Proc. Indian Acad. Sci. (Math. Sci.) Vol. 117, No. 4, November 2007, pp. 505–515.

Extreme points of the convex set of joint probability distributions with fixed marginals

K R PARTHASARATHY

Indian Statistical Institute, Delhi Centre, 7, S.J.S. Sansanwal Marg,
New Delhi 110 016, India
E-mail: krp@isid.ac.in

MS received 19 February 2007

Abstract. By using a quantum probabilistic approach we obtain a description of the extreme points of the convex set of all joint probability distributions on the product of two standard Borel spaces with fixed marginal distributions.

Keywords. C^* algebra; covariant bistochastic maps; completely positive map; Stinespring's theorem; extreme points of a convex set.

1. Introduction

It is a well-known theorem of Birkhoff [3] and von Neumann [6], that the extreme points in the convex set of all $n \times n$ bistochastic (or doubly stochastic) matrices are precisely the n-th order permutation matrices [1,2]. Here we address the following problem: If G is a standard Borel group acting measurably on two standard probability spaces $(X_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2$ where μ_i is invariant under the G-action for each i then what are the extreme points of the convex set of all joint probability distributions on the product Borel space $(X_1 \times X_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ which are invariant under the diagonal action $(x_1, x_2) \mapsto (gx_1, gx_2)$ where $x_i \in X_i$, $i = 1, 2$ and $g \in G$?

Our approach to the problem mentioned above is based on a quantum probabilistic method arising from Stinespring's [5] description of completely positive maps on C^* algebras. We obtain a necessary and sufficient condition for the extremality of a joint distribution in the form of a regression condition. This leads to examples of extremal nongraphic joint distributions in the unit square with uniform marginal distributions on the unit interval. The Birkhoff-von Neumann theorem is deduced as a corollary of the main theorem.

2. The convex set of covariant bistochastic maps on C^* algebras

For any complex separable Hilbert space $\mathcal{H}$, express its scalar product in the Dirac notation $\langle \cdot | \cdot \rangle$ and denote by $\mathcal{B}(\mathcal{H})$ the C^* algebra of all bounded operators on $\mathcal{H}$. Let G be a group with fixed unitary representations $g \mapsto U_g$, $g \mapsto V_g$, $g \in G$ in Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ respectively and let $\mathcal{A}_i \subset \mathcal{B}(\mathcal{H}_i)$, $i = 1, 2$ be unital C^* algebras invariant under the respective conjugations by U_g, V_g for every g in G. Let ω_i be a fixed state in $\mathcal{A}_i$ for each i, satisfying the invariance conditions:

$$\omega_1(U_g X U_g^{-1}) = \omega_1(X), \omega_2(V_g Y V_g^{-1})$$
$$= \omega_2(Y) \quad \forall X \in \mathcal{A}_1,\ Y \in \mathcal{A}_2, g \in G. \tag{2.1}$$

Consider a linear, unital and completely positive map $T: \mathcal{A}_1 \to \mathcal{A}_2$ satisfying the following:

$$\omega_2(T(X)) = \omega_1(X) \quad \forall X \in \mathcal{A}_1, \tag{2.2}$$

$$T(U_g X U_g^{-1}) = V_g T(X) V_g^{-1} \quad \forall X \in \mathcal{A}_1, g \in G. \tag{2.3}$$

Then we say that T is a *G-covariant bistochastic map* with respect to the pair of states ω_1, ω_2 and representations $U_., V_.$. Denote by $\mathbb{K}$ the convex set of all such covariant bistochastic maps from $\mathcal{A}_1$ into $\mathcal{A}_2$. We shall now present a necessary and sufficient condition for an element T in $\mathbb{K}$ to be an extreme point of $\mathbb{K}$.

To any $T \in \mathbb{K}$ we can associate a Stinespring triple $(\mathcal{K}, j, \Gamma)$ where $\mathcal{K}$ is a Hilbert space, j is a C^* homomorphism from $\mathcal{A}_1$ into $\mathcal{B}(\mathcal{K})$ and Γ is an isometry from $\mathcal{H}_2$ into $\mathcal{K}$ satisfying the following properties:

(i) $\Gamma^\dagger j(X)\Gamma = T(X) \quad \forall X \in \mathcal{A}_1$;
(ii) The linear manifold generated by $\{j(X)\Gamma u | u \in \mathcal{H}_2, X \in \mathcal{A}_1\}$ is dense in $\mathcal{K}$.

Such a Stinespring triple is unique up to a unitary isomorphism, i.e., if $(\mathcal{K}', j', \Gamma')$ is another triple satisfying the properties (i) and (ii) above then there exists a unitary isomorphism $\theta: \mathcal{K} \to \mathcal{K}'$ such that $\theta j(X) = j'(X)\theta \ \forall X \in \mathcal{A}_1$ and $\theta\Gamma v = \Gamma' v \ \forall v \in \mathcal{H}_2$ (see [5]).

