Supports of Extremal Doubly and Triply Stochastic Measures - Master's Project

Nathan Killoran

September 9, 2007

Abstract

Doubly stochastic measures are Borel probability measures on the unit square which push forward via the canonical projections to Lebesgue measure on each axis. The set of doubly stochastic measures is convex, so its extreme points are of particular interest. I review necessary and sufficient conditions for a set to support an extremal doubly stochastic measure, and include a proof that such a set can be decomposed into a countable collection of graphs and antigraphs of functions, called a 'limb-numbering system.' I also investigate how this structure partially generalizes to triply stochastic measures on the unit cube.

A doubly stochastic matrix is a real matrix whose entries are positive and whose rows and columns individually sum to one. A classical theorem first due to Birkhoff [1], but also attributed to von Neumann [2], states that the set of doubly stochastic matrices is the convex hull of the set of $n \times n$ permutation matrices. This, along with the Krein-Milman theorem [3], tells us that a matrix is doubly stochastic if and only if it is a convex combination of permutation matrices. In his 1949 book *Lattice Theory*, Birkhoff proposed the problem of extending this to an infinite dimensional analog [4]. Known as Birkhoff's Problem 111, this project has been taken up at various points since it's formulation; one approach is to consider doubly stochastic measures, probability measures on the unit square which project to the Lebesgue measure on each axis [5][6][7][8].

This paper is structured as follows. In the first section, we introduce basic definitions and survey relavent results from the literature relating to doubly stochastic measures. In the second section we ask and examine analagous questions for triply stochastic measures. Proofs of two of the major theorems from section 1 can found in an appendix.

1 Extremal doubly stochastic measures

Definition 1.1.

- *Let γ be a Borel measure on a product of topological spaces, $U \times V$, and let π^U and π^V be the projection maps onto each coordinate. Then γ pushes forward to a measure $\mu = \pi^U_\# \gamma$ on U where for each Borel set $B \in \mathcal{B}_U$,*

  $$\mu(B) = \pi^U_\# \gamma(B) = \gamma((\pi^U)^{-1}(B)) = \gamma(B \times V).$$

  We can define $\nu = \pi^V_\# \gamma$ similarly. We call μ and ν the projections *or* marginals *of γ. For a measure on $U \times V$ represented by γ, its marginals will always be denoted μ and ν respectively.*

- *Let γ be a non-negative probability measure on the unit square $U \times V = I \times I$. We say that γ is a* doubly stochastic measure *if $\mu = \nu = \lambda$, Lebesgue measure restricted to the unit interval I. The set of doubly stochastic measures will be called DSM.*

It is easily verified that the set DSM is convex, and as with the case of doubly stochastic matrices, we are led to consider extreme points.

Definition 1.2.

- *A measure γ in a convex set M is an* extreme point *of M if it cannot be written as a convex combination of measures in M. In other words, if $\gamma = t\alpha + (1-t)\beta$ for $t \in (0,1)$ and $\alpha, \beta \in M$, then γ extreme in M implies that $\alpha = \beta = \gamma$. We will write $EDSM$ for the extreme points of DSM.*

- *A measure γ on a product space $U \times V$ is called a* simplicial measure *if it is an extreme point of the set $M_\gamma(U,V) = \{\text{Measures } \omega \text{ on } U \times V \mid \pi^U_\# \omega = \pi^U_\# \gamma, \pi^V_\# \omega = \pi^V_\# \gamma\}$, consisting of all measures with the same projections as γ. Notice that if γ is doubly stochastic, then $M_\gamma(U,V) = DSM$.*

Doubly stochastic measures and extreme points of DSM are interesting objects to study for several reasons. For instance, all joint probability distributions can be represented using doubly stochastic measures. They are also important when determining which cost functions on the square solve the optimal transport problem uniquely (see e.g. [9],[10],[11]). It would be useful, then, to have simple conditions for determining whether or not a given doubly stochastic measure is extremal. Characterizations of extremal doubly stochastic measures were originally given by Douglas [5] and Lindenstrauss [6] in 1964-65. The statement of their result shown below is from [8].

