*The Annals of Probability*
1991, Vol. 19, No. 4, 1781–1797

# EXISTENCE OF PROBABILITY MEASURES WITH GIVEN MARGINALS

BY SAM GUTMANN, J. H. B. KEMPERMAN, J. A. REEDS AND L. A. SHEPP

*Northeastern University, Rutgers University and AT & T Bell Laboratories*

We show that if $f$ is a probability density on $R^n$ wrt Lebesgue measure (or any absolutely continuous measure) and $0 \le f \le 1$, then there is another density $g$ with only the values 0 and 1 and with the same $(n-1)$-dimensional marginals in any finite number of directions. This sharpens, unifies and extends the results of Lorentz and of Kellerer.

Given a pair of independent random variables $0 \le X, Y \le 1$, we further study functions $0 \le \phi \le 1$ such that $Z = \phi(X, Y)$ satisfies $E(Z \mid X) = X$ and $E(Z \mid Y) = Y$. If there is a solution then there also is a nondecreasing solution $\phi(x, y)$. These results are applied to tomography and baseball.

**1. Introduction.** Let $\mu, \nu$ be a given pair of finite measures on $R^1$. In 1949, Lorentz [12] gave a necessary and sufficient condition on $\mu, \nu$ so that there is a set whose marginals are $\mu$ and $\nu$. Here, a set is also regarded as a 0, 1-valued density $g$ on $R^2$. In 1961, Kellerer [6], page 340, gave a necessary and sufficient condition on $\mu, \nu$ so that there exists a density $f$ on $R^2$ with $0 \le f \le 1$, whose marginals are $\mu$ and $\nu$ (see Strassen [16], page 432, and Jacobs [4] for different proofs). It is not hard to verify (see [3]) that the Lorentz and Kellerer conditions are actually equivalent. Thus it follows that for any integrable density $0 \le f \le 1$ on $R^2$ there is a set with the same marginals.

This carries over to $R^n$. For instance, Kellerer [9], Theorem 1.7, already showed that if $0 \le f \le 1$ is a density in $R^n$, then there is always a set (i.e., a 0, 1-valued function $g$), with the same $(n-1)$-dimensional marginals as $f$. This remains true if the reference Lebesgue measure $\lambda(dx) = dx$ for $f$ and $g$ is replaced by any absolutely continuous measure $\lambda(dx) = q(x)\,dx$ on $R^n$. Employing a very different proof, we will prove a more general result, of considerable importance in tomography. (Romanovskii and Sudakov gave a very similar proof for a related theorem; see [11].) Here, the above set of $(n-1)$-dimensional marginals is generalized to the collection of measures induced by any fixed finite set of linear functions onto lower-dimensional spaces, such as the set of marginals (of the measures $f\,d\lambda$ and $g\,d\lambda$) in any finite number of directions. The general statement is given in Theorem 2.

There is a similar result, [5], page 265, that if a *finite* number of moments of a function $f$, $0 \le f \le 1$, are given, then there is a $g = 0$ or 1 with these same moments. Our condition that $f$ have given marginals amounts to an infinite set of moment conditions on $f$. Our theorems are also closely related

Received June 1988; revised July 1990.
*AMS* 1980 *subject classifications*. 52A40, 28A35, 60A10.
*Key words and phrases*. Baseball, tomography, marginals.

to Lyapounov's theorem, [10] and [5], page 266, that the range of a nonatomic vector-valued measure is convex. See also [7, 8] for related results.

An application of Theorem 2 is given in Section 3 to the following problem which was our initial motivation. Given a pair of *independent* real random variables $X, Y$ with $0 \le X, Y \le 1$, when does there exist a function $\phi(x, y)$, for which $Z = \phi(X, Y)$ has the properties

$$E[Z \mid X] = X, \tag{1.1a}$$

$$E[Z \mid Y] = Y, \tag{1.1b}$$

$$0 \le Z \le 1. \tag{1.1c}$$

Note that the existence of $Z$ is a condition only on the distribution functions $F$ and $G$ of $X$ and $Y$. Taking expectations in (1.1a) and (1.1b) shows that $EX = EY$ is necessary for (1.1). The problem arose in a model for computerized baseball, where $X$ represents the batting average of a random batter and $Y$ the "batting average" of a random pitcher. Here, the batting average of a given pitcher is interpreted as the chance that a randomly chosen batter will get a hit against him. The above problem asks us to determine for what pairs $F, G$ there exists a compatible (randomized) rule $Z = \phi(X, Y)$. It could be used in constructing a somewhat realistic computer game. One would interpret $\phi(x,y)$ as a possible version of the (unknown) chance function $\phi(x, y)$ that a hit will be obtained, in a situation where the batter has batting average $X = x$ and the pitcher has "batting average" $Y = y$.

The problem is discussed in detail in Section 3 where a necessary and sufficient condition on the pair $F, G$ is obtained using the Kellerer–Strassen theorem. Provided $F$ and $G$ are continuous and (1.1) has a solution, one can even achieve (1.1) with a $Z$ or $\phi$ which assumes only the values 0 or 1. Such a deterministic $Z$ would be somewhat pathological in the corresponding computer game, since it would cause hits and outs to be completely predictable. Note that the assumption that $X$ and $Y$ are chosen independently is reasonable unless one wants to allow the subtleties of pinch-hitting with lefty batters against righty pitchers.

We further show (Theorem 5), for general $F$ and $G$, that there always exists a solution $\phi(X, Y)$ of (1.1) having the additional property

$$\phi(x, y) \text{ increases in each variable separately,} \tag{1.2}$$

whenever a solution to (1.1) exists at all. However, we will see that in general there is no $\phi$ which satisfies both (1.1) and (1.2) and also $\phi = 0$ or 1.

Note that the main theorem (Theorem 2) has a somewhat negative and paradoxical application to CT (tomography). It implies that for any human object and corresponding projection data there exist many different reconstructions, in particular, a reconstruction consisting only of bone and air (density 1 or 0) but still having the same projection data as the original object. Related nonuniqueness results are familiar [15] in tomography and are usually ignored because CT machines seem to produce useful images. It is likely that the "explanation" of this apparent paradox is that a *point* reconstruction in tomography is impossible. CT machines produce useful images because all functions $0 \leq f \leq 1$ with the same line integrals have (essentially) the same integrals over "nice" sets. In other words, it is likely that all functions $f$ with $0 \leq f \leq 1$ and with the same line integrals have nearly identical integrals over pixels which are not too small. But we have neither a proof nor a precise statement of this heuristic idea. (L. Khalfin and L. Klebanov have recently proved $f * \phi$ and $g * \phi$ are close when $\phi$ is Gaussian with large enough variance, in a quantitative and precise statement.)

**2. Main theorem.** In general, let $S$ be a fixed measurable space supplied with a *finite* measure $\lambda$. Let $J$ be an arbitrary finite or infinite index set. For each $j \in J$, let $Y_j$ be a fixed measurable space and $\pi_j$: $X \to Y_j$ a fixed measurable function, also called a projection from $X$ onto $Y_j$. We will be interested in (finite) measures $d\mu = g\,d\lambda$ with $0 \leq g \leq 1$ on $S$ having given marginals $\pi_j\mu$, $j \in J$.

