

A CHARACTERIZATION AND EQUATIONS
FOR
MINIMAL PROJECTIONS AND EXTENSIONS

B.L. CHALMERS and F.T. METCALF

Communicated by Norberto Salinas

ABSTRACT. In this paper we develop first a theory providing a characterization theorem (Theorem 1) for finite-rank minimal projections. In order to demonstrate the usefulness of this characterization we provide several examples. More generally, the theory characterizes operators of minimal norm which extend a fixed linear action on a given finite-dimensional subspace. Secondly, a characterization theorem (Theorem 2) is given for minimal linear operators in a general setting. This setting includes, as examples, minimal and co-minimal projections, optimal recovery and linear estimation, linear n -widths, and best linear approximation to continuous proximity maps. These characterization theorems lead to concrete equations from which the minimal operators can be obtained and, in some important cases, described geometrically.

KEYWORDS: *Minimal projections, linear operators.*

AMS SUBJECT CLASSIFICATION: Primary 41A35; Secondary 41A65.

Let $\mathcal{B} = \mathcal{B}(X, V)$ be the space of all bounded linear operators from a real or complex normed space X into a finite-dimensional subspace V and let \mathcal{P} be the family of all operators in \mathcal{B} with a given action on V (e.g., the identity action corresponds to the family of bounded projections onto V).

DEFINITION. $(x, y) \in S(X^{**}) \times S(X^*)$ will be called an *extremal pair* for $Q \in \mathcal{B}$ if $\langle Q^{**}x, y \rangle = \|Q\|$, where $Q^{**} : X^{**} \rightarrow V$ is the second adjoint extension of Q to X^{**} (S denotes unit sphere).

NOTATION. Let $\mathcal{E}(Q)$ be the set of all extremal pairs for Q . To each $(x, y) \in \mathcal{E}(Q)$ associate the rank-one operator $y \otimes x$ from X to X^{**} given by $(y \otimes x)(z) = \langle z, y \rangle x$ for $z \in X$.

THEOREM 1. (Characterization). P has minimal norm in \mathcal{P} if and only if the closed convex hull of $\{y \otimes x\}_{(x,y) \in \mathcal{E}(P)}$ contains an operator for which V is an invariant subspace.

Proof. The problem is equivalent to best approximating, in the operator norm, a fixed operator $P_0 \in \mathcal{P}$ from the space of operators $\mathcal{D} = \{\Delta \in \mathcal{B} : \Delta = 0 \text{ on } V\} = \text{sp} \{\delta \otimes v : \delta \in V^\perp, v \in V\}$. Let $K = B(X^{**}) \times B(X^*)$ endowed with the product topology, where $B(\cdot)$ denotes the unit ball with its weak* topology. Associate with any operator $Q \in \mathcal{B}$ the bilinear form $\hat{Q} \in C(K)$ via $\hat{Q}(x, y) = \langle Q^{**}x, y \rangle$, and let $\hat{\mathcal{D}} = \{\hat{\Delta} : \Delta \in \mathcal{D}\}$. Then, making use of standard duality theory for $C(K)$, K compact (see e.g., [23], Theorem 1.1 (p.18) and Theorem 1.3 (p.29)), we have that $\hat{P} = \hat{P}_0 - \hat{\Delta}_0$ is of minimal norm if and only if there exists a finite, non-zero (total mass one) signed measure $\hat{\mu}$ supported on the critical set

$$\mathcal{C}(\hat{P}) = \{(x, y) \in S(X^{**}) \times S(X^*) : |\hat{P}(x, y)| = \|\hat{P}\|_\infty\}$$

such that $\text{sgn } \hat{\mu}\{(x, y)\} = \text{sgn } \hat{P}(x, y)$ and $\hat{\mu} \in \hat{\mathcal{D}}^\perp$, i.e.,

$$0 = \int_{\mathcal{C}(\hat{P})} \hat{\Delta} d\hat{\mu} \quad \text{for all } \hat{\Delta} \in \hat{\mathcal{D}}.$$

But now, since any $\hat{Q} \in \{\hat{P}\} \cup \hat{\mathcal{D}}$ is a bilinear function, we can replace the signed measure $\hat{\mu}$, supported in $\mathcal{C}(\hat{P})$, by a positive measure μ supported on $\mathcal{E}(P) \subset \mathcal{C}(\hat{P})$ by noting that

$$\mathcal{C}(\hat{P}) = \{(x, e^{i\theta}y) : (x, y) \in \mathcal{E}(P) \text{ and } \theta \in T\},$$

where $T = [0, 2\pi)$ in the complex case and $T = \{0, \pi\}$ in the real case, and setting

$$\mu\{(x, y)\} = |\hat{\mu}|\{(x, e^{i\theta}y) : \theta \in T\}.$$

For then $\text{sgn } \mu\{(x, y)\} = \text{sgn } \hat{P}(x, y) = 1$, for $(x, y) \in \mathcal{E}(P)$, and

$$0 = \int_{\mathcal{E}(P)} \hat{\Delta} d\mu \quad \text{for all } \Delta \in \mathcal{D},$$

since

$$\begin{aligned} \int_{C(\hat{\mathcal{P}})} \hat{\Delta} d\hat{\mu} &= \int_{\substack{(x,y) \in \mathcal{E}(P) \\ \theta \in \mathcal{T}}} \hat{\Delta}(x, e^{i\theta}y) d\hat{\mu}(x, e^{i\theta}y) \\ &= \int_{\substack{(x,y) \in \mathcal{E}(P) \\ \theta \in \mathcal{T}}} e^{-i\theta} \hat{\Delta}(x, y) e^{i\theta} d|\hat{\mu}|(x, e^{i\theta}y) = \int_{\mathcal{E}(P)} \hat{\Delta} d\mu. \end{aligned}$$

