On coordinate-permutation-invariant signed Radon measures on Cartesian powers of compact Hausdorff spaces: a look at de Finetti's theorem

Tim D. Austin

June 26, 2008

Abstract We give two proofs of a representation theorem for those totally finite signed Radon measures on an infinite power K^J of a compact Hausdorff space K (with its product topology) with the property of invariance under permutation of coordinates, in terms of powers of Radon probability measures on K. This is essentially a version of de Finetti's theorem on exchangeable sequences of random variables. Finally we will also use it to deduce an interesting measure algebra isomorphism.

1 Notation and definitions

Before proceeding we fix some notation. If (Ω, Σ, μ) is a probability space and J is a set, we write $(\Omega^J, \Sigma^{\otimes J}, \mu^{\otimes J})$ for the standard product of a J-indexed family of copies of (Ω, Σ, μ). Henceforth we will suppose J is infinite. If $(\Omega^J, \mathrm{T}, \mu)$ is a measure space such that T and μ are invariant under permutation of coordinates, we will refer to this σ-algebra and measure as **permutation-invariant**. If L is a compact Hausdorff space we write $C(L)$ for Banach space of real-valued continuous functions on L, $C_{\mathbb{C}}(L)$ for the corresponding space of complex-valued continuous functions and $\mathcal{B}(L)$ for the Borel σ-algebra of L. For any set J we will write $\mathcal{P}J$ for the power set of J and for any cardinal α we write $[J]^\alpha$ (respectively, $[J]^{<\alpha}$) for the subset of $\mathcal{P}J$ containing those subsets of J of cardinality equal to (respectively, less than) α; the only cases we will need are α finite and countably infinite, where we will write ω in place of α. We also note that the set $[J]^{<\omega}$ is naturally upwards directed by inclusion. For K a compact Hausdorff space and $i \in J$ we denote by π_i the i^{th} coordinate map $K^J \to K$, and more generally for $I \subseteq J$ we write π_I for the canonical projection $K^J \to K^{\#(I)}$ given by taking the I-indexed coordinates (we will find throughout that the order in which these coordinates are taken is irrelevant due to permutation invariance, and so may be chosen arbitrarily in each case). For $I \in [J]^{<\omega}$ we write $A_{K,J,I}$ for the subalgebra of $C(K^J)$ containing the finite sums of functions of the form $\prod_{i\in I} f_i \circ \pi_i$, and set $A_{K,J} = \bigcup_{I\in[J]^{<\omega}} A_{K,J,I}$. If J is infinite then an elementary application of the Stone-Weierstrass theorem shows that $A_{K,J}$ is norm-dense in $C(K^J)$.

Given a compact Hausdorff space L, the Riesz representation theorem justifies identifying $C(L)^*$ with the space of totally finite signed Radon measures on L in the natural way. We write $\mathrm{Pr}(L)$ for the subset of $C(L)^*$ containing the Radon probability measures, that is

$$\mathrm{Pr}(L) = \left\{\lambda \in C(L)^* : \ \|\lambda\| = 1, \textstyle\int_L \chi_L \, d\lambda = 1\right\}.$$

As is standard, this is a compact Hausdorff space with its vague topology (that induced by the weak*-topology on $C(L)^*$), and we shall henceforth take it with this topology unless otherwise stated.

If $\lambda \in \mathrm{Pr}(K)$ then we see that $\lambda^{\otimes J} \in \mathrm{Pr}(K^J)$, that is, it is a Radon probability measure on K^J. Hence for any $f \in C(K^J)$ we may consider the function $\Psi f : \mathrm{Pr}(K) \to \mathbb{R}$ given by $\Psi f(\lambda) = \int_{K^J} f \, d(\lambda^{\otimes J})$. In fact Ψf is continuous on $\mathrm{Pr}(K)$, as may readily be seen using an approximation argument from the dense subalgebra $A_{K,J}$ of $C(K^J)$. Also we have $|\int_{K^J} f\,d(\lambda^{\otimes J})| \leq \|f\|$, since $\lambda^{\otimes J}$ is a probability measure, and taking $f = f_0 \circ \pi_0$ for $f_0 \in C(K)$ we have $|\int_{K^J} f\,d(\lambda^{\otimes J})| = \int_K f_0\,d\lambda = \|f_0\| = \|f\|$ for some appropriate $\lambda \in \Pr(K)$, so in fact Ψ is a norm-one linear operator $C(K^J) \to C(\Pr(K))$. Analogously, for $d \geq 1$ we may define a norm-one linear operator $\Psi_d : C(K^d) \to C(\Pr(K))$ by $\Psi_d f(\lambda) = \int_{K^d} f\,d(\lambda^{\otimes d})$.

We will now see that if J is infinite then $\Psi : C(K^J) \to C(\Pr(K))$ has dense image, by showing that $\Psi(A_{K,J})$ is dense in $C(\Pr(K))$. The functions $\lambda \mapsto \int_K f_0\,d\lambda$ for $f_0 \in C(K)$ are continuous on $\Pr(K)$ by the definition of the vague topology, and it is clear that they separate points. Also, $\int_K \chi_K\,d\lambda = 1$ for $\lambda \in \Pr(K)$. Hence the algebra B_K of finite sums of finite products of functions of this form is norm-dense in $C(\Pr(K))$. However, given $f_0, f_2, \ldots, f_m \in C(K)$, we may choose distinct $j_0, j_1, \ldots, j_m \in J$ (because J is assumed infinite), and now setting $f = \prod_{i \leq m} f_i \circ \pi_{j_i}$ we see that $f \in A_{K,J}$ and

$$\int_{K^J} f\,d(\lambda^{\otimes J}) = \prod_{i \leq m} \big(\int_K f_i\,d\lambda \big)$$

for any $\lambda \in \Pr(K)$. It follows from this and the linearity of Ψ that $\Psi(A_{K,J}) = B_K$, so Ψ has dense image in $C(\Pr(K))$ as claimed. Clearly the corresponding result may fail for Ψ_d.

Hence $\Psi^* : C(\Pr(K))^* \to C(K^J)^*$ is an isometric embedding, given explicitly by

$$\int_{K^J} f\,d(\Psi^*\nu) = \int_{\Pr(K)} \big(\int_{K^J} f\,d(\lambda^{\otimes J}) \big)\,d\nu(\lambda)$$

for $f \in C(K^J)$. The main goal of this work is to classify the image of this adjoint as precisely the subspace of $C(K^J)^*$ of permutation-invariant measures. Since any product $\lambda^{\otimes J}$ is clearly permutation-invariant, it is clear that the image is a subspace of this space; the non-trivial part of the result comes in proving the opposite inclusion. It clearly suffices to show that any permutation-invariant Radon probability measure on K^J is the image under Ψ^* of some Radon probability measure on $\Pr(K)$. We will give two alternative proofs of this. The first is very analytic, whereas the second involves a more probabilistic approach: it is interesting to note that the second is substantially shorter than the first.

Finally we will demonstrate an isomorphism between the measure algebras of the probability spaces $(\Pr(K), \mathcal{B}(\Pr(K)), \nu)$ and $(K^{\mathbb{N}}, \mathrm{T}, \mu|_{\mathrm{T}})$ where T is the tail σ-subalgebra of $\mathcal{B}(K^{\mathbb{N}})$.

Note that we will use the Roman letters i and e to denote the square root of -1 and the base of natural logarithms respectively.

2 First proof

The first approach builds up to the main result by first considering the case of a finite set and then reducing the full theorem to this case.

2.1 The case of a finite set

We now consider the case with K a finite set and $J = \mathbb{N}$; without loss of generality we will work with K of the form $\{0, 1, \ldots, n-1\}$, $n \in \mathbb{N} \setminus \{0\}$. First we must set up some more notation. For $n \in \mathbb{N} \setminus \{0\}$ set

$$\Delta_n = \big\{(y_0, y_1, \ldots, y_n) \in [0, \infty)^{n+1} :\ \textstyle\sum_{i \leq n} y_i = 1\big\},$$

the set of stochastic vectors in $\mathbb{R}^{n+1}$, and

$$\Xi_n = \big\{(y_0, y_1, \ldots, y_n) \in \mathbb{R}^{n+1} :\ \textstyle\sum_{i \leq n} y_i = 1\big\}.$$

It is clear that the map $\phi_n : \{\lambda \in C(\{0, 1, \ldots, n-1\})^* :\ \lambda(\{0, 1, \ldots, n-1\}) = 1\} \to \Xi_n$ given by

$$\phi_n(\lambda) = (\lambda(\{0\}), \lambda(\{1\}), \ldots, \lambda(\{n-1\}))$$

is an affine homeomorphism and that $\phi_n(\Pr(\{0,1,\ldots,n-1\}))=\Delta_n$. For the special case $n=2$ we will write λ_t for the measure $\phi_2^{-1}((t,1-t))$ for $t\in\mathbb{R}$.