We will work with the spaces $L^1(\lambda)$ and $L^\infty(\lambda)$; in particular, functions on $S$ that are equal a.e. $\lambda(dx)$ will be identified. Recall, [2], page 289, that $L^\infty(\lambda)$ is precisely the dual of $L^1(\lambda)$. We will assign to $L^\infty(\lambda)$ the $L^1(\lambda)$-topology of $L^\infty(\lambda)$, usually called the weak* topology of $L^\infty(\lambda)$. It is the coarsest topology on $L^\infty(\lambda)$ such that

$$g \to \int g(x)h(x)\lambda(dx) \tag{2.1}$$

is a continuous function on $L^\infty(\lambda)$, for each choice of $h \in L^1(\lambda)$. It is an easy result due to Alaoglu, [2], page 424, that any closed ball in $L^\infty(\lambda)$ is compact. In particular, the closed and bounded set $K = \{g\colon 0 \leq g \leq 1\}$ is (weak*) compact. That $K$ is closed follows from the fact that $g \in K$ is equivalent to

$$0 \leq \int g(x)h(x)\lambda(dx) \leq \int h(x)\lambda(dx) \quad \text{for all } h \in L^1(\lambda),\ h \geq 0. \tag{2.2}$$

Definition. Let the system of projections $\pi_j$, $j \in J$, be fixed. We will say that the (finite) measure $\lambda$ is *rich* if for any choice of the measurable function $f$: $S \to R$ such that $0 \leq f \leq 1$ there exists a measurable function $g$: $S \to R$, taking only values 0 and 1, such that, for all $j \in J$, the $\pi_j$-marginal (also called the $\pi_j$-projection) of the measure $g\,d\lambda$ on $S$ is equal to the $\pi_j$-marginal of the measure $f\,d\lambda$ on $S$. The latter means that

$$\int 1_B(\pi_j x)g(x)\lambda(dx) = \int 1_B(\pi_j x)f(x)\lambda(dx), \tag{2.3}$$

for all $j \in J$ and all measurable subsets $B$ of $Y_j$. Equivalently,

$$\int \phi_j(\pi_j x)g(x)\lambda(dx) = \int \phi_j(\pi_j x)f(x)\lambda(dx), \tag{2.4}$$

for all $j \in J$ and all functions $\phi_j$: $Y_j \to R$, provided either $\phi_j \geq 0$ or else $\phi_j(\pi_j x)f(x) \in L^1(\lambda)$. In particular, $\int g\,d\lambda = \int f\,d\lambda$.

In this direction, Kellerer [7], Theorem 6.1, already established the following important result. Let the measure space $(S, \rho)$ be the direct product of finitely many nonatomic $\sigma$-finite measure spaces $(S_i, \rho_i)$, $i = 1, \dots, n$, and take $\lambda(dx) = q(x)\rho(dx)$, where $q(x) \geq 0$. Further, choose $J = \{1, \dots, n\}$ and $\pi_j(x) = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$, $j \in J$. Then to any measurable function $0 \leq f \leq 1$ on $S$ there corresponds a measurable subset $A$ of $S$ such that $f\,d\lambda$ has, for each $j \in J$, the same $\pi_j$-marginals as $1_A\,d\lambda$.

DEFINITION. Let $f$ be a fixed measurable function $f\colon S \to R$ satisfying $0 \leq f \leq 1$; thus $f \in L^\infty(\lambda) \subset L^1(\lambda)$. By $M(f)$ we denote the collection of all $g \in L^\infty(\lambda)$, such that $0 \leq g \leq 1$, and that further $g\,d\lambda$ has the same $\pi_j$-marginal as $f\,d\lambda$, for all $j \in J$. One has $f \in M(f)$; thus $M(f)$ is nonempty. Moreover, $\int g\,d\lambda = c$ for all $g \in M(f)$, where $c = \int f\,d\lambda \leq \lambda(S) < \infty$.

The above marginals condition can be expressed by (2.3) with $B$ as a measurable subset of $Y_j$. Here, $1_B(\pi_j x) \in L^\infty(\lambda) \subset L^1(\lambda)$; thus (2.3) defines a closed subset $M(f)$ of $L^\infty(\lambda)$. Consequently, $M(f)$ is a *compact and convex* subset of $L^\infty(\lambda)$.

From the Krein–Milman theorem, [2], page 440, $M(f)$ is the closed convex hull of the set $E(f)$ of extreme points of $M(f)$; in particular, $E(f)$ is nonempty. Our hope, which is true if the number of $\pi$'s is finite and they are each linear projections as in Theorem 3, is that each $g \in E(f)$ only takes the values 0 or 1 (a.e. $[\lambda]$). In any case, each $g \in M(f)$ with that property (if any) must belong to $E(f)$.

REMARK. In this connection, it is of interest to consider the strictly convex functional $N(g) = \int g(x)^2 \lambda(dx)$ on $M(f)$. Since $\sqrt{N(g)} = \sin\{\int gw\,d\lambda\colon \int |w|^2\,d\lambda = 1\}$, the function $g \to N(g)$ is (weak*) lower semicontinuous. Consequently, its restriction to the compact set $M(f)$ assumes its smallest value $\inf\{N(g)\colon g \in M(f)\}$ at a unique point $g_0 \in M(f)$.

On the other hand, it is not clear why there should exist any maximal element $g^\#$ in $M(f)$, by which we mean any $g^\# \in M(f)$ such that $N(g^\#) = \sup\{N(g)\colon g \in M(f)\}$. Necessarily $g^\# \in E(f)$, because $N(g)$ is strictly convex. Thus a maximal element (if it exists) is merely a special type of extreme point of $M(f)$.

Maximal elements do exist when $\lambda$ is rich. For then there is a 0, 1-valued $g^\# \in M(f)$ and such an element $g^\#$ is always maximal. After all, $N(g^\#) = \int g^\#\,d\lambda = c$, where $c = \int f\,d\lambda$. Any other $g \in M(f)$ has $0 \leq g \leq 1$; thus $N(g) \leq \int g\,d\lambda = c$. The latter inequality holds with equality if and only if $g$ is 0, 1. Note that minimizing (maximizing) $N(g)$ is the same as minimizing (maximizing) $\int [g(x) - 1/2]^2 \lambda(dx)$.

Let $g \in E(f)$ be fixed and suppose that $g$ were *not* a 0, 1 function. Then there exists $0 < \varepsilon < 1/2$ such that

$$\lambda(D) > 0, \quad \text{where } D = \{x \in S\colon \varepsilon \leq g(x) \leq 1 - \varepsilon\}. \tag{2.5}$$

Let $h\colon S \to R$ be any *bounded* measurable function, not equal to 0 a.e. $\lambda(dx)$, which is supported by $D$ [i.e., $h(x) = 0$ if $x \in D^c = S \setminus D$]. We claim that it is impossible that the associated signed measure $h\,d\lambda$ has zero $\pi_j$-marginal, for all $j \in J$; equivalently,

$$\int 1_B(\pi_j x) h(x) \lambda(dx) = 0 \quad \text{for all } j \in J,\ B \subset Y_j. \tag{2.6}$$

After all, otherwise, if $|h| \le C =$ constant and $0 < \delta \le \varepsilon/C$ then $g + \delta h \in M(f)$; $g - \delta h \in M(f)$ while $g = (1/2)((g + \delta h) + (g - \delta h))$ and $g$ would not be extreme. These considerations lead to the following definition.