Hence,

$$\begin{aligned} 0 &= \int_{\mathcal{E}(P)} \hat{\Delta} d\mu = \int_{\mathcal{E}(P)} \langle \Delta^{**}x, y \rangle d\mu(x, y) = \int_{\mathcal{E}(P)} \langle x, \delta \rangle \langle v, y \rangle d\mu(x, y) \\ &= \left\langle \int_{\mathcal{E}(P)} \langle v, y \rangle x d\mu(x, y), \delta \right\rangle \end{aligned}$$

for all $\Delta = \delta \otimes v$ ($\delta \in V^\perp, v \in V$), where, for $z \in X$, $\int_{\mathcal{E}(P)} \langle z, y \rangle x d\mu(x, y)$ is the element $w \in X^{**}$ defined by $\langle x^*, w \rangle = \int_{\mathcal{E}(P)} \langle z, y \rangle \langle x^*, x \rangle d\mu(x, y)$ for all $x^* \in X^*$. P is minimal, therefore, if and only if $\int_{\mathcal{E}(P)} \langle v, y \rangle x d\mu(x, y) \in (V^\perp)^\perp = V$, i.e., if and only if there exists an operator (from X into X^{**})

$$(1) \quad E_P = \int_{\mathcal{E}(P)} y \otimes x d\mu(x, y) : V \rightarrow V. \quad \blacksquare$$

REMARK 1. The identification of a minimal norm operator as the error of a best approximation problem in $C(K)$, in the proof of Theorem 1, is useful for examining various other aspects of minimal projections and extensions. For example, if in the proof of Theorem 1, we apply the Kolmogorov criterion for best approximation (see, e.g., [23], Theorem 1.16, p.69), we also get the following characterization.

THEOREM 1.A. P has minimal norm in \mathcal{P} if and only if there does not exist $\Delta \in \mathcal{D} = \{\Delta \in \mathcal{B} : \Delta = 0 \text{ on } V\}$ such that

$$\sup_{(x,y) \in \mathcal{E}(P)} \operatorname{Re} \langle P^{**}x, y \rangle \overline{\langle \Delta^{**}x, y \rangle} < 0.$$

Theorem 1.A was proved in [13] in the case $X = L^1$ and \mathcal{P} consists of projections, where it was then used successfully in the landmark determination of a minimal projection from $L^1[-1, 1]$ onto the lines (see Example 4.a below). See also [20] for examples where Theorem 1.A is proved and used in other settings.

NOTE 1. The existence of a minimal operator characterized in Theorem 1 follows from the fact that V is a finite-dimensional subspace of X and [19], where the existence of a minimal projection is shown in the more general case of V being a dual space, and by noting that the argument of [19] applies just as well to the case of minimizing over the more general class of operators in Theorem 1.

As a first example of the use of Theorem 1, we have the following sufficient condition for the adjoint of a minimal projection to be itself minimal. E_P will refer to the (not necessarily unique) operator in (1). For further discussion of the nature of E_P see Note 4 below.

COROLLARY 1. ([10]). *Let P be a minimal projection. Then P^* is a minimal projection from X^* (onto $(\ker P)^\perp$) if $P^{**} \circ E_P = E_P \circ P$.*

Proof. (Sketch). $X = V \oplus U^\perp$, where $U^\perp = \ker P$ ($U = \text{range } P^*$). Then E_P (as well as P) takes V into V , and it follows that $P^{**} \circ E_P = E_P \circ P$ if and only if E_P takes U^\perp into U^\perp . But the latter occurs if and only if $(E_P)^*$ takes U into U . Finally, (x, y) is an extremal pair for P implies that (y, x) is an extremal pair for P^* , and hence $P^{**} \circ E_P = E_P \circ P$ implies that $(E_P)^* = E_{P^*}$. ■

In the examples and discussion below it is helpful to introduce a fixed basis $\mathbf{v} = (v_1, \dots, v_n)$ for V ; we will write $V = [\mathbf{v}] = [v_1, \dots, v_n]$. Then the necessary and sufficient condition (1) can be rewritten as a system of n equations

$$(2) \quad \int_{\mathcal{E}(P)} \langle \mathbf{v}, \mathbf{y} \rangle x \, d\mu(x, \mathbf{y}) = M \mathbf{v} \text{ for some matrix } M.$$

Let $\sum_{i=1}^n u_i \otimes v_i \in X^* \check{\otimes} V$ (the injective tensor product of X^* and V) represent Q (and Q^{**} , where we set $\langle z, u_i \rangle = \langle u_i, z \rangle$ for $z \in X^{**}$), i.e., $Qx = \sum_{i=1}^n \langle x, u_i \rangle v_i$. Then $\mathcal{E}(Q) = \{(x, \mathbf{y}) \in S(X^{**}) \times S(X^*) : \sum_{i=1}^n \langle x, u_i \rangle \langle v_i, \mathbf{y} \rangle = \|Q\|\}$ and we can write

$$\mathcal{P} = \left\{ \sum_{i=1}^n u_i \otimes v_i : \langle u_i, u_j \rangle = A_{ij} \text{ for } A \text{ a given fixed } n \times n \text{ matrix} \right\}.$$

NOTE 2. M in (2) may be regarded as a function of A above. Hence, (2) may be regarded as determining a minimal P up to the n^2 entries of M , which are in turn determined by the n^2 entries of A .

NOTATION. If $z \in Z$ and $z^* \in Z^*$ are such that $\langle z, z^* \rangle = \|z\| \|z^*\| \neq 0$, then z^* is an extremal for z and we write $z^* = \text{ext } z$. (Then also $z = \text{ext } z^*$.) Note that $\text{ext } z$ is determined only up to a non-zero scalar factor.

For purposes of illustration, observe the following simple examples (Examples 1.a and 1.b) of Theorem 1, where minimal P has an extremal pair (x, y) with $x \in V$, and therefore we can take $E_P = y \otimes x$.

EXAMPLE 1.a. ($\dim V = 1 = \text{rank } P$). Let $P = u_1 \otimes v_1$. Then $(x, y) \in \mathcal{E}(P)$ if and only if $\langle x, u_1 \rangle \langle v_1, y \rangle = \|P\|$, i.e., $(x, y) = (\text{ext } u_1, \text{ext } v_1)$. Then $E_P = y \otimes x : V \rightarrow V$ if and only if $\text{ext } u_1 = v_1$, i.e., $u_1 = \text{ext } v_1$ (the solution given by the Hahn-Banach extension theorem). This example extends to $\dim V > 1 = \text{rank } P$ by applying this $n = 1$ example to $X/V \cap \ker P$.