Now let $K=\{0,1\}$. We will need the following two lemmas about finite Cartesian powers.

Lemma 1 *Suppose V,W are vector spaces over $\mathbb{R}$, $d\in\mathbb{N}\setminus\{0\}$ and $\theta:V^d\to W$ is a symmetric multilinear map. Then for any $v_0,v_1,\ldots,v_{d-1}\in V$ we have*

$$\begin{aligned}&\theta(v_0,v_1,\ldots,v_{d-1})\\&=\tfrac{1}{2^d d!}\textstyle\sum_{S\subseteq\{0,1,\ldots,d-1\}}(-1)^{d-\#(S)}\theta\big(\sum_{j<d}\xi_S(j)v_j,\sum_{j<d}\xi_S(j)v_j,\ldots,\sum_{j<d}\xi_S(j)v_j\big),\end{aligned}$$

writing $\xi_S=\chi_S-\chi_{\{0,1,\ldots,d-1\}\setminus S}$.

Proof This is a standard polarization identity; for a treatment in the language of symmetric algebras of vector spaces, see [1], section 1.9.3.

$\square$

Lemma 2 *If μ is a permutation-invariant probability measure on $\{0,1\}^{2d}$, then μ lies in*

$$\Psi_{2d}^*(\Pr(\Pr(\{0,1\})))=\mathrm{co}(\{\lambda^{\otimes 2d}:\ \lambda\in\Pr(\{0,1\})\})$$

if and only if

$$\textstyle\int_{\{0,1\}^{2d}}(g\circ\pi_{\{2c,2c+1,\ldots,2d-1\}})\cdot(f\circ\pi_{\{0,1,\ldots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\ldots,2c-1\}})\,d\mu\geq 0$$

for any $c\leq d$ and functions $f:\{0,1\}^c\to\mathbb{R}$, $g:\{0,1\}^{2(d-c)}\to[0,\infty)$.

Proof Suppose first that $\mu=\Psi_{2d}^*(\nu)$, and c, f, g are as above. Then we have immediately

$$\begin{aligned}&\textstyle\int_{\{0,1\}^{2d}}(g\circ\pi_{\{2c,2c+1,\ldots,2d-1\}})\cdot(f\circ\pi_{\{0,1,\ldots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\ldots,2c-1\}})\,d\mu\\&=\textstyle\int_{\Pr(\{0,1\})}\big(\int_{\{0,1\}^{2(d-c)}}g\,d(\lambda^{\otimes 2(d-c)})\big)\big(\int_{\{0,1\}^c}f\,d(\lambda^{\otimes c})\big)^2\,d\nu(\lambda)\geq 0,\end{aligned}$$

since the integrand with respect to ν is non-negative.

Now we suppose that μ_0 is a permutation-invariant probability measure on $\{0,1\}^{2d}$ which does not lie in $\Psi_{2d}^*(\Pr(\Pr(\{0,1\})))$, and will show that we can find c, f and g such that the condition of the lemma is violated. Consider the subspace M_0 of $C(\{0,1\}^{2d})^*$ consisting of those signed measures μ on $\{0,1\}^{2d}$ which are permutation-invariant, and the affine subspace $M\subset M_0$ of those permutation-invariant signed μ such that $\mu(\{0,1\}^{2d})=1$. Write $G=\Psi_{2d}^*(\Pr(\Pr(\{0,1\})))$. It is clear that G is a closed convex subset of M. We will show that it has non-empty relative interior in M. Suppose $\mu\in M_0$. This μ is specified uniquely by the values is takes on the singleton subsets of $\{0,1\}^{2d}$, and if two members $x_0=(x_{0,i})_{i<2d}$, $x_1=(x_{1,i})_{i<2d}$ of $\{0,1\}^{2d}$ have the same number of coordinates equal to 1, then by permutation-invariance we must have $\mu(\{x_0\})=\mu(\{x_1\})$. Letting $\theta:(C(\{0,1\})^*)^{2d}\to M$ be the symmetric multilinear map given by the averaging of products over permutations of coordinates,

$$\theta(\lambda_0,\lambda_1,\ldots,\lambda_{2d-1})=\tfrac{1}{(2d)!}\textstyle\sum_{\sigma\in S_{2d}}\lambda_{\sigma(0)}\otimes\lambda_{\sigma(1)}\otimes\cdots\otimes\lambda_{\sigma(2d-1)},$$

we see that any $\mu\in M$ can be written as a linear combination of measures of the above form with each λ_i equal to either δ_0 or δ_1; for if c of the λ_i are equal to δ_0 and the remaining $2d-c$ are equal to δ_1 and $x=(x_i)_{i<2d}\in\{0,1\}^{2d}$, then we have that $\theta(\lambda_0,\lambda_1,\ldots,\lambda_{2d-1})(\{x\})$ equals $\frac{1}{(2d)!}$ if c of the coordinates of x are equal to 0 and equals 0 otherwise. Now the polarization formula of Lemma 1 shows that this measure $\theta(\lambda_0,\lambda_1,\ldots,\lambda_{2d-1})$ can be written as a linear combination of product measures $\lambda^{\otimes 2d}$ with $\lambda\in C(\{0,1\})^*$. Perturbing these λ slightly and using homogeneity, we find that the set of linear combination of measures of the form $\lambda^{\otimes 2d}$ with $\lambda\in C(\{0,1\})^*$ and $\lambda(\{0,1\})=1$ is a dense subspace of M_0; since we are working in finite dimensions, it must be the whole of M_0. Hence the affine subspace of linear combinations of such measures with coefficients summing to 1 can have codimension at most 1 in M_0, and so must be precisely M. It follows that the convex hull of the curve $\{\lambda_t^{\otimes 2d}:\ t\in\mathbb{R}\}$ must have non-empty relative interior in M; since this curve is described by polynomials in t, the same must then be true of any non-trivial arc of it. In particular, G has non-empty relative interior in M, as desired.

Consider $G_0 = \mathrm{co}(G\cup\{\mu_0\})$. Letting μ_1 be an interior point of G (relative to M) and considering the line segment ℓ from μ_1 to μ_0, we see that $\ell\subset G_0$. In particular, we may let μ_2 be the unique point of $\ell\cap\partial G$. Let H be a supporting hyperplane of G (as a subset of M) at μ_2 separating G from μ_0. Since $\mu(\{0,1\}^{2d})=1$ for any $\mu\in M$, this H is of the form

$$\big\{\mu\in M:\ \textstyle\int_{\{0,1\}^{2d}} f_0\,d\mu = 0\big\}$$

for some $f_0:\{0,1\}^{2d}\to\mathbb{R}$.

We have that μ_2 may be written as a convex combination of extreme points of G, and then the necessary extreme points must also lie in H, hence μ_2 may be written as a convex combination of the members of

$$\big\{\lambda^{\otimes 2d}:\ \lambda\in\Pr(\{0,1\}),\ \textstyle\int_{\{0,1\}^{2d}} f_0\,d(\lambda^{\otimes 2d})=0\big\}.$$

Under ϕ the set of those $\lambda\in\Pr(\{0,1\})$ with $\int_{\{0,1\}^{2d}} f_0\,d(\lambda^{\otimes 2d})=0$ is mapped to

$$\big\{t\in[0,1]:\ \textstyle\sum_{i\le 2d} a_i t^i(1-t)^{2d-i}=0\big\},$$

where the coefficients a_i for $i\le 2d$ are given by

$$a_i=\textstyle\sum_{S\subseteq\{0,1,\dots,2d-1\},\#(S)=i} f_0((\chi_S(j))_{j<2d}).$$

This is the set of roots of the polynomial $p(t)=\sum_{i\le 2d} a_i t^i(1-t)^{2d-i}$ which lie in $[0,1]$. This p has degree at most $2d$. That H is a hyperplane not containing the whole of M clearly requires that p is nonconstant, and that H is a supporting hyperplane for G is equivalent to the statement $p(t)\ge 0$ for all $t\in[0,1]$. It follows that any roots in $(0,1)$ must be repeated. We now have various cases to consider.

