On Schrödinger’s bridge problem

S. Friedland

Abstract. In the first part of this paper we generalize Georgiou-Pavon’s result that a positive square matrix can be scaled uniquely to a column stochastic matrix which maps a given positive probability vector to another given positive probability vector. In the second part we prove that a positive quantum channel can be scaled to another positive quantum channel which maps a given positive definite density matrix to another given positive definite density matrix using Brouwer’s fixed point theorem. This result proves the Georgiou-Pavon conjecture for two positive definite density matrices, made in their recent paper. We show that the fixed points are unique for certain pairs of positive definite density matrices.

Bibliography: 15 titles.

Keywords: scaling of matrices, scaling of quantum channels, Schrödinger’s bridge problem, fixed points.

§1. Introduction

A column vector \( a \) is called a probability vector if \( a \) is a nonnegative vector and the sum of its coordinates is 1. An \((n \times n)\)-matrix \( A \) is called a stochastic matrix if each column of \( A \) is a probability vector. (Sometimes \( A \) is called column stochastic.) A square matrix \( A \) is called doubly stochastic if \( A \) and \( A^\top \) are stochastic matrices. The classical Schrödinger bridge problem, studied by Schrödinger in [12] and [13], seeks the most likely probability law for a diffusion process, in path space, that matches marginals at an initial and end point in time. The discrete version of Schrödinger’s bridge problem for Markov chains, as I understand, can be stated succinctly as follows.

Problem 1.1. Let \( A \) be a stochastic matrix. Assume that \( a \) and \( b \) are two positive probability vectors. Does there exists a scaling of \( A \), that is, \( B = D_1 A D_2 \), where \( D_1 \) and \( D_2 \) are two diagonal \((n \times n)\)-matrices with positive diagonal entries, such that \( B \) is stochastic and \( Ba = b \)?

See [6] for the motivation of this problem and also §2.

For \( a = b = 1, \ 1 = (1, \ldots, 1)^\top \), this problem is equivalent to the well-known problem: when can a nonnegative \( A \in \mathbb{R}^{n \times n}, A \geq 0 \), be scaled to a doubly stochastic matrix? This problem was answered by Sinkhorn [14]. Namely, it is always possible to scale a matrix \( A \) to a doubly stochastic matrix if all the entries of \( A \) are positive,
that is, $A > 0$. In that case $D_1$ and $D_2$ are unique (up to $tD_1$ and $t^{-1}D_2$ for $t > 0$). The unique scaling of a fully indecomposable $A \geq 0$ to a doubly stochastic matrix was proved in [11], [3] and [15]. ($A \geq 0$ is fully indecomposable if $PAQ$ is irreducible for any pair of permutation matrices $P$ and $Q$.) Necessary and sufficient conditions for the case $a = 1$ were given in [2] and [9]. Theorem 3.2 in [5] proves the existence of a unique scaling of a fully indecomposable $A$ for $a = b$.

In a recent paper Georgiou and Pavon [6] proved that Problem 1.1 is uniquely solvable for $A > 0$. Their proof is based on the strict contraction of a corresponding map with respect to the Hilbert metric. In the first part of our paper we generalize this result to a nonnegative $A$, such that $A$ has no zero column and $AA^\top$ is irreducible. We believe that our proof is shorter and more straightforward than the proof in [6], and this is the reason that we were able to generalize Georgiou and Pavon’s result above.

Georgiou and Pavon considered in [6] an analogue of Schrödinger’s bridge problem for quantum channels. The simplified version of this problem is: denote the set of complex-valued $(n \times n)$-matrices by $\mathbb{C}^{n \times n}$, and the subsets of Hermitian, positive semi-definite, and positive definite matrices in $\mathbb{C}^{n \times n}$ by $H_n$, $H_n,+$ and $H_n,++$, $H_n \supset H_n,+$ $\supset H_n,++$, respectively. Let $H_{n,+1}$ and $H_{n,++1}$, $H_{n,+1} \supset H_{n,++1}$ be the subsets of positive semi-definite and positive definite matrices of trace 1, respectively, in $H_{n,+}$. Let $\text{GL}(n, \mathbb{C}) \subset \mathbb{C}^{n \times n}$ be the group of invertible matrices. Recall that $Q: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ is called a completely positive operator if

$$Q(X) = \sum_{i=1}^{k} A_i X A_i^*, \quad A_i \in \mathbb{C}^{n \times n}, \quad i = 1, \ldots, k, \quad X \in \mathbb{C}^{n \times n}. \quad (1.1)$$

Assume that $Q$ is completely positive. $Q$ is called positive if $Q(H_{n,+1}) \subset H_{n,++}$. $Q$ is called a quantum channel if

$$\sum_{i=1}^{k} A_i^* A_i = I_n. \quad (1.2)$$

Note that a quantum channel is trace preserving, which is equivalent to $Q(H_{n,+1}) \subset H_{n,+1}$. An operator $R: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ is called a scaling of $Q$ if

$$R(X) = SQ(TXT^\top)^*S^* \quad \text{for fixed } S, T \in \text{GL}(n, \mathbb{C}) \text{ and all } X \in \mathbb{C}^{n \times n}. \quad (1.3)$$

The simplified version of Schrödinger’s bridge problem for quantum channels is the following.

**Problem 1.2.** Let $Q: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ be a positive quantum channel. Assume that $\alpha, \beta \in H_{n,++1}$. Does there exists a scaling of $Q$ to a quantum channel $R$ that satisfies $R(\alpha) = \beta$?

Problem 1.2 for $\alpha = \beta = \frac{1}{n} I_n$ was solved by Gurvits [7] and in [6] by different methods. (This is the analogue of Sinkhorn’s theorem.) Georgiou and Pavon also showed that Problem 1.2 is solvable when $\alpha$ and $\beta$ are two density matrices of rank one (pure states). However, they did not show that Problem 1.2 holds in other cases. Conjecture 1 in [6] implies the solution of Problem 1.2. In this paper we
show that Problem 1.2 is solvable using Brouwer’s fixed point theorem, similar to the methods in [6]. This result also yields a solution of the Georgiou-Pavon conjecture for two positive definite density matrices. We also show that for $\alpha$ and $\beta$ in some neighbourhood of $\frac{1}{n}I_n$, depending on $Q$, the scaling of $Q$ is ‘unique’ in a certain sense. That is, the corresponding map has a unique fixed point.

We now summarize briefly the contents of this paper. We discuss Schrödinger’s bridge problem for stochastic matrices and its generalization to certain nonnegative matrices in §2. Section 3 discusses some known results on quantum channels that are used in the next sections. In §4 we give a solution to Problem 1.2 using Brouwer’s fixed point theorem. We also show that a solution to Problem 1.2 is equivalent to Conjecture 1 in [6] for two positive definite density matrices. In §5 we show that the map constructed to solve Problem 1.2 has a unique fixed point if the two density matrices are in a neighbourhood of the uniform density matrix.

§2. Schrödinger’s bridge problem for stochastic matrices

Denote $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}^{++} = (0, \infty)$. For column vectors $\mathbf{u} = (u_1, \ldots, u_n)^\top$ and $\mathbf{v} = (v_1, \ldots, v_n)^\top$ let

$$\mathbf{u} \circ \mathbf{v} = (u_1v_1, \ldots, u_nv_n)^\top \quad \text{and} \quad D(\mathbf{u}) = \text{diag}(u_1, \ldots, u_n) \in \mathbb{R}^{n \times n}.$$ 

Schrödinger’s bridge problem is stated as follows (see [6]). Let $A = [a_{ij}] \in \mathbb{R}_+^{n \times n}$, and let two probability vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}_+^n$ be given. Do there exist $\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ such that

$$\mathbf{v} = A^\top \mathbf{u}, \quad \mathbf{y} = A\mathbf{x}, \quad \mathbf{a} = \mathbf{v} \circ \mathbf{x} \quad \text{and} \quad \mathbf{b} = \mathbf{u} \circ \mathbf{y}? \quad (2.1)$$

A straightforward calculation yields that if (2.1) holds then the matrix $B = D(\mathbf{u})AD(\mathbf{v})^{-1}$ satisfies:

$$B^\top 1 = 1, \quad Ba = b, \quad B \in \mathbb{R}_+^{n \times n}, \quad a, b \in \mathbb{R}_+^n, \quad 1^\top a = 1^\top b = 1. \quad (2.2)$$

Conversely, if $B = D(\mathbf{u})AD(\mathbf{v})^{-1}$ satisfies the above equalities then (2.1) holds, where $\mathbf{x}$ and $\mathbf{y}$ are determined uniquely by the last two conditions in (2.1).

Note that any $A \in \mathbb{R}_+^{n \times n}$ is uniquely scaled from the right to a stochastic matrix. That is there exists a unique $D_2 = D(A^\top 1)^{-1}$ such that $AD_2$ is stochastic. Hence without loss of generality we can assume that $A \in \mathbb{R}_+^{n \times n}$ is stochastic.

