JOURNAL OF MULTIVARIATE ANALYSIS 23, 233–256 (1987)

Multivariate Liouville Distributions

RAMESHWAR D. GUPTA*

University of New Brunswick, St. John, New Brunswick, Canada E2L 4L5

AND

DONALD ST. P. RICHARDS†

University of North Carolina

Communicated by P. R. Krisnaiah

A random vector $(X_1, ..., X_n)$, with positive components, has a Liouville distribution if its joint probability density function is of the form $f(x_1 + \cdots + x_n)$ $x_1^{a_1-1} \cdots x_n^{a_n-1}$ with the a_i all positive. Examples of these are the Dirichlet and inverted Dirichlet distributions. In this paper, a comprehensive treatment of the Liouville distributions is provided. The results pertain to stochastic representations, transformation properties, complete neutrality, marginal and conditional distributions, regression functions, and total positivity and reverse rule properties. Further, these topics are utilized in various characterizations of the Dirichlet and inverted Dirichlet distributions. Matrix analogs of the Liouville distributions are also treated, and many of the results obtained in the vector setting are extended appropriately. © 1987 Academic Press, Inc.

1. *Introduction*

Recently, increasing attention has been paid to families of probability distributions which are defined through functional form assumptions, both on density functions (cf. Johnson and Kotz [13, p. 294 ff]), and on characteristic functions (cf. Cambanis *et al.* [1, 2]; Fang and Fang [7]; Richards [29]). These studies develop unified treatments of all distributions in some

Received November 15, 1984; revised August 1, 1986.

AMS 1980 subject classifications: Primary 62H10; Secondary 62E15, 28A01.

Key words and phrases: Dirichlet distribution, regression, total positivity, reverse rule, stochastic representation, characterizations, fractional integral.

* Partially supported by the National Scientific and Engineering Research Council, grant A-4850.

† Partially supported by the National Science Foundation (grant MCS-8403381) and by the Research Council of the University of North Carolina. Current address: Department of Mathematics, University of Virginia, Charlottesville, VA 22903.

0047-259X/87 $3.00

particular class, thereby developing their common features. For example, the papers [1, 3] prove that certain important sampling properties of the multivariate normal distributions remain valid for the larger class of elliptically contoured distributions.

In this paper, we study another family, the Liouville distributions, defined through functional form restrictions on the density functions. An (absolutely continuous) random vector $(X_1, ..., X_n)$ has a (multivariate) Liouville distribution if its joint density function is proportional to

$$f\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} x_i^{a_i - 1}, \tag{1.1}$$

where the variables range over the generalized octant $R_+^n = \{(x_1, ..., x_n): x_i > 0, i = 1, ..., n\}$; $a_1, ..., a_n$ are positive numbers; and the function $f(\cdot)$ is positive, continuous, and satisfies the integrability condition (2.1) (with $m = 1$) below. If $f(\cdot)$ has noncompact support, we say that $(X_1, ..., X_n)$ has a Liouville distribution of the first kind. If the support of $f(\cdot)$ is compact then we may assume, after a scaling, that $f(\cdot)$ is suported on (0, 1); then the variables range over the simplex $\mathscr{S}_n = \{(x_1, ..., x_n): x_i > 0, i = 1, ..., n; \sum_{i=1}^{n} x_i < 1\}$, and we say that $(X_1, ..., X_n)$ has a Liouville distribution of the second kind.

Brief treatments of the Liouville distributions appear in [22; p. 308], where they are related with Schur-convex functions; and also in [31], where some results on marginal distributions and transformation properties are presented.

1.1. EXAMPLE. (correlated gamma variables [22; p. 308]). If $f(t) = t^{a-1} e^{-bt}$, $t > 0$, $a > 0$, $b > 0$, then $(X_1, ..., X_n)$ has a Liouville distribution of the first kind, with joint density function proportional to

$$\left(\sum_{i=1}^{n} x_i\right)^{a-1} \prod_{i=1}^{n} x_i^{a_i - 1} e^{-bx_i}. \tag{1.2}$$

1.2. EXAMPLE (the Dirichlet distribution). If $f(t) = (1-t)^{a_{n+1}-1}$, $0 < t < 1$, $a_{n+1} > 0$, then $(X_1, ..., X_n)$ has a Liouville distribution of the second kind with joint density proportional to

$$\left(1 - \sum_{i=1}^{n} x_i\right)^{a_{n+1}-1} \prod_{i=1}^{n} x_i^{a_i - 1}. \tag{1.3}$$

1.3. EXAMPLE (the inverted Dirichlet distribution). If $f(t) =$

$(1+t)^{-(a_1+\cdots+a_{n+1})}$, $t>0$, $a_{n+1}>0$, then we have a Liouville distribution of the first kind, with joint density function proportional to

$$\left(1+\sum_{i=1}^{n} x_i\right)^{-(a_1+\cdots+a_{n+1})} \prod_{i=1}^{n} x_i^{a_i-1}. \tag{1.4}$$

In this paper, we provide a comprehensive treatment of the Liouville distributions, extending the range of topics considered in [22, 31]. Our results pertain to stochastic representations and transformation properties (Section 3); marginal and conditional distributions, and regression functions (Section 4); total positivity, reverse rule, and dependence properties (Section 5); and characterizations based on the above topics (Section 6). Examples 1.1–1.3 provide motivation for our work, consistently pointing towards new results. For instance, we extend the results of [16] on the multivariate reverse rule properties of the Dirichlet distributions to a large subset of the Liouville distributions of the second kind. In return, we characterize the three examples using independence, regression, and neutrality properties.

We also consider matrix analogs of the Liouville distributions. The positive definite (symmetric) $m\times m$ matrices $A_1, ..., A_n$ are said to have a Liouville distribution of the first kind if their (continuous) joint density function exists and is proportional to

$$f\left(\sum_{i=1}^{n} A_i\right) \prod_{i=1}^{n} |A_i|^{a_i-p}. \tag{1.5}$$

Here, $|A|$ denotes the determinant of A; $f(\cdot)$ is positive, continuous, supported on all of $R_+^{m\times m}$ (the cone of positive definite $m\times m$ matrices), and satisfies (2.1) below; $p=\frac{1}{2}(m+1)$ and $a_i>p-1$, $i=1, ..., n$. Further, the matrix Liouville distributions of the second kind are those for which $I-\sum_{i=1}^{n} A_i$ is also positive definite, where I denotes the $m\times m$ identity matrix.

1.4. Example. Let $S_0, ..., S_n$ be independent $m\times m$ Wishart matrices having common covariance matrix Σ, and degrees of freedom $a_0, ..., a_n$, respectively, where $a_i>p-1$, $i=0, ..., n$. Let $S^{1/2}$ denote the unique, positive definite square root of $S=\sum_{i=0}^{n} S_i$. Then Olkin and Rubin [24] (cf. Khatri [18], Mitra [23]) show that the matrices $A_i=S^{-1/2}S_iS^{-1/2}$, $i=1, ..., n$ have the joint density which is proportional to

$$\left|I-\sum_{i=1}^{n} A_i\right|^{(1/2)a_0-p} \prod_{i=1}^{n} |A_i|^{(1/2)a_i-p} \tag{1.6}$$

for $A_1, \ldots, A_n$, $I-\sum_{i=1}^{n} A_i$ all positive definite. This is a Liouville distribution of the second kind, generalizing Example 1.2.

1.5. EXAMPLE. Let $S_0, \ldots, S_n$ be defined as above. It is shown in [24] that the matrices $A_i = S_0^{-1/2} S_i S_0^{-1/2}$, $i=1, \ldots, n$, have the joint density which is proportional to

$$\left| I + \sum_{i=1}^{n} A_i \right|^{-(1/2)(a_0 + \cdots + a_n)} \prod_{i=1}^{n} |A_i|^{(1/2)a_i - p} \qquad (1.7)$$

This is a Liouville distribution of the first kind, generalizing Example 1.3.

Most of the results obtained for Liouville vector distributions will be extended to their matrix counterparts. However, it appears that the completion of this program will require highly sophisticated techniques (see the concluding remarks below Section 9). Section 7 contains stochastic representations and transformations for the matrix distributions, extending Section 3; while Sections 8 and 9 are appropriate extensions of Sections 4 and 6, respectively.

2. *Preliminaries*

Throughout, all random entities are assumed to be absolutely continuous with continuous density functions. However, we note that some of our stochastic representations can be obtained in a density-free way using the methods of Khatri [18], Kumar [21], and Mitra [23]. We shall denote normalizing constants by the symbols c, c_0, c_1, c_j, etc., since their precise expressions are not needed.

We invariably write $(X_1, \ldots, X_n) \sim L_n[f(\cdot); a_1, \ldots, a_n]$ whenever $(X_1, \ldots, X_n)$ has a Liouville distribution. If necessary, we write $(X_1, \ldots, X_n) \sim L_n^{(k)}[f(\cdot); a_1, \ldots, a_n]$, $k=1$ or 2, according as the distribution is of the first or second kind. We denote the Dirichlet distributions (1.3) and (1.4) by $D(a_1, \ldots, a_n; a_{n+1})$ and $ID(a_1, \ldots, a_n; a_{n+1})$, respectively; however, if $n=1$, we use $B(a_1; a_2)$ and $IB(a_1; a_2)$ for the beta and inverted beta distributions. Precisely the same notation will be used for the matrix distributions, since the context always eliminates any possible confusion.