We now claim that the covariance property of T ensures the existence of a unitary representation $g \mapsto W_g$ of G in $\mathcal{K}$ satisfying the relations:

$$W_g j(X)\Gamma u = j(U_g X U_g^{-1})\Gamma V_g u \quad \forall X \in \mathcal{A}_1, g \in G, u \in \mathcal{H}_2, \tag{2.4}$$

$$W_g j(X) W_g^{-1} = j(U_g X U_g^{-1}) \quad \forall X \in \mathcal{A}_1, g \in G. \tag{2.5}$$

Indeed, for any X, Y in $\mathcal{A}_1$ $u, v \in \mathcal{H}_2$ and $g \in G$ we have from the properties (i) and (ii) above and (2.3)

$$\begin{aligned}
&\langle j(U_g X U_g^{-1})\Gamma V_g u | j(U_g Y U_g^{-1})\Gamma V_g v\rangle \\
&\quad = \langle u | V_g^{-1}\Gamma^\dagger j(U_g X^\dagger Y U_g^{-1})\Gamma V_g v\rangle \\
&\quad = \langle u | V_g^{-1} T(U_g X^\dagger Y U_g^{-1}) V_g v\rangle \\
&\quad = \langle u | T(X^\dagger Y) | v\rangle \\
&\quad = \langle j(X)\Gamma u | j(Y)\Gamma v\rangle.
\end{aligned}$$

In other words, the correspondence $j(X)\Gamma u \mapsto j(U_g X U_g^{-1})\Gamma V_g u$ is a scalar product preserving map on a total subset of $\mathcal{K}$, proving the claim.

Theorem 2.1. *Let $T \in \mathbb{K}$ and let $(\mathcal{K}, j, \Gamma)$ be a Stinespring triple associated to T. Let $g \mapsto W_g$ be the unique unitary representation of G satisfying the relations* (2.4) *and* (2.5).

Then T is an extreme point of $\mathbb{K}$ if and only if there exists no nonzero hermitian operator Z in the commutant of the set $\{j(X), X \in \mathcal{A}_1\} \cup \{W_g, g \in G\}$ satisfying the following two conditions:

(i) $\Gamma^\dagger Z \Gamma = 0$;
(ii) $\Gamma^\dagger Z j(X)\Gamma \in \mathcal{A}_2$ *and* $\omega_2(\Gamma^\dagger Z j(X)\Gamma) = 0 \qquad \forall X \in \mathcal{A}_1$.

Proof. Suppose T is not an extreme point of $\mathbb{K}$. Then there exists $T_1, T_2 \in \mathbb{K}$, $T_1 \neq T_2$ such that $T = \frac{1}{2}(T_1 + T_2)$. Let $(\mathcal{K}_1, j_1, \Gamma_1)$ be a Stinespring triple associated to T_1. Then by the argument outlined in the proof of Proposition 2.1 in [4] there exists a bounded operator $J\colon \mathcal{K} \to \mathcal{K}_1$ satisfying the following properties:

(i) $Jj(X)\Gamma u = j_1(X)\Gamma_1 u, \qquad \forall X \in \mathcal{A}_1, u \in \mathcal{H}_2$;
(ii) The positive operator $\rho := J^\dagger J$ is in the commutant of $\{j(X), X \in \mathcal{A}\}$ in $\mathcal{B}(\mathcal{K})$;
(iii) $T_1(X) = \Gamma^\dagger \rho j(X)\Gamma$.

Since $T_1 \neq T_2$ it follows that $T_1 \neq T$ and hence ρ is different from the identity operator. We now claim that ρ commutes with W_g for every g in G. Indeed, for any X, Y in $\mathcal{A}_1$, u, v in $\mathcal{H}_2$ we have from the definition of ρ and J, equation (2.4) and the covariance of T_1,

$$\begin{aligned}
&\langle j(X)\Gamma u|\rho W_g|j(Y)\Gamma v\rangle \\
&\quad= \langle j(X)\Gamma u|J^\dagger J|j(U_g Y U_g^{-1})\Gamma V_g v\rangle \\
&\quad= \langle j_1(X)\Gamma_1 u|j_1(U_g Y U_g^{-1})\Gamma_1 V_g v\rangle \\
&\quad= \langle u|\Gamma_1^\dagger j_1(X^\dagger U_g Y U_g^{-1})\Gamma_1|V_g v\rangle \\
&\quad= \langle u|T_1(X^\dagger U_g Y U_g^{-1})|V_g v\rangle \\
&\quad= \langle u|V_g T_1(U_g^{-1} X^\dagger U_g Y)|v\rangle.
\end{aligned}$$

On the other hand, by the same arguments, we have

$$\begin{aligned}
&\langle j(X)\Gamma u|W_g \rho|j(Y)\Gamma v\rangle \\
&\quad= \langle j(U_g^{-1} X U_g)\Gamma V_g^{-1} u|J^\dagger J|j(Y)\Gamma v\rangle \\
&\quad= \langle j_1(U_g^{-1} X U)\Gamma_1 V_g^{-1} u|j_1(Y)\Gamma_1 v\rangle \\
&\quad= \langle u|V_g T_1(U_g^{-1} X^\dagger U_g Y)|v\rangle.
\end{aligned}$$

