Extreme Affine Transformations

Vittorio Gorini
Istituto di Fisica dell’Università, I-20133 Milano, Italy, and
Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy

E. C. G. Sudarshan
Department of Physics, Center for Particle Theory, The University of Texas, Austin, Texas 78712, USA,
and Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560012, India

Abstract. We classify the extreme points of the compact convex set of affine maps of $\mathbb{R}^n$ which map into itself the closed unit ball. This work is a preliminary step towards solving the problem of finding the extreme points of the compact convex set of affine maps of the $N \times N$ density matrices (dynamical maps of an $N$-level system) and for $n = 3$ furnishes the solution of the problem in the simplest case of a two-level system.

1. Introduction

Let $D_n(n=1, 2, 3, \ldots)$ denote the set of affine maps $\mathbb{R}^n \to \mathbb{R}^n$ which map into itself the closed unit ball $B_n$. $D_n$ is convex, compact and finite-dimensional, hence each point of $D_n$ can be written as a finite convex combination of extreme points of $D_n$. In this note we prove a theorem which classifies the extreme points of $D_n$. The theorem was stated and commented upon in [1] and is a first step towards solving the problem of finding the extreme points of the compact convex set $F_N$ of the affine maps $K_N \to K_N$, where $K_N = \{\rho|\rho \text{ an } N \times N \text{ complex matrix, } \rho \geq 0, \text{Tr}(\rho) = 1\}$ is the convex set of $N \times N$ density matrices. Indeed, $F_N$ can be identified to $D_3$ through the identification of $K_2$ to $B_2$ by means of the representation of a $2 \times 2$ density matrix as $\rho = (1/2)(1_2 + \sum_{i=1}^{2} \sigma_i \rho \sigma_i) = \{\sigma_1, \sigma_2, \sigma_3\}$, where $\{\sigma_1, \sigma_2, \sigma_3\}$ are the familiar Pauli matrices or, more generally, any maximal set of $2 \times 2$ self-adjoint traceless matrices satisfying $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$. The structure analysis of $F_N$ is of interest in connection with the study of the dynamics of an $N$-level quantum mechanical open system, since the dynamical evolution of such a system is represented by a one parameter family $t \to A_t$, $t \in [0, \infty)$, $A_t \in F_N$, $A_0 = 1$, whereby the density matrix (state) $\rho_0$ of the system at time $t$ is given in terms of the initial state $\rho_0$ by $\rho_t = A_t \rho_0 A_t^*$ (for this reason, we refer to the elements of $F_N$ as dynamical maps [2]). Familiar examples are encountered in spin magnetic resonance and relaxation [3, 4] and in quantum optics [5, 6].

After the completion of this work we became aware that, as a particular case of our theorem, a result equivalent to the classification of the extreme points of

* The bulk of this work was performed while the first author was visiting the Center for Particle Theory of the University of Texas at Austin under the partial support of the U.S.A.E.C. under contract ORO-40-1) 3992. A travel grant under the Fulbright-Hays program is acknowledged.
$D_n$ had been previously obtained by Størmer [7]. However, the geometrical aspect of the problem and the symmetry properties of the extreme points are not readily apparent in Størmer’s treatment, since he works in a dual context. On the other hand, we feel that symmetry considerations should play an important role in the determination of the extreme points of $F_n$. We refer to [1] for a discussion thereof and for an explicit (though as yet unproved) conjecture in this connection.

In Section 2 we collect a few notations. In Section 3 we give two instrumental parametrizations of $D_n$ (Theorem 1). In Section 4 we determine the extreme points of $D_n$ (Theorem 2). In Section 5 we briefly comment upon the geometrical meaning of Theorems 1 and 2.

2. Notations

If $n$ is a positive integer, $\mathbb{R}^n = \{x|x = \{x_i\}_{i=1}^{n}; x_i \in \mathbb{R}, i = 1, \ldots, n\}$ is the $n$-dimensional euclidean space and we denote by $M(n)$ [respectively, by $\text{AF}(n)$] the real algebra of linear maps (respectively of affine maps) of $\mathbb{R}^n$ into itself. An element $A$ of $\text{AF}(n)$ acts on $\mathbb{R}^n$ as $A: x \mapsto Ax + b = (b, T)x, x \in \mathbb{R}^n, b \in \mathbb{R}^n, T \in M(n)$ and we can identify $A$ to the pair $(b, T)$, where $T$ can in turn be identified to an $n \times n$ matrix with real entries $(T_{ij})_{i,j=1,\ldots,n}$ (we refer to $b$ and $T$ respectively as the translation and the linear parts of $A$). This establishes a canonical topological vector space isomorphism between $\text{AF}(n)$ [respectively, $M(n)$] and $\mathbb{R}^{n(n+1)}/(n)$ (respectively $\mathbb{R}^n$). We use the standard notations for the real orthogonal group in $n$ dimensions and for its connected component, respectively $O(n) = \{Q|Q \in M(n), QQ^T = I_n\}$ and $\text{SO}(n) = \{Q|Q \in O(n), \det Q = 1\}$ ($A^T$ denotes the transpose of a matrix $A$). Whenever $Q \in O(n)$, we write $Q$ in place of $(0, Q)$ and if $G$ is a subgroup of $O(n)$ and $x \in \mathbb{R}^n$ we denote by $G_x$ the stabilizer of $x$ relative to the canonical action of $G$ on $\mathbb{R}^n$. $1_n$ and $0_n$ denote respectively the identity and the zero map of $\mathbb{R}^n$ and $\text{diag}\{a_i\}_{i=1}^{n}$ denotes a diagonal matrix with diagonal elements $a_1, \ldots, a_n$. If $X$ is a convex subset of $\mathbb{R}^n$ we denote by $\text{extr} X$ the set of the extreme points of $X$. $B_n = \{x|x \in \mathbb{R}^n, \|x\| = \|x\| = 1\}$ and $S_n = \text{extr} B_n = \{x|x \in \mathbb{R}^n, \|x\| = 1\}$ are respectively the closed unit ball and the unit sphere in $\mathbb{R}^n$. We define $D_n = \{A|A \in \text{AF}(n), x \in B_n \Rightarrow A x \in B_n\}$. $D_n$ is a compact convex subset of $\text{AF}(n)$, whose boundary is given by $D_n = \{A|A \in D_n, x \in S_n\}$ for some $x \in S_n$. We call an element $A = (a, A)$ of $\text{AF}(n)$ canonical if $a_i \geq 0, i = 1, \ldots, n$, and $A = \text{diag}\{\lambda_i\}_{i=1}^{n}$, $\lambda_i \geq \lambda_j \geq \cdots \geq \lambda_n \geq 0$. If $Y$ is a subset of $\text{AF}(n)$, we define $\bar{Y} = \{A|A \in Y, A$ canonical$\}$.