EXAMPLE 1.b. (Hilbert space). Let X be a Hilbert space, let the basis \mathbf{v} for V be orthonormal, let \mathcal{P} have a fixed "diagonal" action, i.e., $A_{ij} = d_i \delta_{ij}$, where $\mathbf{d} = (d_1, \dots, d_n)$ is a fixed n -tuple of scalars, and let $J = \{j : |d_j| = \max |d_i|\}$. Then $P = \sum \bar{d}_i v_i \otimes v_i$ is minimal, where $E_P = y \otimes x$ for any choice of $(x, y) = (z, z)$ with z an arbitrary norm-1 element of the eigenspace corresponding to a maximum eigenvalue $d_j, j \in J$.

The minimality of the Fourier projection in the context of compact abelian groups is a simple consequence of Theorem 1 as demonstrated by the following example.

EXAMPLE 2. ([8]). Let T (with "+") be a compact abelian group with Haar measure ν , \hat{T} its dual, $\{v_\gamma\}_{\gamma \in \hat{T}}$ the set of all characters, N a finite part of \hat{T} , V the linear span of the characters $v_\tau, \tau \in N$ and let $X = L^p(T), 1 \leq p < \infty$ or $X = C(T), p = \infty$. Then the Fourier projection $F = \sum_{\tau \in N} v_\tau \otimes v_\tau$ is minimal among all projections from X onto V .

Proof. (Sketch). Let (x, y) be any extremal pair for F . Then $(x_t, y_t) = (x(\cdot + t), y(\cdot + t))$ is an extremal pair for each $t \in T$. Thus, $E_F = \int_T y_t \otimes x_t \, d\nu(t) : V \rightarrow V$, since $\langle E_F v_\tau, v_\gamma \rangle = \langle v_\tau, y \rangle \langle x, v_\gamma \rangle \langle v_\gamma, v_\tau \rangle$ and $\langle v_\gamma, v_\tau \rangle = \delta_{\gamma\tau} = 0$ if $\tau \in N$ and $\gamma \in \hat{T} \sim N$. ■

NOTATION. Write $\langle \mathbf{v}, y \rangle \cdot \mathbf{u}$ for $\sum_{i=1}^n \langle v_i, y \rangle u_i$ in the following.

NOTE 3. For $Q = \sum_{i=1}^n u_i \otimes v_i, (x, y) \in \mathcal{E}(Q)$ implies $x = \text{ext}(\langle \mathbf{v}, y \rangle \cdot \mathbf{u})$ and $y = \text{ext}(\langle x, \mathbf{u} \rangle \cdot \mathbf{v})$, which shows that, to find extremal pairs in general, we must solve the (non-linear) equation

$$(3) \quad \mathbf{d} = \langle \mathbf{v}, y \rangle = \langle \mathbf{v}, \text{ext}(\langle \text{ext}(\mathbf{d} \cdot \mathbf{u}), \mathbf{u} \rangle \cdot \mathbf{v}) \rangle$$

for n -tuples of scalars $\mathbf{d} = (d_1, \dots, d_n) = (\langle v_1, y \rangle, \dots, \langle v_n, y \rangle)$. For $X = C(T)$ and $L^1(T)$, respectively, however, the extremal pairs of Q have, on the support of \mathbf{u} and the support of \mathbf{v} , respectively, the simple forms $(x_t, y_t) = (\text{sgn}(\mathbf{v}(t) \cdot \mathbf{u}), \delta_t)$ and $(\delta_t, \text{sgn}(\mathbf{u}(t) \cdot \mathbf{v}))$, respectively, since $\|Q\| = \sup_{t \in T} L(t)$, where $L(t) = \langle x_t, \mathbf{v}(t) \cdot \mathbf{u} \rangle$, and $\langle \mathbf{u}(t) \cdot \mathbf{v}, y_t \rangle$, respectively, is the so-called "Lebesgue function" of Q . This fact makes these important cases relatively easy to consider.

That P^* in Corollary 1 is not always minimal (and that therefore E_P does not always commute with P) is shown by the following example, which also serves as an example of Theorem 1.

EXAMPLE 3.a. ([21]). Consider $X = \ell_6^\infty$ and $V = [v_1, v_2]$, where $v_1 = (10\alpha\beta\alpha\beta)$, $v_2 = (01\beta\alpha^{-1}\beta^{-1}\alpha)$, with $\alpha = (2 + \sqrt{2})/4$, $\beta = \sqrt{2}/4$. Then the "interpolating projection" $P = \sum_{i=1}^2 u_i \otimes v_i$, where $u_i = \varepsilon_i$ (the i^{th} standard basis element in \mathbb{R}^6), $i = 1, 2$, is minimal. This is seen by first checking that $\{(x_i, y_i)\}_{i=1}^4 \subset S(\ell_6^\infty) \times S(\ell_6^1)$ are extremal pairs, where $x_1 = (111110)$, $x_2 = (11110^{-1})$, $x_3 = (1^{-1}1011)$, $x_4 = (1^{-1}0^{-1}11)$, and $y_i = \varepsilon_{i+2}$, $i = 1, 2, 3, 4$. Indeed, $L(t) = \sum_{s=1}^6 |\mathbf{u}(s) \cdot \mathbf{v}(t)| = \langle x_{t-2}, \mathbf{u} \rangle \cdot \langle \mathbf{v}, y_{t-2} \rangle = \alpha + \beta$ for $t = 3, 4, 5, 6$, while $L(1) = L(2) = 1$. Secondly, check that (2) holds:

$$\sum_{i=1}^4 \langle \mathbf{v}, y_i \rangle \frac{x_i}{4} = M\mathbf{v}, \quad \text{where } M = \frac{\alpha + \beta}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus P is minimal, with norm $\alpha + \beta > 1$, but the projection $\sum_{i=1}^2 \varepsilon_i \otimes u_i$ from ℓ_6^1 onto $[u_1, u_2]$ has norm 1, and thus $P^* : \ell_6^1 \rightarrow [u_1, u_2]$ is minimal ($\|P^*\| = \|P\|$).