It is shown in [5], Theorem 3.2, that for a given fully indecomposable $A \in \mathbb{R}_+^{n \times n}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^n$ there exists a unique scaling of $A$ such that $D_1AD_2\mathbf{u} = \mathbf{u}$ and $D_2A^\top D_1\mathbf{v} = \mathbf{v}$. Choose $\mathbf{a} = \mathbf{b} = \mathbf{u}$ and $\mathbf{v} = 1$ to deduce the solution of Schrödinger’s bridge problem in this case. It is straightforward to show that a solution of Schrödinger’s bridge problem for $\mathbf{a} = \mathbf{b}$ implies the result in [5] cited above by considering the matrix $D(\mathbf{v})^{-1}D_1AD_2D(\mathbf{v})$ and $\mathbf{a} = \mathbf{b} = \mathbf{u} \circ \mathbf{v}$.

Denote the projective space associated with the open cone of positive vectors in $\mathbb{R}^n$ by $\mathbb{P}\mathbb{R}_+^{++}$. That is, $\mathbb{P}\mathbb{R}_+^{++}$ is the set of open rays $R(\mathbf{u}) = \{t\mathbf{u}, \ t > 0, \ \mathbf{u} \in \mathbb{R}_+^n\}$. Denote the simplex of probability vectors by $\Pi_n \subset \mathbb{R}^n$. Clearly, $\mathbb{P}\mathbb{R}_+^{++}$ is isomorphic to the interior of $\Pi_n$, denoted by $\Pi_{n,++} = \Pi_n \cap \mathbb{R}_+^{++}$. A one point compactification of $\Pi_{n,++}$ is the identification of the points on $\partial \Pi_{n,++}$, the boundary of $\Pi_{n,++} \subset \mathbb{R}^n$, to one point, denoted as $\infty$. Let $\overline{\Pi}_n = \Pi_{n,++} \cup \{\infty\}$ denote the above
one point compactification of $\Pi_n$. It is well known that $\overline{\Pi}_n$ is homeomorphic to the
$(n - 1)$-dimensional sphere $S^{n-1} = \{ x \in \mathbb{R}^n, \; x^\top x = 1 \}$. Moreover, $S^{n-1}$ can be viewed as the one point compactification of $\mathbb{R}^{n-1}$: $\overline{\mathbb{R}}^{n-1} = \mathbb{R}^{n-1} \cup \{ \infty \}$. Similarly, the one point compactification of $\mathbb{P}\mathbb{R}^n_++, \mathbb{P}\mathbb{R}^n_+ = \mathbb{P}\mathbb{R}^n_+ \cup \{ \infty \}$, is homeomorphic to $S^{n-1}$, equivalently homeomorphic to $\overline{\Pi}_n$.

Denote the convex set of stochastic matrices by $\Sigma_n \subset \mathbb{R}^{n \times n}_+$. Let $\Sigma_{n,++} = \Sigma_n \cap \mathbb{R}^{n \times n}_{++}$ be the interior of $\Sigma_n$.

Assume that $A \in \mathbb{R}^{n \times n}_+$ is a matrix with no zero column. Let

$$\Phi_A : \mathbb{R}^n_+ \to \Sigma_n, \quad \Phi_A(x) = D(x)AD(A^\top x)^{-1} \quad \text{for} \ x \in \mathbb{R}^n_+.$$  (2.3)

Clearly, for each $t > 0$ and $x \in \mathbb{R}^n_+$ we have $\Phi_A(tx) = \Phi_A(x)$. Hence $\Phi_A$ can be viewed as a smooth map $\overline{\Phi}_A : \mathbb{P}\mathbb{R}^n_+ \to \Sigma_n$. Clearly, $\overline{\Phi}_A$ can also be viewed as the restriction of $\Phi_A$ to $\Pi_n,++$. Note that if $A > 0$ then $\Phi_A(\mathbb{R}^n_+) \subset \Sigma_{n,++}$.

The following theorem generalizes Theorem 3 in [6].

**Theorem 2.1.** Let $A \in \mathbb{R}^{n \times n}_+$ and $a \in \Pi_{n,++}$ be given. Assume that $A$ does not have a zero column or zero row. Consider the map

$$\Phi_{A,a} : \Pi_{n,++} \to \Pi_{n,++}, \quad \Phi_{A,a}(x) = \Phi_A(x)a \quad \text{for} \ x \in \Pi_{n,++}. \quad (2.4)$$

1. Assume that $A$ is a positive matrix. Then the map $\Phi_{A,a}$ extends to a continuous map of $\Pi_n$. It maps the boundary of $\Pi_n$ to itself. Furthermore, $\Phi_{A,a}$ is a self-diffeomorphism of $\Pi_{n,++}$.

2. Assume that $AA^\top$ is an irreducible matrix. Then $\Phi_{A,a}$ is a diffeomorphism of $\Pi_{n,++}$ and $\Phi_{A,a}(\Pi_{n,++})$.

**Proof.** As $A$ is a nonnegative matrix with no zero column the map (2.4) is a well-defined smooth map. As $A$ does not have a zero row, $\Phi_A(x)$ does not have zero row for $x \in \Pi_{n,++}$. Therefore, $\Phi_A(x)a \in \Pi_{n,++}$. Hence $\Phi_{A,a}$ is a self-smooth map of $\Pi_{n,++}$, that is, $\Phi_{A,a}(\Pi_{n,++}) \subseteq \Pi_{n,++}$.

1. Assume that $A > 0$. Then $\Phi_{A,a} : \Pi_n \to \Sigma_n$ is continuous. In particular, $\Phi_{A,a}(x)a \in \Sigma_n$ for each $x \in \Pi_n$. Therefore, $\Phi_{A,a}$ is a continuous map of $\Pi_n$ to itself. Let $x = (x_1, \ldots, x_n)^\top \in \partial \Pi_n$. Thus $x_i = 0$ for some $i$. Hence the $i$th row of $\Phi_A(x)$ is zero. Therefore, the $i$th coordinate of $\Phi_{A,a}(x)a$ is zero. Thus $\Phi_{A,a}(x)a \in \partial \Pi_{n,++}$. Hence $\Phi_{A,a}(\partial \Pi_{n,++}) \subseteq \partial \Pi_{n,++}$. These results yield that $\Phi_{A,a} : \Pi_n \to \Pi_n$ induces a continuous map $\overline{\Phi}_{A,a} : \overline{\Pi}_n \to \overline{\Pi}_n$. More precisely the map $\Phi_{A,a} : \Pi_{n,++} \to \Pi_{n,++}$ is a proper map. That is, given a sequence $x_m \in \Pi_{n,++}$, $m \in \mathbb{N}$ which converges to $\partial \Pi_{n,++}$, the sequence $\Phi_{A,a}(x_m), m \in \mathbb{N}$, converges to $\partial \Pi_{n,++}$.

We now show that $\Phi_{A,a} : \Pi_{n,++} \to \Pi_{n,++}$ is a local diffeomorphism. Assume that $x \in \Pi_{n,++}$. Then a neighbourhood of $x$ in $\Pi_n$ consists of points of the form $x + tw$, $w^\top w = 1$, $w^\top w = 0$, $|t| < \varepsilon$ for some small $\varepsilon > 0$. Equivalently, let $W \subset \mathbb{R}^n$ be the subspace of all vectors orthogonal to $1$. Then any neighbourhood of $x$ is diffeomorphic to an open ball of radius $\varepsilon > 0$ centred at $0$ in $W$.

Suppose first that $A$ is stochastic. Assume that $x = \frac{1}{n}1$. Thus

$$A^\top(x + tw) = x + tA^\top w, \quad (A^\top(x + tw))_i = n(1 - ntA^\top w) + O(t^2),$$

$$D(A^\top(x + tw))^{-1} = n(I - ntD(A^\top w)) + O(t^2),$$

$$\Phi_A(x + tw) = A + nt(D(w)A - AD(A^\top w)) + O(t^2).$$
Assume that $Aa = b$. Then
\[
\Phi_{A,a}(x + tw) = b + nt(D(w)b - AD(A^\top w)a) + O(t^2)
\]
\[
= b + nt(D(b) - AD(a)A^\top)w + O(t^2).
\]
Let
\[
F(A, a) = D(b) - AD(a)A^\top = D(Aa) - AD(a)A^\top.
\]
Clearly $F(A, a)$ is a symmetric matrix satisfying
\[
F(A, a)1 = D(b)1 - AD(a)A^\top 1 = b - AD(a)1 = b - Aa = 0.
\]
Hence $W$ is an invariant subspace of $F(A, a)$. Furthermore, $nF(A, a)|W$ is
the Jacobian of $\Phi_{A,a}$ at $x = \frac{1}{n}1$. We claim that the $n - 1$
eigenvalues of $F(A, a)|W$ are positive. This claim follows from the observation
that $F(A, a)$ a symmetric irreducible singular $M$-matrix (see [4], §6.6). Indeed,
$rI - F(A, a) > 0$ for $r \gg 1$. Clearly $(rI - F(A, a))1 = r1$. Hence
$r$ is the spectral radius of $rI - F(A, a)$, $r$ is a simple eigenvalue of
$rI - F(A, a)$ and all other eigenvalues of $rI - F(A, a)$ are strictly less than $r$.
Hence $-F(A, a)$ is a singular symmetric matrix, which has $n - 1$ negative
values. In particular, $F(A, a)|W$ is an invertible transformation. Hence $\Phi_{A,a}$ is
a local diffeomorphism at $x = \frac{1}{n}1$. For a general $x \in \Pi_{n,++}$ it follows that
the Jacobian of $\Phi_{A,a}$ at $x$ is $nF(\Phi_{A}(x), a)|W$. Hence $\Phi_{A,a}$ is a local
diffeomorphism on $\Pi_{n,++}$.