3. Two Parametrizations of $D_n$

The following theorem establishes two parametrizations of $D_n$ which will be used in the following section.

**Theorem 1.**

i) $D_n = \{(b, T), b \in \mathbb{R}^n; T \in M(n); (b, T) = (Q_1 a, Q_1 A Q_2); Q_1, Q_2 \in O(n)\}$

\[
\begin{align*}
& a_i = \beta \xi_i (1 - \omega \xi_i^2), i = 1, \ldots, n; \\
& A = \text{diag}\{a \beta \omega \xi_i^2 \sum_{j=1}^{m} \xi_j^2 \omega \xi_j^2\}_{i=1}^{n};
\end{align*}
\]

ii) $D_n = \{A|A = (a, A) \text{ canonical}; a_i \geq 0, i = 1, \ldots, n; A = \text{diag}\{\lambda_i\}_{i=1}^{n}, \lambda_i \geq \lambda_j \geq \cdots \geq \lambda_n \geq 0\}$.
0 \leq \alpha \leq 1; 0 \leq \beta \leq 1; 0 \leq \omega_0 \leq \ldots \leq \omega_1 = 1; 0 \leq \xi_i \leq 1, r = 1, \ldots, n;

\sum_{r=1}^n \xi_r = 1\}

ii) \quad D_n = \{(b, T) | b \in \mathbb{R}^n; T \in M(n); (b, T) = (Q_1 a, Q_1 A Q_2), Q_1, Q_2 \in O(n)\};

\begin{align*}
& a_i = \beta \xi_i (1 - \alpha \eta_i^2), i = 1, \ldots, n; \\
& A = \text{diag} \{x \beta \eta_1^2, \ldots, x \beta \eta_n^2\}, 0 \leq x \leq 1; 0 \leq \beta \leq 1; v > 0; 0 \leq \eta_0 \leq \ldots \leq \eta_1 = v^{-1}; \\
& 0 \leq \xi_i \leq 1, r = 1, \ldots, n; \sum_{r=1}^n \xi_r = 1, \sum_{r=1}^n \xi_r \eta_r = 1\}.
\end{align*}

Proof. Using the polar decomposition of a matrix \( A \in M(n) \) as \( A = QS, Q \in O(n), \) symmetric and positive [8], any element \( A \) of \( AF(n) \) can be written in the form \( A = (Q_1 a, Q_1 A Q_2), \) where \( (a, A) \) is canonical. Write

\[ D(x; \beta; \xi_1, \ldots, \xi_n; \omega_0, \ldots, \omega_n) = (\beta \xi_i (1 - \alpha \eta_i^2))_{i=1}^n, \text{diag} \{x \beta \eta_1^2, \ldots, x \beta \eta_n^2\}_{i=1}^n \]  

(3.1)

Then, in order to prove i), it is enough to show that

\[ D_x = \{D(A) | A \in AF(n); D(x; \beta; \xi_1, \ldots, \xi_n; \omega_0, \ldots, \omega_n) \}

\begin{align*}
& 0 \leq x \leq 1; 0 \leq \omega_0 \leq \ldots \leq \omega_1 = 1; 0 \leq \xi_i \leq 1, r = 1, \ldots, n; \sum_{r=1}^n \xi_r = 1\}.
\end{align*}

(3.2)

To this purpose, we first note that if \( x, y \) and \( z \) are elements of \( \mathbb{R}^n \) such that \( \|x\| = \|y\| = 1 \) and \( z^2 = 1 \), then the following identity holds

\[ \sum_{i=1}^n \left[ \left( \sum_{j=1}^n y_j z_i^2 \right)^2 x_i + y_i (1 - z_i^2) \right] = 1 - \sum_{i=1}^n \left( 1 - z_i^2 \right) \left( \sum_{j=1}^n y_j z_i^2 \right)^2 x_i - y_i z_i, \]

(3.3)

as can be readily verified by expanding the squares. Hence, under the conditions

\[ 0 \leq \omega_0 \leq \ldots \leq \omega_1 = 1; 0 \leq \xi_i \leq 1, l = 1, \ldots, n; \sum_{r=1}^n \xi_r = 1, \]

(3.4)

it follows from (3.3) setting \( y = \xi \) and \( z = \omega \) that

\[ D(1; 1; \xi; \eta) = (1; 1; \xi, \ldots, \xi_n; \omega_1, \ldots, \omega_n) \in D_x. \]

Note that if \( (1, 1; \xi; \eta) = (1, 1; \xi; \omega) \) then \( \xi = \omega \). With this in mind, if \( \xi = \omega \), then \( (1, 1; \xi; \omega) \) is an element of \( D_x \) and one has

\[ \{x | x \in S_m, \alpha(1, 1; \xi; \omega) = x \in S_m\} \]

\[ = \{x | x \in S_m, x_l = \xi \omega (\sum_{i=1}^n \xi_i^2 \omega_i) \} = x_i = 1, l = 1, \ldots, n, \]

