Compact Causal Data Interpolation

Niels Juul Munch*

Matematisk Institut, Aarhus Universitet,
Ny Munkegade, 8000 Aarhus C, Denmark

Submitted by A. Schumitzky

Received November 11, 1986

Consider a Hilbert space $\mathcal{H}$ equipped with a time-structure, i.e., a resolution $E$ of the identity on $\mathcal{H}$ defined on subsets of some linearly ordered set $A$. For which $x$ and $y$ in $\mathcal{H}$ is it possible to find a causal (time respecting) compact operator $T$, so that $Tx = y$? When $T$ is required to be a Hilbert-Schmidt operator and $(A, E)$ is sufficiently regular, this question is answered in terms of the “time-densities” of $x$ and $y$. The condition is that the integral $\int_A \mu_x([s \leq t])^{-1} d\mu_x(t)$ should be finite, where $\mu_x$ and $\mu_y$ are the measures on $A$ given by $\mu_x(\Omega) = \|E(\Omega)x\|^2$ and $\mu_y(\Omega) = \|E(\Omega)y\|^2$. Further a solution is given for the related problem of minimizing the sum of $\|Tx - y\|^2$ and the squared Hilbert-Schmidt norm $\|T\|^2$ of $T$.

INTRODUCTION

The setup considered in this paper is the following. Let some otherwise unknown physical transformation $T$ be given which acts boundedly on states $x$ in a Hilbert space $\mathcal{H}$. Also assume given a linearly ordered set $A$, compact in the order topology, along with a countably additive resolution $E$ of the identity on $\mathcal{H}$ defined on the Borel subsets of $A$. Denote by 0 and 1 the minimal and maximal elements of $A$ and for ease of notation let $E'$ and $E$, be the projections determined by the intervals $[0, t]$ and $[t, 1]$, respectively, i.e., $E' = E([0, t])$ and $E = E([t, 1]) = I - E'$. It is assumed throughout this paper that $E'$ is a strongly continuous function of $t$.

If $E'$ is interpreted as representing the past and present for some time $t$, it is for physical reasons clear that $T$ must satisfy $E'TE' = E'T$, $t \in A$. In other words, $T$ should be causal relative to the nest $\{E'\}_{t \in A}$; cf. [1]. This set of causal operators will be denoted $\mathcal{C}$.

The causality constraint on the $T$'s limits somewhat the set of input-output pairs $(x, y)$ compatible with a relation $Tx = y$. In [2] it was shown

* Research supported by the Danish NSRC.
by Lance that given \( x \) and \( y \) in \( \mathcal{H} \), there is a bounded causal operator \( T \) with \( Tx = y \) if and only if the quantity

\[
K = \sup_{T \in \mathcal{A}} \frac{\|E'Ty\|}{\|E'Tx\|}
\]

is finite. (We understand \( 0/0 \) to be zero.) Further this \( K \) is exactly the minimum value of the norms \( \|T\| \), where \( T \) is causal with \( Tx = y \).

Given an input–output pair \((x, y)\), there may be a host of possible choices for the interpolating operator \( T \). It seems natural to look for \( T \)'s, which are not unnecessarily large, preferring bounded operators to unbounded operators whenever possible, and, likewise, preferring compact operators to general bounded ones. The first section of this paper analyses an explicit construction which yields the smallest possible causal Hilbert–Schmidt operator \( T \) with \( Tx = y \).