1. Suppose that neither of 0, 1 is a root of p. Then if $t_0,t_1,\dots,t_{c-1}$ are the roots of p in $[0,1]$, since each of them is repeated we need $2c\le 2d$, so $c\le d$. For $i<c$, let f_i be a function $\{0,1\}\to\mathbb{R}$ with uniform norm 1 and with

$$\textstyle\int_{\{0,1\}} f_i\,d\lambda_{t_i} = f(0)t_i + f(1)(1-t_i)=0.$$

Set $f_i=\prod_{i<c}(f\circ\pi_i):\{0,1\}^c\to\mathbb{R}$. We see that

$$\begin{aligned}&\textstyle\int_{\{0,1\}^{2d}}(f\circ\pi_{\{0,1,\dots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\dots,2c-1\}})\,d(\lambda_t^{\otimes 2d})\\ &=\textstyle\prod_{i<c}\big(\int_{\{0,1\}} f_i\,d\lambda_t\big)^2\ge 0,\end{aligned}$$

with equality for $t\in[0,1]$ if and only if $t=t_i$ for some i. Since μ_1 is in the relative interior of G, when expressed as a convex combination of measures of the form $\lambda_t^{\otimes 2d}$ we must require some λ_t with $t\neq t_i$ for any i, and so we must have

$$\textstyle\int_{\{0,1\}^{2d}}(f\circ\pi_{\{0,1,\dots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\dots,2c-1\}})\,d\mu_1>0$$

and

$$\textstyle\int_{\{0,1\}^{2d}}(f\circ\pi_{\{0,1,\dots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\dots,2c-1\}})\,d\mu_2=0,$$

and hence

$$\textstyle\int_{\{0,1\}^{2d}}(f\circ\pi_{\{0,1,\dots,c-1\}})\cdot(f\circ\pi_{\{c,c+1,\dots,2c-1\}})\,d\mu_0<0$$

(since μ_2 lies strictly between μ_0 and μ_1 on ℓ), giving the desired contradiction.

2. Now suppose that exactly one of 0, 1 is a root of p; without loss of generality we suppose that 0 is a root. If p has $c-1$ roots in $(0,1)$, these must all be repeated, and so $2(c-1)+1 \leq 2d$, again using the fact that p has degree at most $2d$. It follows that $c \leq d$, so letting $t_0, t_1, \ldots, t_{c-1}$ be the roots of p in $[0,1]$, including 0, we may argue as in the above case.

3. Finally suppose both 0, 1 are roots p. Arguing as before we find that in this case p can have at most $d-1$ roots in $(0,1)$, and hence at most $d+1$ roots in $[0,1]$. If it has at most d roots overall, we can argue as above, so we are left with the case of exactly $d-1$ roots in $(0,1)$, say $t_0, t_1, \ldots, t_{d-2}$. For $i < d-1$ let $f_i : \{0,1\} \to \mathbb{R}$ be functions of uniform norm 1 with

   $$\int_{\{0,1\}} f_i \, d\lambda_{t_i} = f(0)t_i + f(1)(1-t_i) = 0$$

   for each i and define f from these f_i as before. Also let $g_0, g_1 : \{0,1\} \to [0,\infty)$ be given by $g_0(a) = \delta_{1,a}$, $g_1(a) = \delta_{0,a}$ for $a \in \{0,1\}$ (where $\delta_{i,j}$ is the Kronecker delta). Now we argue as above with the function

   $$(g_0 \circ \pi_{2d-2}) \cdot (g_1 \circ \pi_{2d-1}) \cdot (f \circ \pi_{\{0,1,\ldots,d-2\}}) \cdot (f \circ \pi_{\{d-1,d,\ldots,2d-3\}}).$$

   in place of $(f \circ \pi_{\{0,1,\ldots,c-1\}}) \cdot (f \circ \pi_{\{c,c+1,\ldots,2c-1\}})$.

□

Remark We observe that the above lemma amounts to identifying a particular family of linear functionals on M which are enough to separate $\Psi^*_{2d}(\mathrm{Pr}(\mathrm{Pr}(\{0,1\})))$ from the points of its complement.

The utility of the above result is immediately apparent from the content of the next lemma.

Lemma 3 *Let K be a compact Hausdorff space, μ a permutation-invariant Radon probability measure on $K^{\mathbb{N}}$, I_0, I_1, I_2 pairwise disjoint finite subsets of $\mathbb{N}$ with $\#(I_1) = \#(I_2)$ and $f \in C(K^{\#(I_1)})$, $g \in C(K^{\#(I_0)})$ with $g \geq 0$. Then*

$$\int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (f \circ \pi_{I_1}) \cdot (f \circ \pi_{I_2}) \, d\mu \geq 0.$$

Proof Using permutation invariance and considering $K^{\#(I_1)}$ instead of K if necessary, it clearly suffices to prove the result in the case $I_0 = \{0,1,\ldots,c\}$, $I_1 = \{c+1\}$, $I_2 = \{c+2\}$. Let $N \geq 2$. The permutation-invariance of μ now gives that

$$\begin{aligned}
&\int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (f \circ \pi_{c+1}) \cdot (f \circ \pi_{c+2}) \, d\mu \\
&= \int_{K^{\mathbb{N}}} \tfrac{1}{\binom{N}{2}} (g \circ \pi_{I_0}) \textstyle\sum_{1 \leq i < j \leq N} (f \circ \pi_{c+i}) \cdot (f \circ \pi_{c+j}) \, d\mu \\
&\geq \tfrac{1}{\binom{N}{2}} \min_{x \in K^{\mathbb{N}}} \big((g \circ \pi_{I_0}(x)) \textstyle\sum_{1 \leq i < j \leq N} (f \circ \pi_{c+i}(x))(f \circ \pi_{c+j}(x)) \big),
\end{aligned}$$

since μ is a probability measure. Take $x \in K^{\mathbb{N}}$, say $x = (x_i)_{i \in \mathbb{N}}$. An elementary computation now gives

$$\begin{aligned}
&\tfrac{1}{\binom{N}{2}} (g \circ \pi_{I_0}(x)) \textstyle\sum_{1 \leq i < j \leq N} (f \circ \pi_{c+i}(x))(f \circ \pi_{c+j}(x)) \\
&= \tfrac{1}{\binom{N}{2}} (g \circ \pi_{I_0}(x)) \textstyle\sum_{1 \leq i < j \leq N} f(x_{c+i}) f(x_{c+j}) \\
&= \tfrac{1}{N(N-1)} (g \circ \pi_{I_0}(x)) \big((f(x_{c+1}) + f(x_{c+2}) + \cdots + f(x_{c+N}))^2 \\
&\qquad\qquad - (f(x_{c+1})^2 + f(x_{c+2})^2 + \cdots + f(x_{c+N})^2) \big) \\
&\geq -\tfrac{1}{N(N-1)} (g \circ \pi_{I_0}(x)) \big(f(x_{c+1})^2 + f(x_{c+2})^2 + \cdots + f(x_{c+N})^2 \big) \geq -\tfrac{1}{N} \|g\| \|f\|^2,
\end{aligned}$$

since $g \geq 0$. This is independent of x and converges to 0 as we let $N \to \infty$, and so our estimate for the original integral now gives the result.