Theorem 1.3. *(Lindenstrauss & Douglas 1964-65) Let $\gamma \in DSM$ and let $L^1(\lambda) \oplus L^1(\lambda)$ be the subspace of $L^1(\gamma)$ consisting of all functions of the form $f(x) + g(y)$ where $f, g \in L^1(\lambda)$. Then $\gamma \in EDSM$ if and only if the subspace $L^1(\lambda) \oplus L^1(\lambda)$ is dense in $L^1(\gamma)$.*

Unfortunately, this approach is framed in a functional analytic language which doesn't give a simple test for extremality; nor is it obvious how this criterion could be reduced to a condition on the support of μ in $U \times V$. An alternative starting point was eventually found by Beneš & Štěpàn [7] in 1987, which provides a relatively simple necessary condition for sets to support simplicial measures. We first need some geometric notions in order to describe these supporting sets.

Definition 1.4. *Let U, V be two spaces and let $f : U \to V$ and $g : V \to U$. Define*

$$T(x) = \begin{cases} (g \circ f)(x) & x \in Dom(f) \cap \{f^{-1}(Dom(g))\} = D(T) \\ x & x \notin D(T) \end{cases} .$$

The maps f, g are aperiodic *if $x \in D(T) \Rightarrow T^n(x) \neq x$ for any $n > 0$.*

Denote the graphs of f and g by:

$$Graph(f) = (x, f(x)) \in U \times V,$$
$$Antigraph(g) = (g(y), y) \in U \times V.$$

If $S = Graph(f) \cup Antigraph(g)$, $Graph(f) \cap Antigraph(g) = \emptyset$ and f, g are aperiodic, then this is called an aperiodic decomposition *of S.*

In what follows, we say that γ is *concentrated* on a set S if the outer measure of its complement is zero, i.e. $\gamma^*(S^c) = 0$. This condition allows that the set S may not completely be in the domain of γ. Note that if we have a Borel set S which has the same measure as the full product space, $\gamma(S) = \gamma(U \times V)$, then γ is concentrated on this Borel set.

Theorem 1.5. *(Beneš & Štěpàn, 1987) Let U and V be complete separable Borel metric spaces. Let γ be a simplicial measure on the product space $U \times V$. Then γ is concentrated on a set which admits an aperiodic decomposition.*

In particular, if $\gamma \in EDSM$, then γ is concentrated on a set $S \subset I^2$ which admits an aperiodic decomposition.

Beginning with the theorem of Lindenstrauss & Douglas, Hestir & Williams [8] provided an alternate proof of Beneš & Štěpàn's result, while further refining the structure these graphs should take. While Hestir & Williams were specifically interested in the case of doubly stochastic measures, their result holds in a more general setting. Again, we will first introduce a geometric concept before their theorem.

Definition 1.6. *A set $A \subset U \times V$ is called a (U,V)* limb-numbering system *if*

$$A = \cup_{k=1}^{\infty}[Graph(h_{2k-1}) \cup Antigraph(h_{2k})],$$

where $h_{2k-1} : Dom(h_{2k-1}) \subset U \to V$ and $h_{2k} : Dom(h_{2k}) \subset V \to U$ are maps subject to the following conditions:

$$\begin{cases} Ran(h_i) \subset Dom(h_{i-1}) \ \forall i > 1 \\ Dom(h_i) \cap Dom(h_j) = \emptyset \text{ for } i - j \text{ even} \\ Ran(h_1) \cap Dom(h_{2i}) = \emptyset \ \forall i > 0 \end{cases} .$$

Theorem 1.7. *(Hestir & Williams, 1995) Let* U*,* V *and* γ *satisfy the hypotheses of the previous theorem. Then* γ *is concentrated on a Borel set which is a limb-numbering system.*

Thus if $\gamma \in EDSM$*, then* γ *is concentrated on a Borel set* S *which is a* (I, I) *limb-numbering system.*

For a proof that Theorem 1.7 follows from Theorem 1.5, see the appendix. Having this necessary condition, one might ask whether it is also a sufficient condition. Both Beneš & Štěpàn and Hestir & Williams took some steps in this direction. Hestir & Williams managed to prove that, provided all disjoint graphs and antigraphs were Borel measurable, a doubly stochastic measure supported on a limb-numbering system is indeed extremal. This cumbersome measurability requirement turns out to be unnecessary, as shown with the following theorem [11].