For general \( x \) and \( y \) in \( \mathcal{H} \) it is of interest to minimize the norm of the difference \( Tx - y \). However, in many cases where \( Tx = y \) is impossible it is in fact possible to obtain arbitrarily close approximations provided that the Hilbert–Schmidt norm of \( T \) is allowed to grow unrestricted. To prevent this kind of triviality and with a view towards potential applications we therefore consider in Section 2 the minimization of the function \( \varepsilon(T) = \|Tx - y\|^2 + \|T\|^2 \) with \( T \) ranging over the causal Hilbert–Schmidt operators. This approach is similar in spirit to the "parametric projection filter" of [5] on which it has some bearing. In the setup of [5] the quantity \( \varepsilon(T) \) is a measure of the error that results when \( T \) is used to reconstruct a signal \( y \) from a degraded version \( x \) in additive, zero-mean random noise. The minimization of \( \varepsilon(T) \) balances the two natural objectives of image reconstruction and noise suppression, corresponding to the first and second part of \( \varepsilon(T) \), respectively.

Our conditions on \((\mathcal{A}, E)\) are not restrictive. To see this, consider a general set \( \mathcal{A}_0 = \{P\} \) of orthogonal projections on \( \mathcal{H} \) each representing some time \( t \). For obvious reasons \( \mathcal{A}_0 \) should be linearly ordered. Now take for \( \mathcal{A} \) the weak, or equivalently strong, closure of \( \mathcal{A}_0 \) in the set of projections on \( \mathcal{H} \) and adjoin if necessary the zero and identity operator. Then \( \mathcal{A} \) will meet the conditions stated and we may take for \( E \) the map determined by \( E([P, Q]) = \frac{Q - P}{2}, P, Q \in \mathcal{A}. \) For the present discussion, the sets \( \mathcal{A} \) and \( \mathcal{A}_0 \) are effectively equivalent, since they give rise to the same set of causal operators.

1. Data Interpolation

Some further notation is now required. Define the vector valued measures \( M_x, x \in \mathcal{H}, \) by \( M_x(\Omega) = E(\Omega)x. \) These are c.a.o.s. (completely
additive, orthogonally scattered) measures in the sense of [3] with quadratic measures \( \mu_x, \mu_x(\Omega) = \| E(\Omega) x \|^2 \). Denote the opposite Hilbert space of \( \mathcal{H} \) by \( \mathcal{H}^\perp \). The c.a.o.s. measure \( M_{x_1} \otimes_2 M_{x_2} \) with values in the tensor product \( \mathcal{H} \otimes_2 \mathcal{H} \) is given by extension of the application

\[
\Omega_1 \times \Omega_2 \rightarrow M_{x_1}(\Omega_1) \otimes_2 M_{x_2}(\Omega_2)
\]

to a c.a.o.s. measure on the product space \( A \times A \) with quadratic measure \( \mu_x \times \mu_{x_2} \). Recall that the Hilbert space \( F_2 = F_2(\mathcal{H}) \) of Hilbert–Schmidt operators on \( \mathcal{H} \) with inner product \( \langle T, S \rangle_2 = \text{Tr}(S^*T) \) is isomorphic to \( \mathcal{H} \otimes_2 \mathcal{H} \). The vector \( x_1 \otimes x_2 \) in \( \mathcal{H} \otimes_2 \mathcal{H} \) corresponds to the operator \( (x_1 \otimes x_2) z = (z, x_1) x_2 \) in \( F_2 \).

The main result of this section is the following.

**Theorem 1.1.** Let \( x \) and \( y \) be vectors in \( \mathcal{H} \). There is an operator \( T \in F \cap F_2 \) so that \( Tx = y \) if and only if the integral

\[
\int_A \mu_x([0, t])^{-1} \, d\mu_y(t)
\]

is finite. In this case \((*)\) is the minimum value of \( \| T \|^2 \) under the conditions stated and this value is attained by the operator \( T_0 \) associated with the vector

\[
T_0 = \int_A \int_A 1_{\{s \leq t\}} \mu_x([0, t])^{-1} \, d(M_x \otimes_2 M_y(s, t))
\]

in \( \mathcal{H} \otimes_2 \mathcal{H} \).