DEFINITION. The measure $\lambda$ on $S$ is said to be *strongly rich* (relative to the given system of projections $\pi_j\colon S \to Y_j$, $j \in J$) if the following holds for each measurable subset $D$ of $S$ with $\lambda(D) > 0$. Namely, there must exist a *bounded* measurable function $h\colon S \to R$, not 0 a.e. $\lambda(dx)$, which is supported by $D$, and such that the signed measure $h(x)\lambda(dx)$ has its $\pi_j$-marginal equal to 0, for all $j \in J$.

The same idea appears in Kingman and Robertson [10], who use the term thin rather than rich. Note that the zero measure is trivially strongly rich. The above argument proves the following theorem, which is due to Kingman and Robertson [10].

THEOREM 1. *A sufficient condition for the finite measure $\lambda$ to be rich is that it be strongly rich.*

PROPOSITION. *Let $\lambda$ and $\mu$ be finite measures on $S$. If they have the same sets of measure 0, then either both $\lambda$ and $\mu$ are strongly rich or neither $\lambda$ and $\mu$ are strongly rich. More generally, suppose $\lambda$ is strongly rich and $\mu$ is absolutely continuous relative to $\lambda$. Thus $\mu(dx) = q(x)\lambda(dx)$ with $q(x) \ge 0$, $q \in L^1(\lambda)$. Then $\lambda$ strongly rich implies that $\mu$ is strongly rich.*

PROOF. Let $\lambda$ and $\mu$ be finite with $\mu(dx) = q(x)\lambda(dx)$ and $\lambda$ strongly rich. We must prove that $\mu$ is strongly rich. Let $D$ be a measurable subset of $S$ with $\mu(D) = \int_D q(x)\lambda(dx) > 0$. Let $D_1 = \{x \in D\colon q(x) \ge \varepsilon\}$. For $\varepsilon > 0$ sufficiently small, we have that $\lambda(D_1) > 0$. Because $\lambda$ is strongly rich, there exists a bounded measurable function $h\colon S \to R$, not equal to 0 a.e. $[\lambda]$, which is supported by $D_1$ and such that the signed measure $\eta(dx) = h(x)\lambda(dx)$ has all its marginals equal to 0 (i.e., $\pi_j \eta = 0$ for all $j \in J$). Now consider the function $H(x) = h(x)/q(x)$ on $S$. It is supported by $D_1$, and thus by $D$, while $H$ is not 0 a.e. relative to $\mu(dx) = q(x)\lambda(dx)$. Moreover, $H$ is bounded, since $h$ is bounded and $q(x) \ge \varepsilon$ whenever $h(x) \ne 0$, always with $x \in D_1$. Finally, $\eta(dx) = H(x)\mu(dx)$ has all its marginals equal to 0. This proves that $\mu$ is strongly rich. □

In most applications, each $x \in S$ is uniquely determined by the set of images $\pi_j x$, $x \in J$, and each one-point subset $\{x\}$ of $S$ and $\{y_j\}$ of $Y_j$ is measurable ($j \in J$). In such a situation, in order that a measure $\lambda$ on $S$ be rich, it is necessary that $\lambda$ be nonatomic. That is, a measure $\lambda$ having a positive mass $\lambda(\{x_0\}) > 0$ at a single point $x_0$ cannot possibly be rich (and certainly not strongly rich either). After all, consider $d\mu = f\,d\lambda$ with $f(x_0) = 1/2$ and $f(x) = 0$, otherwise, and suppose that $dv = g\,d\lambda$ satisfies $\pi_j v = \pi_j \mu$, for all $j \in J$. Since $\mu$ is carried by $\{x_0\}$, we have, for all $j \in J$, that $\pi_j v = \pi_j \mu$ is carried by $\{\pi_j x_0\}$; hence $v$ is carried by $A_j = \{x \in S\colon \pi_j x = \pi_j x_0\}$. But $\bigcap_{j \in J} A_j = \{x_0\}$, showing that thus also $v$ is carried by $\{x_0\}$, which forces that $v = \mu$; thus $g(x_0) = f(x_0) = 1/2$.

On the other hand, $\lambda$ can be nonatomic without being rich. For example, take $S = R^n$ together with the one-dimensional projections $\pi_j(x_1, \dots, x_n) = x_j$, $j = 1, \dots, n$. Furthermore, let $\lambda$ be one-dimensional Lebesgue measure on a fixed finite line segment $L$ in $R^n$; thus $\lambda$ is nonatomic. Nevertheless, $\lambda$ is not rich. For, consider $f \equiv 1/2$. As is easily seen, in order that $dv = g\,d\lambda$ have the same projections as $f\,d\lambda$, it is necessary that $g \equiv f$ (a.e. $[\lambda]$) in which case $g$ is not 0, 1-valued. The basic idea here is that no subset $A$ of $L$ can satisfy $\lambda(A \cap B) = \lambda(B)/2$ for every subinterval $B$ of $L$.

DEFINITION. The following situation will be referred to as the *classical case* [with $(n-1)$-dimensional projections]. Here, we take $S = R^n$ and each $\pi_j$, $j \in J$, as a *parallel projection* of $R^n$ along lines of fixed direction $w_j$. Here, $w_j \in R^n$, $w_j \neq 0$. Naturally, $w_j' = t_j w_j$ is entirely equivalent to $w_j$, $t_j \in R$; $t_j \neq 0$. One may take $Y_j$ as any hyperplane in $R^n$ which is not parallel to $w_j$, thus $\dim(Y_j) = n - 1$, $j \in J$.

Further $m(dx) = dx$ will denote $n$-dimensional Lebesgue measure on $R^n$. In the results below it is assumed that $J = \{1, \dots, N\}$ is finite. These results do not carry over to the case where $J$ is denumerably infinite.

THEOREM 2. *Suppose one is in the above classical case, with $S = R^n$ and $J = \{1, \dots, N\}$ finite. Furthermore, let $q(x) \geq 0$ be a fixed Lebesgue integrable function on $R^n$. Then the (finite) measure $\lambda(dx) = q(x)\,dx$ on $R^n$ is strongly rich; hence it is rich.*

COROLLARY. *The same conclusion obtains when instead $\pi_j$, $j \in J$, is an arbitrary finite collection of linear projections $\pi_j\colon R^n \to Y_j$, provided the $Y_j$ are linear spaces of dimension less than or equal to $n - 1$.*

PROOF OF THEOREM 2. Let $\lambda(dx) = q(x)\,dx$ and let $D$ be any subset of $R^n$ such that $\lambda(D) > 0$. We must show that there is a bounded measurable function $h\colon R^n \to R$, supported by $D$, not 0 a.e. $\lambda(dx)$, such that $h(x)\lambda(dx)$ has all its marginals equal to 0. Replacing $D$ by a bounded subset, we may assume that $D$ is contained in a finite cube $K$. It suffices to prove the above for $\lambda$ replaced by its restriction to $K$. And in view of the above proposition, it therefore suffices to consider the case that $\lambda$ is precisely the restriction of $n$-dimensional Lebesgue measure to $K$.