As $\Phi_{A,a} : \Pi_{n,++} \to \Pi_{n,++}$ is a proper map and a local diffeomorphism it
follows that $\Phi_{A,a}$ is a proper cover of $\Pi_{n,++}$ by $\Pi_{n,++}$. As $\Pi_{n,++}$ is simply
connected we deduce that $\Phi_{A,a}$ is a diffeomorphism of $\Pi_{n,++}$. (This can also be
proved using degree theory [10].)

2. It remains to discuss the theorem when $A$ is a nonnegative matrix with some
zero entries, with no zero columns and $AA^\top$ irreducible. (This condition
yields that $A$ has no zero rows.) We claim first that $\Phi_{A,a}$ is a local diffeomorphism.
As above, the Jacobian of $\Phi_{A,a}$ at $x \in \Pi_{n,++}$ is $n$ times the restriction of
\[
F(\Phi_{A}(x), a) = D(\Phi_{A}(x)a) - D(a)\Phi_{A}(x)^\top
\]
to $W$. Clearly, $F(\Phi_{A}(x), a)1 = 0$. We claim that $F(\Phi_{A}(x), a)$ is an irreducible
singular $M$-matrix. Indeed, an $(i, j)$th off-diagonal entry of $-F(\Phi_{A}(x), a)$ is
positive if and only if the $(i, j)$th entry of $AA^\top$ is positive. Since $AA^\top$ is an irreducible
matrix it follows that $F(\Phi_{A}(x), a)$ is a symmetric irreducible singular $M$-matrix.
Hence the eigenvalues of $F(\Phi_{A}(x), a)|W$ are positive and the Jacobian of $\Phi_{A,a}$ at $x$
is invertible.

It remains to show that $\Phi_{A,a} : \Pi_{n,++} \to \Phi_{A,a}(\Pi_{n,++})$ is one-to-one. Assume
on the contrary that there exists $x_1, x_2 \in \Pi_{n,++}$, $x_1 \neq x_2$ such that $b = \Phi_{A,a}(x_1) = \Phi_{A,a}(x_2)$. Since $\Phi_{A,a}$ is a local diffeomorphism, the following conditions hold. There
exists a closed ball of a small radius $\varepsilon$ centred at $0$: $B(0, \varepsilon) \subset W$, such that
\[
x_1 + B(0, \varepsilon), x_2 + B(0, \varepsilon) \subset \Pi_{n,++} \quad \text{and} \quad (x_1 + B(0, \varepsilon)) \cap (x_2 + B(0, \varepsilon)) = \emptyset,
\]
and there exists $\varepsilon' > 0$ such that
\[
\Phi_{A,a}(x_1 + B(0, \varepsilon)) \cap \Phi_{A,a}(x_2 + B(0, \varepsilon)) \supset b + B(0, \varepsilon').
\]
Assume that $t > 0$ is small and let $A(t) = A + t11^\top > 0$. Then $A(t)$ is a positive perturbation of $A$. Therefore, $\Phi_{A(t), a}$ is a perturbation of the map $\Phi_{A, a}$ on the closed sets $x_1 + B(0, \varepsilon), x_2 + B(0, \varepsilon)$. The above condition yields that

$$\Phi_{A(t), a}(x_1 + B(0, \varepsilon)) \cap \Phi_{A(t), a}(x_2 + B(0, \varepsilon)) \supset b + B(0, \varepsilon_1)$$

for some $\varepsilon_1 \in (0, \varepsilon')$ for some very small and positive $t$. This contradicts the fact that $\Phi_{A(t), a}$ is a self-diffeomorphism of $\Pi_{n,++}$.

Theorem 2.1 is proved.

We now briefly discuss the results of Theorem 2.1. Clearly, part 1 of this theorem is equivalent to Georgiou-Pavon’s result for positive matrices [6]. We now consider part 2 of this theorem. First note that part 2 fails if $AA^\top$ is not irreducible. Indeed, suppose that $A$ is a permutation matrix. Then $\Phi_A(x) = A$ for each $x \in \Pi_{n,++}$. Therefore, $\Phi_{A, a}$ is a constant map. Note that $AA^\top = I$.

Suppose that $A$ is fully indecomposable. This is equivalent to the statement that $A = PBQ$, where $P$ and $Q$ are permutation matrices and $B = D + C$ such that $D$ is a diagonal matrix with positive diagonal entries and $C \geq 0$ is an irreducible matrix [3]. Then $AA^\top = P(D^2 + DC^\top + CD + CC^\top)B^\top$ is irreducible. It is easy to give an example of a fully indecomposable matrix $A$ such that $\partial \Phi_{A, a}(\Pi_{n,++}) \cap \Pi_{n,++} \neq \emptyset$.

Assume that

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad x(t) = \begin{bmatrix} tc_1 \\ tc_2 \\ 1 - tc_1 - tc_2 \end{bmatrix}, \quad c_1, c_2, t > 0, \quad c_1 + c_2 = 1, \quad (2.6)$$

and $t \searrow 0$. Then

$$\lim_{t \searrow 0} \Phi_A(x(t)) = B = \begin{bmatrix} 0 & c_1 & 0 \\ 0 & c_2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \lim_{t \searrow 0} \Phi_{A, a}(x(t)) = Ba. \quad (2.7)$$

Hence $0 < Ba \in \partial \Phi_{A, a}(\Pi_{3,++})$.

Theorem 3.2 in [5] is equivalent to the statement that if $A$ is fully indecomposable and $a \in \Pi_{n,++}$ then $a \in \Phi_{A, a}(\Pi_{n,++})$.

The results in [2] and [9] can be stated as follows. Given a fully indecomposable matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}_+$ and $b, c \in \Pi_{n,++}$ there exist two diagonal matrices $D_1, D_2 \in \mathbb{R}^{n \times n}_+$ with positive diagonal entries such that $D_1 AD_2 1 = nb$ and $D_2 A^\top D_1 1 = nc$ if and only if the following condition holds. There exists a matrix $C = [c_{ij}] \in \mathbb{R}^{n \times n}_+$ having the same pattern (that is, $a_{ij} > 0$ if and only if $c_{ij} > 0$, $i, j = 1, \ldots, n$) such that $C1 = nb$ and $C^\top 1 = nc$. Hence, if $A$ is fully indecomposable then $b \in \Phi_{A, \frac{1}{3}}(\Pi_{n,++})$ if and only if there exists a stochastic matrix $C \in \mathbb{R}^{n \times n}_+$ having the same pattern as $A$ such that $C1 = nb$.

Let $A$ and $B$ be defined as in (2.6) and (2.7) respectively. Clearly, $A$ is fully indecomposable. Let $b = B(\frac{1}{3}1) = \left(\frac{c_1}{3}, \frac{c_2}{3}, \frac{2}{3}\right)^\top$. It is straightforward to show that there is no stochastic matrix $C \in \mathbb{R}^{3 \times 3}_+$ with the same pattern as $A$ such that $C1 = 3b$. So $b \notin \Phi_{A, \frac{1}{3}}(\Pi_{3,++})$ as we claimed.
§ 3. Preliminary results on quantum channels

For $X \in H_n$ arrange the eigenvalues of $X$ in a nonincreasing order $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$. Let $\text{tr} \ X$ be the trace of $X$: $\text{tr} \ X = \sum_{i=1}^{n} \lambda_i(X)$. Denote the Frobenius norm of $X$ by $\|X\|_F = \sqrt{\text{tr} \ F^2} = \sqrt{\sum_{i=1}^{n} \lambda_i(X)^2}$. For two real numbers $a \leq b$ set

$$H_n(a, b) = \{X \in H_n, \ a \leq \lambda_n(X), \ \lambda_1(X) \leq b\}.$$

It is well known that $H_n(a, b)$ is a convex set. (Recall that $\lambda_1(X)$ and $-\lambda_n(X)$ are convex functions on $H_n$ [4].) Let

$$H_n(a, b, 1) := \{X \in H_n(a, b), \ \text{tr} \ X = 1\}. \quad (3.1)$$

Note that $H_n(a, b, 1) \neq \emptyset$ if and only if $a \leq 1/n \leq b$. Furthermore, given $a \leq 1/n$, for each $X \in H_n(a, b, 1)$ it follows that $\lambda_1(X) \leq 1 - (n - 1)a$. Clearly, for

$$0 < a \leq \frac{1}{n} \leq b \leq 1 - (n - 1)a \quad (3.2)$$

the convex set $H_n(a, b, 1)$ is a convex compact set in $H_{n,++}$. For a completely positive $Q$ let

$$a(Q) = \min_{X \in H_{n,++}} \lambda_n(QX) \quad \text{and} \quad b(Q) = \max_{X \in H_{n,++}} \lambda_1(QX). \quad (3.3)$$

Clearly, $Q(H_{n,++}) \subseteq H_n(a(Q), b(Q))$. Thus $Q$ is positive if and only if $a(Q) > 0$. For a positive quantum channel $Q$ the constants $a(Q)$ and $b(Q)$ satisfy (3.2) and $Q(H_{n,++}) \subseteq H_n(a(Q), b(Q), 1)$.