**Remark.** We point out that the expression \((**)\) for \( T_0 \) has been derived from a remark in [1, pp. 222–223].

**Proof.** First assume the existence of such a \( T \). For each finite partition \( \mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_n = 1\} \) of \( A \) let \( f_\mathcal{P} \) be the function

\[
f_\mathcal{P} = \sum_{i=1}^n \mu_x([0, t_i])^{-1} 1_{[t_{i-1}, t_i]}.
\]

Compute

\[
\int_A f_\mathcal{P} \, d\mu_y = \sum_{i=1}^n \mu_x([0, t_i])^{-1} \mu_y([t_{i-1}, t_i])
\]

\[
= \sum_{i=1}^n \| E^i x \|^{-2} \| (E^i - E^{i-1}) y \|^2.
\]
Now by the causality of $T$

$$(E^u - E'^u) y = (E^u - E'^u) T x - (E^u - E'^u) T E'^u x$$

and thus

$$\int_A f_\varphi d\mu_x = \sum_{i=1}^n \left( T^*(E^u - E'^u) T \frac{E'^u x}{\|E'^u x\|} \right)$$

$$\leq \sum_{i=1}^n \text{Tr}(T^*(E^u - E'^u) T)$$

$$= \text{Tr}(T^* T)$$

By an easy compactness argument it is possible to find a sequence of the $f_\varphi$ converging upwards towards the desired integrand. Hence the monotone convergence theorem assures that the integral (*) is less than $\text{Tr}(T^* T)$.

For the proof in the converse direction let $x$ and $y$ be given so that $\mu_x([0, t])^{-1}$ is integrable with respect to $\mu_y$ and define $T_0$ by (**). The procedure of iterated integration demonstrates that the integrand of (**) lies in $L^2(\mu_x \times \mu_y)$. Consequently (**) is a valid definition of an element $T_0$ in $\mathcal{H} \otimes_2 \mathcal{H}$ with squared norm given by the above integral.

We proceed to verify $T_0 x = y$ and the causality of $T_0$. First, for any choice of $t_0$ in $A$,

$$(T_0 x, E'^0 y) = \langle T_0, x \otimes E'^0 y \rangle_2$$

$$= (T_0, x \otimes 2 E'^0 y)_{\mathcal{H} \otimes_2 \mathcal{H}}$$

$$= \left(T_0, \int_A \int_A 1_{\{t \leq t_0\}} d(M_x \times M_y(s, t))\right)$$

$$= \int_A \mu_x([0, t])^{-1} 1_{\{t \leq t_0\}} \int_A 1_{\{s \leq t\}} d(M_x \times M_y(s, t))$$

$$= \int_A 1_{\{t \leq t_0\}} d\mu_x(t)$$

$$= (y, E'^0 y).$$
The range of $T_0$ is clearly contained in the closure of the range of $M_y$, therefore $T_0x = y$.

In a like manner the following expression is obtained for any $t_0, t_1, t_2$ in $A$:

$$(E^n T_0 E^n x, E^n y)$$

$$= \int_A \int_A \mu_x([0, t])^{-1} 1_{x \leq t_1} 1_{x \leq t_2} d(\mu_x \times \mu_y(s, t)).$$

Here $t_1$ can clearly be substituted by $t_0 \wedge t_1$ and the resulting integral is seen to equal

$$(E^n T_0 E^n E^n x, E^n y).$$

Simple range and domain considerations now yield the conclusion $E^n T_0 = E^n T_0 E^n$.

In order to exploit fully the minimality result of the theorem, we introduce an order relation on $\mathcal{F}_2$. For $T$ and $S$ in $\mathcal{F}_2$ define $T >_E S$, if $\text{Tr}(T^* E(\Omega) T) > \text{Tr}(S^* E(\Omega) S)$ for all intervals $\Omega$ in $A$. The quality of this relation is of course highly dependent on the size of the range of $E$. In the most trivial circumstances $T >_E S$ reduces to $\text{Tr}(T^* T) > \text{Tr}(S^* S)$.