The nomenclature "Liouville distribution" is adopted from Liouville's extension of Dirichlet's integral [6; p. 160]. On the space $R_+^{m \times m}$, this integral is as follows: if $p = \frac{1}{2}(m+1)$; $a_i > p-1$ for $i=1, \ldots, n$; $a = a_1 + \cdots + a_n$; and

$$\int_{R_+^{m \times m}} |T|^{a-p} f(T)\, dT < \infty \qquad (2.1)$$

then

$$\int_{R_+^{m\times m}} \cdots \int_{R_+^{m\times m}} f\left(\sum_{i=1}^{n} A_i\right) \prod_{i=1}^{n} |A_i|^{a_i - p} \, dA_i$$
$$= \frac{\prod_{i=1}^{n} \Gamma_m(a_i)}{\Gamma_m(a)} \int_{R_+^{m\times m}} |T|^{a-p} f(T) \, dT. \tag{2.2}$$

In (2.1) and (2.2), dA_i is the Lebesgue measure on $R_+^{m\times m}$ and $\Gamma_m(\cdot)$ is the multidimensional gamma function [11].

If a continuous function $f: R_+^{m\times m} \to R$ satisfies (2.1) and $\alpha > p-1$, then the Weyl fractional integral of order α of $f(\cdot)$ is

$$W^\alpha f(T) = \frac{1}{\Gamma_m(\alpha)} \int_{S>T} |S-T|^{\alpha-p} f(S) \, dS, \tag{2.3}$$

where "$S > T$" means that $S-T$ is positive definite. Detailed properties of W^α are available from Rooney [31] when $m=1$ and for general m from Richards [28]. The main properties needed are (i) if a continuous function $f(\cdot)$ satisfies (2.1) then there is a one–one correspondence between $f(\cdot)$ and its "fractional derivative" $W^\alpha f(\cdot)$; (ii) W^α satisfies the semigroup property $W^{\alpha+\beta} = W^\alpha W^\beta$, $\alpha > p-1$, $\beta > p-1$.

In Section 9, we shall apply the theorem of Deny [4]. Let G be a locally compact Abelian group with a countable basis, and ν be a Radon measure on the Borel subsets of G such that the smallest closed subgroup of G generated by the support of ν is G itself. Further, let $\hat{G}$ be the set of all continuous functions $\phi: G \to R_+$ for which $\phi(s+t) = \phi(s)\,\phi(t)$ for all s, t in G, and

$$\hat{G}(\nu) = \left\{ \phi \in \hat{G} : \int_G \phi(t) \, d\nu(t) = 1 \right\}.$$

With the topology of uniform convergence on compact subsets of G, $\hat{G}$ is a locally compact space and $\hat{G}(\nu)$ is a Borel subset of $\hat{G}$.

2.1. THEOREM (Deny [4]). *Let $f: G \to R_+$ be continuous and satisfy*

$$f(s) = \int_G f(s+t) \, d\nu(t), \qquad s \in G. \tag{2.4}$$

Then there exists a unique positive measure μ on $\hat{G}(\nu)$ such that

$$f(s) = \int_{\hat{G}(\nu)} \phi(s) \, d\mu(\phi). \tag{2.5}$$

Other applications of Deny's theorem are given by Rao [26] and Richards [27].

I. LIOUVILLE DISTRIBUTIONS ON R^n_+

3. *Transformations and Stochastic Representations*

We begin by establishing a one–one correspondence between the two kinds of Liouville distributions.

3.1. PROPOSITION. *Suppose that* $(Y_1, ..., Y_n) \sim L_n^{(2)}[g(\cdot); a_1, ..., a_n]$. *Define*

$$X_i = Y_i \bigg/ \left(1 - \sum_{j=1}^{n} Y_j\right), \qquad i = 1, ..., n. \tag{3.1}$$

Then $(X_1, ..., X_n) \sim L_n^{(1)}[f(\cdot); a_1, ..., a_n]$, *where*

$$f(t) = (1+t)^{-(a_1 + \cdots + a_n + 1)} g(t/(1+t)), \qquad t > 0. \tag{3.2}$$

In particular, the correspondence between $f(\cdot)$ *and* $g(\cdot)$ *is one–one.*

Proof. Solving Eq. (3.1), we obtain

$$Y_i = X_i \bigg/ \left(1 + \sum_{j=1}^{n} X_j\right), \qquad i = 1, ..., n, \tag{3.3}$$

and it is straightforward to show that the Jacobian of the transformation (3.3) is $(1 + \sum_{i=1}^{n} x_i)^{-(n+1)}$. Then, the result follows by routinely applying (3.3) to the density function of $(Y_1, ..., Y_n)$. It is clear that the converse result is also valid.

3.2. THEOREM. *Let* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$. *Then the following stochastic representations are valid*:

(i) $(X_1, ..., X_n) \stackrel{\mathscr{L}}{=} (Y_1, ..., Y_{n-1}, 1 - \sum_{i=1}^{n-1} Y_i) Y_n$, *where* $(Y_1, ..., Y_{n-1})$ *and* Y_n *are mutually independent, and* $(Y_1, ..., Y_{n-1}) \sim D(a_1, ..., a_{n-1}; a_n)$;

(ii) $(X_1, ..., X_n) \stackrel{\mathscr{L}}{=} (\prod_{i=1}^{n-1} Y_i, (1 - Y_1) \prod_{i=2}^{n-1} Y_i, ..., 1 - Y_{n-1}) Y_n$, *where* $Y_1, ..., Y_n$ *are mutually independent, and* $Y_i \sim B(\sum_{j=1}^{i} a_j; a_{i+1})$, $i = 1, ..., n-1$;

(iii) $(X_1, ..., X_n) \stackrel{\mathscr{L}}{=} (\prod_{i=1}^{n-1} (1 + Y_i)^{-1}, Y_1 \prod_{i=2}^{n-1} (1 + Y_i)^{-1}, ..., Y_{n-1}(1 + Y_{n-1})^{-1}) Y_n$, *where* $Y_1, ..., Y_n$ *are mutually independent, and* $Y_i \sim IB(a_{i+1}; \sum_{j=1}^{i} a_j)$, $i = 1, ..., n-1$;

(iv) $(X_1, ..., X_n) =^{\mathscr{L}} (Y_1, Y_2(1-Y_1), ..., Y_{n-1} \prod_{i=1}^{n-2}(1-Y_i), \prod_{i=1}^{n-1}(1-Y_i)) Y_n$, *where* $Y_1, ..., Y_n$ *are mutually independent, and* $Y_i \sim B(a_i; \sum_{j=i+1}^{n} a_j)$, $i=1, ..., n-1$.

In all four cases, $Y_n =^{\mathscr{L}} \sum_{i=1}^{n} X_i \sim L_1[f(\cdot); \sum_{i=1}^{n} a_i]$.

Before proving the theorem, we review two implications. With fixed $a_1, ..., a_n$, the function $f(\cdot)$ and hence the distribution of $(X_1, ..., X_n)$ is uniquely determined by the distribution of Y_n. As an example, Y_n can have a gamma distribution only if $(X_1, ..., X_n)$ has the distribution given in Example 1.1. Next, for an application of the theorem, suppose that $(X_1, ..., X_n) \sim D(a_1, ..., a_n; a_{n+1})$; i.e., $f(t)=(1-t)^{a_{n+1}-1}$, $0<t<1$. Then, in Theorem 3.2(ii), the mutually independent variables $Y_i \sim B(a_1+\cdots+a_i; a_{i+1})$, $i=1, ..., n$, and $\prod_{i=1}^{n} Y_i =^{\mathscr{L}} X_1$. Then (by Proposition 4.1 below) $\prod_{i=1}^{n} Y_i =^{\mathscr{L}} X_1 \sim B(a_1; a_2+\cdots+a_{n+1})$. This is a new proof of a result of Rao [25; p. 168] on certain products of independent beta variables.

3.3. *Proof of Theorem* 3.2. Each stochastic representation defines a transformation from $(X_1, ..., X_n)$ to $(Y_1, ..., Y_n)$. For (i), substitute $x_i = y_i y_n$ $(i=1, ..., n-1)$ and $x_n = (1-\sum_{i=1}^{n-1} y_i) y_n$ into (1.1); the corresponding Jacobian is y_n^{n-1}. After some routine algebra, the joint density of $(Y_1, ..., Y_n)$ is seen to be a multiple of

$$\left(\prod_{i=1}^{n-1} y_i^{a_i-1}\right)\left(1-\sum_{i=1}^{n-1} y_i\right)^{a_n-1} y_n^{a_1+\cdots+a_n-1} f(y_n)$$

which gives the desired result. Finally, the other parts are proven similarly; note that the Jacobians of the transformations in (ii), (iii), and (iv) are, respectively, $\prod_{i=2}^{n} y_i^{i-1}$, $y_n^{n-1} \prod_{i=1}^{n-1}(1+y_i)^{-(i+1)}$, and $y_n^{n-1} \prod_{i=1}^{n-1}(1-y_i)^{n-i-1}$.

3.4. Proposition. *If* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$, *then* $(\sum_{i=1}^{r} X_i)/(\sum_{i=1}^{n} X_i) \sim B(\sum_{i=1}^{r} a_i; \sum_{i=r+1}^{n} a_i)$, $r<n$.