Thus let $D \subset K$ satisfy $m(D) > 0$. Furthermore, let $V$ denote the set of $2^N$ vectors $s = (s_1, \dots, s_N)$ with $s_j = 0$ or 1. Define further $w(s) = s_1 t_1 w_1 + \cdots + s_N t_N w_N$, $s \in V$. Here, the $t_j \neq 0$ are fixed real numbers such that all the $2^N$ vectors $w(s)$ are different. This is always possible; in fact, $t = (t_1, \dots, t_N) \in R^N$ merely needs to avoid a certain set of $N$-dimensional Lebesgue measure 0. Replacing $w_j$ by $w_j' = t_j w_j$, one may as well assume that $t_j = 1$ for all $j$; thus $w(s) = s_1 w_1 + \cdots + s_N w_N$. Let

$$\delta = \min\{|w(s) - w(s')| : s \neq s';\, s, s' \in V\};$$

thus $\delta > 0$.

For $\rho \in R$, define

$$D(\rho) = \bigcap_{s \in S} (D - \rho w(s)) \quad \text{and} \quad \phi(\rho) = m(D(\rho)).$$

Using the regularity of Lebesgue measure, the function $\phi(\rho)$ is easily seen to be continuous. Since $\phi(0) = m(D) > 0$, we have $\phi(\rho) > 0$ when $|\rho|$ is sufficiently small. Let $\rho > 0$ be fixed such that $\phi(\rho) > 0$. Next, choose $A$ as a measurable subset of $D(\rho)$ such that $m(A) > 0$ and that $A$ has its diameter less than $\rho\delta$. For $s \in V$, define

$$A(s) = A + \rho w(s) = \{x \in R^n : x - \rho w(s) \in A\}.$$

These $2^N$ translates of $A$ are disjoint, since the $2^N$ distinct vectors $\rho w(s)$, $s \in V$, are at least $\rho\delta$ apart. Now define $h\colon R^n \to R$ by

$$h(x) = \begin{cases} (-1)^{|s|}, & \text{if } x \in A(s),\ s \in V, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $|s| = s_1 + \cdots + s_N$. Since $A \subset D(\rho) \subset D - \rho w(s)$, we have $A(s) = A + \rho w(s) \subset D$, $s \in V$, showing $h$ is supported by $D$. It is obvious that $h$ is bounded and not 0 a.e; thus it only remains to verify that the $\pi_j$-marginal of $h(x)\,dx$ equals 0, $j = 1, \dots, N$. When $j = 1$ this means that

$$\int_{-\infty}^{+\infty} h(x_0 + t\rho w_1)\,dt = 0 \quad \text{for all } x_0 \in R^n.$$

But note that $x = x_0 + t\rho w_1$ is in $A(s^0) = A(0, s_2, \dots, s_N)$ if and only if $x^* := x + \rho w_1$ is in $A(s^1) = A(1, s_2, \dots, s_N)$ while in that case $h(x) + h(x^*) = 0$. This shows that, for each $s \in V$, the contributions of $A(s^0)$ and $A(s^1)$ to the latter integral precisely cancel each other. This completes the proof of Theorem 2. □

SUMMARY. Consider the classical case as defined above. We have shown that any finite measure $\lambda(dx) = q(x)\,dx$ on $R^n$ is rich and even strongly rich. Let $f\colon S \to R$ be measurable with $0 \le f \le 1$ and consider the associated (weak*) compact and convex set $M(f)$, consisting of all $0 \le g \le 1$ such that $g\,d\lambda$ has the same marginals as $f\,d\lambda$. Then $M(f)$ has at least one extreme point and, moreover, each extreme point $g$ of $M(f)$ is of the form $g = 1_A$ with $A$ as a (Lebesgue) measurable subset of $R^n$. In particular, there exists at least one set $A$ such that $1_A(x)\lambda(dx) = 1_A(x)q(x)\,dx$ has the same marginals as $f(x)q(x)\,dx$.

The previous result can be considerably generalized as follows.

THEOREM 3. *Consider any finite system* ($\pi_j$, $j = 1, \dots, N$) *of central projections in* $R^n$. *Furthermore, let* $q(x) \geq 0$ *be a Lebesgue integrable function on* $R^n$. *Then the measure* $q(x)\,dx$ *on* $R^n$ *is strongly rich; hence it is rich. Consequently, for each measurable function* $0 \leq f \leq 1$ *in* $R^n$ *there is a measurable* $g$ *with only the values* 0 *and* 1, *such that the associated measure* $g(x)q(x)\,dx$ *has the same* $\pi_j$*-marginal as the measure* $f(x)q(x)\,dx$, $j = 1, \dots, N$.

Here, we do allow the center $P_j$ of the central projection $\pi_j$ to be a point at $\infty$. In that case, $\pi_j$ is to be interpreted as a parallel projection along the lines of direction $P_j$, onto a fixed hyperplane $Y_j$ not parallel to $P_j$. Below we present a proof of Theorem 3 for the case of two central projections in $R^2$. The full proof is deferred to a separate paper, to be written by J. H. B. Kemperman.

COROLLARY 1. *Given finitely many central projections* $\pi_j$ *in* $R^n$ *and integrable functions* $0 \leq q_1(x) \leq q_2(x)$ *on* $R^n$, *there always exists a measurable subset* $A$ *of* $R^n$ *such that the measure* $1_A(x)q_2(x)\,dx$ *has, relative to each of the* $\pi_j$, *exactly the same marginals as the measure* $q_1(x)\,dx$.

COROLLARY 2. *Let* $\lambda(dx) = q(x)\,dx$ *with* $q(x) \geq 0$ *an integrable function on* $R^n$. *Given any finite number of central projections in* $R^n$ *and any measurable subset* $D$ *of* $R^n$ *with* $\lambda(D) > 0$, *there exist disjoint measurable subsets* $A_1$ *and* $A_2$ *of* $D$ *of positive* $\lambda$*-measure and possessing exactly the same projections, when regarded as measures* $1_{A_i}(x)\,d\lambda$.

REMARK 1. One moral of Corollary 2 is that additional central projections tend to yield additional information. It is true that some very special sets $A$ (such as sets which are additive, relative to the given set of projections $\pi_j$, $j \in J$; see [3]) are sets of uniqueness in the sense that they are uniquely determined by the associated set of projections $\pi_j\mu$, $j \in J$, of the measure $\mu(dx) = 1_A(x)\,d\lambda$. And in that very special case, additional projections do not supply any new information about $A$. However, whatever the given set $A$, we infer from Corollary 2 that, for any preassigned arbitrarily small set $D$ with $\lambda(D) > 0$, there exists a modification of $A$ within the set $D$ only, such that the resulting perturbation $A'$ of $A$ is no longer determined by its projections. If $A$ happens to be a set of uniqueness and $D$ is sufficiently small then, as is easily seen, $A'$ is at least "nearly" determined by its projections.

REMARK 2. Conversely, Corollary 2 implies Theorem 2. Actually, one could formulate Corollary 2 as still another characterization of strongly rich measures. After all, the following proof makes no use whatsoever of the special structure of the projections on hand, nor of the fact that there are only finitely many projections.