□

Corollary 4 *Let K be a compact Hausdorff space, μ a permutation-invariant Radon probability measure on $K^{\mathbb{N}}$, I_0, I_1, I_2 pairwise disjoint finite subsets of $\mathbb{N}$ with $\#(I_1) = \#(I_2)$ and $f \in C_{\mathbb{C}}(K^{\#(I_1)})$, $g \in C(K^{\#(I_0)})$ with $g \geq 0$. Then*

$$\int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (f \circ \pi_{I_1}) \cdot \overline{(f \circ \pi_{I_2})}\, d\mu \geq 0.$$

Proof Writing $f = \operatorname{Re} f + \mathrm{i} \operatorname{Im} f$ and expanding we have

$$\begin{aligned}&\int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (f \circ \pi_{I_1}) \cdot \overline{(f \circ \pi_{I_2})}\, d\mu \\ &= \int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (\operatorname{Re} f \circ \pi_{I_1}) \cdot (\operatorname{Re} f \circ \pi_{I_2})\, d\mu + \int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (\operatorname{Im} f \circ \pi_{I_1}) \cdot (\operatorname{Im} f \circ \pi_{I_2})\, d\mu \\ &\quad + \mathrm{i}\Big(\int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (\operatorname{Im} f \circ \pi_{I_1}) \cdot (\operatorname{Re} f \circ \pi_{I_2})\, d\mu - \int_{K^{\mathbb{N}}} (g \circ \pi_{I_0}) \cdot (\operatorname{Re} f \circ \pi_{I_1}) \cdot (\operatorname{Im} f \circ \pi_{I_2})\, d\mu \Big).\end{aligned}$$

The first two terms in the above expression are both non-negative, by Lemma 3, and the imaginary term vanishes by the permutation-invariance of μ.

□

Proposition 5 *If $K = \{0, 1\}$ then every permutation-invariant Radon probability measure on $K^{\mathbb{N}}$ lies in the image of Ψ^*.*

Proof For $d \geq 2$ let V_d be the algebra of subsets of $K^{\mathbb{N}}$ depending only on the first d coordinates. Combining Lemmas 2 and 3 we see that for any even $d \geq 2$ the weak* compact subset of $\Pr(\Pr(K))$ given by

$$\{\nu \in \Pr(\Pr(K)) : \Psi^*(\nu)|_{V_d} = \mu|_{V_d}\}$$

is non-empty; since these form a descreasing sequence of sets, it follows that their intersection is non-empty. Any member of this intersection will do.

□

We now start to deal with arbitrary finite K.

Lemma 6 *Suppose $K = \{0, 1, \ldots, n-1\}$, μ is a permutation-invariant Radon probability measure on $K^{\mathbb{N}}$ and $f : K \to \mathbb{R}$ is a function. Then there is a unique compactly-supported Radon probability measure ν_f on $\mathbb{R}$ such that*

$$\int_{K^{\mathbb{N}}} \prod_{i<d} (f \circ \pi_i)\, d\mu = \int_{\mathbb{R}} t^d\, d\nu_f(t)$$

for every $d \in \mathbb{N}$. Furthermore, ν_f is concentrated on $[\min_{x \in K} f(x), \max_{x \in K} f(x)]$, and if $f_1 = a + b f_0$ for $f_0, f_1 : K \to \mathbb{R}$, $a, b \in \mathbb{R}$, $b \neq 0$, then $\nu_{f_0}(E) = \nu_{f_1}(a + bE)$.

Proof We prove this by a little sleight of hand. If f is constant it is clear that we must have $\nu_f = \delta_{f(0)}$. So now suppose first that $\min_{x \in K} f(x) = 0$, $\max_{x \in K} f(x) = 1$. Define

$$\mu_f(\{0\}^p \times \{1\}^q \times \{0,1\} \times \{0,1\} \times \cdots) = \int_{K^{\mathbb{N}}} \prod_{i<p} (f \circ \pi_i) \prod_{p \leq j < p+q} ((1-f) \circ \pi_j)\, d\mu$$

for $p, q \in \mathbb{N}$; it is easy to verify that this extends uniquely to a permutation-invariant Radon probability measure μ_f on $\{0,1\}^{\otimes \mathbb{N}}$ (by appealing to the Daniell-Kolmogorov theorem, if the necessary raw manipulation does not appeal), since $0 \leq f, 1-f \leq 1$. Letting ν_f be the measure on $[0,1]$ given by Proposition 5 and using the identification of $\Pr(\{0,1\})$ with $[0,1]$ via ϕ_2, the existence result follows. Uniqueness is now also clear for this case.

Finally the full result follows by applying translations and dilations and observing that everything falls into place.

□

Lemma 7 *Suppose K is a compact Hausdorff space, μ is a permutation-invariant Radon probability measure on $K^{\mathbb{N}}$ and $f \in C(K)$ is a function. Then the sequence of functions*

$$\left(\sum_{i \le N} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j)\right)_{N \in \mathbb{N}}$$

converges uniformly in $C(K^{\mathbb{N}})$.

Proof Let $x = (x_i)_{i \in \mathbb{N}} \in K^{\mathbb{N}}$ and $N_1 > N_0$. We have the estimate

$$\begin{aligned} \left| \sum_{i \le N_1} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j)(x) - \sum_{i \le N_0} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j)(x) \right| &= \left| \sum_{N_0 < i \le N_1} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j)(x) \right| \\ &\le \sum_{N_0 < i \le N_1} \frac{1}{i!} \|f\|^i. \end{aligned}$$

We see immediately that this must converge to zero as $N_0 \to \infty$ from the identity

$$\sum_{i \in \mathbb{N}} \frac{1}{i!} \|f\|^i = \mathrm{e}^{\|f\|}.$$

□

Corollary 8 *Under the conditions of the previous lemma the limit*

$$\lim_{N \to \infty} \int_{K^{\mathbb{N}}} \sum_{i \le N} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j) \, d\mu$$

exists in $\mathbb{C}$.

□

Remark In fact we will need the above lemma and corollary only for the case with K a finite set, but as can be seen this has no bearing on the proof.

To proceed we will need the following result due to Bochner:

Lemma 9 *A continuous function $\psi : \mathbb{R}^d \to \mathbb{C}$ is said to be* ***positive definite*** *if*

$$\sum_{r,s \le m} c_r \overline{c_s} \psi(a_r - a_s) \ge 0$$

for any $c_0, c_1, \ldots, c_m \in \mathbb{C}$, $y_0, y_1, \ldots, y_m \in \mathbb{R}^d$. A continuous function $\psi : \mathbb{R}^d \to \mathbb{C}$ is the characteristic function of a Radon probability measure on $\mathbb{R}^d$ if and only if it is positive definite and $\psi(0) = 1$.

Proof This is a classic result of Fourier analysis; see, for example, [5].

□

Lemma 10 *Suppose $K = \{0, 1, \ldots, n-1\}$ and μ is a permutation-invariant Radon probability measure on $K^{\mathbb{N}}$. Then there is a Radon probability measure ν on Ξ_n such that for any function $f : K \to \mathbb{R}$*

$$\lim_{N \to \infty} \int_{K^{\mathbb{N}}} \sum_{i \le N} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j) \, d\mu = \int_{\Xi_n} \mathrm{e}^{\mathrm{i} \int_K f \, d(\phi_n^{-1}(y))} \, d\nu(y).$$

Proof Define $\psi : C(K) \to \mathbb{C}$

$$\psi(f) = \lim_{N \to \infty} \int_{K^{\mathbb{N}}} \sum_{i \le N} \frac{\mathrm{i}^i}{i!} \prod_{j<i} (f \circ \pi_j) \, d\mu.$$

The proof of Lemma 7 shows that ψ is continuous for the norm topology of $C(K)$. Identifying Ξ_d affinely with $\mathbb{R}^d$, we see that we want ψ to be the characteristic function of a Radon probability measure, and so an application of Bochner's theorem on positive definite functions (Lemma 9) shows that such a measure ν exists if and only if

1. $\psi(0) = 1$, and

2. we have

$$\textstyle\sum_{r,s\leq m} c_r\overline{c_s}\psi(f_r - f_s) \geq 0$$

   whenever $c_0, c_1, \ldots, c_m \in \mathbb{C}$, $f_0, f_1, \ldots, f_m \in C(K)$.