NOTE 4. (See, e.g., [22] or [24] for definitions and notation). The operator E_P of (1) can be viewed as a norm-one integral operator in $(X^* \otimes V)^*$ separating P from $\mathcal{D} = \{\Delta \in \mathcal{B} : \Delta = 0 \text{ on } V\}$, i.e.,

$$\begin{aligned} \langle P, E_P \rangle &= \text{tr} \left(\sum_1^n u_i \otimes E_P v_i \right) = \sum_1^n \langle E_P v_i, u_i \rangle = \sum_1^n \int \langle v_i, y \rangle \langle x, u_i \rangle d\mu \\ &= \int \langle P^{**} x, y \rangle d\mu = \|P\|, \end{aligned}$$

$\langle \mathcal{D}, E_P \rangle \equiv 0$ as in the proof of Theorem 1, and $\nu(E_P) \leq \int_{\mathcal{E}(P)} \|y\| \|x\| d\mu(x, y) = 1$, where ν denotes the norm of E_P in the space of integral operators $I_1(X, X^{**})$. Note also from the above that $\|P\| = \text{tr}(MA)$, for which upper bounds are determined in [9], extending those in [16] for projections ($A = I$).

By use of Note 4 and the observation that $\langle P, E_P \rangle = \text{tr}(E_P|_V \circ P)$ we have the following known corollary, used for example in [15] and [16], and also used in [21] for obtaining the projection of Example 3.a.

COROLLARY 2. *The relative projection constant (of V relative to X) $\lambda(V, X) = \inf_{P \in \mathcal{P}} \|P\|$, where \mathcal{P} is the family of projections, is given by*

$$\lambda(V, X) = \sup\{\text{tr}(Q|_V) : Q \in I_1(X, X^{**}), Q : V \rightarrow V, \nu(Q) = 1\}.$$

Note that if X is finite-dimensional then $I_1(X, X^{**}) = X^* \widehat{\otimes} X^{**}$ (the projective tensor product of X^* and X^{**}) and ν in Corollary 2 is the nuclear norm.

COROLLARY 3. *Let P have minimal norm in \mathcal{P} . If $\mathcal{E}_1(P) = \{x : (x, y) \in \mathcal{E}(P) \cap \text{supp } \mu\}$ is an independent set, then, for each $x \in \mathcal{E}_1(P)$, let $x^0 \in (\text{span } \mathcal{E}_1(P))^*$ such that $\langle x, x^0 \rangle = 1$ and $\langle z, x^0 \rangle = 0$ for all $z \neq x$ in $\mathcal{E}_1(P)$, and act on (2) with x^0 to get*

$$(2') \quad \langle \mathbf{v}, \mathbf{y} \rangle \mu\{(x, y)\} = M \langle \mathbf{v}, x^0 \rangle.$$

In Examples 4-5 below, we restrict ourselves to minimal projections (i.e., $A = I$).

EXAMPLE 4.a. ([13]). $X = L^1[-1, 1] \supset V = [\mathbf{v}]$, $\mathbf{v} = (1, t)$. Use Note 3 ($(x_t, y_t) = (\delta_t, \text{sgn}(u(t) \cdot \mathbf{v}))$), Corollary 3, the clear fact that $x_t^0|_V = \chi_{\{t\}}$, and symmetry considerations (whence $M = \kappa \text{diag}(1, m)$ for some scalar κ) to write (2') as follows (cancelling the (infinitesimal) scalar multipliers $\mu\{(x_t, y_t)\}$ and dt):

$$\begin{pmatrix} \langle v_1, y_t \rangle \\ \langle v_2, y_t \rangle \end{pmatrix} = \kappa \begin{pmatrix} v_1(t) \\ m v_2(t) \end{pmatrix} = \kappa \begin{pmatrix} 1 \\ m t \end{pmatrix},$$

i.e.,

$$(2'') \quad \frac{1}{\langle v_1, y_t \rangle} = \frac{m t}{\langle v_2, y_t \rangle},$$

where m is a scalar and

$$\begin{aligned} \langle \mathbf{v}, \mathbf{y}_t \rangle &= \int_{-1}^1 \mathbf{v}(s) \text{sgn}(u(t) \cdot \mathbf{v}(s)) ds = \varepsilon(t) \left[\int_{-1}^{r(t)} - \int_{r(t)}^1 \mathbf{v}(s) ds \right] \\ &= \varepsilon(t)(2r(t), r^2(t) - 1), \end{aligned}$$

where $r(t)$ is defined (uniquely since V is a Chebyshev system) by $u(t) \cdot \mathbf{v}(r(t)) = 0$ and $\varepsilon(t) = \pm 1$. Equation (2'') is then easily solved for $r(t)$, yielding the admissible solution (in $[-1, 1]$) $r(t) = m t - \text{sgn}(m t) \sqrt{m^2 t^2 + 1}$, and then u is obtained from

the linear relations $\mathbf{u}(t) \cdot \mathbf{v}(r(t)) = 0$ and $L(t) = \mathbf{u}(t) \cdot \langle \mathbf{v}, \mathbf{y}_t \rangle \equiv \lambda, t \in [-1, 1]$. I.e., $P = \sum_{i=1}^2 u_i \otimes v_i$ is minimal where

$$\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} 2r(t) & r^2(t) - 1 \\ 1 & r(t) \end{pmatrix}^{-1} \begin{pmatrix} \lambda \sigma(t) \\ 0 \end{pmatrix},$$

$\sigma(t) = \text{sgn}((r^2(t) - 1)/mt)$ and m and λ are determined to meet the remaining normality conditions, esp., $\lambda = \|P\| = 1.22040\dots$

This L^1 -example and Example 4.b which follows demonstrate how Theorem 1 provides a direct formula for P via Corollary 3. (In the L^1 -cases, $\{x_t = \delta_t : t \in T\}$ is an independent set (see also [5])).