Let $\|X\| = \max_{\|x\|=1} \|Xx\|$ be the spectral norm of $X \in \mathbb{C}^{n \times n}$. Recall that

$$\|X\| = \lambda_1(X) \quad \text{and} \quad \lambda_n(X) = \frac{1}{\lambda_1(X^{-1})} = \frac{1}{\|X^{-1}\|} \quad \text{for} \ X \in H_{n,++}. \quad (3.4)$$

Furthermore, for each $X \in H_{n,+}$ there exists a unique positive semi-definite matrix $X^{1/2} \in H_{n,+}$, the square root of $X$, so that $(X^{1/2})^2 = X$.

**Lemma 3.1.** Assume that $\alpha \in H_{n,+} \setminus \{0\}$. Let $D_\alpha : H_{n,+} \rightarrow H_{n,++}$ be given by

$$D_\alpha(X) = \frac{1}{\text{tr} \ X^{-1}}X^{-1/2}\alpha X^{-1/2}. \quad (3.5)$$

If $\alpha \in H_{n,++}$ then $D_\alpha(H_{n,++}) \subseteq H_{n,++}$. Assume that $a$ and $b$ satisfy (3.2). Then

$$D_\alpha(H(a, b, 1)) \subseteq H(c, d, 1), \quad (3.6)$$

where

$$c = \frac{a\lambda_n(\alpha)}{a\lambda_n(\alpha) + (n - 1)b\lambda_1(\alpha)} \quad \text{and} \quad d = \frac{b\lambda_1(\alpha)}{b\lambda_1(\alpha) + (n - 1)a\lambda_n(\alpha)}.$$

**Proof.** Clearly, for $X \in H_{n,++}$ and $\alpha \in H_{n,+} \setminus \{0\}$ we have $X^{-1/2}\alpha X^{-1/2} \in H_{n,+} \setminus \{0\}$. Hence $D_\alpha(H_{n,++}) \subseteq H_{n,++}$. Furthermore, if $\alpha \in H_{n,++}$ then $D_\alpha(H_{n,++}) \subseteq H_{n,++}$. 


Lemma 3.2. Assume that $\alpha \in H_{n,+}$. Clearly,
\[
\lambda_1(X^{-1/2} \alpha X^{-1/2}) = \|X^{-1/2} \alpha X^{-1/2}\| \leq \|X^{-1/2}\|^2 \|\alpha\| = \frac{\lambda_1(\alpha)}{\lambda_n(X)} \leq \frac{\lambda_1(\alpha)}{\alpha}.
\]
As $\lambda_n(X^{-1/2} \alpha X^{-1/2}) = 1/\lambda_1(X^{1/2} \alpha^{-1} X^{1/2})$ it follows that
\[
\lambda_n(X^{-1/2} \alpha X^{-1/2}) \geq \frac{\lambda_n(\alpha)}{\lambda_1(X)} \geq \frac{\lambda_n(\alpha)}{b}.
\]
Observe next that if $0 < f \leq x_n \leq \cdots \leq x_1 \leq g$ then
\[
\frac{f}{f + (n-1)g} \leq \frac{x_n}{\sum_{i=1}^{n} x_i} \leq \cdots \leq \frac{x_1}{\sum_{i=1}^{n} x_i} \leq \frac{g}{g + (n-1)f}.
\]
Combine the above results to deduce (3.6). The continuity argument yields (3.6) for $\alpha \in H_{n,+}$. 

For a positive integer $k$ let $[k] := \{1, \ldots, k\}$. Recall that there is an inner product $\langle X, Y \rangle = \text{tr} XY^*$ on $\mathbb{C}^{n \times n}$. For a completely positive operator $Q$ given by (1.1) let
\[
Q'(X) = \sum_{i=1}^{k} A_i^* X A_i, \quad A_i \in \mathbb{C}^{n \times n}, \quad i \in [k].
\] (3.7)
Then $Q'$ is the dual of $Q$. Indeed, $\langle Q(X), Y \rangle = \langle X, Q'(Y) \rangle$. Clearly, $Q'$ is completely positive. Thus $Q$ is a quantum channel if and only if $Q'(I_n) = I_n$. $Q$ is called a unital channel if $Q'(I_n) = Q(I_n) = I_n$. A unital channel is an analogue of a doubly stochastic matrix, and is sometimes called a doubly stochastic quantum channel (see [8]).

For $A, B \in H_n$ we say $A \succeq B$, $A \succeq B$ or $A \succ B$ if $A - B$ is positive semi-definite, nonzero positive semi-definite or positive definite, respectively. Clearly, any completely positive operator $Q: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ is order preserving on $H_n$, that is,
\[
X \succeq Y \implies Q(X) \succeq Q(Y).
\]
Hence $Q'(X) \succeq 0$ if $X \succeq 0$.

If we assume that $X, Y \in H_{n,+}$ then $X^{1/2}YX^{1/2} \in H_{n,+}$. Therefore $\text{tr} XY = \text{tr} X^{1/2}YX^{1/2} \geq 0$. Hence if $U \preceq V$ it follows that $\text{tr} UY \preceq \text{tr} VY$.

Lemma 3.2. Assume that $Q: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ is a completely positive operator. Then
\[
a(Q) = a(Q') \quad \text{and} \quad b(Q) = b(Q').
\] (3.8)
In particular $Q$ is positive if and only if $Q'$ is positive.

Proof. From the definition of $a(Q)$ and $b(Q)$ given by (3.3) it follows that $a(Q)I_n \leq Q(X) \leq b(Q)I_n$ for each $X \in H_{n,+}$. From the definition of $a(Q')$ and $b(Q')$ it follows that there exists $U, V \in H_{n,+}$ and $u, v \in \mathbb{C}^n$, $\|u\| = \|v\| = 1$ such that
\[
a(Q') = \lambda_n(Q'(U)) = u^*Q'(U)u \quad \text{and} \quad b(Q') = \lambda_1(Q'(V)) = v^*Q'(V)v.
\]
Hence
\[ a(Q') = \text{tr} uu^* Q'(U) = \text{tr} Q(uu^*)U \geq \text{tr} a(Q) I_n U = a(Q) \]
and
\[ b(Q') = \text{tr} vv^* Q'(V) = \text{tr} Q(vv^*)V \leq \text{tr} b(Q) I_n V = b(Q). \]

As \((Q')' = Q\), it follows that \(a(Q) = a((Q')') \geq a(Q')\) and \(b(Q') \geq b((Q')') = b(Q)\). Hence (3.8) holds. In particular, \(a(Q) > 0\) if and only if \(a(Q') > 0\).

Lemma 3.2 and the arguments of the proof of Lemma 3.1 yield the following result.

**Corollary 3.1.** Assume that \(Q\) is a positive, completely positive operator given by (1.1). Let \(Q'\) be given by (3.7). Let \(\tilde{Q} : H_{n,+1} \to H_{n,+1}^+\) denote the following nonlinear operator
\[ \tilde{Q}(X) = \frac{1}{\text{tr} Q'(X)} Q'(X), \quad X \in H_{n,+1}. \] (3.9)

Then
\[ \tilde{Q}(H_{n,+1}) \subseteq H_n(e, f, 1), \] (3.10)
where
\[ \frac{a(Q)}{a(Q) + (n-1)b(Q)} \leq e \leq f \leq \frac{b(Q)}{b(Q) + (n-1)a(Q)}. \]

Assume that \(Q\) is a unital quantum channel. In this case \(\tilde{Q} = Q'\). Lemma 3.2 yields that in (3.10) we can assume that \(e = a(Q)\) and \(f = b(Q)\).

**§ 4. Existence of scaling for a positive quantum channel**

**Theorem 4.1.** Let \(Q : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}\) be a positive quantum channel. Assume that \(\alpha, \beta \in H_{n,+1}\). Consider the following continuous nonlinear transformation \(\Phi_{\alpha,\beta} : H_{n,+1} \to H_{n,+1}\):
\[ \Phi_{\alpha,\beta} = D_\alpha \circ \tilde{Q} \circ D_\beta \circ Q. \] (4.1)
Then \(\Phi_{\alpha,\beta}\) has a fixed point \(U(\alpha, \beta) \in H_{n,+1}\). If \(\alpha \in H_{n,+1}\) then \(U(\alpha, \beta) \in H_{n,+1}\).