Theorem 1.8. *(Chiappori, McCann & Nesheim, 2007) Let* U*,* V *and* γ *be as above, and let* γ *also be* σ*-finite. Let* γ *be concentrated on a limb-numbering system* S*, i.e.* $\gamma^*(S^c) = 0$ *and* $S = \cup_{k=1}^{\infty}[Graph(h_{2k-1}) \cup Antigraph(h_{2k})]$ *for functions* $h_{2k} : Dom(h_{2k}) \subset U \to V$*,* $h_{2k-1} : Dom(h_{2k-1}) \subset V \to U$*. If* S *has finitely many limbs or* $\gamma(U \times V) < \infty$*, then* γ *is the unique measure concentrated on* S *with marginals* μ *and* ν*.*

Again, for a proof of this theorem, see the appendix.

Corollary 1.9. *Let* γ *satisfy the hypothesis of the above theorem. Then* γ *is simplicial. In particular, if* $\gamma \in DSM$ *is concentrated on a limb-numbering system, then* $\gamma \in EDSM$*.*

Proof. Let γ_1, γ_2 be positive measures which have the same marginals as γ and let $\gamma = t\gamma_1 + (1-t)\gamma_2$ for $t \in (0, 1)$. Then $\gamma_1^*(S^c) \le \frac{1}{t}\gamma^*(S^c) = 0$, so γ_1 is also concentrated on S. By uniqueness, we conclude $\gamma_1 = \gamma = \gamma_2$. □

Thus, γ is an extremal doubly stochastic measure if and only if γ is concentrated on a limb-numbering system.

2 Extremal triply stochastic measures

We can extend these ideas to the unit cube with a few modifications.

Definition 2.1. *Let* γ *be a Borel measure on the unit cube* I^3*. We say that* γ *is* triply stochastic *if it pushes forward under the three coordinate projections to Lebesgue measure. In other words, if*

$$\pi^X(x, y, z) = x,\ \pi^Y(x, y, z) = y,\ \text{and}\ \pi^Z(x, y, z) = z$$

are the canonical projections, then $\pi^X_{\#}\gamma = \pi^Y_{\#}\gamma = \pi^Z_{\#}\gamma = \lambda$*. The set of triply stochastic measures is denoted by* TSM*.* TSM *is also convex, and we denote its extreme points by* $ETSM$*.*

We would like to make use of the same general theorems as the doubly stochastic case, so we must recast the problem as a 'two-dimensional' one. To do this, we denote the unit x-, y-, and $z-$intervals by I_x, I_y, and I_z and let $U = I_x$, $V = I_y \times I_z$. Then $\nu = \pi^{YZ}_{\#}\gamma$, where $\pi^{YZ} : I^3 \to I^2$ is the projection

$$\pi^{YZ}(x, y, z) = (y, z).$$

It is evident that γ is triply stochastic if and only if both $\mu = \lambda$ and ν is doubly stochastic. In two dimensions we had $M_\gamma(I, I) = DSM$, but in three dimensions there will exist triply stochastic measures whose $I_y \times I_z$ marginal differs from that of γ. Thus $M_\gamma(I_x, I_y \times I_z)$ is a strict subset of TSM. Fortunately, if γ is an extreme point in TSM, it is also an extreme point in the convex set $M_\gamma(I_x, I_y \times I_z)$, so extremal triply stochastic measures are still simplicial. The converse, however, may not be true in general.

Note that the way we decomposed the set I^3 into a product $U \times V$ was arbitrary. We can make use of the theorem of Hestir & Williams with each of these three possible decompositions to get stronger geometric conditions on the supporting set.

Theorem 2.2. *If γ is extremal in TSM, then γ is concentrated on a set $S \subset I^3$ which forms an $(I, I \times I)$ limb-numbering system in three different ways, i.e.*

$$\begin{aligned} S &= \cup_{k=1}^{\infty}[Graph(f_{2k-1}) \cup Antigraph(f_{2k})] \\ &= \cup_{k=1}^{\infty}[Graph(g_{2k-1}) \cup Antigraph(g_{2k})] \\ &= \cup_{k=1}^{\infty}[Graph(h_{2k-1}) \cup Antigraph(h_{2k})], \end{aligned}$$

where

$$\begin{aligned} f_{2k-1} &: Dom(f_{2k-1}) \subset I_x \to I_y \times I_z, \ f_{2k} : Dom(f_{2k}) \subset I_y \times I_z \to I_x \\ g_{2k-1} &: Dom(g_{2k-1}) \subset I_y \to I_z \times I_x, \ g_{2k} : Dom(g_{2k}) \subset I_z \times I_x \to I_y \\ h_{2k-1} &: Dom(h_{2k-1}) \subset I_z \to I_x \times I_y, \ h_{2k} : Dom(h_{2k}) \subset I_x \times I_y \to I_x \end{aligned}$$

and for $k = f, g,$ or h:

$$\begin{cases} Ran(k_i) \subset Dom(k_{i-1}) \ \forall i > 1 \\ Dom(k_i) \cap Dom(k_j) = \emptyset \ for \ i - j \ even \\ Ran(k_1) \cap Dom(k_{2i}) = \emptyset \ \forall i > 0 \end{cases}.$$