(3.5)

\[ \text{and } \xi_i \geq \omega_i, i = 1, \ldots, n. \]

Now let \( A = (a, \alpha(1, 1; \xi; \omega)) \in D_x \) and distinguish two cases, according to whether \( \lambda_1 = 0 \) or \( \lambda_1 > 0 \). The first case implies \( \|a\| = 1 \) and is obtained by setting \( \omega = 0 \) in (3.2). If \( \lambda_1 > 0 \) define

\[ \lambda \chi = \lambda \chi_1, \quad j = 1, \ldots, n \]

(3.6)

and let \( \xi \in A(S_m) \), with \( \xi_i \geq 0, i = 1, \ldots, n \), and let \( \xi_i = \xi_{i-1} = 0 \) and \( \xi_i > 0, i = 2, \ldots, n \). Then \( \xi_i = \ldots = \xi_{i-1} = 0 \) and \( \xi_i > 0, i = 2, \ldots, n \). Then

\[ \xi = \ldots = \xi_{n-1} = 0, \]

and hence \( \sum_{i=1}^n \xi_i^2 = \sum_{i=1}^n \xi_i \omega_i = 1 \). This
implies $a_i = b_i = 0$, $i = 1, \ldots, s - 1$, which contradicts the hypothesis. Then set
\[ \alpha = \lambda_i \left( \sum_{i=1}^{s} \xi_i \omega_i^2 \right)^{-1}, \]
whence $\lambda_i = \alpha \omega_i \left( \sum_{i=1}^{s} \xi_i \omega_i^2 \right)^{1/s}$, $i = 1, \ldots, n$, and consider the affine map $A(\alpha; 1; \xi; \omega)$. One has
\[ A(\alpha; 1; \xi; \omega) = A(1; 1; \xi; \omega) + (1 - \alpha) A(0; 1; \xi; \omega) \]  
and setting
\[ v_i = \xi_i \omega_i \left( \sum_{i=1}^{n} \xi_i \omega_i^2 \right)^{-1/2}, \quad i = 1, \ldots, n, \]  
one gets
\[ (A(\alpha; 1; \xi; \omega)v_i) = \xi_i, \quad i = 1, \ldots, n. \]  
Therefore, the two affine maps $A$ and $A(\alpha; 1; \xi; \omega)$ have the same linear part, the point $\xi$ belongs to $S_\alpha \cap d(S_\alpha) \cap S_\alpha$, $S_{\alpha}$, $A(S_{\alpha})$ and $A(\alpha; 1; \xi; \omega)(S_{\alpha})$ all lie in one and the same, say $\sigma$, of the two closed half-spaces determined by the hyperplane $\pi$ which is tangent to $S_{\alpha}$ at $\xi$. Let $c$ and $d$ be $\{ \xi_i (1 - \omega_i) \}_{i=1}^{n}$ denote the translation parts of $A$ and, respectively, of $A(\alpha; 1; \xi; \omega)$ and set $e = -c - d$. We have $A(\alpha; 1; \xi; \omega) = \xi$ and let $x \in S_{\alpha}$ such that $Ax = \xi$. Then $Ae = -\xi e \in \sigma$ and $A(\alpha; 1; \xi; \omega)x = \xi + \xi e \in \sigma$. This implies $\xi - e \in \sigma$ which, in turn, implies $e = 0$ since, by hypothesis, $A \in D_n$. Hence $A = A(\alpha; 1; \xi; \omega)$. By (3.7) and since $D_n$ is convex we have $A(\alpha; 1; \xi; \omega) \subseteq D_n'$ if $\alpha \in [0, 1]$. On the other hand, it is easy to check that $A(\alpha; 1; \xi; \omega)x \notin B_r$ for some $x \in B_r$ if $\alpha > 1$. Indeed, set
\[ u_1 = -\xi_1 \omega_1 \sum_{i=1}^{s} \xi_i \omega_i^2)^{-1/2}, \quad u_i = \xi_i \omega_i \left( \sum_{i=1}^{n} \xi_i \omega_i^2 \right)^{-1/2}, \quad i = 2, \ldots, n \]  
and $\alpha = 1 + \epsilon$, $\epsilon > 0$. Then $\|A(\alpha; 1; \xi; \omega) - u_i\|^2 = 1 + 4d(\epsilon + 1)\xi_i > 0$ if $\xi_i > 0$. If $\xi_i = 0$, let $r$ be the smallest integer for which $\xi_r > 0$ and note that $\omega_r > 0$ since $\sum_{i=1}^{n} \xi_i \omega_i^2 > 0$. Consider the intersections $C = S_{\alpha} \cap \mathcal{Q}$ and $E = A(\alpha; 1; \xi; \omega)(S_{\alpha}) \cap \mathcal{Q}$, where $\mathcal{Q}$ is the 2-plane $\{ x \in \mathbb{R}^n; x_{2} = \cdots = x_{n-1} = 0, x_{1} = \xi_{1}, l = r + 1, \ldots, n \}$. $C$ and $E$ are respectively a circle and an ellipse whose equations are $C: x_1^2 + x_2^2 = \xi_1^2$ and $E: [x_1 - \xi_1 (1 - \omega_1^2)]^2/(\xi_2 \omega_2^2)^2 + [x_2/(\xi_2 \omega_2)]^2 = 1$. At their common point $(0, \xi)$ the second derivatives are respectively $C:(d^2x_2/dx_1^2)_{x_1=0} = -1/\xi_2$ and $E:(d^2x_2/dx_1^2)_{x_1=0} = -1/\xi_2$. In order that $A(\alpha; 1; \xi; \omega)(S_{\alpha}) \subseteq S_{\alpha}$ one must have $1/\xi_2 \leq 1/\xi_2$, or $\alpha \leq 1$. This completes the proof i). In order to prove ii) take without loss of generality $\sum_{i=1}^{n} \xi_i \omega_i^2 = 0$ in the parametrization i) and set $v = \sum_{i=1}^{n} \xi_i \omega_i^2$ and $\eta_1 = \omega_1 v^{-1/2}$, $l = 1, 2, \ldots, n$.