**Proof.** Let \(\Psi_1 = D_\beta \circ Q, \Psi_2 = \tilde{Q} \circ D_\beta \circ Q\). As \(Q\) is positive, \(Q(H_{n,+1}) \subseteq H_n(a(Q), b(Q), 1)\). Hence \(\Psi_1\) is continuous on \(Q(H_{n,+1})\). Similarly, \(\Psi_2(H_{n,+1}) \subseteq \tilde{Q}(H_{n,+1}) \subseteq H_n(a(Q), b(Q), 1)\). Thus \(\Phi_{\alpha,\beta}\) is continuous on \(H_{n,+1}\). Brouwer’s fixed point theorem yields that \(\Phi_{\alpha,\beta}\) has a fixed point \(U(\alpha, \beta) \in H_{n,+1}\).

Assume finally that \(\alpha \in H_{n,+1}\). Then \(D_\alpha(H_n(a(Q), b(Q), 1)) \subseteq H_{n,+1}\). Hence \(U(\alpha, \beta) = H_{n,+1}\).

**Theorem 4.2.** Let \(Q : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}\) be a positive quantum channel. Assume that \(\alpha, \beta \in H_{n,+1}\). Then each fixed point of \(\Phi_{\alpha,\beta}\), given by (4.1) induces a scaling of the quantum channel \(Q\) to a quantum channel \(R\) satisfying \(R(\alpha) = \beta\). Conversely, each scaling of the quantum channel \(Q\) to a quantum channel \(R\) satisfying \(R(\alpha) = \beta\) induces a fixed point of \(\Phi_{\alpha_1,\beta_1}\), for \(\alpha_1 = O\alpha O^*\) and \(\beta_1 = P\beta P^*\) for some unitary \(O\) and \(P\).
Proof. Suppose that $\Phi_{\alpha,\beta}(U) = U$ for some $U \in \mathbb{H}_{n,+1}$. Set $V = Q(U)$, $W = D_\beta(V)$ and $Z = Q(W)$. Then $D_\alpha(Z) = U$. Observe that $Q(V) = t^{-2}Z$ for some $t > 0$. Let $T = tZ^{-1/2}$ and $S = W^{1/2}$. Let $R_1: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ be given by $R_1(X) = SQ(TXT^*)S^*$. Clearly, $R_1$ is a scaling of $R$. Furthermore

$$R'_1(I_n) = T^*Q'(S^*S)T = (tZ^{-1/2})Q'(W)(tZ^{-1/2}) = (tZ^{-1/2})(t^{-2}Z)(tZ^{-1/2}) = I_n.$$ 

Hence $R_1$ is a quantum channel. 

We now claim that there exists a unitary $O$ such that $R_1(\alpha) = O\beta O^*$. Indeed,

$$Q(U) = Q(D_\alpha(Z)) = Q\left(\frac{1}{\text{tr} Z^{-1/2}}Z^{-1/2}\alpha Z^{-1/2}\right) = \frac{1}{t^2\text{tr} Z^{-1/2}}Q(T\alpha^*) = V.$$ 

Hence $R_1(\alpha) = (st)^2W^{1/2}VW^{1/2}$ for some $s = \sqrt{\text{tr} Z^{-1/2}}$. The equality $W = D_\beta(V)$ is equivalent to $W = r^2V^{-1/2}\beta V^{-1/2}$ for $r = 1/\sqrt{\text{tr} V^{-1/2}}$. That is,

$$(W^{1/2})^2 = (rV^{-1/2}\beta^{1/2})(rV^{-1/2}\beta^{1/2})^*$$

$$\implies (rW^{-1/2}V^{-1/2}\beta^{1/2})(rW^{-1/2}V^{-1/2}\beta^{1/2})^* = I_n.$$ 

Therefore, $O := rW^{-1/2}V^{-1/2}\beta^{1/2}$ is a unitary matrix. Clearly,

$$V^{1/2}W^{1/2} = r\beta^{1/2}O^*.$$ 

Finally,

$$R_1(\alpha) = (st)^2W^{1/2}VW^{1/2} = (st)^2(W^{1/2}V^{1/2})(W^{1/2}V^{1/2})^* = (rst)^2O\beta O^*.$$ 

As $\alpha, \beta \in \mathbb{H}_{n,+1}$ and $R_1$ is trace preserving, it follows that $(rst)^2 = 1$. Hence $R_1(\alpha) = O\beta O^*$. Define $R(X) = O^*R_1(X)O$. Clearly, $R$ is a scaling of $Q$, $R$ is a quantum channel and $R(\alpha) = \beta$.

Assume now that $R$ is a scaling of the quantum channel $Q$ such that $R$ is a quantum channel, and $R(\alpha) = \beta$. Let $R$ be given by (1.3). Define $W := SS^*S$, where $s = 1/(\text{tr} S^*S)$. Let $Z = Q(W)$. As $R$ is a quantum channel, we deduce that $Z = t(T^*)^{-1}T^{-1}$, where $t = 1/(\text{tr}(T^*)^{-1}T^{-1})$. From the above arguments it follows that

$$Z^{1/2} = t^{1/2}(T^*)^{-1}O^* = t^{1/2}OT^{-1}$$

for some unitary matrix $O$. Let $\alpha_1 = O\alpha O^*$. Then

$$U = D_{\alpha_1}(Z) = r^{-2}TO^*\alpha_1OT^* = r^{-2}T\alpha T^*,$$

$$r^2 = \text{tr} T\alpha T^*.$$ 

Then $U = r^{-2}T\alpha T^*$. As $R(\alpha) = \beta$ it follows that $V = Q(U) = r^{-2}S^{-1}\beta(S^*)^{-1}$. Hence $V^{1/2} = r^{-1}S^{-1}\beta^{1/2}P^* = r^{-1}P\beta^{1/2}(S^*)^{-1}$ for some unitary $P$. Let $\beta_1 = P\beta P^*$. Then

$$D_{\beta_1}(V) = qS^*\beta^{-1/2}P^*\beta_1P\beta^{-1/2}S = qS^*S,$$

$$q = \frac{1}{\text{tr} S^*S}.$$ 

Therefore, $D_{\beta_1}(V) = W$, and $U$ is a fixed point of $\Phi_{\alpha_1,\beta_1}$. 
Note that if $Q$ is a quantum channel that satisfies $Q(\alpha) = \beta$, then the scaled channel $R$ given by (1.3), with $S$ and $T$ unitary, is also a quantum channel with $R(T^*\alpha T) = S\beta S^*$. This observation explains the second part of Theorem 4.2.

Theorem 4.2 proves the Georgiou-Pavon conjecture (see [6], Conjecture 1) for two positive definite density matrices.

**Conjecture 4.1** (Georgiou and Pavon). Given a positive quantum channel $Q: \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ and two density matrices $\rho_0$ and $\rho_T$, there exist $\phi_0, \phi_T, \hat{\phi}_0, \hat{\phi}_T \in H_{n,+}$ such that

\[
Q(\phi_T) = \phi_0, \quad Q'(\hat{\phi}_0) = \hat{\phi}_T,
\]

\[
\rho_0 = \chi_0 \phi_0 \chi_0^*, \quad \rho_T = \chi_T \phi_T \chi_T^*,
\]

\[
\phi_0 = \chi_0 \phi_0 \chi_0, \quad \phi_T = \chi_T \phi_T \chi_T.
\]

Furthermore, $\chi_0$ and $\chi_T$ can be chosen to be in $H_{n,+}$.

It is shown in [6] that this conjecture holds for $\rho_0 = \rho_T = \frac{1}{n} I_n$, see also [7], and for rank one matrices $\rho_0, \rho_T \in H_{n,+}$.