Unfortunately, this necessary condition may no longer be sufficient in three dimensions. We can use the theorem of Chiappori, McCann & Nesheim to show that, if γ is concentrated on an (I, I^2) limb-numbering system, then γ is extremal in $M_\gamma(I, I^2)$. But since $M_\gamma(I, I^2) \subset TSM$, we cannot conclude that γ is extreme in TSM. Fortunately, there is an additional condition which can guarantee this.

Proposition 2.3. *Let γ be a triply stochastic measure supported on an (I, I^2) limb-numbering system S. Assume the marginal ν is also concentrated on a limb-numbering system, in this case on (I, I). Then γ is extremal in TSM.*

Proof. To find a contradiction, suppose there exists a second measure $\omega \in TSM$ concentrated on the same limb-numbering system S as γ. Since γ is the unique measure in $M_\gamma(I, I^2)$ concentrated on S, then $\omega \notin M_\gamma(I, I^2)$, which tells us $\pi^V_\# \omega \neq \nu$.

Let $S_V = \pi^V(S) \subset V = I^2$ be the projection of S onto V. Since $I \times S_V^c \subset S^c$, we have $\mu^*(S_V^c) = \gamma^*(I \times S_V^c) \leq \gamma^*(S^c) = 0$. Thus, μ is concentrated on S_V. Similarly, $\pi^V_\# \omega$ must also be concentrated on S_V. Therefore both ν and $\pi^V_\# \omega$ are doubly stochastic measures concentrated on the same set S_V, which is a limb-numbering system by hypothesis. By Theorem 1.8, we must have $\nu = \pi^V_\# \omega$. This contradiction implies γ is the unique triply stochastic measure concentrated on S. By the same argument as in the earlier corollary, γ is therefore extremal in TSM. □

In order to make the necessary condition sufficient in three dimensions, one might conjecture that a subset of the unit cube which can be decomposed into three different (I, I^2) limb-numbering systems must project to one of the faces as an (I, I) limb-numbering system. This, however, remains to be investigated.

A Appendix - Proofs of some theorems

Proof of Theorem 1.7

Proof. There are two possible starting points for this proof. Hestir & Williams began with the characterization of extremal doubly stochastic measures due to Lindenstrauss & Douglas (see [8] for this). We will begin with the theorem of Beneš & Štěpàn and then show that a set with Beneš & Štěpàn's aperiodic decomposition is equivalent to a limb-numbering system.

First, it is easy to see that a limb-numbering system will have an aperiodic decomposition. Let $S = \cup_{k=1}^{\infty}[Graph(h_{2k-1}) \cup Antigraph(h_{2k})]$ be the limb-numbering system on the product $U \times V$. By disjointness of domains, we can define the single-valued functions $f(x) = h_{2k-1}(x)$ when $x \in Dom(h_{2k-1})$ and $g(y) = h_{2k}(y)$ when $y \in Dom(h_{2k})$ as the disjoint unions of functions from the limb-numbering system. We are immediately guaranteed that $Antigraph(h_1)$ is disjoint from $Graph(h_{2i})$ $\forall i > 0$. For $i, j > 0$, since $Ran(h_{2j+1}) \subset Dom(h_{2j})$, then $Ran(h_{2j+1}) \cap Dom(h_{2i}) = \emptyset$, provided $j \neq i$. For $i = j$, we use the fact that $Ran(h_{2i}) \subset Dom(h_{2i-1})$, which is disjoint from $Dom(h_{2i+1})$. In any case, $Antigraph(h_{2j+1}) \cap Graph(h_{2i}) = \emptyset$ $\forall i, j > 0$. Thus $S = Graph(f) \cup Antigraph(g)$ and $Graph(f) \cap Antigraph(g) = \emptyset$.