Comparing the last two identities and using property (ii) of the Stinespring triple we conclude that ρ commutes with W_g. Putting $Z = \rho - I$, we have

$$\Gamma^\dagger Z j(X)\Gamma = T_1(X) - T(X), \quad \forall X \in \mathcal{A}_1. \tag{2.6}$$

Clearly, the right-hand side of this equation is an element of $\mathcal{A}_2$ and

$$\omega_2(\Gamma^\dagger Z j(X)\Gamma) = \omega_1(X) - \omega_1(X) = 0, \quad \forall X \in \mathcal{A}_1.$$

Putting $X = I$ in (2.6) we have $\Gamma^\dagger Z \Gamma = 0$. Then Z satisfies properties (i) and (ii) in the statement of the theorem, proving the sufficiency part.

Conversely, suppose there exists a nonzero hermitian operator Z in the commutant of $\{j(X), X \in \mathcal{A}_1\} \cup \{W_g, g \in G\}$ satisfying properties (i) and (ii) in the theorem. Choose and fix a positive constant ε such that the operators $I \pm \varepsilon Z$ are positive. Define the maps $T_\pm: \mathcal{A}_1 \to \mathcal{A}_2$ by

$$T_\pm(X) = \Gamma^\dagger(I \pm \varepsilon Z) j(X)\Gamma, \qquad X \in \mathcal{A}_1. \tag{2.7}$$

Since

$$(I \pm \varepsilon Z) j(X) = \sqrt{I \pm \varepsilon Z}\, j(X) \sqrt{I \pm \varepsilon Z}$$

it follows that $T_\pm$ are completely positive. By putting $X = I$ in (2.7) and using property (i) of Z in the theorem we see that $T_\pm$ are unital. Furthermore, we have from equations (2.4) and (2.5), for any $g \in G$, $X \in \mathcal{A}_1$,

$$\begin{aligned} T_\pm(U_g X U_g^{-1}) &= \Gamma^\dagger(I \pm \varepsilon Z) W_g j(X) W_g^{-1}\Gamma \\ &= V_g \Gamma^\dagger(I \pm \varepsilon Z) j(X)\Gamma V_g^{-1} \\ &= V_g T_\pm(X) V_g^{-1}. \end{aligned}$$

Also, by property (ii) in the theorem we have

$$\omega_2(T_\pm(X)) = \omega_2(T(X)) = \omega_1(X), \quad \forall X \in \mathcal{A}_1.$$

Thus $T_\pm \in \mathbb{K}$. Note that

$$\langle u|\Gamma^\dagger Z j(X^\dagger Y)\Gamma|v\rangle = \langle j(X)\Gamma u|Z|j(Y)\Gamma v\rangle$$

cannot be identically zero when X and Y vary in $\mathcal{A}_1$ and u and v vary in $\mathcal{H}_2$. Thus $\Gamma^\dagger Z j(X)\Gamma \not\equiv 0$ and hence $T_+ \neq T_-$. But $T = \frac{1}{2}(T_+ + T_-)$. In other words, T is not an extreme point of $\mathbb{K}$. This proves the necessity. □

3. The convex set of invariant joint distributions with fixed marginal distributions

Let $(X_i, \mathcal{F}_i, \mu_i), i = 1, 2$ be standard probability spaces and let G be a standard Borel group acting measurably on both X_1 and X_2 preserving μ_1 and μ_2. Denote by $\mathbb{K}(\mu_1, \mu_2)$ the convex set of all joint probability distributions on the product Borel space $(X_1 \times X_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ invariant under the diagonal G action $(g, (x_1, x_2)) \mapsto (gx_1, gx_2)$, $x_i \in X_i$, $g \in G$ and having the marginal distribution μ_i in X_i for each i. Choose and fix $\omega \in \mathbb{K}(\mu_1, \mu_2)$. Our present aim is to derive from the quantum probabilistic result in Theorem 2.1, a necessary and sufficient condition for ω to be an extreme point of $\mathbb{K}(\mu_1, \mu_2)$. To this end we introduce the Hilbert spaces $\mathcal{H}_i = L^2(\mu_i)$, $\mathcal{K} = L^2(\omega)$ and the abelian von Neumann algebras $\mathcal{A}_i \subset \mathcal{B}(\mathcal{H}_i)$ where $\mathcal{A}_i = L^\infty(\mu_i)$ is also viewed as the algebra of operators of multiplication by functions from $L^\infty(\mu_i)$. For any $\varphi \in L^\infty(\mu_i)$ we shall denote by the same symbol φ the multiplication operator $f \mapsto \varphi f$, $f \in L^2(\mu_i)$. For any $\varphi \in \mathcal{A}_1$, define the operator $j(\varphi)$ in $\mathcal{K}$ by

$$(j(\varphi)f)(x_1, x_2) = \varphi(x_1) f(x_1, x_2), \quad f \in \mathcal{K}, x_i \in X_i. \tag{3.1}$$