EXAMPLE 4.b. ([5]). $X = L^1[-1, 1] \supset V = [\mathbf{v}], \mathbf{v} = (1, t, t^2)$. Analogously as in Example 4.a, (2') becomes

$$(2''') \quad \frac{m_{11} + m_{13}t^2}{2(r_1(t) - r_2(t) + 1)} = \frac{t}{r_1^2(t) - r_2^2(t)} = \frac{m_{31} + m_{33}t^2}{\frac{2}{3}(r_1^3(t) - r_2^3(t) + 1)},$$

where the $r_i(t)$ are defined to be the roots of the quadratic equation $\mathbf{u}(t) \cdot \mathbf{v}(r) = 0$, i.e., $\mathbf{u}(t) \cdot \mathbf{v}(r_i(t)) = 0, i = 1, 2$. Next solve equation (2''') for $r_1(t)$ and $r_2(t)$ and then \mathbf{u} is obtained from the linear relations $\mathbf{u}(t) \cdot \mathbf{v}(r_i(t)) = 0, i = 1, 2$, and $L(t) = \mathbf{u}(t) \cdot \langle \mathbf{v}, \mathbf{y}_t \rangle \equiv \lambda, t \in [-1, 1]$. I.e., $P = \sum_{i=1}^3 u_i \otimes v_i$ is minimal where

$$\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} = \begin{pmatrix} 2(r_1(t) - r_2(t) + 1) & r_1^2(t) - r_2^2(t) & \frac{2}{3}(r_1^3(t) - r_2^3(t) + 1) \\ 1 & r_1(t) & r_1^2(t) \\ 1 & r_2(t) & r_2^2(t) \end{pmatrix}^{-1} \begin{pmatrix} \lambda \sigma(t) \\ 0 \\ 0 \end{pmatrix},$$

$\sigma(t) = \text{sgn}(\frac{2}{3}(r_1^3(t) - r_2^3(t) + 1)/(m_{31} + m_{33}t^2))$, and $m_{11}, m_{13}, m_{31}, m_{33}$ and λ are determined to meet the remaining normality conditions, esp., $\lambda = \|P\| = 1.35948\dots$

In the case $X = L^1$, there is a remarkably simple geometric solution (Corollary 4 below) to the problem of minimal projections and extensions, which accounts for the relative simplicity of Examples 4.a and 4.b above.

NOTATION. Denote the underlying real or complex field by \mathbf{F} and introduce the norm on \mathbf{F}^n given by $\|\mathbf{a}\| = \|\mathbf{a} \cdot \mathbf{v}\|_X$. The following proposition demonstrates a very useful geometric connection in \mathbf{F}^n between the two components of an extremal pair for any $P \in \mathcal{P}$.

PROPOSITION 1. For any extremal pair (x, y) of $P = \sum_{i=1}^n u_i \otimes v_i$,

$$(4) \quad \langle x, u \rangle = \|P\| \frac{\overline{\alpha \langle v, y \rangle} - c^*}{\|\overline{\alpha \langle v, y \rangle} - c^*\|},$$

where α is any positive scalar and $c^* \in F^n$ yields $\min \|\overline{\alpha \langle v, y \rangle} - c\|$ subject to $c \cdot \langle v, y \rangle = 0$.

Proof. Fix $x \in \mathcal{E}(P)$ and let $C = \{c \in F^n : c \cdot \langle v, y \rangle = 0\}$. Then there exists a positive α ($= |\langle v, y \rangle|^2 / \|P\|$) and $c^* \in C$ such that $\overline{\langle x, u \rangle} = \overline{\alpha \langle v, y \rangle} - c^*$. Hence, $0 = c \cdot \langle v, y \rangle = \langle c \cdot v, \text{ext}((x, u) \cdot v) \rangle = \langle c \cdot v, \text{ext}(\overline{\alpha \langle v, y \rangle} \cdot v - c^* \cdot v) \rangle$, for all $c \in C$, which implies that $c^* \cdot v$ is a best approximation to $\overline{\alpha \langle v, y \rangle} \cdot v$ from $\{c \cdot v : c \in C\}$ with respect to the norm of X . Hence c^* yields the minimum of $\|\overline{\alpha \langle v, y \rangle} - c\|$. Further, $\|P\| = \langle x, u \rangle \cdot \langle v, y \rangle = \langle (\overline{\alpha \langle v, y \rangle} - c^*) \cdot v, y \rangle = \|(\overline{\alpha \langle v, y \rangle} - c^*) \cdot v\|_X = \|\overline{\alpha \langle v, y \rangle} - c^*\|$, since $y = \text{ext}((x, u) \cdot v)$. Finally, note that α can be replaced by any other positive quantity by scaling simultaneously the numerator and denominator in (4). ■

NOTE 5. Geometrically, (4) says that $\langle x, u \rangle / \|P\|$ is a point of intersection of the unit $\|\cdot\|$ -sphere in F^n and its tangent plane perpendicular (in the ordinary Euclidean sense) to the direction of $\langle v, y \rangle$.

THEOREM 1.B. Under the hypotheses of Corollary 3, $P = \sum_{i=1}^n u_i \otimes v_i$ is minimal, where

$$\langle x, u \rangle = \|P\|z(x),$$

with $z(x)$ being a point of intersection of the unit $\|\cdot\|$ -sphere in F^n ($\|a\| = \|a \cdot v\|_X$) and its tangent plane perpendicular to $M(v, x^0)$.

Proof. Apply (2') to (4) and use Note 5. ■

COROLLARY 4. (Geometric interpretation for L^1) ([5]). Let $X = L^1(T) \supset V = [v]$. Then $P = \sum_{i=1}^n u_i \otimes v_i$ is minimal where

$$(5) \quad u(t) = \|P\|z(t),$$

with $z(t)$ being a point of intersection of the unit $\|\cdot\|$ -sphere in F^n ($\|a\| = \|a \cdot v\|_{L^1}$) and its tangent plane perpendicular to $Mv(t)$ for some M .

Proof. Take $x_t = \delta_t$ and $x_t^0|_V = \chi_{\{t\}}$ in Theorem 1.B. ■

REMARK 2. Corollary 4 has been especially useful in [7].