It is instructive to examine the relation $>_E$ in the standard matrix case. Here $\mathcal{H} = \mathbb{C}^n$, $A = \{0, 1, 2, \ldots, n\}$, $E(\{i\})$ is the projection onto the $i$th coordinate for $1 \leq i \leq n$, and the causal operators $\mathcal{C}$ are the lower triangular matrices. Two matrices $T = (t_{ki})$ and $S = (s_{ki})$ stand in the relation $T >_E S$ if and only if

$$\sum_{i=1}^n |t_{ki}|^2 > \sum_{i=1}^n |s_{ki}|^2, \quad k = 1, 2, \ldots, n.$$  

The next statement is a localized version of Theorem 1.1.

**Corollary 1.2.** Let $T \in \mathcal{C} \cap \mathcal{F}_2$ satisfy $Tx = y$. Then $T >_E T_0$.

**Proof.** For $Tx = y$ consider any interval $\Omega$. Due to $E(\Omega) y = (E(\Omega) T) x$ we may apply the theorem to the vectors $y_1 = E(\Omega) y$, $x_1 = x$, and the operator $T_1 = E(\Omega) T$. Hence

$$\text{Tr}(T^* E(\Omega) T) = \text{Tr}(T^* T_1)$$

$$\geq \int_A \mu_{y_1}([0, t])^{-1} d\mu_{y_1}(t)$$

$$= \int_\Omega \mu_x([0, t])^{-1} d\mu_x(t).$$
This last integral is precisely the value of $\text{Tr}(T_0^*E(\Omega)T_0)$. It follows that $T > E T_0$.

By a similar calculation $T_0$ is even minimal in the uniform sense that each $E'T_0 = E'T_0 E'$ has the minimal Hilbert-Schmidt norm among all operators in $\mathcal{C} \cap \mathcal{F}_2$ sending $E'x$ to $E'y$.

Apparently it is a harder question to find conditions that characterize the existence of a causal, nuclear operator $T$ so that $Tx = y$. Regarding the operator $T_0$ one finds

(i) $T_0^* T_0 = \int_A \mu_\gamma([0, t])^{-2} E'x \hat{\otimes} E'x \, d\mu_\gamma(t)$

and

(ii) $T_0 T_0^* = \int_A \int_A \mu_\gamma([0, s \vee t])^{-1} \, d(M_\gamma \hat{\otimes} M_\gamma(s, t))$.

Hence $T_0$ will be nuclear if either of these integral expressions admits a nuclear square root. If $\mathcal{F}_2(\mathcal{H})$ is identified with $\mathcal{H} \hat{\otimes} \mathcal{H}$ and $\mathcal{F}(L^2(\mu_\gamma))$ with $L^2(\mu_\gamma \times \mu_\gamma)$ the map $\rho: \mathcal{F}_2(L^2(\mu_\gamma)) \to \mathcal{F}_2(\mathcal{H})$ given by

$$\rho(f) = \int_A \int_A f(s, t) \, d(M_\gamma \hat{\otimes} M_\gamma(s, t))$$

becomes an isometric $\ast$-homomorphism. As now

$$T_0 T_0^* = \rho(\mu_\gamma([0, s \vee t])^{-1}),$$

one criterion for the nuclearity of $T_0$ is that the operator on $L^2(\mu_\gamma)$ associated with the kernel

$$\mathcal{H}(s, t) = \mu_\gamma([0, s \vee t])^{-1} = (E_{x \vee y}, x)^{-1}$$

admits a nuclear square root. Regrettably it seems fairly difficult to interpret this information.