Then the correspondence $\varphi \mapsto j(\varphi)$ is a von Neumann algebra homomorphism from $\mathcal{A}_1$ into $\mathcal{B}(\mathcal{K})$. Define the isometry $\Gamma\colon \mathcal{H}_2 \to \mathcal{K}$ by

$$(\Gamma v)(x_1, x_2) = v(x_2), \quad v \in \mathcal{H}_2. \tag{3.2}$$

Then, for $f \in \mathcal{K}$, $v \in \mathcal{H}_2$ we have

$$\begin{aligned}\langle f|\Gamma v\rangle &= \int_{X_1\times X_2} \bar{f}(x_1, x_2)v(x_2)\omega(\mathrm{d}x_1\mathrm{d}x_2)\\ &= \int_{X_2} \mu_2(\mathrm{d}x_2)\int_{X_1}[\bar{f}(x_1, x_2)\nu(\mathrm{d}x_1, x_2)]v(x_2),\end{aligned}$$

where $\nu(E, x_2)$, $E \in \mathcal{F}_1$, $x_2 \in X_2$ is a measurable version of the conditional probability distribution on $\mathcal{F}_1$ given the sub σ-algebra $\{X_1 \times F,\ F \in \mathcal{F}_2\} \subset \mathcal{F}_1 \otimes \mathcal{F}_2$. Thus the adjoint $\Gamma^\dagger\colon \mathcal{K} \to \mathcal{H}_2$ of Γ is given by

$$(\Gamma^\dagger f)(x_2) = \int_{X_1} f(x_1, x_2)\nu(\mathrm{d}x_1, x_2). \tag{3.3}$$

Hence

$$(j(\varphi)\Gamma v)(x_1, x_2) = \varphi(x_1)v(x_2), \qquad \varphi \in \mathcal{A}_1,\ \ v \in \mathcal{H}_2, \tag{3.4}$$

$$(\Gamma^\dagger j(\varphi)\Gamma v)(x_2) = \left[\int \varphi(x_1)\nu(\mathrm{d}x_1, x_2)\right]v(x_2). \tag{3.5}$$

In other words,

$$\Gamma^\dagger j(\varphi)\Gamma = T(\varphi), \tag{3.6}$$

where $T(\varphi) \in \mathcal{A}_2$ is given by

$$T(\varphi)(x_2) = \int_{X_1} \varphi(x_1)\nu(\mathrm{d}x_1, x_2). \tag{3.7}$$

Equations (3.1)–(3.7) imply that T is a linear, unital and positive (and hence completely positive) map from the abelian von Neumann algebra $\mathcal{A}_1$ into $\mathcal{A}_2$ and $(\mathcal{K}, j, \Gamma)$ is, indeed, a Stinespring triple for T. Furthermore, the unitary operators U_g, V_g and W_g in $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{K}$ respectively defined by

$$\begin{aligned}(U_g u)(x_1) &= u(g^{-1}x_1), \quad u \in \mathcal{H}_1,\\ (V_g v)(x_2) &= v(g^{-1}x_2), \quad v \in \mathcal{H}_2,\\ (W_g f)(x_1, x_2) &= f(g^{-1}x_1,\ g^{-1}x_2), \quad f \in \mathcal{K}\end{aligned}$$

satisfy the relations (2.4) and (2.5).

Our next lemma describes operators of the form Z occurring in Theorem 2.1.

Lemma 3.1. *Let Z be a bounded hermitian operator in $\mathcal{K}$ satisfying the following conditions*:

(i) $Zj(\varphi) = j(\varphi)Z, \quad \forall\varphi \in \mathcal{A}_1,$
(ii) $ZW_g = W_g Z, \quad \forall g \in G,$
(iii) $\Gamma^\dagger Zj(\varphi)\Gamma \in \mathcal{A}_2, \quad \forall\varphi \in \mathcal{A}_1.$

Then there exists a function $\zeta \in L^\infty(\omega)$ satisfying the following properties:

(a) $\zeta(gx_1, gx_2) = \zeta(x_1, x_2)$ *a.e.* $(\omega), \quad \forall g \in G,$
(b) $(Zf)(x_1, x_2) = \zeta(x_1, x_2) f(x_1, x_2), \qquad \forall f \in \mathcal{K}.$

Proof. Let

$$\zeta(x_1, x_2) = (Z1)(x_1, x_2),$$

where the symbol 1 also denotes the function identically equal to unity. For functions u, v on X_1, X_2 respectively denote by $u \otimes v$ the function on $X_1 \times X_2$ defined by $u \otimes v(x_1, x_2) = u(x_1)v(x_2)$. By property (i) of Z in the lemma we have

$$\begin{aligned}(Z\varphi \otimes 1)(x_1, x_2) &= (Zj(\phi)1)(x_1, x_2)\\ &= (j(\phi)Z1)(x_1, x_2)\\ &= \varphi(x_1)\zeta(x_1, x_2), \qquad \forall\varphi \in \mathcal{A}_1.\end{aligned} \tag{3.8}$$