Proof. In Theorem 3.2(i), it was shown that $(X_1, ..., X_n) =^{\mathscr{L}} (Y_1, ..., Y_{n-1}, 1-\sum_{i=1}^{n-1} Y_i) Y_n$, where $(Y_1, ..., Y_{n-1}) \sim D(a_1, ..., a_{n-1}; a_n)$ independently of $Y_n \sim L_1[f(\cdot); \sum_{i=1}^{n} a_i]$. Since $(Y_1, ..., Y_r) \sim D(a_1, ..., a_r; \sum_{i=r+1}^{n} a_i)$ (by Proposition 4.1 below), then $(\sum_{i=1}^{r} X_i)/(\sum_{i=1}^{n} X_i) =^{\mathscr{L}} \sum_{i=1}^{r} Y_i \sim B(\sum_{i=1}^{r} a_i; \sum_{i=r+1}^{n} a_i)$, $r<n$.

The special case when $a_i \equiv \frac{1}{2}$ can be alternatively obtained by first relating the Liouville and the spherically symmetric distributions [9] and then applying results from [17].

4. *Marginal and Conditional Distributions*

4.1. PROPOSITION. *If* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$ *then* $(X_1, ..., X_r) \sim L_r[f_r(\cdot); a_1, ..., a_r]$, $r<n$, *where* $a=\sum_{i=r+1}^{n} a_i$, *and* $f_r(t)=W^a f(t)$ *is the fractional integral of order* a *of* $f(\cdot)$.

Proof. By definition, the marginal density function of $(X_1, ..., X_r)$ is proportional to

$$\left(\prod_{i=1}^{r} x_i^{a_i-1}\right)\int_{R_+^{n-r}} f\left(\sum_{i=1}^{r} x_i+\sum_{i=r+1}^{n} x_i\right)\prod_{i=r+1}^{n} x_i^{a_i-1}\,dx_i. \tag{4.1}$$

Applying Liouville's integral (2.2) (with $m=1$) to the function $f(t+\sum_{i=1}^{r} x_i)$, we see that (4.1) is proportional to

$$\left(\prod_{i=1}^{r} x_i^{a_i-1}\right)\int_0^{\infty} f\left(t+\sum_{i=1}^{r} x_i\right)t^{a-1}\,dt. \tag{4.2}$$

From this, the result follows readily.

4.2 *Remark.* (i) In the case of the class $L_n^{(2)}$, the marginal distribution of $(X_1, ..., X_r)$ was previously derived in [31].

(ii) From the unicity of the fractional integral operators, it follows that for fixed $a_1, ..., a_n$, there is a one-one correspondence between $f(\cdot)$ and $f_r(\cdot)$. In the extreme case $r=1$, we find that the distribution of $(X_1, ..., X_n)$ is uniquely determined by the distribution of X_1.

(iii) A curious property of the class $L_n^{(2)}$ is that at most one of the univariate marginals can be uniformly distributed on (0, 1) (cf. [13, p. 305] for the case of the Dirichlet distributions). In proving this result, it suffices to assume $n=2$; thus, suppose that $(X_1, X_2) \sim L_2^{(2)}[f(\cdot); a_1, a_2]$, where $a_1 \geqslant a_2$ without loss of generality. Then, $X_1 \sim L_1^{(2)}[W^{a_2} f(\cdot); a_1]$ and $X_2 \sim L_1^{(2)}[W^{a_1} f(\cdot); a_2]$. If X_1 is uniformly distributed on (0, 1) then $t^{a_1-1} W^{a_2} f(t)=c$, equivalently, $W^{a_2} f(t)=ct^{-a_1+1}$, $0<t<1$. Then, the density of X_2 is proportional to

$$t^{a_2-1} W^{a_1} f(t)=t^{a_2-1} W^{a_1-a_2}(W^{a_2} f(t))=ct^{a_2-1} W^{a_1-a_2}(t^{-a_1+1})$$

which is not constant. That is, X_2 is not uniformly distributed on (0, 1).

4.3. COROLLARY. *If* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$ *then the conditional distribution of* $(X_{r+1}, ..., X_n)$, *given* $\{X_1=x_1, ..., X_r=x_r\}$, $r<n$, *is* $L_{n-r}[g_r(\cdot); a_{r+1}, ..., a_n]$, *where* $g_r(t)=f(t+\sum_{i=1}^{r} x_i)/f_r(\sum_{i=1}^{r} x_i)$.

The proof is straightforward from Proposition 4.1. Since the conditional distribution depends on $x_1, ..., x_r$ only through $\sum_{i=1}^{r} x_i$, we lose no generality by conditioning on $\{\sum_{i=1}^{r} X_i = y\}$, where $y = \sum_{i=1}^{r} x_i$. We shall use this result repeatedly in developing the multiple regression properties of the class L_n.

4.4. PROPOSITION. *If* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$, $r < n$, *and the expectation below exists, then*

$$\mathscr{E}\left(\prod_{i=r+1}^{n} X_i^{j_i} \,\middle|\, \sum_{i=1}^{r} X_i = t\right) = cW^{j+a} f(t) / W^{a} f(t), \tag{4.3}$$

where $a = \sum_{i=r+1}^{n} a_i$, $j = \sum_{i=r+1}^{n} j_i$. *In particular,* (4.3) *remains valid if* $\prod_{i=r+1}^{n} X_i^{j_i}$ *is replaced by* X_i^j, $r+1 \leqslant i \leqslant n$.

Proof. By Corollary 4.3,

$$\mathscr{E}\left(\prod_{i=r+1}^{n} X_i^{j_i} \,\middle|\, \sum_{i=1}^{r} X_i = t\right) f_r(t) = c_1 \int_{R_+^{n-r}} f\left(t + \sum_{i=r+1}^{n} x_i\right) \prod_{i=r+1}^{n} x_i^{j_i + a_i - 1}\, dx_i.$$

On applying Liouville's integral (2.2), we see that the last integral above equals

$$c_2 \int_0^{\infty} y^{j+a-1} f(y+t)\, dy = cW^{j+a} f(t).$$

Since $f_r(t) = W^a f(t)$, then we have proven (4.3). The corresponding result for $\mathscr{E}(X_i^j)$ is proven similarly.

4.5. EXAMPLE. If $(X_1, ..., X_n) \sim D(a_1, ..., a_n; a_{n+1})$ then it follows from (4.3) that $\mathscr{E}(\prod_{i=r+1}^{n} X_i^{j_i} \mid X_1 = x_1, ..., X_r = x_r)$ is proportional to $(1 - \sum_{i=1}^{r} x_i)^j$. See [13; p. 304] for a special case of this result.

4.6. PROPOSITION. *Under the same hypotheses assumed in Proposition* 4.4,

$$\mathscr{E}\left(h\left(\sum_{i=r+1}^{n} X_i\right) \,\middle|\, \sum_{i=1}^{r} X_i = t\right) f_r(t) = c \int_t^{\infty} (y-t)^{a-1} h(y-t) f(y)\, dy, \tag{4.4}$$

where $h(\cdot)$ *is a real-valued function for which the expectation exists.*

The proof is similar to the proof of Proposition 4.4. In particular, if $h(t) = t^j$ then we find that for $1 \leqslant r < k \leqslant n$,

$$\mathscr{E}\left(\left(\sum_{i=r+1}^{n} X_i\right)^j \,\middle|\, \sum_{i=1}^{r} X_i = t\right) = c_j \mathscr{E}\left(X_k^j \,\middle|\, \sum_{i=1}^{r} X_i = t\right). \tag{4.5}$$

5. *Total Positivity and Dependence Properties*

In this section, we use the results of Karlin [14] and Karlin and Rinott [15, 16] to obtain the totally positive and reverse rule properties of the Liouville distributions. Then, we make applications to derive some probability and correlation inequalities and positive dependence properties.

On R^n, introduce the lattice operations $\vee$ and $\wedge$: if $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ are in R^n, then

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), ..., \max(x_n, y_n)),$$

$$\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), ..., \min(x_n, y_n)).$$

5.1. DEFINITION [15]. A function $g: R^n \to R_+$ is *multivariate totally positive of order* 2 (MTP$_2$) if for all $\mathbf{x}$, $\mathbf{y}$ in R^n,

$$g(\mathbf{x})\, g(\mathbf{y}) \leqslant g(\mathbf{x} \vee \mathbf{y})\, g(\mathbf{x} \wedge \mathbf{y}). \tag{5.1}$$

A random vector $(X_1, ..., X_n)$ is MTP$_2$ if its density function is MTP$_2$.

In order to verify (5.1), it is sufficient [15, p. 469] to check that $g(\mathbf{x}) > 0$ is MTP$_2$ in every pair of variables while the remaining variables are held fixed. With regard to the Liouville distributions, we have the following result.

5.2. PROPOSITION. *Let* $(X_1, ..., X_n) \sim L_n[f(\cdot);\ a_1, ..., a_n]$. *Then the following are equivalent*:

(i) $(X_1, ..., X_n)$ *is* MTP$_2$;

(ii) $f(x+y)$ *is* TP$_2$ *in* (x, y) *on* R_+^2;

(iii) $f(\cdot)$ *is logarithmically convex on* R_+.