However, when $E$ is essentially a discrete chain with finite-dimensional jumps some progress is contained in Proposition 1.3 below. Denote by $t^x$ for $x \in \mathcal{H}$ the time

$$t^x = \max \{ t \in A | \mu_\gamma([0, t]) = 0 \}.$$

**Proposition 1.3.** If the integral

$$\left( \int_{[t^x, 1]} \left\{ \int_{[0, 1]} \mu_\gamma([0, t])^{-1} \, d\mu_\gamma(t) \right\}^{1/2} \, d(\text{Tr} \circ E(s)) \right)^{1/2}$$

is finite, the operator $T_0$ will belong to $\mathcal{C} \cap \mathcal{F}_1$ with $\mathcal{F}_1$-norm dominated by $(***)$. 

Proof. For any function $f$ in $L^2(\mu_x)$ the reader may verify that
\[
T_0 \int_A f(s) \, dM_x(s) = \int_A \mu_x([0, t])^{-1} \left\{ \int_{[0, t]} f(s) \, d\mu_x(s) \right\} \, dM_y(t).
\]
Hence
\[
\|T_0 \int_A f(s) \, dM_x(s)\|^2
\]
\[
= \int_A \mu_x([0, t])^{-2} \left\{ \int_{[0, t]} |f(s)|^2 \, d\mu_x(s) \right\} \, d\mu_y(t)
\]
\[
\leq \int_A \mu_x([0, t])^{-1} \left\{ \int_{[0, t]} |f(s)|^2 \, d\mu_x(s) \right\} \, d\mu_y(t)
\]
\[
= \int_A \left\{ 1_{[r', 1]}(s) \int_{[x, 1]} \mu_x([0, t])^{-1} \, d\mu_x(t) \right\} |f(s)|^2 \, d\mu_x(s).
\]
Let $h(s)$ be the square root of the quantity in brackets and denote by $H$ the bounded, positive operator $H = \int_A h \, dE$. The calculation above demonstrates that $\|T_0 z\|^2 \leq \|Hz\|^2$ for any $z$ in the subspace generated by $M_x$. By the construction of $T_0$ and $H$ this holds for any $z$ in $\mathcal{H}$. Hence $T_0^* T_0 \leq H^2$ in the standard operator sense and consequently, due to the monotonicity of the square root operation, $|T_0| = (T_0^* T_0)^{1/2} \leq H$. It follows that
\[
\|T_0\|_1 \leq \|H\|_1 = \int_A h(s) \, d(T(\circ E)(s)).
\]
In [4] for each probability measure $\mu$ on $\mathcal{A}$ a class $L^\mu$ of operators $T$ on $\mathcal{H}$ is defined by the property that there exists an $\alpha > 0$ so that $\|E(\Omega) T\|^2 \leq \alpha \mu(\Omega)$ for all $\Omega$. The estimate
\[
\|E(\Omega) T_0\|^2 \leq \|E(\Omega) T_0\|^2_2 = \int_{\Omega} \mu_x([0, t])^{-1} \, d\mu_y(t)
\]
demonstrates that $T_0 \in L^\mu$ for a measure $\mu_1$ proportional to $\mu_x([0, \cdot])^{-1} \mu_y$. Similarly $T_0^* \in L^\mu$, where $\mu_2$ is proportional to
\[
\Omega \rightarrow \int_{\Omega} \int_{[r, 1]} \mu_x([0, s])^{-2} \, d\mu_y(s) \, d\mu_x(t).
\]
Using (straightforward versions of) Theorem 3.3 and Corollary 3.4 in [4] it follows that $T_0$ is quasi-nilpotent, or equivalently uniformly strictly causal, if (and only if) either $\mu_x$ or $\mu_y$ is completely non-atomic.

In the next section there will be occasion again to draw on the results of [4] regarding the classes $L^\mu$.

2. APPROXIMATE DATA INTERPOLATION

For the remainder of this paper attention is focussed on the case where $T_x = y$ cannot be achieved for $T \in \mathcal{C} \cap \mathcal{F}_2$. As stated in the Introduction we then seek to minimize the function $\varrho(T)$. To further motivate this approach the following observation is inserted.