**Proof of the Georgiou-Pavon conjecture for $\rho_0, \rho_T \in H_{n,+}$ and non-Hermitian $\chi_0$ and $\chi_T$.** Let $\alpha = \rho_T$ and $\beta = \rho_0$. Theorem 4.1 implies that $\Phi_{\alpha,\beta}$ has a fixed point in $H_{n,+}$. In the proof of Theorem 4.2 we constructed the quantum channel $R_1(X) = SQ(TXT^*)S^*$ such that $R_1(\alpha) = O\beta O^*$ for a corresponding unitary $O$. Furthermore, $S, T \in H_{n,+}$. Thus

\[
Q(\rho_T) = S^{-1} O\beta O^* S^{-1}.
\]

Since $R_1$ is a quantum channel, it follows that $Q'(S^2) = T^{-2}$. Choose

\[
\phi_T = TaT, \quad \phi_0 = S^{-1} O\beta O^* S^{-1}, \quad \hat{\phi}_0 = S^2, \quad \hat{\phi}_T = T^{-2}, \quad \chi_0 = \beta^{1/2} O^* S^{-1} \quad \text{and} \quad \chi_T = \alpha^{1/2} T.
\]

Then

\[
Q(\phi_T) = \phi_0, \quad Q'(\hat{\phi}_0) = \hat{\phi}_T,
\]

\[
\chi_0^* \chi_0 = S^{-1} O\beta^{1/2} \beta^{1/2} O^* S^{-1} = \phi_0, \quad \chi_T^* \chi_T = T \alpha^{1/2} \alpha^{1/2} T = \phi_T,
\]

\[
\rho_0 = \beta = \beta^{1/2} O^* S^{-1} S^2 S^{-1} O^{1/2} = \chi_0^* \phi_0 \chi_0^*,
\]

and

\[
\rho_T = \alpha = \alpha^{1/2} T T^{-2} T^{-1/2} = \chi_T^* \hat{\phi}_T \chi_T^*.
\]

We now note that Conjecture 4.1, under the assumption that $\rho_T$ and $\rho_0$ are positive definite, implies that $Q$ can be scaled to a quantum channel $R$ satisfying $R(\rho_T) = \rho_0$. Observe first that (4.3) yields that $\chi_0$ and $\chi_T$ are invertible. Let

\[
T = \chi_T^* \rho_T^{-1/2}, \quad S = \rho_0^{1/2} (\chi_0^*)^{-1} \quad \text{and} \quad R(X) = SQ(TXT^*)S^*.
\]

Then

\[
R(\rho_T) = SQ(T \rho_T T^*) S^* = SQ(\chi_T^* \rho_T^{-1/2} \rho_T \rho_T^{-1/2} \chi_T) S^* = SQ(\chi_T^* \chi_T) S^* = SQ(\phi_T) S^* = S \phi_0 S^* = S \chi_0^* \chi_0 S^* = \rho_0^{1/2} (\chi_0^*)^{-1} \chi_0^* \chi_0 \chi_0^{-1} \rho_0^{1/2} = \rho_0
\]
and

\[ R'(I_n) = T^*Q'(S^*S)T = T^*Q'(\chi_0^{-1}\rho_0(\chi_0^*)^{-1}) = T^*Q'(\hat{\phi}_0)T \\
= T^*\hat{\phi}_T T^* = \rho_T^{-1/2}\chi_T\hat{\phi}_T \chi_T^* \rho_T^{-1/2} \rho_T^{-1/2} = \rho_T^{-1/2} \rho_T^{-1/2} = I_n. \]

§ 5. Uniqueness of fixed points

Denote the space of all rays \( tX, X \in H_{n,++}, t > 0 \), by \( \mathbb{P}H_{n,++} \). Clearly, we can identify \( \mathbb{P}H_{n,++} \) with \( H_{n,++} \). The Hilbert metric on \( \mathbb{P}H_{n,++} \) is given as follows:

\[ \text{dist}(X, Y) = \log \lambda_1(XY^{-1}) - \log \lambda_n(XY^{-1}), \quad X, Y \in H_{n,++}. \tag{5.1} \]

To justify that \( \text{dist}(X, Y) \) is a metric on \( \mathbb{P}H_{n,++} \) we recall the following facts:

\[ \lambda_1(XY^{-1}) = \lambda_1(Y^{-1/2}XY^{-1/2}) = \|Y^{-1/2}XY^{-1/2}\|, \]
\[ \lambda_n(XY^{-1}) = \frac{1}{\|Y^{1/2}X^{-1}Y^{1/2}\|}. \tag{5.2} \]

Hence

\[ \lambda_1(XZ^{-1}) = \lambda_1(Y^{-1/2}XY^{-1/2}Y^{1/2}Z^{-1}Y^{1/2}) \leq \|Y^{-1/2}XY^{-1/2}Y^{1/2}Z^{-1}Y^{1/2}\| \]
\[ \leq \|Y^{-1/2}XY^{-1/2}\| \|Y^{1/2}Z^{-1}Y^{1/2}\| = \lambda_1(XY^{-1})\lambda_1(YZ^{-1}). \]

By replacing \( X, Y \) and \( Z \) by their inverses and using (3.4) we deduce that

\[ \lambda_n(XZ^{-1}) \geq \lambda_n(XY^{-1})\lambda_n(YZ^{-1}). \]

Combine these two inequalities to deduce that

\[ 0 \leq \text{dist}(X, Z) \leq \text{dist}(X, Y) + \text{dist}(X, Z) \quad \text{for } X, Y, Z \in H_{n,++}. \]

If \( \lambda_1(XY^{-1}) = \lambda_n(XY^{-1}) \), then \( \lambda_1(Y^{-1/2}XY^{-1/2}) = \lambda_n(Y^{-1/2}XY^{-1/2}) \). Therefore,

\[ Y^{-1/2}XY^{-1/2} = tI \implies X = tY \quad \text{for some } t > 0. \]

Assume that \( \alpha \in H_{n,++} \). As \( Q, Q', \tilde{Q} \) and \( D_\alpha \) are homogeneous maps of degree 1, 1, 0 and 0, respectively, it follows that the maps \( Q, Q', \tilde{Q}, D_\alpha : H_{n,++} \to H_{n,++} \) induce the corresponding maps from \( \mathbb{P}H_{n,++} \) to itself. By abusing the notation we denote these maps by \( Q, Q', \tilde{Q}, D_\alpha : \mathbb{P}H_{n,++} \to \mathbb{P}H_{n,++} \), and no ambiguity will arise. Clearly, the maps \( Q' \) and \( \tilde{Q} \) from \( \mathbb{P}H_{n,++} \) to itself are identical. Furthermore, we identify the maps \( Q, \tilde{Q}, D_\alpha : \mathbb{P}H_{n,++} \to \mathbb{P}H_{n,++} \) with the corresponding maps \( Q, \tilde{Q}, D_\alpha : H_{n,++} \to H_{n,++} \). We also will view \( H_{n,++} \) as a metric space with respect to the Hilbert metric.

Theorem 5.1. Let \( Q \) be a positive quantum channel given by (1.1). Then

1. The maps \( Q \) and \( \tilde{Q} \) are strict contractions on \( \mathbb{P}H_{n,++} \):

\[ \text{dist}(Q(X), Q(Y)) \leq \kappa \text{dist}(X, Y), \quad \text{dist}((\tilde{Q}(X), \tilde{Q}(Y))) \leq \kappa \text{dist}(X, Y), \tag{5.3} \]

\[ \kappa = \frac{b(Q) - a(Q)}{b(Q) + a(Q)}. \tag{5.4} \]
2. The map $D_{tI_n}: H_{n,++} \rightarrow H_{n,++}$ preserves the Hilbert metric on $\mathbb{P}H_{n,++}$ for $t > 0$.
3. The map $\Theta = \Phi_{\frac{1}{n}I_n, \frac{1}{n}I_n}: H_{n,++}, \rightarrow H_{n,++}$ is a contraction with respect the Hilbert metric. The contraction constant is bounded by $\kappa^2$.
4. $\Theta$ has a unique fixed point $F \in H_{n,++}$ which lies in the interior of $H_{n,++}$. For each $X \in H_{n,++}$ the iterations $\Theta^m(X)$, $m \in \mathbb{N}$, converge to $F$.
5. There exist open balls $B_1, B_2 \subset H_{n,++}$, in the Frobenius norm, centred at $\frac{1}{n}I_n$ and $F$, respectively, and with positive radii $\varepsilon_1$ and $\varepsilon_2$, respectively, with the following properties. Assume that $\alpha, \beta \in B_1$. Then $\Phi_{\alpha, \beta}$ has a unique fixed point $U = U(\alpha, \beta)$ which lies in $B_2$. Furthermore, for each $X \in B_2$ the iterations $\Theta^m(X)$, $m \in \mathbb{N}$, converge to $U(\alpha, \beta)$.

Proof. 1. We start by proving the first inequality of (5.3). Recall that the map $Q: H_{n,+} \rightarrow H_{n,+}$ is linear. Hence we can apply Birkhoff’s theorem [1], which gives an upper bound on the contraction coefficient of $Q$ on $\mathbb{P}H_{n,++}$. Namely, let

$$
\Delta(Q) := \max\{\text{dist}(Q(X), Q(Y)) \mid X, Y \in H_{n,++}\}.
$$

Then

$$
\text{dist}(Q(X), Q(Y)) \leq \tanh \left( \frac{1}{4} \Delta(Q) \right) \text{dist}(X, Y) \quad \text{for } X, Y \in H_{n,++}.
$$

As $a(Q)I_n \leq Q(X)$ and $Q(Y) \leq Q(b)I_n$, it follows that

$$
\frac{a(Q)}{b(Q)} Q(Y) \leq Q(X) \leq \frac{b(Q)}{a(Q)} Q(Y).
$$

Hence $\text{dist}(Q(X), Q(Y)) \leq 2 \log(b(Q)/a(Q))$. Therefore $\Delta(Q) \leq 2 \log(b(Q)/a(Q))$. Note that

$$
\tanh \left( \frac{1}{4} \Delta(Q) \right) \leq \tanh \left( \frac{1}{2} \log \frac{b(Q)}{a(Q)} \right) = \frac{\sqrt{b(Q)/a(Q)} - \sqrt{a(Q)/b(Q)}}{\sqrt{b(Q)/a(Q)} + \sqrt{a(Q)/b(Q)}} = \frac{b(Q)/a(Q) - 1}{b(Q)/a(Q) + 1} = \frac{b(Q) - a(Q)}{b(Q) + a(Q)}
$$

This shows the first part of the inequality (5.3).