The first of these is immediate, since $\mu(K^{\mathbb{N}}) = 1$. There is a little work to do in proving the second. We will in fact show it for an arbitrary compact Hausdorff space K, by showing that the terms

$$\textstyle\sum_{r,s\leq m} c_r\overline{c_s} \int_{K^{\mathbb{N}}} \sum_{i\leq N} \frac{\mathrm{i}^i}{i!} \prod_{j<i}((f_r - f_s)\circ\pi_j)\, d\mu$$

and

$$\textstyle\int_{K^{\mathbb{N}}} \big(\sum_{r\leq m} c_r \sum_{i_0\leq N} \frac{\mathrm{i}^{i_0}}{i_0!} \prod_{j_0<i_0}(f_r\circ\pi_{j_0})\big) \cdot \overline{\big(\sum_{s\leq m} c_s \sum_{i_1\leq N} \frac{\mathrm{i}^{i_1}}{i_1!} \prod_{j_1<i_1}(f_s\circ\pi_{N+j_1})\big)}\, d\mu$$

converge to the same value as $N \to \infty$; since the latter is non-negative, by Corollary 4, the result will follow. To see the above, we expand and consider the coefficient of $c_r\overline{c_s}$ in the N^{th} term of the second sequence. We find that it is equal to

$$\begin{aligned}&\textstyle\int_{K^{\mathbb{N}}} \big(\sum_{i_0\leq N} \frac{\mathrm{i}^{i_0}}{i_0!} \prod_{j_0<i_0}(f_r\circ\pi_{j_0})\big) \cdot \big(\sum_{i_1\leq N} \frac{(-\mathrm{i})^{i_1}}{i_1!} \prod_{j_1<i_1}(f_s\circ\pi_{N+j_1})\big)\, d\mu\\ &\textstyle= \sum_{i\leq 2N} \sum_{p+q=i,p,q\leq N} \frac{\mathrm{i}^p(-\mathrm{i})^q}{p!q!} \int_{K^{\mathbb{N}}} \prod_{j_0<p}(f_r\circ\pi_{j_0}) \prod_{j_1<q}(f_s\circ\pi_{N+j_1})\, d\mu.\end{aligned}$$

Using the permutation-invariance of μ, for $i \leq N$ the i^{th} term of this outer sum is

$$\begin{aligned}&\textstyle\sum_{p+q=i,p,q\leq N} \frac{\mathrm{i}^p(-\mathrm{i})^q}{p!q!} \int_{K^{\mathbb{N}}} \prod_{j_0<p}(f_r\circ\pi_{j_0}) \prod_{j_1<q}(f_s\circ\pi_{N+j_1})\, d\mu\\ &\textstyle= \sum_{p+q=i,p,q\leq N} \frac{(-1)^q\mathrm{i}^{p+q}}{(p+q)!}\binom{p+q}{p} \int_{K^{\mathbb{N}}} \prod_{j_0<p}(f_r\circ\pi_{j_0}) \prod_{j_1<q}(f_s\circ\pi_{N+j_1})\, d\mu\\ &\textstyle= \sum_{p+q=i,p,q\leq N} \frac{\mathrm{i}^{p+q}}{(p+q)!}\binom{p+q}{p} \int_{K^{\mathbb{N}}} \prod_{j_0<p}(f_r\circ\pi_{j_0}) \prod_{j_1<q}(-f_s\circ\pi_{N+j_1})\, d\mu\\ &\textstyle= \sum_{P\cup Q=\{0,1,\ldots,i-1\},P\cap Q=\emptyset} \frac{\mathrm{i}^i}{i!} \int_{K^{\mathbb{N}}} \prod_{j_0\in P}(f_r\circ\pi_{j_0}) \prod_{j_1\in Q}(-f_s\circ\pi_{j_1})\, d\mu\\ &\textstyle= \int_{K^{\mathbb{N}}} \frac{\mathrm{i}^i}{i!} \prod_{j<i}((f_r - f_s)\circ\pi_j)\, d\mu.\end{aligned}$$

Hence

$$\begin{aligned}&\textstyle\lim_{N\to\infty} \int_{K^{\mathbb{N}}} \sum_{i\leq N} \frac{\mathrm{i}^i}{i!} \prod_{j<i}((f_r - f_s)\circ\pi_j)\, d\mu\\ &\textstyle= \lim_{N\to\infty} \sum_{i\leq N} \sum_{p+q=i,p,q\leq N} \frac{\mathrm{i}^p(-\mathrm{i})^q}{p!q!} \int_{K^{\mathbb{N}}} \prod_{j_0<p}(f_r\circ\pi_{j_0}) \prod_{j_1<q}(f_s\circ\pi_{N+j_1})\, d\mu.\end{aligned}$$

On the other hand, for any $x = (x_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}}$ we have

$$\begin{aligned}&\textstyle\big|\sum_{N<i\leq 2N} \sum_{p+q=i,p,q\leq N} \frac{\mathrm{i}^p(-\mathrm{i})^q}{p!q!} \prod_{j_0<p} f_r(x_{j_0}) \prod_{j_1<q} f_s(x_{N+j_1})\big|\\ &\textstyle\leq \sum_{N<i\leq 2N} \sum_{p+q=i,p,q\leq N} \frac{1}{p!q!}\|f_r\|^p\|f_s\|^q\\ &\textstyle= \big(\sum_{i\leq N} \frac{1}{i!}\|f_r\|^i\big)\big(\sum_{i\leq N} \frac{1}{i!}\|f_s\|^i\big) - \sum_{i\leq N} \frac{1}{i!}(\|f_r\| + \|f_s\|)^i\\ &\to \mathrm{e}^{\|f_r\|}\mathrm{e}^{\|f_s\|} - \mathrm{e}^{\|f_r\|+\|f_s\|} = 0,\end{aligned}$$

uniformly as $N \to \infty$. Putting these facts together gives the result.

□

Proposition 11 *If K is a finite set with its discrete topology then every permutation-invariant Radon probability measure on $K^{\mathbb{N}}$ lies in the image of Ψ^*.*

Proof We will prove this by showing that the measure ν from the previous theorem is the measure we want, under our identification of Ξ_n and Δ_n with spaces of measures on $\{0, 1, \ldots, n-1\}$. However, it is not easy to derive the necessary properties (for example, that ν is concentrated on Δ_n, and so gives a measure on $\Pr(K)$) from the characterization of its characteristic function given above. We can overcome these difficulties by using Lemma 6. For if $f : K \to \mathbb{R}$, we see from Lemma 6 and the fact that ν_f is compactly-supported that for any $b \in \mathbb{R}$

$$\begin{aligned}\textstyle\int_{\Xi_n} \mathrm{e}^{\mathrm{i}\int_K bf\,d(\phi_n^{-1}(y))}\,d\nu(y) &= \textstyle\lim_{N\to\infty}\int_{K^{\mathbb{N}}}\sum_{i\leq N}\frac{\mathrm{i}^i}{i!}\prod_{j<i}(bf\circ\pi_j)\,d\mu \\ &= \textstyle\lim_{N\to\infty}\int_{\mathbb{R}}\sum_{i\leq N}\frac{\mathrm{i}^i}{i!}(bt)^i\,d\nu_f(t) = \int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}bt}\,d\nu_f(t),\end{aligned}$$

and hence, using the uniform density of trigonometric polynomials in $C([u,v])$ for any $u, v \in \mathbb{R}$, $u < v$, that

$$\textstyle\int_{\Xi_n} h\big(\int_K f\,d(\phi_n^{-1}(y))\big)\,d\nu(y) = \int_{\mathbb{R}} h(t)\,d\nu_f(t)$$

for any $h \in C_{\mathrm{b}}(\mathbb{R})$. In particular, considering $f = \chi_{\{k\}}$ for $k < n$, so that ν_f is concentrated on $[0,1]$, we find that

$$\textstyle\int_{\Xi_n} h(y_k)\,d\nu(y) = \int_{\mathbb{R}} h(t)\,d\nu_f(t)$$

for any $h \in C_{\mathrm{b}}(\mathbb{R})$. This implies that ν is concentrated on $\{y \in \Xi_n :\ y_k \in [0,1]\}$ for each $k < n$, and so is actually concentrated on Δ_n. It follows that we may drop the boundedness assumption on h above, and that

$$\begin{aligned}\textstyle\int_{K^{\mathbb{N}}}\prod_{j<d}(f\circ\pi_j)\,d\mu &= \textstyle\int_{\mathbb{R}} t^d\,d\nu_f(t) = \int_{\Delta_n}\big(\int_K f\,d(\phi_n^{-1}(y))\big)^d\,d\nu(y) \\ &= \textstyle\int_{\Delta_n}\big(\int_{K^{\mathbb{N}}}\prod_{j<d}(f\circ\pi_j)\,d((\phi_n^{-1}(y))^{\otimes\mathbb{N}})\big)\,d\nu(y)\end{aligned}$$

for any $f : K \to \mathbb{R}$, $d \in \mathbb{N}$. Finally we may use linearity, permutation-invariance and the identity

$$\textstyle\frac{1}{d!}\sum_{\sigma\in S_d}\prod_{i<d} f_i\circ\pi_{\sigma(i)} = \frac{1}{2^d}\sum_{S\subseteq\{0,1,\dots,d-1\}}\big(\prod_{i<d}\xi_S(i)\big(\sum_{j<d}\xi_S(j)f_j\big)\circ\pi_i\big)$$

(obtained from another consideration of Lemma 1) to find that

$$\textstyle\int_{K^{\mathbb{N}}} f\,d\mu = \int_{\Delta_n}\big(\int_{K^{\mathbb{N}}} f\,d((\phi_n^{-1}(y))^{\otimes\mathbb{N}})\big)\,d\nu(y)$$

for all $f \in A_{K,\mathbb{N}}$. Now the result follows from the density of $A_{K,\mathbb{N}}$ in $C(K^{\mathbb{N}})$.