By use of (2) and (3), we can rewrite Theorem 1 in a form which is constructive in terms of creating a minimal P from \mathcal{P} by use only of the Banach space geometry of V as a subset of X .

THEOREM 1.C. (Equations). $P = \sum_{i=1}^n u_i \otimes v_i$ has minimal norm in \mathcal{P} if and only if, for some matrix M ,

$$\int_{\mathcal{E}} \langle v, y \rangle x \, d\mu(x, y) = Mv,$$

where $\mathcal{E} = ES(U^*) \times ES(V^*)$ (ES denotes the extreme points of the unit sphere S), with $U = [u] \subset X^*$, and u, x and y satisfying $y = \text{ext}(\langle x, u \rangle \cdot v)$, $x = \text{ext}(\langle v, y \rangle \cdot u)$, and $\langle v_i, u_j \rangle = A_{ij}$.

COROLLARY 5. In Theorem 1.C, suppose (as in the cases $X = L^1$ or $X = C$) that the $\langle v, y \rangle = d$ are known (up to a scalar multiple). Then u (subject to the conditions $\langle v_i, u_j \rangle = A_{ij}$) is determined from the single equation $\text{ext}(d \cdot u, u) = e$, where $e = \langle x, u \rangle$ is determined from Proposition 1.

As a first example of the use of Theorem 1.C, we can construct the projection of Example 3.a as follows.

EXAMPLE 3.b. Consider $X = \ell_7^\infty$ and $V = [v_1, v_2]$, where $v_1 = (0 \beta \alpha 1 \alpha \beta 0)$, $v_2 = (1 \alpha \beta 0 -\beta -\alpha -1)$ with $\alpha = (2 + \sqrt{2})/4$ and $\beta = \sqrt{2}/4$. Write each v_i as a (convex) combination of the (signed) independent extreme points $x_1 = (1 1 1 1 1 -1)$, $x_2 = (1 1 1 1 1 -1)$, \dots , $x_6 = (1 -1 -1 -1 -1 -1)$ in $ES(\ell_7^\infty)$:

$$\frac{1}{4(\alpha + \beta)} [c_1 x_1 + 2c_2 x_2 + c_3 x_3 + c_4 x_4 + 2c_5 x_5 + c_6 x_6] = v$$

where $c_1 = (\alpha, \beta)$, $c_2 = (\alpha + \beta, \alpha + \beta)/2$, $c_3 = (\beta, \alpha)$, $c_4 = (-\beta, \alpha)$, $c_5 = (-\alpha - \beta, \alpha + \beta)/2$, $c_6 = (-\alpha, \beta)$. Set $\langle v, y_i \rangle = c_i$, $i = 1, \dots, 6$, where $y_1 = \varepsilon_3$, $y_2 = (\varepsilon_2 + \varepsilon_3)/2$, $y_3 = \varepsilon_2$, $y_4 = -\varepsilon_6$, $y_5 = -(\varepsilon_5 + \varepsilon_6)/2$, $y_6 = -\varepsilon_5$. Next, construct the symmetric 12-sided ball $\|a\| \doteq \|a \cdot v\|_{\ell_7^\infty} = 1$ and use Note 5 to conclude that $\langle x_i, u \rangle = (1, 1)$, $i = 1, 2, 3$, and $\langle x_i, u \rangle = (-1, 1)$, $i = 4, 5, 6$. I.e., $u_1 = \varepsilon_4$ and $u_2 = \varepsilon_1$.

NOTE 6. If $X = L^p(T)$, $1 \leq p < \infty$ or $X = C(T)$, $p = \infty$, and V is piecewise continuously differentiable, then for $P = \sum_{i=1}^n u_i \otimes v_i$ minimal, we have the following

necessary *linear* equation for \mathbf{u} (obtained by the first author using different methods in the cases $p = 1, \infty$ in [2] and as a corollary of Theorem 1 of the present paper in [3])):

$$(*) \quad \frac{1}{p} \mathbf{u}' \cdot M \mathbf{v} = \frac{1}{q} \mathbf{u} \cdot M \mathbf{v}' \quad \text{on } T,$$

where M is the matrix in (2) and $1/q + 1/p = 1$. (One may view $(*)$ as an n -dimensional version of the Hölder equality condition.) If $p = 1$, $(*)$ is an easy consequence of (2). If $p = \infty$, $(*)$ is very useful and when used in conjunction with (2), as in Example 5.c below, yields a second *linear* equation for \mathbf{u} .

EXAMPLE 5.a. $X = C[-1, 1] \supset V = [\mathbf{v}]$, $\mathbf{v} = (1, t)$. Then $P = \sum_{i=1}^2 u_i \otimes v_i$ is minimal with norm 1, where

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} A & A \\ -C & C \end{pmatrix} \begin{pmatrix} \delta_{-1} \\ \delta_1 \end{pmatrix},$$

with $A = C = \frac{1}{2}$.

Proof. (Sketch). Check that the orthonormality conditions hold and that (2) holds with $(x_t, y_t) = (\text{sgn}(\cdot - t), (\delta_1 - \delta_{-1})/2)$ and $d\mu(t) = dt/2, -1 \leq t \leq 1, M = \text{diag}(0, 1)$. ($E_P x = \frac{1}{2}[x(1) - x(-1)]v_2$ for all $x \in X$.) ■

In the case of the following example, in [1] the form of P was guessed by use of the above $(*)$ -equation and local constancy of the Lebesgue function, and the minimality of P was proved by showing that its norm was the same as the norm of a known ([13]) minimal projection in an isometric L^1 -setting. In the present paper we check that P is minimal by direct use of Theorem 1. Recall (Note 6 above) that the equation $(*)$ can be derived as a consequence of Theorem 1 ([3]). Also local constancy of the Lebesgue function is a consequence of using knowledge of the form of the extremal pairs in Theorem 1. In other words, Theorem 1 is being used both to obtain the formula for P and to establish its minimality.