Assume some $x_0 \in U$ exists such that $(g \circ f)^n(x_0) = x_0$ for some $n \geq 0$. By disjointness, x_0 will reside in some particular domain, say $Dom(h_{2j-1})$. Then $h_{2j-1}(x_0) \in Ran(h_{2j-1}) \subset Dom(h_{2j-2})$ by one of our limb-numbering conditions. This gives that $(g \circ f)(x_0) = h_{2j-2} \circ h_{2j-1}(x_0) \in Ran(h_{2j-2}) \subset Dom(h_{2j-3})$. By the same reasoning, $(g \circ f)^2(x_0) = (h_{2j-4} \circ h_{2j-3}) \circ (h_{2j-2} \circ h_{2j-1})(x_0) \in Dom(h_{2j-5})$. Working recursively, we get the following: $(g \circ f)^n(x_0) = h_{2k-2n} \circ \cdots \circ h_{2k-1}(x_0) = x_0$. Thus $x_0 \in Dom(h_{2k-1})$ and $x_0 \in$

$Ran(h_{2k-2n}) \subset Dom(h_{2k-2n-1})$. By disjointness, we find that n can only be zero, and the limb-numbering system S has an aperiodic decomposition.

Conversely, for a set S with an aperiodic decomposition, define a map on the power set of S, $T_S : P(S) \to P(S)$ by $T_S(A) = \{(x, y) \in S \mid \exists v \in V$ such that $(x, v) \in A$ or $\exists u \in U$ such that $(u, y) \in A\}$. We can define the orbit of a set under T_S by $\cup_{k=1}^{\infty} T_S^k(A)$, or the orbit of a single point (x_0, y_0) by $\cup_{k=1}^{\infty} T_S^k((x_0, y_0))$. Let $\{(u_j, v_j)\}_{j=1}^m$ be a sequence of points in S. We will say that these points form a path if the points are disjoint and either of the following conditions holds:

1. $v_{2j} = v_{2j-1}$ and $u_{2j+1} = u_{2j}$ (starting 'horizontally')
2. $u_{2j} = u_{2j-1}$ and $v_{2j+1} = v_{2j}$ (starting 'vertically').

If, in addition, m is even, condition 1 holds and $u_m = u_1$, then we say the points form a cycle. Now, let $\{(u_j, v_j)\}_{j=1}^m$ be a cycle in S. Then either: a) $(u_1, v_1) \in Graph(f)$ or b) $(u_1, v_1) \in Antigraph(g)$. If a), then $v_1 = f(u_1)$ and we must have $(u_m, v_m) \in Antigraph(g)$ in order for f to be single valued at the point $u_1 = u_m$. Similarly, (u_{m-1}, v_{m-1}) must lie on $Graph(f)$ so that g is single-valued at $v_m = v_{m-1}$. Working recursively, we determine that $(u_{2j}, v_{2j}) \in Antigraph(g)$ and $(u_{2j-1}, v_{2j-1}) \in Graph(f)\ \forall j$. But, recalling Beneš & Štěpàn's map $T(u) = (g \circ f)(u)$, we have $T^{\frac{m}{2}}(u_1) = u_m = u_1$. If instead b) holds, we can show by a similar argument that again $T^{\frac{m}{2}}(u_1) = u_1$. In either case, if S has any cycles, then S is not aperiodic. Taking the contrapositive, we conclude that if S is aperiodic, it cannot contain any cycles. In fact, the converse is true, but won't be proved here.

Following Hestir & Williams, we use the axiom of choice to pick a maximal subset $A \subset S$ consisting of points with disjoint orbits. The orbit of A will be the set S, and each point $(x, y) \in S$ will be in the orbit of a unique point $(w, z) \in A$. We can create a path of length k joining (x, y) to (w, z) by using the steps in the orbit of (w, z). If there were two distinct paths, of lengths k_1 and k_2, then we would have a cycle within the orbit of (z, w) and hence in S. Since this isn't allowed, the path joining (x, y) to (w, z) is always unique. Define $H_k = \{(x, y) \in S \mid (x, y)$ can be joined to some $(z, w) \in A$ by a path of type 1 with length $k\}$. Define V_k analogously as points which can be joined to A by paths of type 2 with length k.