$\square$

Remark It is perhaps worth giving a brief summary of the ideas leading up to the above proof. First we obtained the result for K a two-point set via a rather hands-on result concerning the case of finite even Cartesian powers (Lemma 2). Then we used this to show that for an arbitrary finite set K we could obtain a family of auxiliary Radon probability measures on $\mathbb{R}$ (the ν_f of Lemma 6). The hope is then that these correspond to the projections of some common Radon probability measure ν on Ξ_n onto the one-dimensional subspaces of Ξ_n, for if so then we can show that ν is the measure we wanted from the properties of these projections. We obtain the existence of such a measure ν from Bochner's theorem (Lemma 10).

Remark It seems to be difficult to give an exact characterization of $\Psi_d^*(\mathrm{Pr}(K))$ for finite d and arbitrary finite K; the ideas in the last few propositions do not seem to extend readily beyond the case of an infinite Cartesian power.

2.2 The main theorem

Theorem 12 *For any compact Hausdorff space K, every permutation-invariant Radon probability measure on $K^{\mathbb{N}}$ lies in the image of Ψ^*; that is, for any such measure μ there is some (necessarily unique) Radon probability measure ν on $\mathrm{Pr}(K)$ such that*

$$\textstyle\int_{K^{\mathbb{N}}} f\,d\mu = \int_{\mathrm{Pr}(K)}\big(\int_{K^{\mathbb{N}}} f\,d(\lambda^{\otimes\mathbb{N}})\big)\,d\nu(\lambda)$$

for any $f \in C(K^{\mathbb{N}})$.

Proof We will prove this by showing that

$$\textstyle|\int_{K^{\mathbb{N}}} f\,d\mu| \leq \max_{\lambda\in\mathrm{Pr}(K)}|\int_{K^{\mathbb{N}}} f\,d(\lambda^{\otimes\mathbb{N}})|$$

for every $f \in A_{K,\mathbb{N}}$. This will show that μ induces a linear functional on the space $\Psi(A_{K,\mathbb{N}})$ of norm at most one; this may therefore be extended by continuity to a continuous linear functional on the whole of $C(\mathrm{Pr}(K))$, that is, a totally finite Radon measure on $\mathrm{Pr}(K)$. Hence μ lies in $\Psi^*(C(\mathrm{Pr}(K))^*)$, as required.

To obtain the desired inequality, let $f \in A_{K,\mathbb{N}}$ and $\varepsilon > 0$. Using the uniform continuity of f and the fact that it depends on only finitely many coordinates, we may take a finite indexed Borel partition $A = (A_i)_{i \leq n}$ of K and a choice $T = (t_i)_{i \leq n}$ in A such that $|f(x_0) - f(x_1)| < \varepsilon$ for $x_0 = (x_{0,i})_{i\in\mathbb{N}}, x_1 = (x_{1,i})_{i\in\mathbb{N}} \in K^{\mathbb{N}}$ whenever $x_{0,i}, x_{1,i}$ lie in the same member of A for each $i \in \mathbb{N}$. Set $K_0 = \{t_i : i \leq n\}$. We define a permutation-invariant Radon probability measure $\tilde{\mu}$ on $K_0^{\mathbb{N}}$ as follows. If $E = \prod_{j\in\mathbb{N}} E_j$ is a finite-dimensional cylinder set in $K_0^{\mathbb{N}}$, say with $E_j = K_0$ for $j > m$, we set

$$\tilde{\mu}(E) = \textstyle\sum_{(x_0,x_1,\ldots,x_m)\in E_0\times E_1\times\cdots\times E_m} \mu(A(x_0) \times A(x_1) \times \cdots \times A(x_m) \times K \times \cdots),$$

where for $x \in K_0$ we write $A(x)$ for the unique member of A containing x. It is elementary to see that this recipe can be extended to give a $\tilde{\mu}$ of the desired form, and also that we must have $|\int_{K^{\mathbb{N}}} f\,d\mu - \int_{K_0^{\mathbb{N}}} (f|_{K_0^{\mathbb{N}}})\,d\tilde{\mu}| \leq \varepsilon$. Since K_0 is finite we know that there is some Radon probability measure $\tilde{\nu}$ on $\mathrm{Pr}(K_0)$ such that

$$\textstyle\int_{K_0^{\mathbb{N}}} (f|_{K_0^{\mathbb{N}}})\,d\tilde{\mu} = \int_{\mathrm{Pr}(K_0)} \big(\int_{K_0^{\mathbb{N}}} (f|_{K_0^{\mathbb{N}}})\,d(\lambda^{\otimes\mathbb{N}})\big)\,d\tilde{\nu}(\lambda) = \int_{\mathrm{Pr}(K)} \big(\int_{K_0^{\mathbb{N}}} f\,d(\lambda^{\otimes\mathbb{N}})\big)\,d\nu(\lambda),$$

where ν is the Radon probability measure pushed forward from $\tilde{\nu}$ by the natural embedding $\mathrm{Pr}(K_0) \to \mathrm{Pr}(K)$. It follows that

$$\textstyle|\int_{K_0^{\mathbb{N}}} (f|_{K_0^{\mathbb{N}}})\,d\tilde{\mu}| \leq \max_{\lambda\in\mathrm{Pr}(K)} |\int_{K^{\mathbb{N}}} f\,d(\lambda^{\otimes\mathbb{N}})|;$$

since ε was arbitrary the desired result follows.

□

Theorem 13 *For any compact Hausdorff space K and any infinite set J, every permutation-invariant Radon probability measure on K^J lies in the image of Ψ^*; that is, for any such measure μ there is some (necessarily unique) Radon probability measure ν on $\mathrm{Pr}(K)$ such that*

$$\textstyle\int_{K^J} f\,d\mu = \int_{\mathrm{Pr}(K)} \big(\int_{K^J} f\,d(\lambda^{\otimes J})\big)\,d\nu(\lambda)$$

for any $f \in C(K^J)$.

First proof If J is countable, this is clearly equivalent to the case $J = \mathbb{N}$. If J is uncountable, let $J_0 \subset J$ be countably infinite. The restriction of μ to sets depending only on coordinates in J_0 gives a permutation-invariant Radon probability measure on K^{J_0}, which we know by the above can be represented by some Radon probability measure ν on $\mathrm{Pr}(K)$ in the way described. It follows that for any $f \in C(K^J)$ depending only on coordinates in J_0 we have

$$\textstyle\int_{K^J} f\,d\mu = \int_{\mathrm{Pr}(K)} \big(\int_{K^J} f\,d(\lambda^{\otimes J})\big)\,d\nu(\lambda).$$

However, any $f \in C(K^J)$ depends on only countably many coordinates, so applying some permutation of these coordinates we can ask that they all be indexed by J_0; since both sides of the above equality are unaffected by this operation, we see that it must hold for arbitrary $f \in C(K^J)$.

□

3 Second proof

The linchpin of the second proof of the main result will be the identification of a weakly*-convergent net of Radon probability measures on $\mathrm{Pr}(K)$ with limit the desired measure ν.