Proof. It follows from Definition 5.1 that the density which is proportional to $(\prod_{i=1}^n x_i^{a_i - 1})\, f(\sum_{i=1}^n x_i)$ is MTP$_2$ if and only if $f(\sum_{i=1}^n x_i)$ is MTP$_2$. Since MTP$_2$ is equivalent here to pairwise MTP$_2$, then we have $f(x_1 + x_2)$ TP$_2$ on R_+^2, proving that (i) and (ii) are equivalent. Finally, the equivalence of (ii) and (iii) was proven by Karlin [14, p 160]; indeed, condition (5.1), when $n = 2$, is precisely the definition of logarithmic convexity.

5.3. EXAMPLE. Let $f(\cdot)$ be either of the functions $f_1(t) = t^{a-1} e^{-bt}$ or $f_2(t) = t^{a-1}(1+t)^{-b}$, $t > 0$, $a > 0$, $b > 0$; the related Liouville distributions were encountered in Examples 1.1 and 1.3. The function $f(\cdot)$ is logarithmically convex and, hence, the corresponding densities are MTP$_2$, if and only if $0 < a \leqslant 1$.

Having determined necessary and sufficient criteria for a Liouville distribution to be MTP_2, we may now apply the results of Karlin and Rinott [15, Section 4] to obtain far-reaching information on probability and expectation inequalities. In order to do this, we introduce a partial order on R^n as dollows: if $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ then $\mathbf{x} \preccurlyeq \mathbf{y}$ if $x_i \leqslant y_i$, $i = 1, ..., n$. A function ϕ: $R^n \to R$ is increasing (decreasing) if $\mathbf{x} \preccurlyeq \mathbf{y}$ implies $\phi(\mathbf{x}) \leqslant \phi(\mathbf{y})$ ($\phi(\mathbf{x}) \geqslant \phi(\mathbf{y})$). Then we obtain the following results from [15, pp. 484–487].

5.4. Theorem. *Let* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$ *and have a* MTP_2 *density function. Then,*

(i) *the marginal distribution of* $(X_1, ..., X_r)$ *is* MTP_2, $1 \leqslant r \leqslant n$;

(ii) *for any increasing function* ϕ: $R^r_+ \to R$, $1 \leqslant r \leqslant n$, *the multiple regression*

$$\mathscr{E}(\phi(X_1, ..., X_r) \mid X_{r+1} = x_{r+1}, ..., X_n = x_n)$$

is increasing in $(x_{r+1}, ..., x_n)$;

(iii) *if* ϕ *and* ψ *are both increasing or both decreasing on* R^n_+, *then* $\operatorname{Cov}(\phi(X_1, ..., X_n), \psi(X_1, ..., X_n)) \geqslant 0$. *More generally, if* $\phi_1, ..., \phi_m$ *are all increasing or all decreasing on* R^n_+, *then*

$$\mathscr{E}\left(\prod_{i=1}^{m} \phi_i(X_1, ..., X_n)\right) \geqslant \prod_{i=1}^{m} \mathscr{E}(\phi_i(X_1, ..., X_n)); \tag{5.2}$$

(iv) *if* $d_0(t) \equiv 1$ *and* $d_r(t) = P(X_1 \leqslant t, ..., X_r \leqslant t)$, $1 \leqslant r \leqslant n$, *then*

$$d_{r-1}(t)\, d_{r+1}(t) \geqslant d_r^2(t), \qquad 1 \leqslant r \leqslant n-1.$$

It is customary to apply (5.2) to the development of inequalities for the distribution function of $(X_1, ..., X_n)$. To this end, choose positive numbers x_i and define

$$\phi_i(X_1, ..., X_n) = \begin{cases} 1, & X_i \leqslant x_i, \\ 0, & \text{otherwise}, \end{cases}$$

$i = 1, ..., n$. Then, (5.2) becomes

$$P(X_1 \leqslant x_1, ..., X_n \leqslant x_n) \geqslant \prod_{i=1}^{n} P(X_i \leqslant x_i). \tag{5.3}$$

Replacing ϕ_i by $1 - \phi_i$ leads to

$$P(X_1 \geqslant x_1, ..., X_n \geqslant x_n) \geqslant \prod^{n} P(X_i \geqslant x_i). \tag{5.4}$$

In the case of the inverted Dirichlet distribution, the inequalities (5.3) and (5.4) were obtained by Kimball [19] (cf. [13, p. 240]).

5.5. DEFINITION [16]. A function $g: R^n \to R_+$ *is multivariate reverse rule of order* 2 (MRR_2) if $g(\cdot)$ satisfies the reverse of (5.1); that is, for all $\mathbf{x}$, $\mathbf{y}$ in R^n,

$$g(\mathbf{x})\, g(\mathbf{y}) \geqslant g(\mathbf{x} \vee \mathbf{y})\, g(\mathbf{x} \wedge \mathbf{y}). \tag{5.5}$$

A random vector $(X_1, ..., X_n)$ is MRR_2 if its density function is MRR_2.

Suppose that a nonnegative function $g(\cdot)$ has the property that $g(\mathbf{x})\, g(\mathbf{y}) \neq 0$ implies $g(\mathbf{z}) \neq 0$ for all $\mathbf{z}$ such that $\mathbf{x} \leqslant \mathbf{z} \leqslant \mathbf{y}$. Then similar to the MTP_2 case, $g(\cdot)$ is MRR_2 if and only if $g(\cdot)$ is RR_2 in every pair of variables while the remaining variables are held fixed (cf. [16, p. 500]).

5.6. PROPOSITION. *Let* $(X_1, ..., X_n) \sim L_n[f(\cdot); a_1, ..., a_n]$, *where* $f(\cdot)$ *is monotone increasing or decreasing. Then* $(X_1, ..., X_n)$ *is* MRR_2 *if and only if* $f(\cdot)$ *is logarithmically concave.*

Proof. It is easy to see from (5.5) that $(X_1, ..., X_n)$ is MRR_2 if and only if the function $f(\sum_{i=1}^n x_i)$ is MRR_2. Since $f(\cdot)$ is monotone, then $f(\sum_{i=1}^n x_i)\ f(\sum_{i=1}^n y_i) \neq 0$ implies $f(\sum_{i=1}^n z_i) \neq 0$ for all $x_i \leqslant z_i \leqslant y_i$, $i = 1, ..., n$. Hence $(X_1, ..., X_n)$ is MRR_2 if and only if $f(\sum_{i=1}^n x_i)$ is pairwise RR_2; that is, $f(x_1 + x_2)$ is RR_2 or $f(\cdot)$ is logarithmically concave.

As is noted in [16], the MRR_2 property does not suffice to imply expectation and probabilistic inequalities analogous to (5.2)–(5.4). To this end, we are forced to strengthen the reverse rule requirements.

5.7. DEFINITION [16]. (i) A function $\phi: R \to R$ is a *Pólya frequency function of order* 2 (PF_2) if $\phi(s - t)$ is TP_2 in the variables s, t, $-\infty < s$, $t < \infty$.

(ii) A random vector $(X_1, ..., X_n)$ or its density function $g(\cdot)$ is *strongly multivariate reverse rule of order* 2 (S-MRR_2) if for any set $\{\phi_i(\cdot)\}$ of PF_2 functions the marginal

$$h(x_{i_1}, ..., x_{i_r}) = \int_{R^{n-r}} g(x_1, ..., x_n) \prod_{i=1}^{n-r} \phi_{j_i}(x_{j_i})\, dx_{j_i} \tag{5.6}$$

is MRR_2 in the variables $x_{i_1}, ..., x_{i_r}$, where $\{i_1, ..., i_r\}$ and $\{j_1, ..., j_{n-r}\}$ are complementary sets of indices drawn from $\{1, ..., n\}$.

An example of a S-MRR_2 distribution is the Dirichlet distribution $D(a_1, ..., a_n;\ a_{n+1})$ with $a_i \geqslant 1$ for all $i = 1, ..., n+1$; this result was established by Karlin and Rinott [16, p. 508]. Indeed, a close inspection of their analysis yields the following general principle.

5.8. THEOREM. *Let* $(X_1, ..., X_n) \sim L_n^{(2)}[f(\cdot); a_1, ..., a_n]$, *where* $a_i \geq 1$, $i = 1, ..., n$, *and* $f(\cdot)$ *is monotone increasing or decreasing. Then* $(X_1, ..., X_n)$ *is* S-MRR$_2$ *if and only if* $f(\cdot)$ *is logarithmically concave.*

Proof. Since S-MRR$_2$ implies MRR$_2$, then the necessity follows from Proposition 5.6. In proving the sufficiency, we modify the argument of [16, Proposition 2.4], presenting the explicit details for the convenience of the reader. If $n = 2$, then (X_1, X_2) is RR$_2$ if and only if $f(x_1 + x_2)$ is RR$_2$; this holds because of Proposition 5.6. Let $t_+^a = t^a$ for $t > 0$ and 0 for $t \leq 0$. When $n = 3$ we have to show that for any PF$_2$ function $\phi(\cdot)$,

$$\int_0^1 \phi(x_1) f(x_1 + x_2 + x_3) \left(\prod_{i=1}^3 x_i^{a_i - 1} \right) dx_1$$
$$= x_2^{a_2 - 1} x_3^{a_3 - 1} \int_0^1 \phi(t - x_2) f(t + x_3)(t - x_2)_+^{a_1 - 1} \, dt \tag{5.7}$$

is RR$_2$ in (x_2, x_3).