EXAMPLE 5.b. ([1]). $X = C[-1, 1] \supset V = [\mathbf{v}]$, $\mathbf{v} = (1 - t^2, t)$. Then $P = \sum_{i=1}^2 u_i \otimes v_i$ is minimal with norm 1.220404917116354... (same as $\|P\|$ in Example 4.a), where

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & B & 0 \\ -C & 0 & C \end{pmatrix} \begin{pmatrix} \delta_{-1} \\ \delta_0 \\ \delta_1 \end{pmatrix} + \begin{pmatrix} b \\ cs \end{pmatrix} \frac{1}{[1 + (\sigma s)^2]^{\frac{3}{2}}}, \quad -1 \leq s \leq 1,$$

with $(B, C, b, c) = (\lambda/2)(1, 1/\sqrt{1+\sigma^2}, 2\sigma, \sigma^2)$, $\lambda = \|P\| = -2\sigma/\log t_0$, $\sigma = (1 - t_0^2)/2t_0 = (t_0^2 - t_0 - 1)\log t_0$.

Proof. (Sketch). Use (*) ($0 = \mathbf{u} \cdot M\mathbf{v}'$) to determine the continuous part of $\mathbf{u}(s)$ to be $w(s)(b, cs)$ for scalars b and c and then determine the scalar function $w(s)$ to be $[1 + (\sigma s)^2]^{-3/2}$ by forcing local constancy of the Lebesgue function. Choose all the remaining parameters so that the (ortho)normality conditions hold and that (2) holds (with $(x_t, y_t) = (\text{sgn}(t)\text{sgn}(\cdot - r(t)), \delta_t)$, $d\mu(t) = (1 + t^2)dt/|t|^3$ on $t_0 \leq |t| \leq 1$ (μ unnormalized), where $r(t) = (t^2 - 1)/2\sigma t$, and $M = (4\sigma^2)\text{diag}(1, 1/\sigma)$. ■

As in the previous example, the form of P in the following example is obtained with the help of the equation (*), which can be derived as a consequence of Theorem 1 ([3]) as noted in Note 6, and by use of local constancy of the Lebesgue function, again a consequence of Theorem 1. In [6] it was checked that P is minimal by direct use of Theorem 1. I.e., Theorem 1 is again being used both to obtain the formula for P and to verify its minimality.

EXAMPLE 5.c. ([6]). $X = C[-1, 1] \supset V = [\mathbf{v}]$, $\mathbf{v} = [1, t, t^2]$. This projection was first found by the authors in 1978, and may now be verified as minimal by use of Theorem 1. $P = \sum_{i=1}^3 u_i \otimes v_i$ is minimal with norm 1.220173064217988..., where

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} A & B & A \\ -C & 0 & C \\ D & -B & D \end{pmatrix} \begin{pmatrix} \delta_{-1} \\ \delta_0 \\ \delta_1 \end{pmatrix} + \sum_{k=1}^2 \begin{pmatrix} b_k + a_k|s| \\ c_k s \\ -b_k + d_k|s| \end{pmatrix} \frac{\chi_{S_k}(|s|)}{(1 + w_k|s|)^3},$$

for $-1 \leq s \leq 1$, with $S_k = [s_{k1}, s_{k2}]$ ($k = 1, 2$). For the equations (and values) for all the parameters see [6].

Proof. (Sketch). Assume extremal pairs (x_t, δ_t) and $(x_t^{(1)}, \delta_{\pm 1})$, for t in the region of constancy of $L(t)$, where $x_t = \text{sgn}(t)\text{sgn}(\cdot - r(t))$, $x_t^{(1)}(s) = \pm x_t(s)$ for $|s| \in S_1 \cup S_2$, and the sign \pm is constant for $|s| \in S_k$, $k = 1, 2$. Then apply x_t^0 (x_t^0 is differentiation at $r(t)$) to (2) to obtain

$$(2'') \quad w_1(t)\mathbf{v}(t) + w_2(t)\mathbf{v}(\pm 1) = M\mathbf{v}'(r(t)),$$

where w_1 and w_2 are positive scalar-valued functions. Next, "dot" both sides of (2'') with $\mathbf{u}(r)$ to get

$$w_1(t)\mathbf{u}(r) \cdot \mathbf{v}(t) + w_2(t)\mathbf{u}(r) \cdot \mathbf{v}(\pm 1) = \mathbf{u}(r) \cdot M\mathbf{v}'(r).$$

But $u(r(t)) \cdot v(t) = 0$ and apply (*) to obtain the linear equation $v(\pm 1) \cdot u = 0$ for u . Next use this equation, and (*), to determine the continuous part of $u(s)$ to be $w(s)(b_k + a_k|s|, c_k s, -b_k + d_k|s|)$, where $c_k = (-1)^k(d_k + a_k)$ and then determine the scalar function $w(s)$ to be $1/(1+w_k|s|)^3, |s| \in S_k, k = 1, 2, w_1 = (a_1 - d_1)/2b_1, w_2 = -w_1$, by forcing local constancy of the Lebesgue function. Finally, choose all the remaining parameters so that the orthonormality conditions hold and so that (2) holds (see [6]). ■

For some further examples which could be obtained by using Theorem 1, see, e.g., [11] and [14]. Theorem 1 is a special case of Theorem 2 which follows. The proof of Theorem 2 is an easy extension of the proof of Theorem 1.

DEFINITION. If X and Y are any two Banach spaces, a continuous homogeneous (not necessarily linear) operator Q from X into Y will be said to be *jointly weak* continuous* if Q has a continuous homogeneous extension $Q^{**} : X^{**} \rightarrow Y$ such that $\widehat{Q}(x, y) = \langle Q^{**}x, y \rangle$ is continuous on the compact set $K = B(X^{**}) \times B(Y^*)$. (E.g., $Q \in X^* \otimes Y$ or Q any compact linear operator.)