If $\varphi \in \mathcal{A}_1$, $v \in \mathcal{H}_2$, we have

$$\begin{aligned}(Z\varphi \otimes v)(x_1, x_2) &= (Zj(\varphi)\Gamma v)(x_1, x_2)\\ &= (j(\varphi)Z\Gamma v)(x_1, x_2)\\ &= \varphi(x_1)(Z1 \otimes v)(x_1, x_2).\end{aligned} \tag{3.9}$$

From properties (i) and (iii) of Z in the lemma and equations (3.3), (3.8) and (3.9) we have

$$\begin{aligned}(\Gamma^\dagger Zj(\varphi)\Gamma v)(x_2) &= \int (Z\varphi \otimes v)\nu(\mathrm{d}x_1, x_2)\\ &= \int \varphi(x_1)(Z1 \otimes v)(x_1, x_2)\nu(\mathrm{d}x_1, x_2)\end{aligned}$$

whereas the left-hand side is of the form $R(\varphi)(x_1)v(x_2)$ for some $R(\varphi) \in L^\infty(\mu_2)$. Thus

$$R(\varphi)(x_2)v(x_2) = \int \varphi(x_1)(Z1 \otimes v)(x_1, x_2)\nu(\mathrm{d}x_1, x_2).$$

Choosing $v = 1$, we have from the definition of ζ

$$R(\varphi)(x_2) = \int \varphi(x_1)\zeta(x_1, x_2)\nu(\mathrm{d}x_1, x_2).$$

Thus, for every $\varphi \in \mathcal{A}_1$,

$$\int \varphi(x_1)\zeta(x_1, x_2)v(x_2)\nu(\mathrm{d}x_1, x_2) = \int \varphi(x_1)(Z1 \otimes v)(x_1, x_2)\nu(\mathrm{d}x_1, x_2)$$

and hence

$$(Z1 \otimes v)(x_1, x_2) = \zeta(x_1, x_2)v(x_2) \text{ a.e. } x_1(\nu(., x_2)) \text{ a.e. } x_2(\mu_2).$$

Applying $j(\varphi)$ on both sides, we get

$$(Z\varphi \otimes v)(x_1, x_2) = \zeta(x_1, x_2)\varphi(x_1)v(x_2) \text{ a.e. } (\omega).$$

In other words, Z is the operator of multiplication by ζ and it follows that $\zeta \in L^\infty(\omega)$. Now property (ii) of Z implies property (a) in the lemma. □

Theorem 3.2. *Let $\omega \in \mathbb{K}(\mu_1, \mu_2)$. Then ω is an extreme point of $\mathbb{K}(\mu_1, \mu_2)$ if and only if there exists no nonzero real-valued function $\zeta \in L^\infty(\omega)$ satisfying the following conditions*:

(i) $\zeta(gx_1, gx_2) = \zeta(x_1, x_2)$ *a.e.* $\omega, \quad \forall g \in G$;
(ii) $\mathbb{E}(\zeta(\xi_1, \xi_2)|\xi_1) = 0$, $\mathbb{E}(\zeta(\xi_1, \xi_2)|\xi_2) = 0$ *where* (ξ_1, ξ_2) *is an* $X_1 \times X_2$*-valued random variable with distribution* ω.

Proof. Let Z be a bounded self-adjoint operator in the commutant of $\{j(\varphi), \varphi \in \mathcal{A}_1\} \cup \{W_g, g \in G\}$ such that $\Gamma^\dagger Z j(\varphi)\Gamma \in \mathcal{A}_2, \forall \varphi \in \mathcal{A}_1$. Then by Lemma 3.1 it follows that Z is of the form

$$(Zf)(x_1, x_2) = \zeta(x_1, x_2)f(x_1, x_2),$$

where $\zeta \in L^\infty(\omega)$ and $\zeta(gx_1, gx_2) = \zeta(x_1, x_2)$ a.e. (ω). Note that

$$(\Gamma^\dagger Z\Gamma v)(x_2) = \left[\int_{X_1} \zeta(x_1, x_2)\nu(\mathrm{d}x_1, x_2)\right] v(x_2) \text{ a.e. } (\mu_2), v \in \mathcal{H}_2.$$

Thus $\Gamma^\dagger Z\Gamma = 0$ if and only if $\mathbb{E}(\zeta(\xi_1, \xi_2)|\xi_2) = 0$. Now we evaluate

$$(\Gamma^\dagger Z j(\varphi)\Gamma v)(x_2) = \int \varphi(x_1)v(x_2)\zeta(x_1, x_2)\nu(\mathrm{d}x_1, x_2) \quad \text{a.e. } (\mu_2).$$