Since $\phi(\cdot)$ is PF$_2$ then $\phi(t - x_2)$ is TP$_2$ in (t, x_2); also, $(t - x_2)_+^{a_1 - 1}$ is TP$_2$ in (t, x_2), $a_1 \geq 1$. Therefore, $\phi(t - x_2)(t - x_2)_+^{a_1 - 1}$ is TP$_2$ in (t, x_2). Next, since $f(\cdot)$ is log-concave, then $f(t + x_3)$ is RR$_2$ in (t, x_3). Consequently, it follows from the basic composition formula [14, p. 98] that the integral (5.7) is RR$_2$ in (x_2, x_3).

Assume by induction that the function

$$\int_0^1 \cdots \int_0^1 \left(\prod_{i=1}^n x_i^{a_i - 1} \right) f\left(\sum_{i=1}^n x_i \right) \prod_{i=1}^{r-1} \phi_i(x_i) \, dx_i \tag{5.8}$$

is RR$_2$ in every pair of variables (x_i, x_j), with $r \leq i < j \leq n$, for any set of PF$_2$ functions $\phi_1, ..., \phi_{r-1}$. The integral in (5.8) is clearly of the form

$$\left(\prod_{i=r}^n x_i^{a_i - 1} \right) h\left(\sum_{i=r}^n x_i \right)$$

for some function $h(\cdot)$. Multiplying (5.8) by a PF$_2$ function $\phi_r(x_r)$ and proceeding as in the case $n = 3$ we find that

$$\int_0^1 \cdots \int_0^1 \left(\prod_{i=1}^n x_i^{a_i - 1} \right) f\left(\sum_{i=1}^n x_i \right) \prod_{i=1}^{r} \phi_i(x_i) \, dx_i$$

is RR$_2$ in (x_i, x_j), $r + 1 \leq i < j \leq n$. This establishes the inductive step and completes the proof.

5.9. COROLLARY. *Under the hypotheses of Theorem 5.8, if $\phi_1, ..., \phi_n$ are all increasing (or decreasing)* PF$_2$ *functions then*

$$\mathscr{E}\left(\prod_{i=1}^{n} \phi_i(X_i)\right) \leqslant \mathscr{E}\left(\prod_{i=1}^{r} \phi_i(X_i)\right) \mathscr{E}\left(\prod_{i=r+1}^{n} \phi_i(X_i)\right), \tag{5.9}$$

$r = 1, ..., n$, *provided the expectations exist. In particular,*

$$\mathscr{E}\left(\prod_{i=1}^{n} \phi_i(X_i)\right) \leqslant \prod_{i=1}^{n} \mathscr{E}(\phi_i(X_i)). \tag{5.10}$$

Note also that $I_{[t,\infty)}(\cdot)$, the indicator function of $[t, \infty)$, is increasing and PF$_2$. Approximating an increasing, nonnegative function $h(t)$ by $\varepsilon \sum_{k=0}^{\infty} I_{[k\varepsilon,\infty)}(t)$, $\varepsilon \to 0$, it follows that (5.9) and (5.10) remain valid for all nonnegative, increasing ϕ_i, *without the* PF$_2$ *condition.*

With the hypotheses of Theorem 5.8, the inequality (5.10) implies that for $x_i > 0$,

$$P(X_1 \geqslant x_1, ..., X_n \geqslant x_n) \leqslant \prod_{i=1}^{n} P(X_i \geqslant x_i),$$

which is analogous to (5.3) (5.4). However, more stringent inequalities can be obtained by appealing to [16, p. 513]; a typical result is that for $0 < x_i \leqslant y_i \leqslant \infty$, $1 \leqslant k < l < n$,

$$\begin{aligned} &P(x_i \leqslant X_i \leqslant y_i,\ 1 \leqslant i \leqslant k;\ X_j \leqslant y_j,\ k < j \leqslant n)\ P(x_i \leqslant X_i \leqslant y_i,\ 1 \leqslant i \leqslant k) \\ &\quad \leqslant P(x_i \leqslant X_i \leqslant y_i,\ 1 \leqslant i \leqslant k;\ X_j \leqslant y_j,\ k < j \leqslant l) \\ &\qquad \times P(x_i \leqslant X_i \leqslant y_i,\ 1 \leqslant i \leqslant k;\ X_j \leqslant y_j,\ l < j \leqslant n). \end{aligned}$$

6. *Characterizations*

In the preceding sections, we have treated several aspects of the Liouville distributions. Here, we see how these distributions may be characterized using the topics considered earlier. However, the proofs of the main statements are placed in Section 9 where more general results are established for the matrix analogs.

6.1. PROPOSITION. *Let* $(X_1, ..., X_n) \sim L_n^{(1)}[f(\cdot);\ a_1, ..., a_n]$. *Then the following are equivalent:*

(i) $X_1, ..., X_n$ *are mutually independent;*

(ii) *there exists* i, j *with* X_i *and* X_j *mutually independent;*

(iii) *for* $1 \leqslant r < k \leqslant n$ *and* $j \in R$, *the regression* $\mathscr{E}(X_k^j \mid \sum_{i=r+1}^n X_i = t)$, *if it exists, is constant* (*a.s.*).

Note that all three statements will be established by proving that $f(t) = e^{-bt}$, $t > 0$, $b > 0$. In particular, we will use Deny's theorem to prove part (iii).

6.2. *Remarks.* (i) From Propositions 4.4 and 4.6, it follows that Proposition 6.1(iii) remains valid if X_k^j is replaced by $(\sum_{i=r+1}^n X_i)^j$ or $h(\sum_{i=r+1}^n X_i)$, where $h(\cdot)$ is any continuous, nonnegative function for which the regression exists.

(ii) If j is a positive integer, then Theorem 6.1(iii) can be established without recourse to Deny's theorem. By Proposition 4.4, the hypothesis of constant regression is equivalent to $W^{j+a} f(t) = c W^a f(t)$, where $a = a_{r+1} + \cdots + a_n$. Differentiating fractionally, we get $f(t) = c W^j f(t)$, $t > 0$, which can be rewritten as

$$f(t) = \frac{c_j}{\Gamma(j)} \int_t^{\infty} (y-t)^{j-1} f(y)\, dy, \tag{6.1}$$

where $c_j > 0$. By repeated differentiation of (6.1), we get

$$(-1)^j f^{(j)}(t) = c_j f(t), \qquad t > 0. \tag{6.2}$$

From (6.2) we deduce that $f(\cdot)$ is infinitely differentiable and even that $f(\cdot)$ is completely monotone; that is,

$$(-1)^i f^{(i)}(t) \geqslant 0, \qquad i = 0, 1, 2, \ldots.$$

By the well-known Hausdorff–Bernstein theorem, there exists a unique, nonnegative, finite, Borel measure μ such that

$$f(t) = \int_0^{\infty} e^{-ty}\, d\mu(y).$$

From the uniqueness of μ and (6.2), it follows that μ is singular; therefore, $f(t) = ce^{-bt}$, $t > 0$, $b > 0$.

(iii) In the case of the Dirichlet distributions $D(a_1, \ldots, a_n; a_{n+1})$, it was noted in Example 4.5 that

$$\mathscr{E}\left(X_k^j \,\middle|\, \sum_{i=1}^r X_i = t\right) = c(1-t)^j, \qquad 0 < t < 1, \tag{6.3}$$

where $j > 0$, $1 \leqslant r < k \leqslant n$. We conjecture that (6.3) characterizes the

Dirichlet distributions among the class $L_n^{(2)}$, and below, we prove this result when j is a positive integer.

By Proposition 4.4 and the definition of the W^a operators, the problem of characterizing the Dirichlet distributions through (6.3) is equivalent to solving the integral equation

$$\frac{1}{\Gamma(a+j)}\int_t^1 (y-t)^{a+j-1} f(y)\,dy = c_j \frac{(1-t)^j}{\Gamma(a)}\int_t^1 (y-t)^{a-1} f(y)\,dy, \tag{6.4}$$

where $a = a_{r+1} + \cdots + a_n$, $0 < t < 1$, subject to $f(t) \geqslant 0$ and the integrability condition (2.1). Substituting $y = (z-1)/z$, $t = (s-1)/s$, and $g(z) = z^{-(a+1)} f((z-1)/z)$, then (6.4) is transformed into

$$\frac{1}{\Gamma(a+j)}\int_s^\infty z^{-j}(z-s)^{a+j-1} g(z)\,dz = \frac{c_j}{\Gamma(a)}\int_s^\infty (z-s)^{a-1} g(z)\,dz,$$

$s > 0$. That is, $W^{a+j}(s^{-j}g(s)) = c_j W^a g(s)$, and by fractional differentiation,

$$W^j(s^{-j}g(s)) = c_j\, g(s). \tag{6.5}$$

If j is a positive integer, then by repeatedly differentiating (6.5), we get the ordinary differential equation

$$g^{(j)}(s) = c_0(-1)^j s^{-j} g(s), \qquad s > 0,\ c_0 = 1/c_j > 0. \tag{6.6}$$

When $j = 2$, (6.6) is known [20] as Cauchy's equation. The standard procedure for solving these equations is by way of the substitution $s = e^x$, which transforms (6.6) into the linear differential equation

$$[D(D-1)(D-2)\cdots(D-j+1) - c_0(-1)^j]\, h(x) = 0, \tag{6.7}$$

where $h(x) = g(e^x)$ and $D = d/dx$. If j is even, the characteristic polynomial, $p(x)$, of (6.7) is strictly decreasing for $x < 0$. Since $p(0) = -c_0 < 0$ and $p(x) \to \infty$ as $x \to -\infty$, then $p(x)$ has a unique negative root $-\lambda$. A similar argument yields the same conclusion if j is odd. Also, any complex roots of $p(x)$ are obviously to be disregarded, while positive roots are eventually ruled out by (2.1). Therefore, (6.7) has the unique solution $h(x) = e^{-\lambda x}$, and hence $f(t) = (1-t)^{\beta-1}$. Moreover, (6.4) implies that $\beta > 0$.