Consider $\tilde{Q}: \mathbb{P}H_{n,++} \rightarrow \mathbb{P}H_{n,++}$. Clearly,

$$
\text{dist}(\tilde{Q}(X), \tilde{Q}(Y)) = \text{dist}(Q(X), Q(Y)).
$$

Hence, it is enough to estimate the contraction coefficient of $Q'$. Apply Lemma 3.2 and the above arguments to deduce the second part of inequality (5.3).

2. Let $\Omega: H_{n,++} \rightarrow H_{n,++}$ be given by $\Omega(X) = X^{-1}$. Clearly, $\text{dist}(D_{tI_n}(X), D_{tI_n}(Y)) = \text{dist}(\Omega(X), \Omega(Y))$ for $X, Y \in H_{n,++}$. Observe that

$$
\lambda_1(\Omega(X)\Omega(Y)^{-1}) = \lambda_n(XY^{-1})^{-1} \quad \text{and} \quad \lambda_n(\Omega(X)\Omega(Y)^{-1}) = \lambda_1(XY^{-1})^{-1}
$$

for $X, Y \in H_{n,++}$. Hence $D_{tI_n}$ preserves the metric on $H_{n,++}$ for $t > 0$.

3. The statement follows straightforwardly from 2 and 1.
4. As \( \Theta(H_{n,+1}) \subset H_{n,+1} \) and \( \Theta \) is a contraction on \( H_{n,+1} \), the Banach Fixed Point Theorem yields that \( F \in H_{n,+1} \) is a unique fixed point of \( \Theta \). Furthermore, for each \( X \in H_{n,+1} \) the iterations \( \Theta^m(X) \), \( m \in \mathbb{N} \), converge to \( F \).

5. Denote the space of real symmetric and real skew symmetric \((n \times n)\)-matrices by \( \mathcal{S}_n, \mathcal{A}_n \subset \mathbb{R}^{n \times n} \). Note that \( \dim \mathcal{S}_n = n(n+1)/2 \) and \( \dim \mathcal{A}_n = n(n-1)/2 \). We now consider \( X \in H_n \) as having the form \( X = X_{\text{Re}} + iX_{\text{Im}} \), where \( X_{\text{Re}} \in \mathcal{S}_n \) and \( X_{\text{Im}} \in \mathcal{A}_n \). That is, we view \( H_n \) as the space \( \mathbb{R}^{n^2} \), where the \( n^2 \) variables are the real entries of the upper diagonal and strict upper diagonal parts of \( X_{\text{Re}} \) and \( X_{\text{Im}} \), respectively. An open ball of radius \( r > 0 \) centred at \( X_0 \in H_n \) is of the form \( B(X_0, r) = \{ X_0 + W, \ W \in H_n, \text{tr} W^2 < r^2 \} \). Thus \( H_{n,+} \) is a domain in \( H_n \).

A map \( \Lambda : H_{n,+} \to H_{n,+} \) is called real analytic if the following conditions hold. Let \( X_0 \in H_{n,+} \) be given. Then there exists an open ball centred at \( X_0 \) of radius \( \varepsilon = \varepsilon(X_0, \Lambda) \), which is contained in \( H_{n,+} \), such that the following conditions hold. For \( X = X_0 + W \), in this ball the real and imaginary parts of the entries of \( \Lambda(X_0 + W) \) are given by convergent power series in the real and complex parts of the entries of \( W \).

A similar definition applies for the map \( \Lambda_1 : H_{n,+1} \to H_{n,+1} \) to be analytic.

Clearly, the maps \( Q, Q' : H_{n,+} \to H_{n,+} \) are linear in \( n^2 \) real variables of \( X \in H_{n,+} \). Hence these maps are analytic. It is straightforward to see that the map \( \tilde{Q} : H_{n,+} \to H_{n,+1} \subset H_{n,+} \) is analytic. Use the formula for \( \Omega(X) \) in terms of the adjoint matrix \( \text{adj} X \) of \( X \), that is, \( \Omega(X) = \frac{1}{\det X} \text{adj} X \), to deduce that \( \Omega \) is analytic on \( H_{n,+} \).

Let \( \Gamma : H_{n,+} \to H_{n,+} \) be given by \( \Gamma(X) = X^{1/2} \). Then \( \Gamma \) is analytic on \( H_{n,+} \). This fact can easily be deduced from the Cauchy integral formula (see [4], §3.4). Indeed, let \( \mathbb{C}_+ = \{ z \in \mathbb{C}, \text{Im} z > 0 \} \) be the right-half complex plane. Let \( f : \mathbb{C}_+ \to \mathbb{C}_+ \) be the analytic function \( f(z) = \sqrt{z}, \sqrt{1} = 1 \). Assume that \( X_0 \in H_{n,+} \) is given. Let \( S \) be the circle centred at \( (\lambda_1(X_0) + \lambda_n(X_0))/2 \) with radius \( \lambda_1(X_0)/2 \). Assume that \( \varepsilon > 0 \) is small enough so that \( \lambda_n(X_0 + W) > \lambda_n(X_0)/2 \) for each \( W \in B(0, \varepsilon) \). Then

\[
\Gamma(X_0 + W) = \frac{1}{2\pi i} \int_S f(z)(zI - (X_0 + W))^{-1} \, dz.
\]

Express \((zI - (X_0 + W))^{-1}\) in terms of the adjoint of \((zI - (X_0 + W))\) to deduce that \( \Gamma \) is analytic.

Hence the map \( D_{\alpha} : H_{n,+} \to H_{n,+1} \) is analytic for \( \alpha \in H_{n,+} \). Let \( \Phi_{\alpha,\beta} \) be the map given by (4.1). Then \( \Phi_{\alpha,\beta} : H_{n,+} \to H_{n,+} \) is analytic. In particular, \( \Theta \) is analytic on \( H_{n,+} \). We now consider \( \Theta \) in a neighbourhood of the fixed point \( F \in H_{n,+} \). Set \( H_{n,0} = \{ W \in H_n, \text{tr} W = 0 \} \). Then the Jacobian of \( \Theta \) at \( F \) is a linear map \( P : H_{n,0} \to H_{n,0} \). That is,

\[
\Theta(F + W) = F + P(W) + \text{higher order terms}.
\]

Denote the spectral radius \( P \) by \( \rho(P) \). We claim that \( \rho(P) < 1 \). By considering a scaled quantum channel we can assume without loss of generality that \( Q \) is a unital channel, that is, \( Q(\frac{1}{n}I_n) = \frac{1}{n}I_n \). Hence \( F = \frac{1}{n}I_n \). Assume that \( W \in H_{n,0} \) and...
consider \( F(t) = \frac{1}{n}I + tW \). Then

\[
Q(F(t)) = \frac{1}{n}I + tQ(W), \quad D_{\frac{1}{n}I_n} \left( \frac{1}{n}I + tQ(W) \right) = \frac{1}{n}I_n - tQ(W) + O(t^2),
\]

\[
\tilde{Q} \left( \frac{1}{n} - tQ(W) + O(t^2) \right) = \frac{1}{n}I_n - t(Q' \circ Q)(W) + O(t^2),
\]

\[
D_{\frac{1}{n}I_n} \left( \frac{1}{n}I - t(Q' \circ Q)(W) + O(t^2) \right) = \frac{1}{n}I_n + t(Q' \circ Q)(W) + O(t^2).
\]

Therefore, \( P = Q' \circ Q \) is a selfadjoint linear operator, which is positive semi-definite with respect to the standard inner product on \( H_{n,0} \). That is, all eigenvalues of \( P \) are nonnegative. We claim that \( Q \) and \( Q' \) are strict contractions on \( H_{n,0} \). Indeed, \( Q(H_{n,+,1}) \subseteq H_{n}(a(Q), b(Q), 1) \) and \( Q(\frac{1}{n}I_n) = \frac{1}{n}I_n \). Hence any ball in \( H_{n,+,1} \) centred at \( \frac{1}{n}I_n \) is mapped to the interior of this ball. Similarly for \( Q' \). Hence \( P \) is a positive semi-definite matrix with \( \|P\| = \rho(P) < 1 \). In particular, \( P \) is a contraction on \( H_{n,0} \) with respect to the Frobenius norm.