PROPOSITION 2.1. Let $x$ and $y$ be vectors in $\mathcal{H}$ so that $T_x$ is unequal to $y$ for any $T$ in $\mathcal{C} \cap \mathcal{F}_2$.

(a) If $T_n$ is a net in $\mathcal{C} \cap \mathcal{F}_2$ satisfying $y = \lim_n T_n x$ then $\lim_n \| T_n \|_2 = +\infty$.

(b) There is a net $T_n$ in $\mathcal{C} \cap \mathcal{F}_2$ satisfying $y = \lim_n T_n x$ if and only if $\mu_x([0, t, \omega]) = 0$, where $t_x = \min\{ t \in A | \mu_x([0, t]) > 0 \}$.

Proof: (a) If the conclusion does not hold it is possible to find a bounded net $T_n$ of operators in $\mathcal{C} \cap \mathcal{F}_2$ so that $y = \lim_n T_n x$ and so that $T_n$ converges weakly to $T$ in the Hilbert space $\mathcal{F}_2$. But then $T \in \mathcal{C}$ and the impossible equality $T x = y$ follow from the identity

$$(Sz_1, z_2) = \langle S, z_1 \otimes z_2 \rangle_{\mathcal{F}_2}, \quad S \in \mathcal{F}_2, \ z_i \in \mathcal{H}.$$  

(b) Assume that $y = \lim_n T_n x$. It must then be shown that $\mu_x([0, t, \omega]) = 0$. For any $t < t_x$

$$\mu_x([0, t]) = \| E^t y \|^2 = \lim_n \| E^t T_n x \|^2 = \lim_n \| E^t T_n E^t x \|^2 \leq \lim_n (\| T_n \| \mu_x([0, t])) = 0.$$  

If $t_x$ has an immediate predecessor $t$ we are done due to $[0, t_x, \omega] = [0, t]$. Otherwise $t_x$ belongs to the closure of $[0, t, \omega]$, and by the strong continuity of $E^t$ even $\mu_x([0, t])$ is zero.

Conversely, assume $\mu_x([0, t_x, \omega]) = 0$. If $\mu_x(t_x)$ is non-zero it follows that
\[ \mu_x([0, t])^{-1} \text{ is } \mu_y\text{-almost everywhere dominated by } \mu_x(\{t_x\})^{-1}. \] By Theorem 1.1 this contradicts our assumption. Hence for each \( t > t_x \) the interval \([t_x, t]\) has positive \( \mu_y\)-measure, and it is possible, again by Theorem 1.1, to find an operator \( T_x \in \mathcal{C} \cap \mathcal{F}_2 \) so that \( T_x x = y_x \), where \( y_x = E_x y \). Since \( y \) is the limit of \( y_x \) over the net \( \{t > t_x\} \) with the given order reversed, this concludes the proof. 

For the sake of generality and to bring out clearly the connection between this material and that in [1, 5, 6], consider for a moment \( \mathcal{H}\text{-valued random variables } x \) so that the expectation \( \mathcal{E}\left\| x \right\|^2 \) is finite. For such \( x \) and \( y \), operators \( Q_{xy} \) can be defined on \( \mathcal{H} \) by

\[
(Q_{xy}z_1, z_2) = \mathcal{E}\left((z_1, x)(y, z_2)\right), \quad z_i \in \mathcal{H}.
\]

All operators \( Q_{xx} \) will be nuclear, indeed \( \text{Tr}(Q_{xx}) = \mathcal{E}\left\| x \right\|^2 \), and all \( Q_{xy} \) will be in \( \mathcal{F}_2 \).

Recall that \( \mathcal{F}_2 \) can be decomposed as the direct sum of \( \mathcal{F}_2 \cap \mathcal{C} \) and \( \mathcal{F}_2 \cap \mathcal{R}^* \), where \( \mathcal{R}^* \) is the space of uniformly strictly anti-causal operators, and denote the projection of an element \( A \) in \( \mathcal{F}_2 \) onto these subspaces by \( A_x \) and \( A_x^* \), respectively.