4. Extreme Points of $D_{n}$

We classify the extreme points of $D_n$ by means of the following

**Theorem 2.**

\[ \text{Extr} D_n = \{(b, T) | b \in \mathbb{R}^n, T \in M(n); (b, T) = (Q_1a, P_1AQ_2); Q_1, Q_2 \in O(n); (a, A) = A(1; 1; 0, \ldots, 0, (1 - \delta^2)^{1/2}, \delta; 1, 1, \ldots, 1); 0 \leq \delta \leq 1 \}. \]

**Proof.** For $n = 1$ the result is trivial, so we assume $n \geq 2$. First note that if $(b, T) \in \text{extr} D_n$ and $Q, Q' \in O(n)$, then $(Qb, QTQ') \in \text{extr} D_n$. Thus it is enough to
look for the extreme points of $D_n$ which are canonical, and these belong to $D_n^e$. If $\sum_{i=1}^n \xi_i^2 \omega_i^2 \neq 0$, we get from (3.7) that $A(\xi; 1; \xi; \omega)$ is not extreme if $0 < \alpha < 1$. Consider $A(0; 1; \xi; \omega)$. It is an extreme since it maps extreme points of $B_n$ to extreme points of $B_n$ and it is obtained by setting $\delta = 1, \kappa = 0$ in (4.1) and by choosing therein $Q$, such that $Q_{\phi} \rho = \xi$, where $\rho$ is the “north pole”,
\begin{equation}
p = \{0, \ldots, 0, 1\}.
\end{equation}

We now prove that $A(1; 1; \xi_1, \ldots, \xi_n; 1, \ldots, 1, \omega_n)$ is extreme if $0 < \xi_n < 1$. First we note that the statement is trivial if $\omega_n = 1$ and that if $\omega_n < 1$ the map
\begin{equation}
A(1; 1; \xi_1, \ldots, \xi_n; 1, 0, \ldots, 1, \omega_n)
\end{equation}
is not extreme since it equals the convex combination $[(1+\omega_n)/2] \Sigma + [(1-\omega_n)/2] \varepsilon$, where
\begin{equation}
P_j = \text{diag} \{e_i\}_{i=1, \ldots, n}, \quad e_i = 1 \quad \text{if} \quad i \neq j, \quad e_j = -1.
\end{equation}

Then, let
\begin{equation}
0 < \xi_n < 1, \quad 0 \leq \omega_n < 1
\end{equation}
and assume $A(\xi_n, \omega_n) = A(1; 1; \xi_1, \ldots, \xi_n; 1, \ldots, 1, \omega_n)$ to be a convex combination
\begin{equation}
A(\xi_n, \omega_n) = y A_1 + (1-y) A_2, \quad A_1, A_2 \in D_n, 0 < y < 1.
\end{equation}
From (4.4) we get $0 < [(1-\xi_n^2) + \xi_n^2 \omega_n^2]^{1/2}$ and
\begin{equation}
0 \leq \omega_n = \xi_n \omega_n [(1-\xi_n^2) + \xi_n^2 \omega_n^2]^{-1/2} < \xi_n.
\end{equation}

Defining $\Sigma = \{x \in \mathbb{R}^n; \|x\| = 1, x_n = \xi_n \}$ and $\Sigma = \{x \in \mathbb{R}^n; \|x\| = 1, x_n = \xi_n \}$ we have $A(\xi_n, \omega_n)(\Sigma) = \Sigma$ and one checks easily that if $A \in D_n$ and $A(\Sigma) = \Sigma$, then $A = A(\xi_n, \omega_n)$. Then, since $S_n = \text{extr} B_n$, we have that
\begin{equation}
u \in S \Rightarrow A(\xi_n, \omega_n) \nu = A_1 \nu = A_2 \nu,
\end{equation}

Write $A_1 = Q_1 A_1 Q_2$ with $A_1$ canonical, $A_1 = A(\xi; 1; \xi_1, \ldots, \xi_n; \omega_1, \ldots, \omega_n)$. From (4.7) we have $A_1 [Q_2(\xi)] = Q_1^{-1} \Gamma(\xi, \omega)$. Then, since $Q_2(\Sigma)$ and $Q_1^{-1}(\Sigma)$ are $(n-2)$-dimensional subspaces of $S_n$ from (3.5) and (3.7) we obtain $\xi = 1$ and $\omega_{n-1} = 1$. $Q_2(\Sigma)$ and $Q_1^{-1}(\Sigma)$ have radii respectively $(1-\xi_n^2)^{1/2}$ and $(1-\xi_n^2)^{1/2}$, where $u_n = \xi_n \omega_n [(1-\xi_n^2) + \xi_n^2 \omega_n^2]^{-1/2}$. Since $Q_1, Q_2 \in O(n)$, there follows $\xi_n = \xi_n$ and $u_n = u_n$. Hence also $\omega_n = \omega_n$. Therefore, we have $A_1 = A(\xi_n, \omega_n)$ and $Q_1 \rho = (-1)^i \rho_i$, where $i = 0$ or $i = 0, 1$ according to whether $\omega_n > 0$ or $\omega_n = 0$. Then
\begin{equation}
A_1 = Q A ((-1)^i \xi_n \omega_n)
\end{equation}
where $Q = Q_1 Q_2$ and, by (4.6), $A(\xi_n, \omega_n) \nu = Q A ((-1)^i \xi_n \omega_n) \nu, \forall \nu \in \Sigma$, which implies $Q = P_n$. Substituting into (4.8) gives $A_1 = A(\xi_n, \omega_n)$ which proves that under conditions (4.4) $A(\xi_n, \omega_n)$ is extreme.