PROOF OF COROLLARY 2. All functions considered here will be 0 outside $D$. From Theorem 3, the measure $\lambda$, and thus its restriction to $D$, are strongly rich. Hence there exists a function $-1 \le h(x) \le 1$ on $D$ which is not 0 a.e. $[\lambda]$ and such that $h(x)\,d\lambda$ has all its projections equal to 0. Now consider the set $M(h_+)$ of measurable functions $0 \le g \le 1$ on $D$ such that $g(x)\,d\lambda$ has the same projections as $h_+(x)\,d\lambda$. It is convex and compact in the weak* topology and contains the distinct members $h_-$ and $h_+$; hence $M(h_+)$ possesses at least two distinct extreme points. From the proof of Theorem 1 (applied to $D$ instead of $S$), these extreme points are characteristic functions of distinct subsets $A_1$ and $A_2$ of $D$ as desired (deleting $A_1 \cap A_2$, one can achieve that $A_1$, $A_2$ are disjoint). □

We will now present a proof of Theorem 3 in the special case of just two central projections $\pi_1$ and $\pi_2$ in $S = R^2$, having corresponding centers $P_1 = (1,0)$ and $P_2 = (0,1)$. The coordinates $x_1, x_2$ of $x \in R^2$ will also be denoted as $u$ and $v$. A straight line $L$ through $P_2$ typically has equation $au + v = 1$ and thus arc length $ds = (1 + a^2)^{1/2}\,du$. Knowing the $\pi_2$-projections of a signed measure $h(x)\,dx = h(u,v)\,du\,dv$ is the same as knowing the function

$$\theta_2(a) = \int_{-\infty}^{+\infty} h(u, 1 - au)\,du.$$

Similarly, knowing the analogous $\pi_1$-projection is the same as knowing the function

$$\theta_1(b) = \int_{-\infty}^{+\infty} h(1 - bv, v)\,dv.$$

We would like to prove that each finite measure $q(x)\,dx$ is strongly rich relative to the pair $\pi_1, \pi_2$. In view of the proposition following Theorem 1, it suffices to show that, for each bounded measurable subset $D$ of $R^2$ with $m(D) > 0$, there is a bounded function $h$: $R^2 \to R$ which is supported by $D$, is not equal to 0 a.e. and has $\theta_1(b) = \theta_2(a) = 0$, for almost all choices of the real numbers $a$ and $b$. Replacing $D$ by a smaller set, one can achieve that $D$ has a positive distance to the line $u + v = 1$ through $P_1$ and $P_2$.

The idea is to use a projective transformation $T$ sending $P_1 = (1,0,1)$ and $P_2 = (0,1,1)$ (in homogeneous coordinates) to the point $Q_1 = (1,0,0)$ and $Q_2 = (0,1,0)$, respectively. Relative to the original Euclidean plane, the latter are points at $\infty$. And the central projection by lines through $P_i$ then becomes a parallel projection by lines in direction $Q_i$, $i = 1,2$. In homogeneous coordinates, one such transformation $T$ is given by the matrix

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{with } T^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

For "ordinary" coordinates, this means that we transform the pair $(u, v) \in R^2$ to the pair $(x, y) = T(u, v)$ by means of the equations

$$x = u/(1-u-v), \qquad y = v/(1-u-v)$$

and

$$u = x/(1+x+y), \qquad v = y/(1+x+y).$$

Note that $(1+x+y)(1-u-v) = 1$. The $T$-marginal of the line $u + bv = 1$ through $P_1 = (1,0)$ is the horizontal line (line through $Q_1$) $y = 1/(b-1)$. Similarly, the $T$-marginal of the line $au + v = 1$ through $P_2 = (0,1)$ is the vertical line (line through $Q_2$) $x = 1/(a-1)$. Let $H(x, y)$ be defined by

$$H(x,y) = h\left(\frac{x}{1+x+y}, \frac{y}{1+x+y}\right)\frac{1}{(1+x+y)^2}$$

[where $(1+x+y)^{-2}$ is not the Jacobian; the latter equals $(1+x+y)^{-3}$ instead]. The function $h$ is carried by $D$ if and only if $H$ is carried by its image $TD$. Since the latter set is bounded (by our condition that $D$ stay away from $u + v = 1$), we have in this case that $h$ is bounded as soon as $H$ is bounded. One easily calculates that

$$\theta_1(b) = \frac{1}{b-1}\int H\left(x, \frac{1}{b-1}\right)dx,$$

$$\theta_2(a) = \frac{1}{a-1}\int H\left(\frac{1}{a-1}, y\right)dy.$$

Thus our aim is now to choose $H$ as a bounded function supported by the bounded set $TD$, not 0 a.e., and such that the $H(x, y)$ has all vertical and horizontal line integrals equal to 0. But since $m(TD) > 0$, we know from Theorem 2 that such a function $H$ always exists.

**3. Baseball.** Given a pair of *independent* random variables $X$ and $Y$ with $0 \le X, Y \le 1$, we want to know when there exists a random variable $Z$ satisfying

$$E(Z \mid X) = X, \tag{3.1a}$$

$$E(Z \mid Y) = Y, \tag{3.1b}$$

$$0 \le Z \le 1. \tag{3.1c}$$

Note that $EX = EY$ is necessary and sufficient for (3.1a) and (3.1b) (take $Z = X + Y - EX$ for sufficiency). Henceforth, we assume that $EX = EY = K$ (say). The added conditions (3.1c) is imposed for the following application to baseball. Let $X$ be the batting average of a random batter, and $Y$ the "batting average" of a random pitcher that is, the probability that a random batter gets a hit against him. Since additional randomness in $Z$ is irrelevant for existence, by replacing $Z$ by $Z' = E(Z \mid X, Y)$, one may as well take $Z$ in (3.1) to be of the form $\phi(X, Y)$. Then $\phi(x, y)$ is an a priori possible version of the probability that a batter of batting average $x$ gets a hit against a pitcher of "batting average" $y$. Such a function $\phi(x, y)$ is also needed in computer programs such as the commercial program called Micro-League Baseball, Micro League Sports Association, 1985. The programs play out games where real hitters and batters are represented by their actual statistics. The independence of $X$ and $Y$ arises from the assumption that the choice of batter and pitcher are independent of each other.

Because $X$ and $Y$ are independent, (3.1) is merely a condition on the (right-continuous) distribution functions $F$ and $G$ of $X$ and $Y$ (where $0 \le X, Y \le 1$). We will use the same symbols to denote the corresponding measures. Clearly, (3.1) with $Z = \phi(X, Y)$ now becomes

$$\int_0^1 \phi(x,y) G(dy) = x \qquad \text{a.e. } (F), \tag{3.2a}$$

$$\int_0^1 \phi(x,y) F(dx) = y \qquad \text{a.e } (G), \tag{3.2b}$$

$$0 \le \phi(x,y) \le 1, \qquad 0 \le x, y \le 1. \tag{3.2c}$$

THEOREM 4. *A solution* $\phi(x, y)$ *of* (3.2) *exists if and only if*

$$U(s) + V(t) \le K + [1 - F(s)][1 - G(t)] \qquad \text{for all } 0 \le s, t \le 1. \tag{3.3}$$