Next, we characterize the Dirichlet distributions among the class $L_n^{(2)}$ using the concept of *complete neutrality* (Doksum [5]). This notion is related to the tailfree distributions of Freedman [8], and has been used by James and Mosimann [12] to derive other characterizations of the Dirichlet distributions.

6.3. Definition. A random vector $(X_1, \ldots, X_n)$ taking values in the simplex $\mathscr{S}_n$ is *completely neutral* if there exist mutually independent, nonnegative random variables $Y_1, ..., Y_n$ such that

$$(X_1, ..., X_n) \overset{\mathscr{L}}{=} \left(Y_1, Y_2(1 - Y_1), ..., Y_n \prod_{i=1}^{n-1} (1 - Y_i) \right).$$

6.4. PROPOSITION. *If $(X_1, ..., X_n) \sim L_n^{(2)}[f(\cdot); a_1, ..., a_n]$, then $(X_1, ..., X_n)$ is completely neutral if and only if $(X_1, ..., X_n) \sim D(a_1, ..., a_n; a_{n+1})$ for some $a_{n+1} > 0$.*

A similar result may be established for the class $L_n^{(1)}$ and the inverted Dirichlet distributions.

6.5. PROPOSITION. *Suppose that $(X_1, ..., X_n) \sim L_n^{(1)}[f(\cdot); a_1, ..., a_n]$. Then there exist mutually independent, positive random variables $Y_1, ..., Y_n$ such that*

$$(X_1, ..., X_n) \overset{\mathscr{L}}{=} \left(Y_1, Y_2(1 + Y_1), ..., Y_n \prod_{i=1}^{n-1} (1 + Y_i) \right)$$

if and only if $(X_1, ..., X_n) \sim ID(a_1, ..., a_n; a_{n+1})$ for some $a_{n+1} > 0$.

II. LIOUVILLE DISTRIBUTIONS ON $R_+^{n \times n}$

7. *Transformations and Stochastic Representations*

Throughout, the unique, positive definite (symmetric) square root of a positive definite matrix T will be denoted by $T^{1/2}$.

7.1. PROPOSITION. *Suppose that $(B_1, ..., B_n) \sim L_n^{(2)}[g(\cdot); a_1, ..., a_n]$. For $i = 1, ..., n$, define*

$$A_i = \left(I - \sum_{j=1}^{n} B_j \right)^{-1/2} B_i \left(I - \sum_{j=1}^{n} B_j \right)^{-1/2}. \tag{7.1}$$

Then $(A_1, ..., A_n) \sim L_n^{(1)}[f(\cdot); a_1, ..., a_n]$, where

$$f(T) = |I + T|^{-(a_1 + \cdots + a_n + p)} g(T(I + T)^{-1}), \qquad T > 0. \tag{7.2}$$

In particular, there is a one–one correspondence between $f(\cdot)$ and $g(\cdot)$.

Proof. The statement and proof of this result are natural extensions of those given in Proposition 3.1. Let $A_0 = \sum_{i=1}^n A_i$, $B_0 = \sum_{i=1}^n B_i$. It follows from (7.1) that $A_0 = (I - B_0)^{-1/2} B_0 (I - B_0)^{-1/2} = (I - B_0)^{-1} B_0$, the last equality holding since B_0 commutes with any rational function of B_0. By simple manipulation of these identities, we get $B_0=(I+A_0)^{-1}A_0$ and $I-B_0=(I+A_0)^{-1}$. Inverting (7.1), we see that

$$B_i=(I+A_0)^{-1/2}A_i(I+A_0)^{-1/2}, \qquad i=1, ..., n. \tag{7.3}$$

The Jacobian of (7.3) may be shown to equal $|I+A_0|^{-(n+1)p}$, and then the rest of the proof is standard.

7.2. *Remarks.* When $g(T)=|I-T|^{a-p}$, $0<T<I$, the transformation (7.3) was used by Olkin and Rubin [24, Theorem 3.4] to transform an inverted Dirichlet distribution into a Dirichlet distribution; their method was somewhat more roundabout, being based on a set of successive transformations.

7.3. PROPOSITION. *Let* $(A_1, ..., A_n)\sim L_n[f(\cdot);\ a_1, ..., a_n]$. *Then the following stochastic representations are valid*:

(i) $(A_1, ..., A_n) \stackrel{\mathscr{L}}{=} B_n^{1/2}(B_1, ..., B_{n-1},\ I-\sum_{i=1}^{n-1}B_i)\,B_n^{1/2}$, *where* $(B_1, ..., B_{n-1})$ *and* B_n *are mutually independent, and* $(B_1, ..., B_{n-1})\sim D(a_1, ..., a_{n-1}; a_n)$;

(ii) $(A_1, ..., A_n) \stackrel{\mathscr{L}}{=} B_n^{1/2}((\prod_{i=1}^{n-1}B_{n-i}^{1/2})(\prod_{i=1}^{n-1}B_i^{1/2}),\ (\prod_{i=2}^{n-1}B_{n+1-i}^{1/2})(I-B_1)(\prod_{i=2}^{n-1}B_i^{1/2}), ..., I-B_{n-1})\,B_n^{1/2}$, *where* $B_1, ..., B_n$ *are mutually independent, and* $B_i\sim B(\sum_{j=1}^{i}a_j;\ a_{i+1})$, $i=1, ..., n-1$;

(iii) $(A_1, ..., A_n) \stackrel{\mathscr{L}}{=} B_n^{1/2}((\prod_{i=1}^{n-1}(I+B_{n-i})^{-1/2})(\prod_{i=1}^{n-1}(I+B_i)^{-1/2}),\ (\prod_{i=2}^{n-1}(I+B_{n+1-i})^{-1/2})B_1(\prod_{i=2}^{n-1}(I+B_i)^{-1/2}), ..., (I+B_{n-1})^{-1/2}B_{n-1}(I+B_{n-1})^{-1/2})\,B_n^{1/2}$, *where* $B_1, ..., B_n$ *are mutually independent, and* $B_i\sim IB(a_{i+1};\ \sum_{j=1}^{i}a_j)$, $i=1, ..., n-1$;

(iv) $(A_1, ..., A_n) \stackrel{\mathscr{L}}{=} B_n^{1/2}(B_1,\ (I-B_1)^{1/2}B_2(I-B_1)^{1/2}, ..., (\prod_{i=1}^{n-2}(I-B_i)^{1/2})B_{n-1}(\prod_{i=1}^{n-2}(I-B_{n-i-1})^{1/2}),\ (\prod_{i=1}^{n-1}(I-B_i)^{1/2})(\prod_{i=1}^{n-1}(I-B_{n-i})^{1/2}))\,B_n^{1/2}$, *where* $B_1, ..., B_n$ *are mutually independent, and* $B_i\sim B(a_i;\ \sum_{j=i+1}^{n}a_j)$, $i=1, ..., n-1$.

In all four cases, $B_n \stackrel{\mathscr{L}}{=} L_1[f(\cdot);\ \sum_{i=1}^{n}a_i]$.

In view of the similarity with Theorem 3.2, a detailed proof is unnecessary. However, it should be noted that the Jacobians of (i)–(iv) are, respectively, $|B_n|^{(n-1)p}$, $\prod_{i=2}^{n}|B_i|^{(i-1)p}$, $|B_n|^{(n-1)p}\prod_{i=1}^{n-1}|I+B_i|^{-(i+1)p}$, and $|B_n|^{(n-1)p}\prod_{i=1}^{n-1}|I-B_i|^{(n-i-1)p}$.

The above representations can also be used to generalize some earlier observations (Theorem 3.2, *infra*); we obtain the following result which is due to Khatri [18].

7.4. COROLLARY (Khatri [18]). *If mutually independent* $B_i\sim B(\sum_{j=1}^{i}a_j;\ a_{i+1})$, $i=1, ..., n$, *then* $(\prod_{i=1}^{n}B_{n-i+1}^{1/2})(\prod_{i=1}^{n}B_i^{1/2})\sim B(a_1;\ \sum_{i=2}^{n+1}a_i)$.

Proof. We use induction, with $n \geqslant 2$. When $n=2$, define (A_1, A_2) in terms of (B_1, B_2) using the transformation in Proposition 7.3(iv). By reversing the argument which establishes that result, 7.3(iv), we obtain $(A_1, A_2) \sim D(a_1; a_2+a_3)$; in particular, $B_2^{1/2} B_1 B_2^{1/2} = A_1 \sim B(a_1; a_2+a_3)$. The inductive step is also proven using similar arguments.

Finally, if we mimic the proof of Proposition 3.4 then we obtain the following result.