Next we show that if $n \geq 3$ and $\sum_{i=1}^n \xi_i^2 \omega_i^2 > 0$, the map $A(1; 1; \xi_1, \ldots, \xi_n; \omega_1, \ldots, \omega_n)$ is not extreme if $\omega_{n-1} < 1$. To this purpose, we use parametrization ii) established in Theorem 1. Then, writing
\begin{equation}
\Gamma(\nu; \xi; \eta) = (\xi_i (1-\nu_i^2))_{i=1, \ldots, n}, \text{diag} \{\nu_i\}_{i=1, \ldots, n},
\end{equation}

\begin{equation}

\end{equation}
we must prove that $\Gamma(v; \xi, \eta)$ is not extreme if $\eta_{n-1} < v^{-\frac{1}{2}}$. First remark that the map (4.9) satisfies the following composition law

$$
\Gamma(v; \xi, \eta)\Gamma(v'; \eta'; \eta'') = \Gamma(v'v; \xi', \eta''),
$$

(4.10)

where we have used the notation $xy = \{x_i y_i\}_{i=1}^n$. Now, let $r$ be the smallest integer for which $\eta_r < v^{-\frac{1}{2}}$ (by hypothesis, $2 \leq r \leq n-1$). If $\eta_r = 0$, we have

$$
\Gamma(v; \xi; \eta) = (1/2)Q^{-1}\Gamma(v; \hat{\xi}, \eta)\hat{\eta} + (1/2)Q^{-1}P_P\Gamma(v; \hat{\xi}, \eta)Q
$$

where,

$$
\hat{\xi} = \{\xi_2, \ldots, \xi_{r-1}, 0, (\xi_r^2 + \xi_{r+1}^2)^{1/2}, \xi_{r+2}, \ldots, \xi_n\}, \hat{\eta} = v^{-\frac{1}{2}}, \hat{\eta}_{r+1} = 0
$$

and

$$
Q\xi = \hat{\xi}, Q \in SO(n), Q^{-1} \text{ diag } \{\eta_i\} = \text{ diag } \{\eta_i\}.
$$

If $\eta_r > 0$, set $\xi = \sum_{i=1}^{r-1} \xi_i^2 + v\eta_r^2 \sum_{i=r}^n \xi_i^2$ and note that

$$
\xi \geq \sum_{i=1}^{r-1} \xi_i^2 + v \sum_{i=r}^n \eta_i^2 \xi_i^2 + \sum_{i=r}^n \eta_i^2 \xi_i^2 = v > 0.
$$

Setting $\lambda = \xi^{-1}$ and $\tau = \lambda^{1/2}$, we have thus by hypothesis $\lambda > \tau > 0$ and we define the vectors $\eta'$ and $\eta''$ as $\eta'_1 = \ldots = \eta'_{r-1} = \lambda$, $\eta'_{r+1} = \ldots = \eta'_n = \tau$, $\eta''_{r+1} = \lambda^{-1}$, $\eta''_{r+1} = \ldots = \eta''_{n+1} = \tau$, $\eta''_{n+2} = \ldots = \eta''_{n+1} = \eta_r$, and $\xi''_{n+2} = \ldots = \xi''_{n+1} = \xi_r$. Then, since $\Gamma(v; \xi, \eta) \in D_2$ by hypothesis, setting $v' = \lambda^{-\frac{1}{2}}$ and $v'' = \lambda^\frac{1}{2}$, it is a straightforward matter to check that the maps $\Gamma(v'; \xi', \eta')$ and $\Gamma(v''; \xi'', \eta'')$ belong to $D_2$ and by (4.10) one gets $\Gamma(v; \xi, \eta) = \Gamma(v; \xi', \eta')\Gamma(v''; \xi'', \eta'')$. From this we obtain

$$
\Gamma(v; \xi, \eta) = [(1 + v^2 \eta_r)/2]A_2 + [(1 - v^2 \eta_r)/2]A_2,
$$

(4.11)

where

$$
A_1 = Q^{-1}\Gamma(v'; \xi', \eta')Q\Gamma(v''; \xi'', \eta''),
$$

(4.12)

$$
A_2 = Q^{-1}P\Gamma(v'; \xi, \eta')Q\Gamma(v''; \xi', \eta''),
$$

(4.13)

and, since $0 < \eta_r < v^{-\frac{1}{2}}, 0 < (1/2)(1 - v^2 \eta_r) < 1/2$. Let $M$ and $N$ denote the linear parts of $Q_A$ and, respectively, of $Q_A$. If $\xi = 0$ we can take $Q = 1_n$, hence $M_r = v^2 = -N_r$, implying $A_1 + A_2$. If $\xi > 0$, we get $M_r = v^2 \xi_{r+1} (\xi_r^2 + \xi_{r+1}^2)^{-1} = -N_r$ and $M_{r+1} = -v^2 \eta_{r+1} (\xi_r^2 + \xi_{r+1}^2)^{-1} = -N_{r+1}$, whence again $A_1 + A_2$ provided that $\xi_{r+1}$ and $\eta_{r+1}$ are not both zero. On the other hand, if $\xi = 0$ and $\xi_{r+1} = \eta_{r+1} = 0$, set $\xi = (\xi_1, \ldots, \xi_{r-1}, 0, \xi_r, \xi_{r+2}, \ldots, \xi_n)$, $\eta = \ldots = \eta_r = v^{-\frac{1}{2}}$, $\eta_{r+1} = \eta_r$, $\eta_{r+2} = \ldots = \eta_{n+1} = 0$ and let $Q$ be the rotation of $\pi/2$ in the $(\xi_r, \xi_{r+1})$-plane. Then $\Gamma(v; \xi, \eta)$ can be expressed as the following non trivial convex combination

$$
\Gamma(v; \xi, \eta) = (1/2)Q\Gamma(v; \xi', \eta')Q^{-1} + (1/2)Q_P\Gamma(v; \xi, \eta')Q^{-1}.
$$

(4.14)