Take $Antigraph(h_{2k-1}) = V_{2k-1} \cup H_{2k}$ and $Graph(h_{2k}) = V_{2k} \cup H_{2k+1}$. To see that this gives single-valued functions, let $\pi_1 = \pi^U$, $\pi_2 = \pi^V$ be the canonical projections, and suppose $\exists u \in U$ such that $u \in \pi_1(V_{2k-1})$ and $u \in \pi_1(H_{2k})$. Then $\exists (u, v_1) \in V_{2k-1}$ and $(u, v_2) \in H_{2k}$ with $v_1 \neq v_2$. But (u, v_2) connects to (u, v_1) on a vertical path of length 1, which we can concatenate with the path to (u, v_2) from H_{2k} to conclude that (u, v_1) also lies in H_{2k+1}, contradicting uniqueness of paths. Thus π_1 is injective on $V_{2k-1} \cup H_{2k}$. Defining $h_{2k-1}(u) = \pi_1^{-1}(u)$ for $u \in \pi_1(V_{2k-1} \cup H_{2k})$ thus gives a single-valued function. Similarly, $h_{2k}(v) = \pi_2^{-1}(v)$ for $v \in \pi_2(V_{2k} \cup H_{2k+1})$ gives a single-valued function.

To show that these functions are a limb-numbering system, we first notice that if $(u_{2k+1}, v_{2k+1}) \in H_{2k+1}$, then there is a point $(u_{2k}, v_{2k}) \in H_{2k}$ which must have $u_{2k} = u_{2k+1}$. Thus, $\pi^U(H_{2k+1}) \subset \pi^U(H_{2k})$. Similarly, we can deduce $\pi^U(V_{2k}) \subset \pi^V(V_{2k-1})$, $\pi^V(H_{2k}) \subset \pi^V(H_{2k-1})$ and $\pi^V(V_{2k+1}) \subset \pi^V(V_{2k})$. From this we conclude that $Ran(h_i) \subset Dom(h_i - 1)$ for $i > 1$. Finally, by uniqueness of paths, we find that for $j \neq k$, $\pi_1(H_{2k}) \cap \pi_1(H_{2j})$, $\pi_1(V_{2k-1}) \cap \pi_1(V_{2j-1})$, $\pi_2(H_{2k-1}) \cap \pi_2(H_{2j-1})$ and $\pi_2(V_{2k}) \cap \pi_2(V_{2j})$ are empty, as are $\pi_1(H_k) \cap \pi_1(V_j)$ and $\pi_2(H_k) \cap \pi_2(V_j)$ for all $k, j > 1$. From this, we conclude that $Dom(h_{2k-1}) \cap Dom(h_{2j-1}) = \emptyset$ and $Dom(h_{2k-1}) \cap Dom(h_{2j-1}) = \emptyset$ for $k, j > 1$. In a similar way, we can determine that $Ran(h_1) \cap Dom(h_{2i}) = \emptyset \ \forall i > 0$. Thus, S is a limb-numbering system. □

Proof of Theorem 1.8

The proof of this theorem is taken from [11] without any significant modifications. First, we will need a lemma.

Lemma A.1. *Let U, V be complete separable metric spaces and γ a positive Borel measure on $U \times V$. If γ is concentrated on the graph of a function $f : Dom(f) \subset U \to V$, then f is measurable with respect to the marginal μ and γ is the push-forward of μ onto $Graph(f)$, $\gamma = (id_U \times f)_{\#}\mu$.*

Proof. This was also proven in [11], using an idea from Villani's new book [10]. If γ is Borel and σ-finite on a complete separable metric space, then γ is regular and σ-compact (see e.g. [12]). Thus, we can find an increasing sequence of compact sets $K_i \subset Graph(f)$ such that the outer measure $\gamma^*(Graph(f) \setminus K_i) < 2^{-i}$. Their union, $K_\infty = \cup_{i=1}^{\infty} K_i$, is a σ-compact set which exhausts the measure of γ.

Since each $K_i \subset Graph(f)$ is compact, then f must be continuous on the compact projection $U_i = \pi^U(K_i)$. Denote the restriction of f to $U_\infty = \pi^U(K_\infty)$ by f_∞. This map is Borel and its Graph is K_∞. Let $A \times B \subset U \times V$ be an arbitrary Borel rectangle. Then

$$\begin{aligned} \gamma(A \times B) &= \gamma((A \times B) \cap Graph(f_\infty)) \\ &= \gamma(((A \cap f_\infty^{-1}(B)) \times V)) \\ &= \mu(A \cap f_\infty^{-1}(B)). \end{aligned}$$