7.5. PROPOSITION. *If* $(A_1, ..., A_n) \sim L_n[f(\cdot); a_1, ..., a_n]$ *and* $r<n$, *then* $(\sum_{i=1}^n A_i)^{-1/2} (\sum_{i=1}^r A_i)(\sum_{i=1}^n A_i)^{-1/2} \sim B(\sum_{i=1}^r a_i; \sum_{i=r+1}^n a_i)$.

8. *Marginal and Conditional Distributions*

As expected, the results for marginal and conditional distributions in the matrix case are entirely analogous to the vector situation. We state the needed results without proof.

8.1. PROPOSITION. *Let* $(A_1, ..., A_n) \sim L_n[f(\cdot); a_1, ..., a_n]$ *and* $r<n$. *Then*

(i) $(A_1, ..., A_r) \sim L_r[f_r(\cdot); a_1, ..., a_r]$, *where* $a = \sum_{i=r+1}^n a_i$ *and* $f_r(T) = W^a f(T)$ *is the fractional integral of order* a *of* $f(\cdot)$;

(ii) *the conditional distribution of* $(A_{r+1}, ..., A_n)$ *given* $\{A_1, ..., A_n\}$ *is* $L_{n-r}[g_r(\cdot); a_{r+1}, ..., a_n]$, *where* $g_r(T) = f(T + \sum_{i=1}^r A_i)/f_r(\sum_{i=1}^r A_i)$;

(iii) *if the expectation below exists, then*

$$\mathscr{E}\left(\prod_{i=r+1}^{n} |A_i|^{j_i} \,\middle|\, \sum_{i=1}^{r} A_i = T\right) = cW^{j+a} f(T)/W^a f(T), \tag{8.1}$$

where $a = \sum_{i=r+1}^n a_i$, $j = \sum_{i=r+1}^n j_i$.

9. *Characterizations*

Here, we extend the results stated in Section 6 by characterizing the Liouville distributions through various properties treated in earlier sections. Since the zero matrix is a limit point of the convex cone $R_+^{m \times m}$, we may use the notation "$T \to 0+$" to mean that $T \to 0$ through $R_+^{m \times m}$. It is assumed throughout that $\lim_{T \to 0+} f(T) = 1$.

9.1. PROPOSITION. *Let* $(A_1, ..., A_n) \sim L_n^{(1)}[f(\cdot); a_1, ..., a_n]$. *Then the following are equivalent*:

(i) $f(T) = \exp(-\operatorname{tr} \Sigma T)$ *for some* $\Sigma > 0$;

(ii) $A_1, ..., A_n$ *are mutually independent*;

(iii) *there exists* i, j *with* A_i *and* A_j *mutually independent.*

Proof. Since (i) implies (ii) and (ii) implies (iii), we need only show that (i) follows from (iii). By Proposition 8.1(i), the pair (A_i, A_j) has a marginal distribution which belongs to the class $L_n^{(1)}$; therefore, in proving that (iii) implies (i), it suffices to assume that $n = 2$. Therefore, suppose that A_1 and A_2 are independent. Then

$$|A_1|^{a_1 - p} |A_2|^{a_2 - p} f(A_1 + A_2) = |A_1|^{a_1 - p} |A_2|^{a_2 - p} h_1(A_1) h_2(A_2),$$

equivalently, $f(A_1 + A_2) = h_1(A_1) h_2(A_2)$ for all $A_1 > 0$, $A_2 > 0$, and continuous, nonnegative functions h_1, h_2. By symmetry, $h_1 = h_2$; further, when $A_1, A_2 \to 0+$ we get $h_1(0+)^2 = 1$, so that $h_1(0+) = 1$. As $A_2 \to 0+$, we even see that $f(T) = h_1(T)$, $T > 0$. Therefore,

$$f(T_1 + T_2) = f(T_1) f(T_2), \tag{9.1}$$

which is the multiplicative analog of Cauchy's equation on $R_+^{m \times m}$. Regarding (9.1) as a functional equation in the $m(m+1)/2$ distinct entries t_{ij} of the matrix T, then we have

$$f(T) = \exp\left(-\sum_{i=1}^{m} \sum_{j=1}^{i} \sigma_{ij} t_{ij}\right) \equiv \exp(-\operatorname{tr} \Sigma T),$$

where the $m \times m$ matrix Σ is symmetric. If there exists $T_0 > 0$ such that $\operatorname{tr}(\Sigma T_0) < 0$ then the sequence $f(iT_0)$, $i = 1, 2, ...,$ is unbounded, contradicting (2.1). Consequently, $\operatorname{tr}(\Sigma T) \geqslant 0$ for all $T > 0$ and therefore [10, p. 478] Σ is positive semidefinite. Finally, the positive definiteness of Σ is guaranteed by (2.1).

9.2. THEOREM. *Let* $(A_1, ..., A_n) \sim L_n[f(\cdot); a_1, ..., a_n]$, $1 \leqslant r < k \leqslant n$, *and* j *be such that the following regression exists. Then* $\mathscr{E}(|A_k|^j \mid \sum_{i=1}^r A_i = T)$ *is constant (a.s.) if and only if*

$$f(T) = \int_{R_+^{m \times m}} \exp(-\operatorname{tr} \Sigma T)\, d\mu(\Sigma), \qquad T > 0, \tag{9.2}$$

where μ *is a probability measure concentrated on a hypersurface of the form* $\{\Sigma > 0 : |\Sigma| = c\}$.

Proof. From Proposition 8.1, the a.s. constancy of the regression function is equivalent to $f(\cdot)$ satisfying the integral equation

$$f(T) = c \int_{R_+^{m \times m}} f(S+T)\,|S|^{j-p}\,dS, \qquad T > 0,$$

which is a special case of Deny's equation (2.4). To apply Deny's theorem, we need to find all continuous, bounded solutions ϕ of the functional equation $\phi(T_1 + T_2) = \phi(T_1)\,\phi(T_2)$, $T_1 > 0$, $T_2 > 0$. In proving Proposition 9.1, it was shown that every such ϕ is of the form $\phi(T) = \exp(-\operatorname{tr} \Sigma T)$, where $\Sigma > 0$. Then, the representation (9.2) follows from Deny's theorem.

9.3. *Remark.* Note that when $n = 1$, the measure μ is singular; then, the regression is constant (a.s.) if and only if $f(t) = e^{-bt}$ for some $b > 0$. Combining this remark with the results in Proposition 9.1, then we have completely proven Proposition 6.1. We also point out that results generalizing Theorem 9.2 may be derived (from Deny's theorem and Proposition 8.1) if $|A_k|^j$ is replaced by $h(\sum_{i=1}^r A_i)$, where $h(\cdot)$ is unbounded, nonnegative, continuous, and such that the conditional expectation exists.

Finally, we develop the appropriate generalization of Proposition 6.4.

9.4. Definition. Let $A_1, \ldots, A_n$ and $I - \sum_{i=1}^n A_i$ be random matrices taking values in $R_+^{m \times m}$. Then $(A_1, \ldots, A_n)$ is *completely neutral* if there exist mutually independent $B_1, \ldots, B_n$, $0 < B_i < I$ $(i = 1, \ldots, n)$, such that

$$(A_1, \ldots, A_n) \overset{\mathscr{L}}{=} \Bigg(B_1, (I - B_1)^{1/2} B_2 (I - B_1)^{1/2}, \ldots, \\ \left(\prod_{i=1}^{n-1} (I - B_i)^{1/2}\right) B_n \left(\prod_{i=1}^{n-1} (I - B_{n-i})^{1/2}\right)\Bigg). \tag{9.3}$$

9.5. Proposition. *Let $(A_1, \ldots, A_n) \sim L_n^{(2)}[f(\cdot); a_1, \ldots, a_n]$. Then $(A_1, \ldots, A_n)$ is completely neutral if and only if $(A_1, \ldots, A_n) \sim D(a_1, \ldots, a_n; a_{n+1})$, for some $a_{n+1} > p - 1$.*

Proof. The Jacobian of the transformation defined through (9.3) is $\prod_{i=1}^{n-1} |B_i|^{(n-i)p}$. Also, it can be shown (using induction, say) that

$$\sum_{i=1}^n A_i \overset{\mathscr{L}}{=} I - \left(\prod_{i=1}^n (I - B_i)^{1/2}\right)\left(\prod_{i=1}^n (I - B_{n+1-i})^{1/2}\right).$$

Since $(A_1, ..., A_n) \sim L_n^{(2)}[f(\cdot);\ a_1, ..., a_n]$, then the joint density of $(B_1, ..., B_n)$ is proportional to

$$\left(\prod_{i=1}^{n} |B_i|^{a_i - p}\right)\left(\prod_{i=1}^{n-1} |I - B_i|^{a_{i+1} + \cdots + a_n}\right)$$
$$\times f\left(I - \left(\prod_{i=1}^{n} (I - B_i)^{1/2}\right)\left(\prod_{i=1}^{n} (I - B_{n+1-i})^{1/2}\right)\right). \tag{9.4}$$

If $f(T) = |I - T|^{a_{n+1} - p}$, $a_{n+1} > p - 1$, then it follows from (9.4) that $B_1, ..., B_n$ are independent multivariate beta matrices; hence, $(A_1, ..., A_n)$ is completely neutral.