It remains to show that $d(x) = d(1; 1; 0, \ldots, 0, 1; 1, \ldots, 1, x)$ is extreme if

$$
0 < x < 1.
$$

(4.15)
To this purpose, for a given $\kappa$ satisfying (4.15) we express $\mathcal{A}(\kappa)$ as a convex combination of extreme points of $D_n$,

$$\mathcal{A}(\kappa) = \sum_{i=1}^{s} \gamma_i \hat{A}_i, \quad 0 < \gamma_i < 1, \quad \forall \gamma_i, \hat{A}_i \in \text{ext} \, D_n, \quad i = 1, \ldots, s, \quad \sum_{i=1}^{s} \gamma_i = 1 \quad (4.16)$$

and we show that this implies $\hat{A}_i = \mathcal{A}(\kappa), \ i = 1, \ldots, s$. If $\mu$ denotes the normalized Haar measure on $\text{SO}(n)_p$, we get from (4.16)

$$\mathcal{A}(\kappa) = \sum_{i=1}^{s} \gamma_i \hat{A}_i + P \hat{A}_i P = \sum_{i=1}^{s} \gamma_i \hat{A}_i, \quad (4.17)$$

where $P = P_{n-1}$,

$$\hat{A}_i = \int \hat{Q} \hat{A}_i \hat{Q}^{-1} d\mu(\hat{Q}), \quad i = 1, \ldots, s \quad (4.18)$$

the integration being extended over $\text{SO}(n)_p$, and

$$\hat{A}_i = (1/2)(\hat{A}_i + P \hat{A}_i P), \quad i = 1, \ldots, s \quad (4.19)$$

The $\hat{A}_i$'s are invariant under $\text{O}(n)_p$, hence they have the form $\hat{A}_i = ((0, \ldots, 0, d_i), \ \text{diag}(b_i, \ldots, b_i))$ and since $\mathcal{A}(\kappa)p = \hat{P}$ and $p \in \text{ext} \, B_n$ we have $\mathcal{A}_i p = \hat{P}_i p$, $i = 1, \ldots, s$. Therefore $d_i = 1 - c_i$ and since $\hat{A}_i \in D_n, \ i = 1, \ldots, s$, the $c_i$'s and the $b_i$'s satisfy the inequalities $0 < c_i \leq 1$ and $c_i \geq b_i$, $i = 1, \ldots, s$. The first inequality follows from $\hat{A}_i(-p) \in B_n$. On the other hand, if it were $c_i < b_i$ one would get $\hat{A}_i x \notin B_n$, for some points $x$ of $B_n$. Then, from (4.17) we have

$$\kappa^2 = \sum_{i=1}^{s} \gamma_i c_i = \sum_{i=1}^{s} \gamma_i b_i^2 = \sum_{i=1}^{s} \gamma_i b_i^2 = \gamma^2 \quad (4.20)$$

Denoting by $\mathcal{A}$ any given $\hat{A}_i$, since by hypothesis $\hat{A} \in \text{ext} \, D_n$ it follows from the hitherto obtained results that it must be of the form

$$\mathcal{A} = Q_1 \mathcal{A}(\xi, \omega) Q_2; Q_1, Q_2 \in \text{O}(n); \ 0 \leq \omega \leq 1; \ 0 < \xi \leq 1 \quad (4.21)$$

where $d(\xi, \omega) = \mathcal{A}(1; 1; 0, \ldots, 0, (1 - \xi^2)^3, \xi, \xi, \ldots, 1, \omega)$. If $\omega = 1$ we have $\mathcal{A}(\xi, \omega) = \mathcal{A}_1$, hence $\mathcal{A} = Q_1 \mathcal{A}_1 Q_2$ and, from (4.18)–(4.20),

$$\mathcal{A}(\kappa) = (1/2) \int \hat{Q} \hat{Q}^{-1} d\mu(\hat{Q}) + (1/2) \int P \hat{Q} \hat{Q}^{-1} P \mu(\hat{Q}) \quad (4.22)$$

Applying both sides to the zero vector we get $1 - \kappa^2 = 0$ which contradicts (4.15).

If $\xi = 1$ we have $\mathcal{A} = \hat{Q}_1 \mathcal{A}(\omega)$, where $\hat{Q}_1 \in \text{O}(n)_p$. Then $\mathcal{A}(\kappa) = (1/2) d(d(\omega)) (\hat{Q} \mathcal{A}(\omega) \hat{Q}^{-1} + P \hat{Q} \hat{Q}^{-1} P \mu(\hat{Q}))$ and applying to the zero vector gives $\omega = \kappa$ so that, since $\mathcal{A}(\kappa)$ is non singular, we get $\mathcal{A} = (1/2) d(d(\omega)) (\hat{Q} \hat{Q}^{-1} + P \hat{Q} \hat{Q}^{-1} P \mu(\hat{Q}))$.

Taking the trace gives $n = \text{Tr}(\hat{Q})$ which implies $\hat{Q} = \hat{I}_n$ and therefore $\mathcal{A} = \mathcal{A}(\kappa)$.