*Here,* $K = EX = EY = U(0-) = V(0-)$. *Further* $U$ *and* $V$ *are the right-continuous functions*

$$U(s) = \int_{s^+}^1 x F(dx) \le \min(K, 1 - F(s))$$

*and*

$$V(t) = \int_{t^+}^1 y G(dy) \le \min(K, 1 - G(t)).$$

PROOF. Consider $S = [0,1] \times [0,1]$ with product measure

$$d\lambda = F(dx) G(dy).$$

Condition (3.2) requires the existence of a function $0 \le \phi \le 1$ on $S$ such that $\phi\, d\lambda$ has the measures $\mu_1(dx) = x\,dF(x)$ and $\mu_2(dy) = yG(dy)$ on $[0,1]$ as its marginals. Since each of $\mu_1, \mu_2$ has mass $K$, it follows from Theorems 5.1 and 2.2 of Kellerer [6] (see also [16], page 437, and [13], pages 541–547) that such $\phi$ exists if and only if $\mu_1(A) \le \mu_2(B^C) + \lambda(A \times B)$ for all $A, B \subset [0,1]$ (the necessity is obvious since $A \times [0,1]$ is contained in the union of $[0,1] \times B^C$ and $A \times B$). Here, $\lambda(A \times B) = F(A)G(B)$ and $\mu_2(B^C) = K - \mu_2(B)$, leading to the condition that

$$\mu_1(A) + \mu_2(B) \le K + F(A)G(B) \quad \text{for all } A, B \subset [0,1]. \tag{3.4}$$

Given $G(B) = t$, one may as well maximize $\mu_1(A) - tF(A) = \int_A (x - t)\, dF(x)$ leading to $A = (t, 1]$. Similarly, given $F(A) = s$, one may as well choose $B = (s, 1]$. Finally, note that condition (3.4) with $A = (s, 1]$ and $B = (t, 1]$ is precisely condition (3.3). □

EXAMPLE 1. Let $X$ and $Y$ be discrete with $P(X = a_i) = p_i$, $i = 1, \dots, M$, and $P(Y = b_j) = q_j$, $j = 1, \dots, N$. Here, $\Sigma p_i = \Sigma q_j = 1$, $0 \le a_1 < \cdots < a_M \le 1$ and $0 \le b_1 < \cdots < b_N \le 1$. Moreover, $\Sigma_i p_i a_i = \Sigma_j q_j b_j = K$ (say). One easily verifies that now condition (3.3) takes the form

$$\sum_{i=m}^{M} p_i a_i + \sum_{j=n}^{N} q_j b_j \le K + \left(\sum_{i=m}^{M} p_i\right)\left(\sum_{j=n}^{N} q_j\right), \tag{3.5}$$

whenever $2 \le m \le M$ and $2 \le n \le N$. In view of the last part of the above proof, one only needs to verify (3.5) for the pairs $(m, n)$ which satisfy

$$a_{m-1} \le \sum_{j=n}^{N} q_j < a_m \quad \text{and} \quad b_{n-1} \le \sum_{i=m}^{M} p_i < b_n.$$

In the special case $M = N = 2$ there is only the single condition $p_2 a_2 + q_2 b_2 \le K + p_2 q_2$. It is trivially satisfied unless $a_1 \le q_2 < a_2$ and $b_1 \le p_2 < b_2$. In the special case $F = G$, that is, $b_i = a_i$ and $q_i = p_i$, $i = 1, 2$, one is led to the condition:

Either $a_1 \ge \alpha^2$ or else $|p - \alpha| \ge \sqrt{\alpha^2 - a_1}$, where $p = p_2$ and $\alpha = (a_1 + a_2)/2$.

Thus $\phi$ would for instance not exist when $a_1 < \alpha^2$ and $p = \alpha$.

PROPOSITION. *Suppose that* (3.2) *has a solution* $\phi$; *equivalently*, (3.3) *holds. Suppose further that F and G are continuous. Then there even exists a solution* $\phi(x, y)$ *of* (3.2) *which satisfies*

$$\phi(x, y) \in \{0, 1\} \quad \textit{for all } 0 \le x, y \le 1. \tag{3.6}$$

PROOF. Using the notation of the proof of Theorem 4, this assertion follows from the fact that presently the marginals $\mu_1(dx) = x\,dF(x)$ and $\mu_2(dy) = yG(dy)$ on $[0, 1]$ are nonatomic, namely, by applying the result of Kellerer [7] mentioned in the paragraph following (2.4). In the special case that $F$ and $G$ are absolutely continuous, the result also follows from Theorem 2 since then $d\lambda = F(dx)G(dy)$ is absolutely continuous relative to two-dimensional Lebesgue measure. □

REMARK. The corollary breaks down if one of $F, G$ is allowed to be discontinuous. Suppose for instance that $F$ is continuous but that $G = \delta_b$, where $b = EX = \int xF(dx)$. Then only the values $\phi(x, b)$ are relevant and, in fact, (3.2) requires that $\phi(x, b) = x$ a.e. $(F)$. It follows that (3.6) fails unless $F$ is supported on $\{0, 1\}$.

EXAMPLE 2. Suppose $F = G =$ uniform on $[0, 1]$. Then (3.3) becomes $\frac{1}{2}(1 - s^2) + \frac{1}{2}(1 - t_2) \le \frac{1}{2} + (1 - s)(1 - t)$, which is true. In this case, we can take $\phi(x, y) = 1$ if $x + y > 1$ and $0$ if $\phi(x, y) > 1$. In view of the latter corollary, the existence of a 0, 1-valued $\phi$ comes as no surprise.

In the present example, the solution $\phi(x,y) = 1_{\{x+y>1\}}$ of (3.2) is even unique. For, suppose $\psi$ were another solution of (3.2). Then

$$0 = \int_0^1 \int_0^1 (\psi(x,y) - \phi(x,y))(x + y - 1)\, dx\, dy.$$

But since $0 \le \psi \le 1$, the latter integrand is a.e. nonpositive implying that $\psi = \phi$ a.e. After all, $\psi(x,y) > \phi(x,y)$ implies $\phi(x,y) = 0$ thus $x + y - 1 < 0$. Similarly, $\psi(x,y) < \phi(x,y)$ implies $\phi(x,y) = 1$ and thus $x + y - 1 > 0$.

The same type of reasoning applies to any 0, 1-function $\phi = 1_A$ for which $A$ is an additive set. In the present setting, a subset $A$ of $S = [0,1]^2$ is said to be additive (see [3]) if there exist functions $a(x), b(y)$, which are integrable relative to the marginals $F(dx)$ and $G(dy)$ of $d\lambda = F(dx)G(dy)$, such that $A = \{(x,y) \in S\colon a(x) + b(y) \ge 0\}$. If so there exists no other $0 \le \psi \le 1$ on $S$ such that $\psi\, d\lambda$ has the same marginals as $\phi\, d\lambda = 1_A\, d\lambda$.

The baseball origin of the problem motivates two additional conditions that we might ask $\phi$ to satisfy

$$\min(x,y) \le \phi(x,y) \le \max(x,y), \tag{3.7}$$

$$\phi(x,y)\uparrow \quad \text{in } x \text{ and } y. \tag{3.8}$$

The first of these contends, for example, that it would be strange for a .300 hitter to be transformed into a .350 hitter when he faces a (tough) .250 pitcher! The second argues that a hitter improves against inferior pitching (and vice versa). But note that Example 2, and the subsequent remarks, show that (3.7) cannot always be satisfied.

EXAMPLE 3. Let $F = G = \frac{1}{2}(\delta_{1/3} + \delta_{2/3})$. To satisfy (3.2), take (for instance) $\phi(\frac{1}{3}, \frac{1}{3}) = \phi(\frac{1}{3}, \frac{2}{3}) = \phi(\frac{2}{3}, \frac{1}{3}) = \frac{1}{3}$ and $\phi(\frac{2}{3}, \frac{2}{3}) = 1$. Note that $\phi$ satisfies (3.8). It is easy to check, however, that for this $F$ and $G$ there is no $\phi'$ satisfying (3.2) and (3.6). Nor does there exist any $\phi'$ satisfying (3.2) and (3.7).