Suppose J is infinite, and for $n \in \mathbb{N} \setminus \{0\}$ let μ_n be the n-dimensional distribution of μ, given by $\mu_n = (\pi_I)_\# \mu$ for any $I \in [J]^n$ (this is well-defined by the permutation-invariance of μ). Also for $I \in [J]^{<\omega}$ let $\eta_I : K^I \to \mathrm{Pr}(K)$ be the map given by $\eta_I((x_i)_{i\in I}) = \frac{1}{\#(I)} \sum_{i \in I} \delta_{x_i}$; this is clearly continuous and has image the family of probability measures on K supported on $\#(I)$ points (counting multiplicity) with each point given equal weight. We will consider the net of pushforwards $((\eta_I)_\# \mu_{\#(I)})_{I \in [J]^{<\omega}}$ in $\mathrm{Pr}(\mathrm{Pr}(K))$; that is, for $I \in [J]^{<\omega}$ the measure $(\eta_I)_\# \mu_{\#(I)}$ is given by

$$\int_{\mathrm{Pr}(K)} g \, d((\eta_I)_\# \mu_{\#(I)}) = \int_{K^I} g \circ \eta_I \circ \pi_I \, d\mu$$

for $g \in C(\mathrm{Pr}(K))$.

Proposition 14 *For any $f \in C(K^J)$ we have*

$$\int_{\mathrm{Pr}(K)} \Psi f \, d((\eta_I)_\# \mu_{\#(I)}) \to \int_{K^J} f \, d\mu$$

as $I \uparrow$ in $[J]^{<\omega}$.

Proof By the uniform density of $A_{K,J}$ in $C(K^J)$ it suffices to consider f a finite sum of products of the form $\prod_{j \in I_0} f_j \circ \pi_j$ for $I_0 \in [J]^{<\omega}$; since permutation-invariance allows us to assume further that these I_0 are pairwise disjoint, in fact we need consider only one such product. Since we are letting $I \uparrow$, for f of this form we may assume that $I \supseteq I_0$, and, letting $m = \#(I_0)$, $n = \#(I)$, a simple computation using permutation-invariance gives

$$\begin{aligned}
\int_{\mathrm{Pr}(K)} \Psi f \, d((\eta_I)_\# \mu_{\#(I)}) &= \int_{K^J} \prod_{j \in I_0} \big(\int_K f_j \, d\big(\tfrac{1}{\#(I)} \textstyle\sum_{i \in I} \delta_{x_i}\big)\big) \, d\mu(x) \\
&= \int_{K^J} \prod_{j \in I_0} \big(\tfrac{1}{n} \textstyle\sum_{i \in I} f_j(x_i)\big) \, d\mu(x) \\
&= \tfrac{1}{n^m} \int_{K^J} \textstyle\sum_{\sigma \in I^{I_0}} \prod_{j \in I_0} f_j(x_{\sigma(j)}) \, d\mu(x) \\
&= \tfrac{1}{n^m} \int_{K^J} \textstyle\sum_{\sigma \in \Xi(I_0, I)} \prod_{j \in I_0} f_j(x_{\sigma(j)}) \, d\mu(x) \\
&\quad + \tfrac{1}{n^m} \int_{K^J} \textstyle\sum_{\sigma \in I^{I_0} \setminus \Xi(I_0, I)} \prod_{j \in I_0} f_j(x_{\sigma(j)}) \, d\mu(x),
\end{aligned}$$

writing $\Xi(I_0, I)$ for the subset of I^{I_0} containing the injections. There are $n(n-1)\cdots(n-m+1)$ of these, so since μ is permutation-invariant the first integral in the above sum contributes

$$\tfrac{n(n-1)\cdots(n-m+1)}{n^m} \int_{K^J} \prod_{j \in I_0} f_j(x_j) \, d\mu(x) = (1 - \tfrac{1}{n}) \cdots (1 - \tfrac{m-1}{n}) \int_{K^J} f \, d\mu \to \int_{K^J} f \, d\mu$$

as $n \to \infty$, and, since

$$\tfrac{\#(I^{I_0} \setminus \Xi(I_0, I))}{n^m} \to 0$$

as $n \to \infty$, we see that the second integral contributes 0. Putting these together gives the result.

$\square$

Remark We note that if we set $J = \mathbb{N}$, $\mu = \lambda^{\otimes \mathbb{N}}$ for some $\lambda \in \mathrm{Pr}(K)$ and take I_0 to be a singleton the above proof simply gives the Weak Law of Large Numbers.

Second proof of Theorem 13 Since $\mathrm{Pr}(K)$ is compact Hausdorff, the net $((\eta_I)_\# \mu_{\#(I)})_{I \in [J]^{<\omega}}$ has nonempty limit. Let ν be any point of this limit. Now the definition of the vague topology and Proposition 14 show that

$$\int_{\mathrm{Pr}(K)} \Psi f \, d\nu = \int_{K^J} f \, d\mu$$

for any $f \in C(K^J)$, so there must be precisely one such limit ν and it is the measure on $\mathrm{Pr}(K)$ that we wanted.

$\square$

4 Application: an interesting measure algebra isomorphism

We assume again that J is infinite. For $I \subseteq J$ we write Σ_I for the σ-subalgebra of $\mathcal{B}(K^J)$ generated by the Borel finite-dimensional cylinder sets independant of the I-indexed coordinates, and we write T for the tail σ-subalgebra of $\mathcal{B}(K^J)$, $\mathrm{T} = \bigcap_{I \in [J]^{<\omega}} \Sigma_I$.

Lemma 15 *For any $E \in \mathcal{B}(K^J)$, the map $\mathrm{Pr}(K) \ni \lambda \to \lambda^{\otimes J}(E) \in [0,1]$ is Borel-measurable.*

Proof Write $\mathcal{A}$ for the family of all Borel subsets E of K^J for which the map $\lambda \to \lambda^{\otimes J}(E)$ is Borel.

Suppose first that E is open, that $a \in [0,1]$ and that $\lambda^{\otimes J}(E) > a$. Then, since $\lambda^{\otimes J}$ is a Radon measure, there is some $f \in C(K^J)$ with $0 \leq f \leq \chi_E$ and $\int_K f \, d(\lambda^{\otimes J}) > a$, so the set $\{\lambda \in \mathrm{Pr}(K) : \ \lambda^{\otimes J}(E) > a\}$ is open, hence Borel. So $\mathcal{A}$ contains all open subsets of K^J.

Suppose now that $E_0, E_1 \in \mathcal{A}$ with $E_0 \subseteq E_1$ and $a \in [0,1]$. Then $\lambda^{\otimes J}(E_1 \setminus E_0) = \lambda^{\otimes J}(E_1) - \lambda^{\otimes J}(E_0)$, a difference of two Borel functions of λ and hence itself Borel, so $E_1 \setminus E_0 \in \mathcal{A}$. Also, if $(E_i)_{i \in \mathbb{N}}$ is a non-decreasing sequence in $\mathcal{A}$, then

$$\lambda^{\otimes J}\big(\textstyle\bigcup_{i \in \mathbb{N}} E_i\big) = \sup_{i \in \mathbb{N}} \lambda^{\otimes J}(E_i),$$

a supremum of Borel functions of λ, so $\bigcup_{i \in \mathbb{N}} E_i \in \mathcal{A}$. Therefore $\mathcal{A}$ is a Dynkin class of subsets of K^J containing the open subsets, and so must include the Borel subsets, by the monotone class theorem.

□

Theorem 16 *Let μ be a permutation-invariant Radon probability measure on K^J and ν the Radon probability measure on $\mathrm{Pr}(K)$ with $\mu = \Psi^* \nu$ (Theorem 13). Then the measure spaces $(K^J, \mathrm{T}, \mu|_{\mathrm{T}})$ and $(\mathrm{Pr}(K), \mathcal{B}(\mathrm{Pr}(K)), \nu)$ have isomorphic measure algebras.*

Proof Given $E \in \mathrm{T}$, define $\alpha(E) \subseteq \mathrm{Pr}(K)$ by $\alpha(E) = \{\lambda \in \mathrm{Pr}(K) : \ \lambda^{\otimes J}(E) = 1\}$; by Lemma 15 this is a Borel subset of $\mathrm{Pr}(K)$. We now see that $\alpha : T \to \mathcal{B}(\mathrm{Pr}(K))$ is a Boolean algebra homomorphism, for if $E_0, E_1 \in \mathrm{T}$ then, because J is infinite and $\lambda^{\otimes J}$ is a probability measure, Kolmogorov's zero-one law implies that

$$\begin{array}{rcl} \lambda^{\otimes J}(E_0) = 1 \text{ and } \lambda^{\otimes J}(E_1) = 1 & \iff & \lambda^{\otimes J}(E_0 \cap E_1) = 1, \\ \lambda^{\otimes J}(E_0) = 1 \text{ or } \lambda^{\otimes J}(E_1) = 1 & \iff & \lambda^{\otimes J}(E_0 \cup E_1) = 1, \end{array}$$

so $\alpha(E_0 \cap E_1) = \alpha(E_0) \cap \alpha(E_1)$ and $\alpha(E_0 \cup E_1) = \alpha(E_0) \cup \alpha(E_1)$.