Thus $\gamma = (id_{U_\infty} \times f_\infty)_{\#}\mu$. Letting $A = U \setminus U_\infty$, $B = V$, we see that $U \setminus U_\infty$ is a μ-null set. Since the functions $id_U \times f$ and $id_{U_\infty} \times f_\infty$ agree μ-a.e. we conclude that f is measurable with respect to the complete domain of μ and $\gamma = (id_U \times f)_{\#}\mu$. □

Proof of theorem 1.8. Let $S = \cup_{k=1}^{\infty}[Graph(h_{2k-1}) \cup Antigraph(h_{2k})]$ be a limb-numbering system with $\gamma^*(S^c) = 0$. Since $Dom(h_{2k-1})$ is disjoint from $Dom(h_{2j-1})$ for $j \neq k$, then the Graphs of all odd functions are disjoint. Similarly, all the even function Antigraphs are disjoint. In fact, the Graphs are disjoint from the Antigraphs: if $j \neq k$, $Ran(h_{2j}) \subset Dom(h_{2j-1})$, which is disjoint from $Dom(h_{2k-1})$, while if $j = k$, then $Ran(h_{2k-1}) \subset Dom(h_{2k-2})$, which is disjoint from $Dom(h_{2k})$.

As in the above thoerem, we can find an increasing sequence of compact sets $K_k^j \subset (Anti)Graph(h_k)$ such that the outer measure $\gamma^*((Anti)Graph(h_k) \setminus K_k^j) < 2^{-(j+k)}$. For each k, the union $K_k^\infty = \cup_{j=1}^\infty K_k^j$ is a σ-compact set which exhausts the measure of $\gamma|_{(Anti)Graph(h_k)}$, and the sets K_k^∞ are disjoint. Finally, the union $K^\infty = \cup_{j=1}^\infty K_k^\infty$ exhausts the full measure of γ on $U \times V$. Decomposing γ into disjoint pieces, $\gamma_k = \gamma|_{K_k^\infty}$, we get $\gamma = \sum_{k=1}^\infty \gamma_k$.

Let $\mu_{2k} := \pi_\#^U \gamma_{2k}$, $\nu_{2k-1} := \pi_\#^V \gamma_{2k-1}$ be the projections of each γ_k onto U and V and let η_i represent the (U or V) marginal indexed by i. From the above lemma, we conclude that h_i is η_i-measurable and that

$$\gamma_{2k} = (id_U \times h_{2k})_\# \eta_{2k}, \tag{1}$$

and

$$\gamma_{2k-1} = (h_{2k-1} \times id_V)_\# \eta_{2k-1}. \tag{2}$$

To determine the marginals, we use the fact that μ_{2k} is non-zero only on $Dom(h_{2i})$, so $\mu_{2k} = (\mu - \sum_{i \neq 2k} \mu_i)|_{Dom(h_{2k})}$. However, each piece μ_i is zero outside of $Dom(f_i)$ (if i even) or $Ran(f_i) \subset Dom(f_{i-1})$ (if i odd). By disjointness of domains, only when $i = 2k+1$ is μ_i non-zero on $Dom(h_{2k})$, so

$$\mu_{2k} = (\mu - \mu_{2k+1})|_{Dom(h_{2k})}. \tag{3}$$

Similarly, we have that

$$\nu_{2k-1} = (\nu - \nu_{2k})|_{Dom(h_{2k-1})}. \tag{4}$$

If the limb-numbering system S has only finitely many limbs, then there is some N such that $\forall i \geq N$, $Dom(h_i) = \emptyset$. In this case, we can compute γ_N explicitly using the above formulae, then work recursively down to $k = 1$. The formulae for each γ_k will specify γ uniquely.

If S instead has countably many limbs, but the space $U \times V$ has finite total measure, assume there are two measures γ and $\bar{\gamma}$, both concentrated on S with the same projections μ and ν. If necessary, take the compact sets K_k^j to be large that their union K_k^∞ exhausts both measures γ and $\bar{\gamma}$ on $(Anti)Graph(h_k)$. Let $\epsilon > 0$ and find N such that both measures assign less than ϵ to the tail $\sum_{k=N+1}^\infty K_k^\infty$. Designate $\gamma^\epsilon := \sum_{k=1}^N \gamma_k$ and $\bar{\gamma}^\epsilon := \sum_{k=1}^N \bar{\gamma}_k$ where $\bar{\gamma}_k := \bar{\gamma}|_{K_k^\infty}$.

