors' research has been supported by NSF grants, DMS-9705288, DMS-9625782, and DMS-9877077, respectively.

References

[AJ1] E. Andersen and B. Jessen. Some limit theorems on integrals in an abstract set. *Danske Vid. Selsk. Mat.-Fys.* **22**(14) (1946), 1–29.

[AJ2] E. Andersen and B. Jessen. Some limit theorems on set-functions. *Danske Vid. Selsk. Mat.-Fys.* **25**(5) (1948), 1–8.

[Cho] G. Choquet. *Lectures on Analysis.* Vol. 2. Benjamin, New York, 1969.

[D] J. Doob. *Stochastic Processes*. Wiley–Interscience, New York, 1953.

[HR] E. Hewitt and K. Ross. *Abstract Harmonic Analysis*. Vol. 2. Springer, New York, 1970.

[HS] E. Hewitt and L. Savage. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** (1955), 470–501.

[Ph] R. Phelps. *Lectures on Choquet's Theorem*. Van Nostrand, New York, 1966.

[R1] J. Rosenblatt. Convergence of series of translations. *Math. Annal.* **230** (1977), 245–272.

[R2] J. Rosenblatt. Almost everywhere convergence of series. *Math. Annal.* **280** (1988), 565–577.

https://doi.org/10.1017/S0143385702000925 Published online by Cambridge University Press