Another application of Kolmogorov's zero-one law shows that, since $\mu = \Psi^* \nu$,

$$\begin{array}{rcl} \mu(E) & = & \int_{\mathrm{Pr}(K)} \lambda^{\otimes J}(E) \, d\nu(\lambda) \\ & = & \nu(\{\lambda \in \mathrm{Pr}(K) : \ \lambda^{\otimes J}(E) = 1\}) = \nu(\alpha(E)), \end{array}$$

so α is a measure-preserving map from T to $\mathcal{B}(\mathrm{Pr}(K))$. Finally we show that for any $F \in \mathcal{B}(\mathrm{Pr}(K))$ there is some $E \in \mathrm{T}$ with $\nu(F \triangle \alpha(E)) = 0$. Suppose first that F is a member of the usual subbase for the weak* topology, say

$$F = \big\{\lambda \in \mathrm{Pr}(K) : \ \textstyle\int_K f \, d\lambda > a\big\}$$

for some $f \in C(K)$, $a \in \mathbb{R}$. Let $(j_i)_{i \in \mathbb{N}}$ be an enumeration of a countably infinite subset of J. Define $g : K^J \to \mathbb{R}$ by

$$g((x_j)_{j \in J}) = \liminf_{n \to \infty} \tfrac{1}{n} \textstyle\sum_{i<n} f(x_{j_i});$$

as a limit infimum of a sequence of continuous functions this is Borel. By the Strong Law of Large Numbers we have for any $\lambda \in \mathrm{Pr}(K)$ that

$$\tfrac{1}{n} \textstyle\sum_{i<n} f(x_{j_i}) \to \int_K f \, d\lambda$$

as $n \to \infty$ for λ-a.e. $(x_j)_{j\in J} \in K^J$, so defining the function $\tilde{g} : \mathrm{Pr}(K) \to \mathbb{R}$ by $\tilde{g}(\lambda) = \int_{K^J} g \, d(\lambda^{\otimes J})$ the dominated convergence theorem shows that

$$\tilde{g}(\lambda) = \lim_{n\to\infty} \int_{K^J} \left(\tfrac{1}{n}\sum_{i<n} f(x_{j_i})\right) d(\lambda^{\otimes J})$$

for any $\lambda \in \mathrm{Pr}(K)$. Hence the function $\tilde{g}$ is also Borel with $\tilde{g}(\lambda) = \int_K f \, d\lambda$, and so

$$F = \left\{\lambda \in \mathrm{Pr}(K) : \; g((x_j)_{j\in J}) > a \text{ for } \lambda^{\otimes J}\text{-a.e. } (x_j)_{j\in J} \in K^J\right\} = \alpha(\{(x_j)_{j\in J} : \; g((x_j)_{j\in J}) > a\})$$

(since clearly $\{(x_j)_{j\in J} : \; g((x_j)_{j\in J}) > a\} \in \mathrm{T}$). Since α is a Boolean algebra homomorphism, it follows that any member of the usual base for the weak* topology on $\mathrm{Pr}(K)$ lies in the image of α. Given an arbitrary open subset $U \subseteq \mathrm{Pr}(K)$ and $\varepsilon > 0$, since ν is Radon we may find a compact set $W \subseteq U$ with $\nu(W) > \nu(U) - \varepsilon$, and now taking a finite cover of W by members of the usual base lying inside U, we see that we may find a finite union V of members of the usual base with $V \subseteq U$ and $\nu(V) > \nu(U) - \varepsilon$, and so, since ε was arbitrary, there is some countable union V_0 of members of this base with $\nu(U \triangle V_0) = 0$. Since we know that V_0 lies in the image of α, we see that the desired conclusion holds for any open subset of $\mathrm{Pr}(K)$, and so for any Borel subset of K. This completes the proof.

□

Closing Remark

The main result described in this work is really a version of the classic result of de Finetti that any exchangeable sequence of random variable is a mixture of i.i.d. sequences of random variables. This is described in many books on elementary probability theory, such as [6], often in roughly the following form:

Theorem 17 ***De Finetti's theorem*** *Suppose $X = (X_n)_{n\in\mathbb{N}}$ is a sequence of integrable real-valued random variables on a probability space (Ω, Σ, P) such that for any finitely-supported permutation σ of $\mathbb{N}$ the distribution of*

$$\Omega \ni \omega \mapsto (X_{\sigma(i)}(\omega))_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$$

under P is the same as that of

$$\Omega \ni \omega \mapsto (X_i(\omega))_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}.$$

For each $n \in \mathbb{N}$ let T_n be the σ-subalgebra of $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ containing those sets which are invariant under permutations of the first n coordinates, and let $\Sigma_n = X^{-1}(\mathrm{T}_n)$. Then writing

$$S_n = \tfrac{1}{n}\sum_{i<n} X_i,$$

we have that the sequence $(S_n)_{n\in\mathbb{N}}$ converges both P-almost surely and in $L^1(P)$-norm, and each S_n is a conditional expectation of X_1 with respect to the σ-subalgebra Σ_n of Σ.

This form of the result is closely related to our second proof of Theorem 13, where setting $J = \mathbb{N}$ we have that the measure $(\eta_{\{0,\dots,n-1\}})_{\#}\mu_n$ corresponds to the law of the n^{th} sample mean S_n (technically we cannot recover the above version of de Finetti's theorem from Theorem 13 immediately, as we have worked throughout with a compact Hausdorff space K; however, this problem is easily solved, for example by replacing $\mathbb{R}$ by its two-point compactification $[-\infty, \infty]$ in the above version and considering a permutation invariant measure on $[-\infty, \infty]^{\mathbb{N}}$ for which $\mathbb{R}^{\mathbb{N}}$ has full outer measure). The ideas behind de Finetti's theorem can be extended further in various ways (as witnessed by the much more topological and functional analytic flavour of this work); for a very general treatment, see [3], sections 458 and 459.

I was not aware of de Finetti's theorem when I first began to investigate permutation-invariant Radon measures on an infinite Cartesian power of a compact Hausdorff space, and all the results in this write-up were discovered independently. I have no references for the first proof, but it turns out that the second has been around for some time; see the important paper [4] of Hewitt and Savage, where the ideas of the second proof above are combined with an explicit characterization of the extreme points of the weak* closed convex set of permutation-invariant Radon probabilities on K^J as being precisely the powers or Radon probabilities on K to obtain the result. (This paper also contains the Hewitt-Savage zero-one law, probably quoted more often than de Finetti's theorem itself, stating that any power of a Radon probability on K takes only the values 0 and 1 on the σ-subalgebra of $\mathcal{B}(K^J)$ containing those sets which are invariant under permutation of coordinates.) To the best of my knowledge Theorem 16 is a new application of these various ideas.

Acknowledgments

I am grateful to D. Aldous for pointing out to me the link between the main results of this work and standard versions of de Finetti's theorem (and referring me to [4]), and particularly to D.H. Fremlin for then explaining this link to me and for many other helpful discussions.

References

[1] Federer H., *Geometric Measure Theory*, Springer, Berlin, 1969;

[2] Fremlin D.H., *Measure Theory*, vol. 2, Torres Fremlin, Colchester, 2001;

[3] Fremlin D.H., *Measure Theory*, vol. 4, Torres Fremlin, Colchester, to appear 2003;

[4] Hewitt E. & Savage L.J.,"Symmetric measures on Cartesian products", *Trans. Amer. Math. Soc.*, 80, 470 – 501, 1955;

[5] Loomis L.H., *Abstract Harmonic Analysis*, Van Nostrand, Princeton, 1953;

[6] Stroock D.W., *Probability Theory, an Analytic View*, Cambridge University Press, Cambridge, 1993.

Tim D. Austin
Trinity College
Cambridge
09 . 2003