Conversely, if $(A_1, ..., A_n)$ is completely neutral, then by (9.4),

$$f\left(I - \left(\prod_{i=1}^{n} (I - B_i)^{1/2}\right)\left(\prod_{i=1}^{n} (I - B_{n+1-i})^{1/2}\right)\right) = \prod_{i=1}^{n} h_i(I - B_i) \tag{9.5}$$

for continuous, nonnegative functions h_i. Using the argument of Proposition 9.1, we find that $h_i(I - T) = f(T)$, $0 < T < I$, $i = 1, ..., n$. As $B_i \to 0$ in (9.5) (for $i \geqslant 3$), we find that the function $g(T) = f(I - T)$ satisfies

$$g(T_1^{1/2} T_2 T_1^{1/2}) = g(T_1)\, g(T_2) \tag{9.6}$$

for $0 < T_1$, $T_2 < I$. The conclusion rests on the following result.

9.6. LEMMA. *Let $g: R_+^{m \times m} \to R_+$ be nontrivial, continuous, and satisfy (9.6). Then $g(T) = |T|^k$ for some k in R.*

Proof. Without loss of generality, we may suppose that $g(\cdot)$ is defined on all of $R_+^{m \times m}$. Since $g(\cdot)$ is nontrivial, then (9.6) implies that $g(I) = 1$. Next, recall that every $m \times m$ nonsingular matrix X has a "polar coordinates" decomposition [10, p. 482] $X = VT_1^{1/2}$, where $T_1 > 0$ and $V \in O(m)$, the group of $m \times m$ orthogonal matrices. Then $T_1 = X'X$ and $V'X = T_1^{1/2} = (T_1^{1/2})' = X'V$. Substituting these relations into (9.6), we obtain

$$g(V'XT_2X'V) = g(T_1^{1/2} T_2 T_1^{1/2}) = g(T_1)\, g(T_2) = g(V'XX'V)\, g(T_2). \tag{9.7}$$

With $X = I$ in (9.7), we get $g(V'T_2V) = g(T_2)$ for all $V \in O(m)$, $T_2 > 0$; that is, $g(\cdot)$ is orthogonally invariant. Hence, (9.7) reduces to $g(XT_2X') = g(XX')\, g(T_2)$, and on setting $T_2 = YY'$ in this last equation we find that the function $p(X) = g(XX')$ satisfies

$$p(XY) = p(X)\, p(Y) \tag{9.8}$$

for all $X, Y \in GL(m)$, the group of all $m \times m$ nonsingular matrices. The

THEOREM A [3, Theorem 3.6]. *Suppose* $n \geqslant 2$. *Suppose* $\mathbf{p} := (p_1, ..., p_n) \in [1, \infty]^n$ *is such that* $\sum_{k=1}^{n} 1/p_k \leqslant 1$. *Define the number* $c = c(\mathbf{p}) := \sum_{\{k: p(k) < \infty\}} 1/p'_k$. *Then there exists a constant* $A = A(\mathbf{p})$ *which is a function only of* $\mathbf{p}$, *such that the following statement holds: If* $(\Omega, \mathscr{M}, P)$ *is a probability space,* $\mathscr{F}_1, ..., \mathscr{F}_n$ *are σ-fields* $\subset \mathscr{M}$, *and* $B: \mathscr{S}(\mathscr{F}_1) \times \cdots \times \mathscr{S}(\mathscr{F}_n) \to \mathbb{C}$ *is an n-linear product form, then* $\|B\|_{\mathbf{p}} \leqslant A \cdot d_{\mathbf{p}}(B) \cdot [1 - \log d_{\mathbf{p}}(B)]^c$.

Here $c := 0$ if $\mathbf{p} = (\infty, ..., \infty)$. The main result of this section is as follows:

THEOREM 3.1. *Suppose* $n \geqslant 2$, *and* $\mathbf{p} := (p_1, ..., p_n) \in [1, \infty]^n$. *Define the number* $c = c(\mathbf{p}) := \sum_{\{k: p(k) < \infty\}} 1/p'_k$. *Then there exists a positive constant* $a = a(\mathbf{p})$ *such that the following statement holds:*

For each t, $0 < t \leqslant 2^{-n}$, *there exists a probability space* $(\Omega, \mathscr{M}, P)$ *and σ-fields* $\mathscr{F}_1, ..., \mathscr{F}_n \subset \mathscr{M}$ *and an n-linear product form* $B: \mathscr{S}(\mathscr{F}_1) \times \cdots \times \mathscr{S}(\mathscr{F}_n) \to \mathbb{C}$ *(defined by* $B(f_1, ..., f_n) := E(f_1 \cdots f_n) - \prod_{k=1}^{n} Ef_k$*), such that* $d_{\mathbf{p}}(B) = t$ *and* $\|B\|_{\mathbf{p}} \geqslant a \cdot t(1 - \log t)^c$.

Remark 3.2. Several comments will be made:

(a) The assumption $\sum_{k=1}^{n} 1/p_k \leqslant 1$ in Theorem A is not required in Theorem 3.1.

(b) The constant $c = c(\mathbf{p})$ in Theorem 3.1 is exactly the same as in Theorem A. Consequently, Theorem 3.1 shows that Theorem A is within a constant factor of being sharp, for any choice of parameters meeting the specifications in Theorem A. (This "constant factor" may depend on the parameters.) Consequently [3, Theorem 4.1(vi)] is sharp in the same sense, by Theorem 3.1 for $n = 2$. Theorem 3.1 also shows indirectly that [3, Theorems 2.1 and 2.2] are sharp in the same sense; for if this were not so, then (see the proof of Theorem A) an improvement in [3, Theorems 2.1 and 2.2] (beyond just a better constant factor) would lead to a similar improvement in Theorem A, contradicting Theorem 3.1.

(c) The n-linear form B in Theorem 3.1 was chosen partly for its simplicity. Because of the extensive role played by cumulants in the study of dependence between more than two random variables, it is natural to consider measures of dependence based on norms of cumulants. For example, Mase [11] studied the measure of dependence $d_{(\infty,\infty,\infty,\infty)}(\mathrm{Cum})$ between four σ-fields, where Cum denotes the 4th-order cumulant. Theorem 3.1 holds with B defined by $B(f_1, ..., f_n) = \mathrm{Cum}(f_1, ..., f_n)$ (the nth-order cumulant). Because of our proof, this will be a trivial corollary of Theorem 3.1 itself; in our proof the construction will be such that any $n - 1$

[7] FANG, K.-T., AND FANG, B.-Q. (1986). A new family of multivariate exponential distributions. Academia Sinica, Beijing.
[8] FREEDMAN, D. A. (1963). On the asymptotic behaviour of Bayes estimates in the discrete case. *Ann. Math. Statist.* **34** 1386–1403.
[9] GUTTMAN, I., AND TIAO, G. C. (1965). The inverted Dirichlet distribution with applications. *J. Amer. Statist. Assoc.* **60** 793–805.
[10] HERZ, C. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.
[11] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.
[12] JAMES, I. R., AND MOSIMANN, J. E. (1980). A new characterization of the Dirichlet distribution through neutrality. *Ann. Statist.* **8** 183–189.
[13] JOHNSON, N. L., AND KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York.
[14] KARLIN, S. (1968). *Total Positivity*, Vol. I. Stanford Univ. Press, Stanford, CA.
[15] KARLIN, S., AND RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467–498.
[16] KARLIN, S., AND RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. II. Multivariate reverse rule distributions. *J. Multivariate Anal.* **10** 499–516.
[17] KELKER, D. (1970). Distribution theory of spherical distributions and a location-scale parameter. *Sankhyā Ser. A* **32** 419–430.
[18] KHATRI, C. G. (1970). A note on Mitra's paper "A density-free approach to the matrix variate beta distribution." *Sankhyā Ser. A* **32** 311–318.
[19] KIMBALL, A. W. (1951). On dependent tests of significance in the analysis of variance. *Ann. Math. Statist.* **22** 600–602.
[20] KREYSZIG, E. (1983). *Advanced Engineering Mathematics*, 5th ed. Wiley, New York.
[21] KUMAR, A. (1970). Abstract no. 141: A matrix variate Dirichlet distribution. *Sankhyā Ser. A* **32** 133.
[22] MARSHALL, A. W., AND OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.
[23] MITRA, S. K. (1970). A density-free approach to the matrix variate beta distribution. *Sankhyā Ser. A* **32** 81–88.
[24] OLKIN, I., AND RUBIN, H. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.
[25] RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.
[26] RAO, C. R. (1983). Deny's theorem and its applications to characterizations of probability distributions, in *Festschrift for E. L. Lehmann*, pp. 348–366. Wadsworth, Belmont, CA.
[27] RICHARDS, D. ST. P. (1982). *Exponential Distributions on Abelian Semigroups.* Mimeo. Series 1387, Institute of Statistics, Univ. of North Carolina.
[28] RICHARDS, D. ST. P. (1984). Hyperspherical models, fractional derivatives, and exponential distributions on matrix spaces. *Sankhyā Ser. A* **46** 155–165.
[29] RICHARDS, D. ST. P. (1986). Positive definite symmetric functions on finite dimensional spaces. I. Applications of the Radon transform. *J. Multivariate Anal.* **19** 280–298.
[30] ROONEY, P. G. (1972). On the ranges of certain fractional integrals. *Canad. J. Math.* **24** 1198–1216.
[31] SIVAZLIAN, B. D. (1981). On a multivariate extension of the gamma and beta distributions. *SIAM J. Appl. Math.* **41** 205–209.
[32] WEYL, H. (1939). *The Classical Groups.* Princeton Univ. Press, Princeton, NJ.