Returning to Example 2 (with $F$ and $G$ uniform), we found $\phi$ which satisfies (3.2), (3.6) and (3.8) simultaneously. Even for general $F$ and $G$, such a $\phi = 1_A$ if it exists must of course be of the form

$$\phi(x,y) = 1 \quad \text{if and only if } y > b(x) \ [\text{or } y \ge b(x)],$$

where $b(x)$ is nonincreasing. Thus $A$ would be additive and $\phi = 1_A$ is then the only solution of (3.2). Using Theorem 4, it is easy to show the following result.

PROPOSITION. *Suppose $F$ and $G$ are continuous. Then there exists $\phi$ satisfying* (3.2), (3.6) *and* (3.8) *if and only if*

$$G(1 - F(x)) = 1 - x, \qquad 0 \le x \le 1;$$

*or, equivalently, $\overline{G}(\overline{F}(x)) \equiv x$, where $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$.*

EXAMPLE 4. Let $F = G$, where $F(x) = x^2$. It is easy to see that (3.3) holds and the hypothesis of the last proposition fails. Thus here we have a case where there is a $\phi$ satisfying (3.2) and (3.6), while there is no $\phi$ satisfying (3.2), (3.6) and (3.8).

THEOREM 5. *If there is a* $\phi$ *satisfying* (3.2), *then there is one satisfying* (3.2) *and* (3.8).

REMARK. Consider Example 4, where (3.2) does have solutions $\phi$. By Theorem 5, there always exists a solution $\phi$ which is increasing (i.e., nondecreasing). By the above corollary, there also exists a solution satisfying $\phi \in \{0, 1\}$. On the other hand, as we have seen, there is no solution $\phi$ satisfying both conditions. The proof of Theorem 5 will be based on the following result, which relates to [14], Chapter 6.

THEOREM 6. *Let* $(\phi_{ij})$ *be an* $m \times n$ *matrix with* $0 \le \phi_{ij} \le 1$ *having nondecreasing row sums and nondecreasing column sums. In other words,*

$$(3.9)\quad \xi_i = \sum_{j=1}^{n} \phi_{ij}, \qquad i = 1, \dots, m; \qquad \eta_j = \sum_{i=1}^{m} \phi_{ij}, \qquad j = 1, \dots, n,$$

*satisfy* $\xi_i \le \xi_{i+1}$, $i = 1, \dots, m-1$, *and* $\eta_j \le \eta_{j+1}$, $j = 1, \dots, n-1$. *Then there exists another* $m \times n$ *matrix* $(\psi_{ij})$ *with* $0 \le \psi_{ij} \le 1$, *which has exactly the same row sums and column sums as* $(\phi_{ij})$, *such that* $\psi_{ij}$ *is nondecreasing in both* $i$ *and* $j$. *Thus*

$$(3.10)\qquad \psi_{ij} \le \psi_{i+1}, j \quad \text{and} \quad \psi_{ij} \le \psi_{i, j+1}.$$

PROOF. Let $(\psi_{ij})$ be the (unique) $m \times n$ matrix with $0 \le \psi_{ij} \le 1$, having the same row sums $\xi_i$ and column sums $\eta_j$ as $(\phi_{ij})$, and such that $\sum_{i,j}(\psi_{ij})^2$ is as small as possible. It suffices to show that $(\psi_{ij})$ satisfies (3.10). Suppose instead that, for some $(i, j)$,

$$(3.11)\qquad 0 \le \psi_{i+1, j} < \psi_{i, j} \le 1.$$

Since $\sum_k \psi_{ik} = \xi_i \le \xi_{i+1} = \sum_k \psi_{i+1, k}$, there exists $1 \le k \le n$ for which

$$(3.12)\qquad 0 \le \psi_{i, k} < \psi_{i+1, k} \le 1.$$

But for small enough $\varepsilon > 0$, if we increase $\psi_{i+1, j}$ and $\psi_{ik}$ by $\varepsilon$, and decrease $\psi_{ij}$ and $\psi_{i+1, k}$ by $\varepsilon$, then we arrive at a new matrix $(\Psi'_{ij})$ with $0 \le \Psi'_{ij} \le 1$ with the same row and column sums such that $\sum_{i,j}(\Psi'_{ij})^2 < \sum_{i,j}(\psi_{ij})^2$, contradicting the minimality of $(\psi_{ij})$. □

The same proof would go through if instead we had minimized $\sum_{i,j} f(\psi_{i,j})$ with $f$ as a fixed strictly convex continuous function on $[0, 1]$, such as $f(u) = u \log u + (1 - u)\log(1 - u)$. In order to prove Theorem 5, we need the following continuous version of Theorem 6.

THEOREM 7. *Let* $\phi = \phi(x, y)$ *be measurable on* $[0, 1] \times [0, 1]$ *and such that*

$$(3.13)\qquad 0 \le \phi(x, y) \le 1,$$

$$(3.14)\qquad \xi(x) = \int_0^1 \phi(x, y)\, dy \text{ is nondecreasing in } x,$$

$$(3.15)\qquad \eta(y) = \int_0^1 \phi(x, y)\, dx \text{ is nondecreasing in } y.$$

*Then there exists* $\psi(x, y)$ *measurable on* $[0,1] \times [0,1]$ *satisfying* $0 \le \psi(x, y) \le 1$, *having the same marginals as* $\phi$, *that is,*

$$\int_0^1 \psi(x, y)\, dy = \xi(x) \quad a.e., \tag{3.16}$$

$$\int_0^1 \psi(x, y)\, dx = \eta(y) \quad a.e. \tag{3.17}$$

*and such that*

$$\psi(x, y) \uparrow \text{ in } x \text{ and } y \text{ separately}. \tag{3.18}$$

PROOF. We will prove Theorem 7 by applying Theorem 6 to a certain discrete approximation. Let $A(n, i)$ denote the half-open interval

$$A(n, i) = [(i-1)2^{-n}, i2^{-n}), \qquad n = 1, 2, \dots,\ i = 1, 2, \dots, 2^n.$$

For fixed $n$, the $A(n, i)$ define a partition of $[0, 1)$ into $2^n$ disjoint intervals each of length $2^{-n}$. Let $\mathscr{F}_n$ denote the finite field consisting of all possible unions of the $A(n, i)$. Note that $\mathscr{F}_n \subset \mathscr{F}_{n+1}$. Also let $E(n, i, j) = A(n, i) \times A(n, j)$; thus the $E(n, i, j)$ define a partition of $S = [0, 1)^2$ into disjoint half-open squares of area $2^{-2n}$. Further consider the averages

$$\phi_{ij}(n) = 2^{2n} \int_{E(n,i,j)} \phi(x, y)\, dx\, dy$$

and

$$\xi_i(n) = 2^n \int_{A(n,i)} \xi(x)\, dx, \qquad \eta_j(n) = 2^n \int_{A(n,j)} \eta(y)\, dy,$$