Looking upon $\Gamma^\dagger Z j(\varphi)\Gamma$ as an element of $\mathcal{A}_2$ and evaluating the state μ_2 on this element we get

$$\begin{aligned}\mu_2(\Gamma^\dagger Z j(\varphi)\Gamma) &= \int \varphi(x_1)\zeta(x_1, x_2)\nu(\mathrm{d}x_1, x_2)\mu(\mathrm{d}x_2)\\ &= \int \varphi(x_1)\zeta(x_1, x_2)\omega(\mathrm{d}x_1\, \mathrm{d}x_2)\\ &= \mathbb{E}_\omega\varphi(\xi_1)\zeta(\xi_1, \xi_2)\\ &= \mathbb{E}_{\mu_1}\varphi(\xi_1)\mathbb{E}(\zeta(\xi_1, \xi_2)|\xi_1).\end{aligned}$$

Thus $\mu_2(\Gamma^\dagger Z j(\varphi)\Gamma) = 0, \forall \varphi \in \mathcal{A}_1$ if and only if $\mathbb{E}(\zeta(\xi_1, \xi_2)|\xi_1) = 0$. Now an application of Theorem 2.1 completes the proof of the theorem. □

We shall now look at the special case when G is the trivial group consisting of only the identity element. Let $(X_i, \mathcal{F}_i, \mu_i), i = 1, 2$ be standard probability spaces and let $T: X_1 \to X_2$ be a Borel map such that $\mu_2 = \mu_1 T^{-1}$. Consider an X_1-valued random variable ξ with distribution μ_1. Then the joint distribution ω of the pair $(\xi, T \circ \xi)$ is an element of $\mathbb{K}(\mu_1, \mu_2)$ and by Theorem 2.1 is an extreme point. Similarly, if $T: X_2 \to X_1$ is a Borel map such that $\mu_2 T^{-1} = \mu_1$ and η is an X_2-valued random variable with distribution μ_2 then $(T \circ \eta, \eta)$ has a joint distribution which is an extreme point of $\mathbb{K}(\mu_1, \mu_2)$. Such extreme points are called *graphic* extreme points. Thus there arises the natural question whether there exist nongraphic extreme points. Our next lemma facilitates the construction of nongraphic extreme points.

Lemma 3.3. *Let $(X, \mathcal{F}, \lambda)$, $(Y, \mathcal{G}, \mu)$, $(Z, \mathcal{K}, \nu)$ be standard probability spaces and let ξ, η, ζ be random variables on a probability space with values in X, Y, Z and distribution λ, μ, ν respectively. Suppose ζ is independent of (ξ, η) and the joint distribution ω of (ξ, η) is an extreme point of $\mathbb{K}(\lambda, \mu)$. Let $\tilde{\lambda}, \tilde{\mu}, \tilde{\omega}$ be the distributions of (ξ, ζ), (η, ζ) and $((\xi, \zeta), (\eta, \zeta))$ respectively in the spaces $X \times Z$, $Y \times Z$ and $(X \times Z) \times (Y \times Z)$. Then $\tilde{\omega}$ is an extreme point of $\mathbb{K}(\tilde{\lambda}, \tilde{\mu})$.*

Proof. Let f be a bounded real-valued measurable function on $(X \times Z) \times (Y \times Z)$ satisfying the relations

$$\mathbb{E}\{f((\xi, \zeta), (\eta, \zeta)) | (\eta, \zeta)\} = 0,$$

$$\mathbb{E}\{f((\xi, \zeta), (\eta, \zeta)) | (\xi, \zeta)\} = 0.$$

If we write

$$F_z(x, y) = f((x, z), (y, z)) \quad \text{where} \quad (x, y, z) \in X \times Y \times Z$$

then we have

$$\mathbb{E}(F_z(\xi, \eta) | \eta) = 0, \quad \mathbb{E}(F_z(\xi, \eta) | \xi) = 0 \;\text{ a.e. } z(\nu).$$

Since ω is extremal it follows that $F_z(\xi, \eta) = 0$ a.e. $z(\nu)$ and therefore $f((\xi, \zeta), (\eta, \zeta)) = 0$. By Theorem 3.1 it follows that $\tilde{\omega}$ is, indeed, an extreme point of $\mathbb{K}(\tilde{\lambda}, \tilde{\mu})$. □

Example 3.4. Let λ be the uniform distribution in the unit interval [0, 1]. We shall use Lemma 3.3 and construct nongraphic extreme points of $\mathbb{K}(\lambda, \lambda)$ which are distributions in the unit square. To this end we start with the two-point space $\mathbb{Z}_2 = \{0, 1\}$ with the probability distribution P where

$$P(\{0\}) = p, \;\; P(\{1\}) = q, \;\; 0 < p < q < 1, \;\; p + q = 1.$$

Now consider $\mathbb{Z}_2$-valued random variables ξ, η with the joint distribution given by

$$P(\xi = 0, \eta = 0) = 0, \;\; P(\xi = 0, \eta = 1) = P(\xi = 1, \eta = 0) = p,$$

$$P(\xi = 1, \eta = 1) = q - p.$$

Note that the joint distribution of (ξ, η) is a nongraphic extreme point of $\mathbb{K}(P, P)$. Now consider an i.i.d sequence $\zeta_1, \zeta_2, \ldots$ of $\mathbb{Z}_2$-valued random variables independent of (ξ, η) and having the same distribution P. Put