The measures \( \mu_x \) have an obvious generalization to \( \mu_x \) for random variables \( x \) with \( \mathcal{E}\left\| x \right\|^2 < +\infty \) by

\[
\mu_x(\Omega) = \mathcal{E}\left\| E(\Omega) x \right\|^2, \quad \Omega \in \mathcal{A}
\]

and it is straightforward to show that \( Q_{xx} \in (L^{\mu_x})^* \cap L^{\mu_x} \) (again) in the terminology of [4]. The existence of the operator \( C \) below follows from this last property and [4, Theorem 3.5].

**Lemma 2.2.** Assume that \( x \) and \( y \) are \( \mathcal{H}\text{-valued random variables so that } E\left(\left\| x \right\|^2 \right) \text{ and } E\left(\left\| y \right\|^2 \right) \text{ are both finite and so that } \mu_x \text{ is completely non-atomic. Let } C \in \mathcal{C} \cap \mathcal{C}^{-1} \text{ satisfy } CC^* = Q_{xx} + I. \text{ Then the function}

\[
\varepsilon(T) = E\left\| Tx - y \right\|^2 + ||T||^2
\]

is minimized over \( T \) in \( \mathcal{C} \cap \mathcal{F}_2 \) by

\[
T_1 = (Q_{xy}C^*^{-1}) \in C^{-1}
\]

with minimum value

\[
\varepsilon_{\text{min}} = \text{Tr}(Q_{yy} - Q_{xy}(Q_{xx} + I)^{-1} Q_{yx}) + \|(Q_{xy}C^*^{-1}) \in C^*\|_2^2.
\]

**Proof** (Cp. [1, p. 253]). Using the orthogonality of \( \mathcal{F}_2 \cap \mathcal{C} \) and \( \mathcal{F}_2 \cap \mathcal{R}^* \) in \( \mathcal{F}_2 \) and the simple relation \( Q_{(Ax)(By)} = BQ_{xy}A^* \) there is no difficulty in computing

\[
\varepsilon(T) = \|TC - (Q_{xy}C^*^{-1})_x\|_2^2 + \varepsilon_{\text{min}}, \quad T \in \mathcal{C} \cap \mathcal{F}_2.
\]
It is an interesting and probably quite difficult problem to find a more enlightening description of $T_1$. Here we restrict attention to stochastically trivial $x$ and $y$. When $x = x$ and $y = y$ are constant, $Q_{xx} = x \otimes x$ and $Q_{xy} = x \otimes y$. In this case a much more tangible result is possible.

**Theorem 2.3.** Let $x$ and $y$ be vectors in $\mathcal{H}$ and assume that $\mu_x$ is completely non-atomic. Then the function

$$\varepsilon(T) = \|Tx - y\|^2 + \|T\|_2^2$$

is minimized over $T$ in $\mathcal{C} \cap \mathcal{F}_2$ by

$$T_1 = \int_A \int_A 1_{\{s \leq t\}} (1 + \mu_x([0, t]))^{-1} d(M_x \otimes_2 M_y(s, t))$$

with minimum value

$$\varepsilon_{\text{min}} = \int_A (1 + \mu_x([0, t]))^{-1} d\mu_x(t).$$

**Remark.** From this result it is seen that for any $\gamma > 0$ the minimization of $\varepsilon'(T) = \|Tx - y\|^2 + \gamma \|T\|_2^2$ over $T \in \mathcal{C} \cap \mathcal{F}_2$ is solved by

$$T_\gamma = \int_A \int_A 1_{\{s \leq t\}} (\gamma + \mu_x([0, t]))^{-1} d(M_x \otimes_2 M_y(s, t))$$

with minimum value

$$\varepsilon_{\text{min}}' = \gamma \int_A (\gamma + \mu_x([0, t]))^{-1} d\mu_y(t).$$

Thus for the case considered in Section 1 where $Tx = y$, $T \in \mathcal{C} \cap \mathcal{F}_2$, is actually possible it follows that $\|T_0\|_2^2 = \lim_{\gamma \to 0 +} \gamma^{-1} \varepsilon_{\text{min}}^\gamma$ and that $T_0 = \lim_{\gamma \to 0 +} T_\gamma$ in $\mathcal{F}_2$.