Finally, consider the case

$$0 < \xi < 1, \quad 0 \leq \omega \leq 1 \quad (4.23)$$

Let $d(1) = (1/2) \int d(\omega) (\omega^2 + (1 - \xi^2)^2) (0, \ldots, 0, (1 - \xi^2)^2, \xi)$ and $d(2) = (1/2) \int d(\omega) (\omega^2 + (1 - \xi^2)^2, \xi)$. Since $\mathcal{A}(\xi, \omega)$ maps $d(1)$ to $d(2)$ [compare (3.9)] whereas $\mathcal{A}(\kappa) = \mathcal{A}(\kappa)$, we have from (4.21)

$$\mathcal{A} = \hat{Q}_2 \mathcal{A}(\xi, \omega, m_1, m_2) \hat{Q}_1, \quad (4.24)$$
where
\[ \bar{Q}_2, \bar{Q}_1 \in \text{SO}(n), \quad m_1 = \text{0 or 1}, \quad m_2 = \text{0 or 1} \]
and
\[ D(\xi; \omega; m_1, m_2) = (c, S), \]
where
\[ c_1 = \ldots = c_{n-2} = 0, \]
\[ c_{n-1} = (-1)^{m_2+1}\xi(1-\xi^2)^{\frac{1}{2}}(1-\omega^2), \]
\[ c_n = \xi(1-\omega^2), S_n = \ldots = S_{n-2, n-2} = \frac{1}{2}(1-\xi^2 + \xi^2\omega^2) \frac{1}{2}, \]
\[ S_{n-1, n-1} = (-1)^{m_1+m_2}\omega, S_{n-1} = (1-\xi^2) + \xi^2\omega^2, \]
\[ S_{n-1, n} = (-1)^{m_2}\xi(1-\xi^2)^{\frac{1}{2}}(1-\omega^2) \]
and \[ S_{ij} = 0 \text{ if } i \neq j \text{ and } (i, j) \neq (n-1, n). \] Hence
\[ M(x) = \frac{1}{2} \int P \bar{Q} D(\xi; \omega; m_1, m_2) Q^{-1} d\mu(Q) \]
\[ + \frac{1}{2} \int P \bar{Q} D(\xi; \omega; m_1, m_2) Q^{-1} P d\mu(Q), \quad \text{(4.24)} \]
where \[ \bar{Q} = \bar{Q}_1 \bar{Q}_2. \] Equating the \((n, n)\) matrix elements of the linear parts of the two sides of (4.24) gives
\[ x^2 = (1-\xi^2) + \xi^2\omega^2. \quad \text{(4.25)} \]
Introducing the \((n-1) \times (n-1)\) matrix
\[ E(\xi; \omega; m_1 + m_2) = \text{diag} \{ [(1-\xi^2) + \xi^2\omega^2]^4, \ldots, [(1-\xi^2) + \xi^2\omega^2]^4, (-1)^{m_1+m_2}\omega \}, \]
we get from (4.24)
\[ (1/2) \int_{\text{SO}(n-1)} P \bar{Q} E(\xi; \omega; m_1 + m_2) Q^{-1} d\mu(Q) \]
\[ + (1/2) \int_{\text{SO}(n-1)} P \bar{Q} E(\xi; \omega; m_1 + m_2) Q^{-1} P d\mu(Q) = x_{n-1}, \quad \text{(4.26)} \]
where we have used the same symbols for the restrictions of \(P\) and \(Q\) to \(\mathbb{R}^{n-1}\.\)
Taking the squares of the traces of both sides of (4.26) and using Schwartz's inequality gives
\[ (n-1)^2 x^2 = [\text{Tr}(\bar{Q} E(\xi; \omega; m_1 + m_2))]^2 \leq [\text{Tr}(\bar{Q}^2 Q)] \]
\[ \times [\text{Tr}(E(\xi; \omega; m_1 + m_2)^2] = (n-1)^2 (1-\xi^2)^2 + \xi^2\omega^2 + \omega^2 \]
whereby, using (4.25), we get \((1-\xi^2) + \xi^2\omega^2 \leq \omega^2\) which contradicts (4.22) \(\blacksquare\)

5. Geometrical Considerations

Among the extreme points of \(D\), are those which map \(S_n\) into itself (in the physical case \(n = 3\) they correspond to the transformations which map pure states to pure states). There are two types of such maps: those of the form \((0, Q), Q \in \text{O}(n)\), and those which map \(B_n\) onto a point of \(S_n\). They are obtained by setting \(x = 1\) and, respectively, \(x = 0\) and \(\delta = 1\) in (4.1). In the physical case \(n = 3\), \((0, Q)\) corresponds to a unitary transformation on the density matrices \(u \rightarrow u v u^*, v a^* = 1, \) if \(Q \in \text{SO}(3)\).
It corresponds to a transformation of the form \( q \rightarrow u q^T u^* \), \( m a^* = 1 \), if \( Q \in O(3) \), \( \det Q = -1 \). Transposition on the density matrices corresponds to the antisymmetry transformation \( \{ x_i \} \rightarrow \{ x_i \} \) on \( \mathbb{Q} \). (Consider the pure states \( q = \{ q_i = x_i x_j \} \), then \( q_{ij} \rightarrow x_i x_j = q_{ji} \) and extend by linearity).