$$\boldsymbol{\varsigma} = (\zeta_1, \zeta_2, \ldots).$$

Then by Lemma 3.3 the joint distribution ω of $((\xi, \varsigma), (\eta, \varsigma))$ is an extreme point of $\mathbb{K}(\nu, \nu)$ where $\nu = P \otimes P \otimes \cdots$ in $\mathbb{Z}_2^{\{0,1,2,\dots\}}$. Furthermore, since (ξ, η) is nongraphic so is $((\xi, \varsigma), (\eta, \varsigma))$. Denote by F_p the common probability distribution function of the random variables

$$\tilde{\xi} = \frac{\xi}{2} + \sum_{j=1}^{\infty} \frac{\zeta_j}{2^{j+1}}, \quad \tilde{\eta} = \frac{\eta}{2} + \sum_{j=1}^{\infty} \frac{\zeta_j}{2^{j+1}}.$$

Then F_p is a strictly increasing and continuous function on the unit interval and therefore the correspondence $t \to F_p(t)$ is a homeomorphism of $[0, 1]$. Put $\xi' = F_p(\tilde{\xi})$, $\eta' = F_p(\tilde{\eta})$. Then the joint distribution ω of (ξ', η') is a nongraphic extreme point of $\mathbb{K}(\lambda, \lambda)$.

Now we consider the case when X_1 and X_2 are finite sets, G is a finite group acting on each X_i, the number of G-orbits in X_1, X_2 and $X_1 \times X_2$ are respectively m_1, m_2 and m_{12} and μ_i is a G-invariant probability distribution in X_i with support X_i for each $i = 1, 2$. For any probability distribution λ in any finite set, denote by $S(\lambda)$ its support set. We first note that Theorem 3.2 assumes the following form.

Theorem 3.5. *A probability distribution $\omega \in \mathbb{K}(\mu_1, \mu_2)$ is an extreme point if and only if there is no nonzero real-valued function ζ on $S(\omega)$ satisfying the following conditions*:

(i) $\zeta(gx_1, gx_2) = \zeta(x_1, x_2) \quad \forall (x_1, x_2) \in S(\omega), g \in G$;
(ii) $\sum_{x_2 \in X_2} \zeta(x_1, x_2)\omega(x_1, x_2) = 0 \quad \forall x_1 \in X_1$;
(iii) $\sum_{x_1 \in X_1} \zeta(x_1, x_2)\omega(x_1, x_2) = 0 \quad \forall x_2 \in X_2$.

Proof. Immediate. □

COROLLARY 3.6

Let ω_1, ω_2 be extreme points of $\mathbb{K}(\mu_1, \mu_2)$ and $S(\omega_1) \subseteq S(\omega_2)$. Then $\omega_1 = \omega_2$. In particular, any extreme point ω of $\mathbb{K}(\mu_1, \mu_2)$ is uniquely determined by its support set $S(\omega)$.

Proof. Suppose $\omega_1 \neq \omega_2$. Put $\omega = \frac{1}{2}(\omega_1 + \omega_2)$. Then $\omega \in \mathbb{K}(\mu_1, \mu_2)$ and ω is not an extreme point. By Theroem 3.5 there exists a nonzero real-valued function ζ satisfying conditions (i)–(iii) of the theorem. By hypothesis $S(\omega) = S(\omega_2)$. Define

$$\zeta'(x_1, x_2) = \frac{\zeta(x_1, x_2)\omega(x_1, x_2)}{\omega_2(x_1, x_2)}, \quad \text{where } (x_1, x_2) \in S(\omega_2).$$

Then conditions (i)–(iii) of Theorem 3.5 are fulfilled when the pair ζ, ω is replaced by ζ', ω_2 contradicting the extremality of ω_2. □

COROLLARY 3.7

For any $\omega \in \mathbb{K}(\mu_1, \mu_2)$ let $N(\omega)$ denote the number of G-orbits in its support set $S(\omega)$. If ω is an extreme point of $\mathbb{K}(\mu_1, \mu_2)$ then

$$\max(m_1, m_2) \leq N(\omega) \leq m_1 + m_2.$$

In particular, the number of extreme points in $\mathbb{K}(\mu_1, \mu_2)$ does not exceed

$$\sum_{\max(m_1,m_2) \leq r \leq m_1+m_2} \binom{m_{12}}{r}.$$

Proof. Let ω be an extreme point of $\mathbb{K}(\mu_1, \mu_2)$. Suppose $N(\omega) > m_1 + m_2$. Observe that all G-invariant real-valued functions on $S(\omega)$ constitute a linear space of cardinality $N(\omega)$. Functions ζ satisfying conditions (i)–(iii) of the theorem constitute a subspace of dimension $\geq N(\omega) - (m_1 + m_2)$, contradicting the extremality of ω. For any distribution ω in $\mathbb{K}(\mu_1, \mu_2)$ we have $N(\omega) \geq m_i$, $i = 1, 2$. This proves the first part. The second part is now immediate from Corollary 3.6. □

COROLLARY 3.8 (Birkhoff–von Neumann theorem)

Let $X_1 = X_2 = X$, $\#X = m$, $\mu_1 = \mu_2 = \mu$ where $\mu(x) = \frac{1}{m}$ $\forall x \in X$. Then any extreme point ω in $\mathbb{K}(\mu, \mu)$ is of the form

$$\omega(x, y) = \frac{1}{m}\delta_{\sigma(x)y}, \quad \forall x, y \in X$$

where σ is a permutation of the elements of X.