**Proof of Theorem 2.3.** As stated above, the problem is (merely) to identify the operator occurring in Lemma 2.2. We indicate only the main steps of the following calculations.

The key observation is that $C$ has the form $C = I + D$, where

$$D = \int_A \int_A 1_{\{s \leq t\}} (1 + \mu_x([0, s]))^{-1} d(M_x \otimes_2 M_x(s, t)).$$

The desired identity $CC^* = x \otimes x + I$ reduces to the condition that the equation
\[(i) \quad \int_{A \times A} (1 + \mu_x([0, s]])^{-1} \, d\mu_x(u) \]

holds for $\mu_x \times \mu_x$-almost all $(s, t)$ in $A \times A$. The diagonal in $A \times A$ has $\mu_x \times \mu_x$-measure zero. Therefore to verify (i) it suffices to consider the case $s > t$. Here $s \wedge t = t$ and

\[
\int_{[0, t]} (1 + \mu_x([0, u]))^{-1} \, d\mu_x(u)
\]

is equal to the integral on the real axis

\[
\int_0^{\mu_x([0, t])} (1 + \eta)^{-2} \, dF(\eta),
\]

where $F(\eta) = \mu_x(K(\eta))$ is the measure of the set $K(\eta) = \{ u \in A \mid \mu_x([0, u]) \leq \eta \}$. But $K(\eta)$ is a closed interval of the form $K(\eta) = [0, t_0]$. The identity $F(\eta) = \eta$ for $0 \leq \eta \leq \|x\|^2$ and the truth of (i) follow.

Next $C^{-1}$, which exists due to quasinilpotence of $D$, has the form $C^{-1} = I - B$, where

\[
B = \int_A \int_A 1_{\{s \leq t\}} (1 + \mu_x([0, t]))^{-1} \, d(M_x \otimes_2 M_x(s, t)).
\]

The equation $I = C^{-1}C = (I - B)(I + D)$ reduces to the trivial matter that

\[
0 = (1 + \mu_x([0, s]))^{-1} - (1 + \mu_x([0, t]))^{-1} - (1 + \mu_x([0, s]))^{-1}(1 + \mu_x([0, t]))^{-1} \mu_x([s, t])
\]

for all $(s, t)$ in $A \times A$ with $s \leq t$.

Now $Q_{xy} C^{-1} = C^{-1} \otimes y = x \otimes y = Bx \otimes y$. For $Bx$ one finds

\[
dM_{Bx}(t) = (1 + \mu_x([0, t]))^{-1} \mu_x([0, t]) \, dM_x(t).
\]

Hence

\[
(Bx \otimes y) \varphi = \int_A \int_A 1_{\{s \leq t\}} d(M_{Bx} \otimes_2 M_x(s, t))
\]

\[
= \int_A \int_A 1_{\{s \leq t\}} (1 + \mu_x([0, s]))^{-1} \, d(M_x \otimes_2 M_x(s, t))
\]
and it is seen that

\[ (Q, C^{*-1})_6 = \int_A \int_{A'} 1_{s \leq t'} (1 + \mu_x([0, s]))^{-1} d(M_x \otimes M_{x'}(s, t)). \]

The formula for \( T_1 \) follows by postmultiplying this with \( C^{-1} = I - B \), which involves an integration procedure similar to the one in the first part of this proof.

Finally, at this point it is easy to calculate \( \epsilon_{\text{min}} \).

REFERENCES