$i, j = 1, \dots, 2^n$. Since $0 \le \phi \le 1$ one has $0 \le \phi_{ij}(n) \le 1$. From (3.14) and (3.15), the $\xi_i(n)2^n$ and $\eta_j(n)2^n$ are precisely the row sums and column sums of the matrix $(\phi_{ij}(n))$. Also note that $\xi_i(n)$ and $\eta_j(n)$ are nondecreasing in $i$ and $j$, respectively. From Theorem 6, there exists a matrix $(\psi_{ij}(n))$ of the same size with $0 \le \psi_{ij}(n) \le 1$, having these same row sums $\xi_i(n)2^n$ and column sums $\eta_j(n)2^n$, with the additional property that $\psi_{ij}(n)$ is nondecreasing in both $i$ and $j$. Now define $\psi_n$: $S \to [0, 1]$ by letting $\psi_n(x, y) = \psi_{ij}(n)$ for all $(x, y) \in E(n, i, j)$. Clearly, $\psi_n(x, y)$ is nondecreasing in each coordinate and satisfies $0 \le \psi_n \le 1$. Let

$$u_n(x) = \int_0^1 \psi_n(x, y)\, dy \quad \text{and} \quad v_n(y) = \int_0^1 \psi_n(x, y)\, dy.$$

Then

$$\int_{A(n,i)} \xi(x)\, dx = 2^{-n}\xi_i(n) = 2^{-2n} \sum_j \phi_{ij}(n)$$

$$= \sum_j \int_{E(n,i,j)} \phi(x, y)\, dx\, dy = \int_{A(n,i)} u_n(x)\, dx.$$

Thus $u_n(x) - \xi(x)$ integrates to 0 over every $A \in \mathscr{F}_n$. Equivalently, $\psi_n(x, y) -$

$\xi(x)$ integrates to 0 over every set $A \times [0,1]$ with $A \in \mathscr{F}_n$. Similarly, $\psi_n(x,y) - \eta(y)$ integrates to 0 over every set $[0,1] \times B$ with $B \in \mathscr{F}_n$.

Now draw a weak* convergent subsequence from the sequence $\{\psi_n\}$ with limit $\psi$. Then $\psi$ has all the required properties. It is obvious that $0 \le \psi \le 1$. Moreover, $\psi(x,y)$ is nondecreasing in each coordinate. For that property can be expressed as $\int_C \psi \le \int_D \psi$ with $C$ and $D$ as arbitrary disjoint subsets of $S = [0,1]^2$ of equal area, such that $D$ is northeast of $C$. Moreover, $\psi(x,y) - \xi(x)$ integrates to 0 over every set $A \times [0,1]$ with $A \in \mathscr{F}$, where $\mathscr{F} = \bigcup_n \mathscr{F}_n$. This in turn implies (3.16). Equality (3.17) follows similarly. □

To prove Theorem 5, note that (3.2) is equivalent to $0 \le \phi \le 1$ together with

$$\int_0^1 \phi(F^{-1}(x), G^{-1}(y))\,dy = F^{-1}(x), \qquad x \in C_x, \tag{3.19a}$$

$$\int_0^1 \phi(F^{-1}(x), G^{-1}(y))\,dx = G^{-1}(y), \qquad y \in C_y, \tag{3.19b}$$

where $C_x$ and $C_y$ are the continuity points of $F^{-1}$ and $G^{-1}$. By Theorem 7, there exists a function $0 \le \psi(x,y) \le 1$ which is nondecreasing in $x$ and $y$ and has its marginals equal to $\xi(x) = F^{-1}(x)$ and $\eta(y) = G^{-1}(y)$. Thus (3.19) holds with $\psi$ replaced by

$$\phi(x,y) = \psi(F(x), G(y)). \tag{3.20}$$

In fact, the latter function $0 \le \phi \le 1$ has all the properties promised in Theorem 5.

Finally, we remark that the above proof of Theorems 6 and 7 immediately carries over to dimension $k \ge 2$, but only for one-dimensional marginals. Thus, given an integrable function $0 \le \phi \le 1$ on $R^k$ having nondecreasing one-dimensional marginals, there exists a nondecreasing function $0 \le \psi \le 1$ on $R^k$ having exactly the same marginals.

That, for instance, Theorem 6 does not carry over to two-dimensional marginals can be seen by choosing $k = 3$ and $\phi_{i,j,k} = \beta(i+j+k)$ with $i,j,k \in \{0,1\}$. Its two-dimensional marginals are nondecreasing as soon as $\beta(0) \le \beta(2)$ and $\beta(1) \le \beta(3)$. On the other hand, choosing $\beta(1) = 0$ and $\beta(j) = 1$ otherwise, the resulting $\phi$ fails to be nondecreasing and, moreover, there is no other $0 \le \psi \le 1$ having the same two-dimensional marginals as $\phi$.

**Acknowledgments.** We conjectured Theorem 2 only after a discussion with Walter Carrington of related results in his thesis [1]. We are grateful to Hans Kellerer for kindly correcting a very serious error in our earlier submitted version and to Benjy Weiss for the interesting remark on the proof of Theorem 3 and which is our first remark in the paper.

## REFERENCES

[1] CARRINGTON, W. A. (1982). Moment problems and ill-posed operator equations with convex constraints. Ph.D. dissertation, Washington Univ.

[2] DUNFORD, N. and SCHWARTZ, J. T. (1963). *Linear Operators*. Interscience, New York.

[3] FISHBURN, P. C., LAGARIAS, J. C., REEDS, J. A. and SHEPP, L. A. (1990). Sets uniquely determined by projections on axes. I. Continuous case. *SIAM J. Appl. Math.* **50** 288–306.

[4] JACOBS, K. (1978). *Measure and Integral* 534–548. Academic, New York.

[5] KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems*. Interscience, New York.

[6] KELLERER, H. G. (1961). Funktionen und Produkträumen mit vorgegeben Marginal-Funktionen. *Math. Ann.* **144** 323–344.

[7] KELLERER, H. G. (1962). Die Schnittmasz-Funktionen meszbarer Teilmengen eines Produktraumes. *Math. Ann.* **146** 103–128.

[8] KELLERER, H. G. (1964). Masztheoretische Marginal Problemen. *Math. Ann.* **153** 168–198.

[9] KELLERER, H. G. (1964). Schnittmasz-Funktionen in mehrfachen Produkträmen. *Math. Ann.* **155** 369–391.

[10] KINGMAN, J. F. C. and ROBERTSON, A. P. (1968). On a theorem of Lyapounov. *J. London Math. Soc.* **43** 347–351.

[11] LINNIK, J. V. (1968). *Statistical Problems with Nuisance Parameters*, *Transl. Math. Monographs* **20** 197–203. Amer. Math. Soc., Providence, R.I.

[12] LORENTZ, G. G. (1949). A problem of plane measure. *Amer. J. Math.* **71** 417–426.

[13] PHELPS, R. R. (1966). *Lectures on Choquet's Theorem*. Van Nostrand, New York.

[14] RYSER, H. J. (1963). *Combinatorial Mathematics*. Wiley, New York.

[15] SHEPP, L. A. and KRUSKAL, J. B. (1978). Computerized tomography: The new medical X-ray technology. *Amer. Math. Monthly* **85** 420–439.

[16] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

SAM GUTMANN
NORTHEASTERN UNIVERSITY
BOSTON, MASSACHUSETTS 02115

J. H. B. KEMPERMANN
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903

J. A. REEDS
AT & T BELL LABORATORIES
MURRAY HILL, NEW JERSEY 07974

L. A. SHEPP
AT & T BELL LABORATORIES
MURRAY HILL, NEW JERSEY 07974