Both γ^ϵ and $\bar{\gamma}^\epsilon$ are concentrated on the same finite limb system, but the differences between their projections $\delta\mu^\epsilon = \bar{\mu^\epsilon} - \mu^\epsilon$, and $\delta\nu^\epsilon = \bar{\nu^\epsilon} - \nu^\epsilon$ have total variation bounded above by 2ϵ. The restrictions $\delta\eta_{2k}^\epsilon = \mu^\epsilon|_{Dom(h_{2k})}$ and $\delta\eta_{2k+1}^\epsilon = \nu^\epsilon|_{Dom(h_{2k+1})}$ are disjoint, so the sum of these total variations is also bounded: $\sum_{k=1}^N ||\delta\eta_k||_{Dom(h_k)} = ||\delta\mu^\epsilon|| + ||\delta\nu^\epsilon|| \leq 4\epsilon$. Making use of equations (1) and (2) above, we find

$$||\bar{\gamma}^\epsilon - \gamma^\epsilon||_{TV(U \times V)} = \begin{cases} ||(id_U \times h_k)_\# \delta\eta_k||_{TV(U \times V)} & \text{k even} \\ ||(h_k \times id_V)_\# \delta\eta_k||_{TV(U \times V)} & \text{k odd} \end{cases} = ||\delta\eta_k||_{TV(Dom(f_k))}.$$

By summing on k, we conclude $||\bar{\gamma}^\epsilon - \gamma^\epsilon||_{TV(U \times V)} \leq 4\epsilon$. Taking the limit $\epsilon \to 0$, we get $\bar{\gamma}^\epsilon \to \bar{\gamma}$, $\gamma^\epsilon \to \gamma$, and hence $\bar{\gamma} = \gamma$, so γ must be unique. $\square$

References

[1] G. Birkhoff. Tres observaciones sobre el algebra lineal. *Univ. nac. Tucumn, Revista Ser. A*, **5**:147-150, 1946. Contains first proof of Birkhoff-von Neumann theorem.

[2] J. von Neumann. A certain zero-sum two-person game equivalent to the optimal assignment problem. *Contributions to the Theory of Games*, **2**:5-12, Princeton, 1953. Contains an alternate proof of Birkhoff-von Neumann theorem.

[3] M. Krein & D. Milman. On extreme points of regular convex sets. *Studia Mathematica*, **9**:133-138, 1940. Reference for Krein-Milman theorem.

[4] G. Birkhoff. *Lattice Theory* (revised edition). Amer. Math. Soc., New York, 1948. Problem 111 of this book originally lead to research in extremal doubly stochastic measures.

[5] R. D. Douglas. On extremal measures and subspace density. *Michigan Math. J.*, **11**:243-246, 1964. See below.

[6] J. Lindenstrauss. A remark on doubly stochastic measures. *Amer. Math. Monthly*, **72**:379-382, 1965. Along with Douglas, provides first necessary condition for a doubly stochastic measure to be extremal, in terms of subspace density.

[7] V. Beneš & J. Štěpàn. The support of extremal probability measures with given marginals. *Mathematical Statistics and Probability Theory*, **A** (M.L Puri et. al. (eds.)), 33-41, 1987. Gives, in a general setting, necessary geometric conditions for measures to be simplicial.

[8] K. Hestir & S.C. Williams. Supports of doubly stochastic measures. *Bernouilli* **1**:217-243, 1995. Extends Beneš & Štěpàn's result to find more structure in the special case of extremal doubly stochastic measuares.

[9] C. Villani. *Topics in Optimal Transportation.* Graduate Studies in Mathematics Volume 58, 2003. Reference for optimal transport problem.

[10] C. Villani. *Optimal Transport, Old and New.* Springer, 2007. See above. Also provides an idea used to prove Lemma A.1.

[11] P.-A. Chiappori, R. McCann & L. Nesheim. Nonlinear hedonic pricing: finding equilibria through linear programming. (in preparation). Provides a sufficient condition for doubly stochastic measures to be extremal.

[12] W. Rudin, *Real and Complex Analysis* 3rd edition. McGraw-Hill, 1986. Used as reference for some elementary properties of measures.

[13] L. Mirsky, Results and problems in the theory of doubly-stochastic matrices, *Z. Wahrscheinlichkeitstheorie* **1**:319-334, 1963. Useful, if dated, overview of the study of doubly stochastic matrices; provides references to Birkhoff & von Neumann's original papers.