Proof. Without loss of generality we assume that $X = \{1, 2, \ldots, m\}$ and view ω as a matrix of order m with nonnegative entries with each row or column total being $1/m$. First assume that in each row or column there are at least two nonzero entries. Then ω has at least $2m$ nonzero entries and by Corollary 3.7 it follows that every row or column has exactly two nonzero entries. We claim that for any $i \neq i'$, $j \neq j'$ in the set $\{1, 2, \ldots, m\}$ at least one among $\omega_{ij}, \omega_{ij'}, \omega_{i'j}, \omega_{i'j'}$ vanishes. Suppose this is not true for some $i \neq i'$, $j \neq j'$. Put

$$p = \min\{\omega_{rs} | (r, s) \colon \omega_{rs} > 0\}.$$

Define

$$\omega_{rs}^{\pm} = \begin{cases} \omega_{rs} \pm p, & \text{if} \quad r = i, s = j \quad \text{or} \quad r = i', s = j', \\ \omega_{rs} \mp p, & \text{if} \quad r = i', s = j \quad \text{or} \quad r = i, s = j', \\ \omega_{rs}, & \text{otherwise.} \end{cases}$$

Then $\omega^{\pm} \in \mathbb{K}(\mu, \mu)$, $\omega^+ \neq \omega^-$ and $\omega = \frac{1}{2}(\omega^+ + \omega^-)$, a contradiction to the extremality of ω. Now observe that permutation of columns as well as rows of ω lead to extreme points of $\mathbb{K}(\mu, \mu)$. By appropriate permutations of columns and rows ω reduces to a tridiagonal matrix of the form

$$\widetilde{\omega} = \begin{bmatrix} p_{11} & p_{12} & 0 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ p_{21} & 0 & p_{23} & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & p_{32} & 0 & p_{34} & \ldots & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \ldots & 0 & p_{n-1\,n-2} & 0 & p_{n-1\,n} \\ 0 & 0 & \ldots & \ldots & \ldots & 0 & 0 & p_{n\,n-1} & p_{nn} \end{bmatrix},$$

where the p's with suffixes are all greater than or equal to p. Now consider the matrices

$$\lambda^{\pm} = \begin{bmatrix} p_{11} \pm p & p_{12} \mp p & 0 & 0 & 0 & \ldots \\ p_{21} \mp p & 0 & p_{23} \pm p & 0 & 0 & \ldots \\ 0 & p_{32} \pm p & 0 & p_{34} \mp p & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix}.$$

Then $\lambda^{\pm} \in \mathbb{K}(\mu, \mu)$ and $\tilde{\omega} = \frac{1}{2}(\lambda^{+} + \lambda^{-})$, contradicting the extremality of $\tilde{\omega}$ and therefore of ω. In other words, any extreme point ω of $\mathbb{K}(\mu, \mu)$ must have at least one row with exactly one nonzero entry. Then by permutations of rows and columns ω can be brought to the form

$$\omega_1 = \left[\begin{array}{c|c} 1/m & 0 \;\; 0 \ldots 0 \\ \hline 0 & \\ \vdots & \widehat{\omega} \\ 0 & \end{array}\right],$$

where $\frac{m}{m-1}\hat{\omega}$ is an extreme point of $\mathbb{K}(\hat{\mu}, \hat{\mu})$, $\hat{\mu}$ being the uniform distribution on a set of $m - 1$ points. Now an inductive argument completes the proof. □

We conclude with the remark that it is an interesting open problem to characterize the support sets of all extreme points of $\mathbb{K}(\mu_1, \mu_2)$ in terms of μ_1 and μ_2.

References

[1] Bapat R B and Raghavan T E S, Nonnegative Matrices and Applications (Cambridge: Cambridge University Press) (1997)

[2] Bhatia R, Matrix Analysis (New York: Springer Verlag) (1996)

[3] Birkhoff G, Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucuman Rev. Ser.* **A5** (1946) 147–150

[4] Parthasarathy K R, Extreme points of the convex set of stochastic maps on a C^* algebra, *Inf. Dim. Analysis, Quantum Probab. and Rel. Topics* **1** (1998) 599–609

[5] Stinespring W F, Positive functions on C^* algebras, *Proc. Am. Math. Soc.* **6** (1955) 211–216

[6] von Neumann J, A certain zero-sum two-person game equivalent to an optimal assignment problem, *Ann. Math. Studies* **28** (1953) 5–12