We now describe the geometrical meaning of the parametrizations of \( D_n \) given in Theorem 1. Let \( A = (b, T) \) be an element of \( D_n \) and write \( (b, T) = (Q, a, \alpha, \lambda, A \alpha) \) with \( Q_1, Q_2 \in O(n) \) canonical, \( A = \text{diag} \{ \lambda_1, \lambda_2, \ldots, \lambda_n \} \). \((a, \alpha)\) maps \( S_n \) to an ellipsoid \( E_n \) whose axes have lengths \( \lambda_1, \lambda_2, \ldots, \lambda_n \) and whose center \( a \) lies in the positive cone. If \( \lambda_1 = 0 \), \( E_n \) degenerates to a point and \( A \) is extreme or not according to whether or not \( \alpha \in S_n \). Assume \( \lambda_1 > 0 \) and write \( a = \beta \xi_1 (1 - \alpha \nu_1) = \beta \xi_1 (1 - \alpha \eta_1) \) and \( \lambda_1 = \alpha \beta \nu \eta \left( \sum_{i=1}^{n-1} \xi_i (1 - \alpha \eta_i) \right)^2 = \alpha \beta \nu \eta \), \( i = 1, 2, \ldots, n \), as in Theorem 1. The geometrical meaning of the parameters \( \omega_1, \omega_2, \ldots, \omega_n = 0 \) is clear from the relation \( \omega_1 = \lambda_1 / \lambda_i \). As regards the vector \( \xi \), take \( \beta = 1 \) and \( \alpha < 1 \). Then \( E_n \cap S_n = \{ \xi \} \). By (3.9), the point \( \xi \) of \( S_n \) which is mapped to \( \xi \) by \( (a, \alpha) = (a; 1; \xi; \omega) \), \( \alpha < 1 \), is given by (3.8) and we have \( \eta_i = v_i \xi_i \). As an illustration, in the case \( n = 3 \), for fixed \( \xi_1 \) and \( \omega_2 = \omega_3 \) range in their domain \( 0 \leq \omega_2, \omega_3 \leq 1 \), the point \( \xi \) sweeps the spherical triangle whose vertices are the points \( \xi_1 (0, 0, 1) \) and \( \xi_1 (\xi_1^2 + \xi_2^2, 0, 0, 0) \). \( \beta \) and \( \alpha \) are parameters of convex combinations. Indeed we have i) \( A(\alpha, \beta; \xi; \omega) = \beta A(\alpha; 1; \xi; \omega) + (1 - \beta) A(\alpha; 0; \xi; \omega) \) [note that \( A(\alpha; 0; \xi; \omega) = (0, 0, 1) \)] and ii) \( A(\alpha; 1; \xi; \omega) = (1 - \alpha) A(\alpha; 1; \xi; \omega) + (1 - \alpha) A(\alpha; 0; \xi; \omega) \) [see (3.7) and note that \( \xi \in A(1; 1; \xi; \omega)(S_n) \cap S_n \) and that \( A(0; 1; \xi; \omega) \) maps \( S_n \) to \( \xi \)]. Now take \( \alpha = \beta = 1 \) and \( \xi_1 > 0 \). Then, it is seen from (3.5), if \( \omega_1 = 0 \) and \( \omega_{n+1} = 1 \) the intersection \( E_n \cap S_n \) is an \((n - 1)\)-dimensional sphere and we obtain an extreme map if \( s = n - 1 \) [\( \delta < 1 \) in (4.1)]. The remaining extreme maps are obtained as the limit of the latter as \( \omega_n \rightarrow 1 \) for which the \((n - 2)\)-dimensional sphere \( E_n \cap S_n \) degenerates to the "north pole" \( p = (0, 0, 0, 0) \) \([\delta = 1 \) in (4.1)]. To be specific, divide \( D_n \) into the two subsets \( A \) and \( B \) which correspond to taking \( \delta = 1 \) and, respectively, \( n < \delta < 1 \), \( \delta < 1 \) in (4.1): \( A = \{ A(1, \alpha) \} \) \( \alpha \leq 1 \) and \( B = \{ A(\delta, \alpha) \} \) \( \delta < 1 \). We have \( \delta (\alpha) = \{ S_n \} \cap S_n \) is the \((n - 2)\)-dimensional hypersphere \( \Sigma = \{ x \in S_n, x_\delta = 0 \} \). Now assume \( \delta \) to be an element of \( D_n \) such that \( \{ S_n \} \cap S_n \) is a point \( q \) and assume that \( A \) can be expressed as a non trivial convex combination \( A = \gamma A_2 + (1 - \gamma) A_2 \) of elements of \( D_n \). Then, there is at least one direction in the hyperplane which is tangent to \( S_n \) at \( q \) along which either \( A_2 \) or \( A_2 \) has a larger curvature than \( A_2 \) has at \( q \) along the same direction. If \( A = A(1, \alpha) \) this is impossible since \( A(1, \alpha) \) has at \( q \) and along all directions the same curvature as \( S_n \). This explains intuitively why the elements of \( A \) are extreme. As to the elements of \( B \), if we write \( A(\delta, \alpha) \) as a convex combination \( A(\delta, \alpha) = \gamma A_1 + (1 - \gamma) A_2 \), we must have that \( A(\delta, \alpha), A_1 \) and \( A_2 \) agree on the \((n - 2)\)-dimensional hypersphere \( \Sigma = \{ x \in S_n, x_\delta = 0 \} \), where \( u_\delta \) is given by (4.6) with \( \xi_\delta = \delta, \omega_\delta = \alpha \). Here, the dimensionality of \( \Sigma \) is just large enough as to imply \( A_1 = A_2 = A(\delta, \alpha). \) On the other hand, it is no more so if \( A_1, A_2 \) and \( A = \gamma A_1 + (1 - \gamma) A_2 \) are to agree on an hypersphere of \( S_n \) whose dimension is less than \( n - 2 \) (except in the case when \( A = Q_1 \) \( Q_2 \) with \( \{ Q_1, Q_2 \} \in O(n) \) and \( \delta A = \delta A \)). Finally, we remark that the extreme elements of \( D_n \) have a high symmetry. Precisely, if \( (b, T) \in D_n \) is extreme, then there exists \( C \in O(n) \) and a subgroup of \( O(n) \), say \( I \), isomorphic to \( O(n - 1) \), such that \( QTC^{-1} Q^{-1} \) is a transformation for
every $Q \in \Gamma$. However, this condition is not sufficient for $(h, T)$ to be extreme, as
the example $\beta = \alpha = \omega_{n-1} = 1, \omega_2 < 1, \xi_n = 0$ shows.

Acknowledgements. One of us (V.G.) is greatly indebted to J.L. Richard for the continuous moral
support, for many fruitful discussions and, in particular, for suggesting that a suitable factorization
of the elements of $D_2$ could be used in order to prove that the map $A(1; 1, \xi_1, \ldots, \xi_n, \omega_1, \ldots, \omega_n)$ is not
an extreme point of $D_2$ if $\omega_{n-1} < 1$.

References

   Phys. Rev. 121, 920--924 (1961)
6. Agarwal, G.S.: Quantum statistical theories of spontaneous emission and their relation to other
   Chapter IX, Section 14, Theorem 9

Communicated by H. Araki

Received July 15, 1974; in revised form June 14, 1975