Consider now \( \Phi_{\alpha, \beta} \) as the family of maps depending on parameters \( \alpha \) and \( \beta \). That is, \( \Phi_{\alpha, \beta} \) is given by \( (X, \alpha, \beta) \mapsto \Phi_{\alpha, \beta}(X) \) for \( (X, \alpha, \beta) \in H_{n,+,+}^3 \). The arguments above show that \( \Phi_{\alpha, \beta} \) is analytic on \( H_{n,+,+}^3 \). Hence, in a neighbourhood of \( (\frac{1}{n}I_n, \frac{1}{n}I_n, \frac{1}{n}I_n) \) in \( H_{n,+,+,1}^3 \) we have:

\[
\Phi_{\frac{1}{n}I_n + \alpha_1, \frac{1}{n}I_n + \beta_1} \left( \frac{1}{n}I_n + W \right)
\]

\[
= \frac{1}{n}I_n + W_1(\alpha_1, \beta_1) + (P(W) + E_1(\alpha_1, \beta_1)(W)) + E_2(W, \alpha_1, \beta_1),
\]

\[
\|W_1(\alpha_1, \beta_1)\|_F \leq K(\|\alpha_1\|_F + \|\beta_1\|_F), \quad \|E_1(\alpha_1, \beta_1)\| \leq K(\|\alpha_1\|_F + \|\beta_1\|_F),
\]

\[
E_2(0, \alpha_1, \beta_1) = 0
\]

and

\[
\|E_2(W_1, \alpha_1, \beta_1) - E_2(W_2, \alpha_1, \beta_1)\|_F \\
\leq K\|W_1 - W_2\|_F(\|W_1\|_F + \|W_2\|_F)(1 + \|\alpha_1\|_F + \|\beta_1\|_F).
\]

Note that \( W_1(\alpha_1, \beta_1), E_2(W, \alpha_1, \beta_1) \in H_{n,0} \) and \( E_1(\alpha_1, \beta_1) : H_{n,0} \to H_{n,0} \) is a linear operator for each \( \alpha_1 \) and \( \beta_1 \) in some open ball of radius \( r \) centred at \( 0 \in H_{n,0} \). Also assume that the above inequalities hold for \( \|\alpha_1\|_F, \|\beta_1\|_F, \|W\|, \|W_1\|, \|W_2\| < r \).

Assume that \( \|P\| < 1 - 2\eta \), for some \( \eta > 0 \), and \( t \in (0, r) \). Suppose that \( \|\alpha_1\|, \|\beta_1\| < t^2/2 \) and \( \|W\|_F < t \). Then

\[
\left\| \Phi_{\frac{1}{n}I_n + \alpha_1, \frac{1}{n}I_n + \beta_1} \left( \frac{1}{n}I_n + W \right) - \frac{1}{n}I_n \right\| < Kt^2 + (1 - 2\eta + Kt^2)t + Kt^2(1 + t^2).
\]

Hence there exists \( \varepsilon \in (0, \min(r, 1)) \) such that

\[
2K(\varepsilon + \varepsilon^2 + \varepsilon^3) < \eta.
\]

In particular,

\[
K\varepsilon^2 + (1 - 2\eta + K\varepsilon^2)\varepsilon + K\varepsilon^2(1 + \varepsilon^2) = (1 - 2\eta + K(2\varepsilon + \varepsilon^2 + \varepsilon^3))\varepsilon < (1 - \eta)\varepsilon.
\]
Therefore,
\[
\Phi_{\alpha,\beta} \left( B \left( \frac{1}{n} I_n, \varepsilon \right) \right) \subset B \left( \frac{1}{n} I_n, \varepsilon \right) \quad \text{for } \alpha, \beta \in B \left( \frac{1}{n} I_n, \frac{\varepsilon^2}{2} \right).
\]

Assume that \( W_1, W_2 \in B(0, \varepsilon) \). Then
\[
\left\| \Phi_{\alpha,\beta} \left( \frac{1}{n} I_n + W_1 \right) - \Phi_{\alpha,\beta} \left( \frac{1}{n} I_n + W_2 \right) \right\|_F \\
\leq (1 - 2\eta + K\varepsilon^2)\|W_1 - W_2\| + 2K\varepsilon(1 + \varepsilon^2)\|W_1 - W_2\| \\
= (1 - 2\eta + K(2\varepsilon + \varepsilon^2 + 2\varepsilon^3))\|W_1 - W_2\| < (1 - \eta)\|W_1 - W_2\|.
\]

That is, for fixed \( \alpha, \beta \in B \left( \frac{1}{n} I_n, \frac{\varepsilon^2}{2} \right) \) and \( X \in B_2 = B \left( \frac{1}{n} I_n, \varepsilon \right) \) the map \( \Phi_{\alpha,\beta} \) is a contraction with respect to the Frobenius norm. Hence \( \Phi_{\alpha,\beta} \) has a unique fixed point \( U = U(\alpha, \beta) \) in \( B_2 \). Furthermore, for each \( X \in B_2 \) the iterations \( \Theta^m(X) \), \( m \in \mathbb{N} \), converge to \( U \).

It remains to show that there exists \( \varepsilon_1 \in (0, \varepsilon^2/2) \) such that for \( \alpha, \beta \in B \left( \frac{1}{n} I_n, \varepsilon_1 \right) \) the map \( \Phi_{\alpha,\beta} \) have a unique fixed point which is \( U(\alpha, \beta) \). Since \( \Theta \) has a unique fixed point \( \frac{1}{n} I_n \), the continuity argument yields that there exists \( \varepsilon_1 \in (0, \varepsilon^2/2) \) such that for each \( \alpha, \beta \in B_1 = B \left( \frac{1}{n} I_n, \varepsilon_1 \right) \) all fixed points of \( \Phi_{\alpha,\beta} \) lie in \( B_2 \). Our previous results show that in \( B_2 \) the map \( \Phi_{\alpha,\beta} \) has a unique fixed point \( U(\alpha, \beta) \).

It is an open problem to prove that \( \Phi_{\alpha,\beta} \) has a unique fixed point in \( H_{n,++} \) for each \( \alpha, \beta \in H_{n,++} \) or to give a counterexample. See [6] for numerical simulations.

Bibliography

[1] G. Birkhoff, “Extensions of Jentzsch’s theorem”, Trans. Amer. Math. Soc. 85:1 (1957), 219–227.
[2] R. A. Brualdi, “Convex sets of non-negative matrices”, Canad. J. Math. 20 (1968), 144–157.
[3] R. A. Brualdi, S. V. Parter and H. Schneider, “The diagonal equivalence of a nonnegative matrix to a stochastic matrix”, J. Math. Anal. Appl. 16:1 (1966), 31–50.
[4] S. Friedland, Matrices: algebra, analysis and applications, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ 2016, xii+582 pp.
[5] S. Friedland and S. Karlin, “Some inequalities for the spectral radius of non-negative matrices and applications”, Duke Math. J. 42:3 (1975), 459–490.
[6] T. T. Georgiou and M. Pavon, “Positive contraction mappings for classical and quantum Schrödinger systems”, J. Math. Phys. 56:3 (2015), 033301, 24 pp.
[7] L. Gurvits, “Classical complexity and quantum entanglement”, J. Comput. System Sci. 69:3 (2004), 448–484.
[8] C. B. Mendl and M. M. Wolf, “Unital quantum channels – convex structure and revivals of Birkhoff’s theorem”, Comm. Math. Phys. 289:3 (2009), 1057–1086.
[9] M. V. Menon, “Matrix links, an extremization problem, and the reduction of a non-negative matrix to one with prescribed row and column sums”, Canad. J. Math. 20 (1968), 225–232.
[10] J. W. Milnor, Topology from the differentiable viewpoint, 2nd rev. ed., Princeton Landmarks Math., Princeton Univ. Press, Princeton, NJ 1997, xii+64 pp.
[11] H. Perfect and L. Mirsky, “The distribution of positive elements in doubly-stochastic matrices”, *J. London Math. Soc.* **40** (1965), 689–698.

[12] E. Schrödinger, “Über die Umkehrung der Naturgesetze”, *Sitzungsber. Preuss. Akad. Wiss. Phys.-Math. Kl.* **1931**:8–9 (1931), 144–153.

[13] E. Schrödinger, “Sur la théorie relativiste de l’électron et l’interprétation de la mécanique quantique”, *Ann. Inst. H. Poincaré* **2**:4 (1932), 269–310.

[14] R. Sinkhorn, “A relationship between arbitrary positive matrices and doubly stochastic matrices”, *Ann. Math. Statist.* **35**:2 (1964), 876–879.

[15] R. D. Sinkhorn and P. J. Knopp, “Concerning nonnegative matrices and doubly stochastic matrices”, *Pacific J. Math.* **21**:2 (1967), 343–348.

**Shmuel Friedland**

Department of Mathematics,  
Statistics and Computer Science,  
University of Illinois at Chicago,  
Chicago, IL, USA  
*E-mail: friedlan@uic.edu*