THEOREM 2. Let P_0 be a continuous homogeneous operator from a Banach space X into a Banach space Y . Let W be a subspace of X^*, V be a subspace of Y , and $\mathcal{D} = \overline{W \otimes V}$, where the closure is taken in $\mathcal{B}(X, Y)$. Then $P = P_0 - \Delta_0$ has minimal norm in $\mathcal{P} = P_0 + \mathcal{D}$ if the closed convex hull of $\{y \otimes x\}_{(x,y) \in \mathcal{E}(P)}$ contains an operator taking V into W^\perp , i.e., for some (total mass one) measure μ supported in $\mathcal{E}(P) \subset S(X^{**}) \otimes S(Y^*)$

$$(6) \quad E_P = \int_{\mathcal{E}(P)} y \otimes x \, d\mu(x, y) : V \rightarrow W^\perp,$$

or equivalently,

$$E_P^* = \int_{\mathcal{E}(P)} x \otimes y \, d\mu(x, y) : W \rightarrow V^\perp.$$

If P_0 is jointly weak* continuous, then (6) is also necessary for P to be minimal.

EXAMPLE A. (Co-minimal (or minimal) projections (in the case V finite dimensional)). Let $Y = X$, take $P_0 = I - P_1$ (or $P = P_1$) where P_1 is a projection with (fixed) range V , and $W = V^\perp$.

$$(6.A) \quad E_P : V \rightarrow V.$$

As an example where P_0 is not necessarily linear, replace I above by a continuous proximity map onto V .

EXAMPLE B. (Minimal extension operators (in the case V finite dimensional)). Let $V \subset S \subset X$, let P_0 be a fixed operator from S to V , and let $P : X \rightarrow V$ be a minimal-norm extension of P_0 . In Theorem 2, let $Y = X$ and $W = S^\perp$.

$$(6.B) \quad E_P : V \rightarrow S.$$

EXAMPLE C. (Optimal recovery (in the case U finite dimensional)). In Theorem 2, let $Y = X$, take $P_0 = I - P_1$ (or $P_0 = P_1$) where P_1 is a projection with (fixed) kernel U^\perp , $W = U$, $V = U^\perp$.

$$(6.C) \quad E_P^* : U \rightarrow U.$$

COROLLARY 6. (Geometric interpretation for optimal recovery in C). Let $X^* = C(T)^* \supset U = [\mathbf{u}]$. Then $P = \sum_{i=1}^n u_i \otimes v_i$ is minimal where

$$\mathbf{v}(t) = \|P\|z(t),$$

with $z(t)$ being a point of intersection of the unit $\|\cdot\|$ -sphere in F^n ($\|\mathbf{a}\| = \|\mathbf{a} \cdot \mathbf{u}\|_{C(T)^*}$) and its tangent plane perpendicular to $M\mathbf{u}(t)$ for some M .

EXAMPLE D. (Linear optimal estimation (in the case U finite dimensional)). In Theorem 2, let P_0 be linear with (fixed) kernel U^\perp , $W = U$, $V = X \subset Y$.

$$(6.D) \quad E_P^* : U \rightarrow 0.$$

EXAMPLE E. (Linear n -widths). Let P_0 be the injection of X into $Y \supset X$, let n be fixed (and finite) and let $\tilde{\mathcal{D}} = \left\{ \sum_{i=1}^n \delta_i \otimes \varepsilon_i, \delta_i \otimes \varepsilon_i \in X^* \otimes Y \right\}$. Now $\tilde{\mathcal{D}}$ is not a subspace, but apply Theorem 2 to subspaces $\mathcal{D}_1 = X_n^* \otimes Y$ and $\mathcal{D}_2 = X^* \otimes Y_n$, where X_n^* and Y_n are n -dimensional subspaces of X^* and Y , respectively, and determine $P = P_1 = P_2$ from

$$(6.E) \quad E_{P_1}^* : X_n^* \rightarrow 0 \quad \text{and} \quad E_{P_2} : Y_n \rightarrow 0.$$

For some specific instances of Examples C, D, and E, and further references see [12], [17], and [18].

For an extension of the theory of this paper to "constraints", see [4].

REFERENCES

1. B.L. CHALMERS, Absolute projection constant of the linear functions in a Lebesgue space, *Constr. Approx.* 4(1988), 107–110.
2. B.L. CHALMERS, The (*)-equation and the form of the minimal projection operator, in *Approximation Theory. IV*, C.K. Chui, et al. Eds., Academic Press, New York 1983, pp. 393–399.
3. B.L. CHALMERS, The n -dimensional Hölder inequality, submitted.
4. B.L. CHALMERS Minimal constrained (shape-preserving) projections and extensions, in preparation.
5. B.L. CHALMERS, F.T. METCALF, The determination of minimal projections and extensions in L^1 , *Trans. Amer. Math. Soc.* 329(1992), 289–305.
6. B.L. CHALMERS, F.T. METCALF, Determination of a minimal projection from $C[-1, 1]$ onto the quadratics, *Numer. Funct. Anal. Optim.* 11 (1990), 1–10.
7. B.L. CHALMERS, F.T. METCALF, A simple formula showing L^1 is a maximal over-space for two-dimensional real spaces, *Ann. Polon. Math.* 56(1992), 303–309.
8. B.L. CHALMERS, F.T. METCALF, Minimal projections and extensions for compact abelian groups, in *Approximation Theory. VI*, C.K. Chui, et al. Eds., Academic Press, New York 1989, pp. 129–132.
9. B.L. CHALMERS, K.C. PAN, Finite-dimensional action constants, submitted.
10. B.L. CHALMERS, K.C. PAN, B. SHEKHTMAN, When is the adjoint of a minimal projection also minimal, in *Proceedings of Memphis Conference*, Lect. Notes in Pure and Applied Math. 138(1991), pp. 217–226.
11. B.L. CHALMERS, B. SHEKHTMAN, Minimal projections and absolute projection constants for regular polyhedral spaces, *Proc. Amer. Math. Soc.* 95(1985), 449–452.
12. S.D. FISHER, *Function theory on planar domains*, John Wiley & Sons, New York 1983.
13. C. FRANCHETTI, E.W. CHENEY, Minimal projections in L_1 spaces, *Duke Math. J.* 43(1976), 501–510.
14. C. FRANCHETTI, G. VOTRUBA, Perimeter, Macphail number and projection constant in Minkowski planes, *Boll. Un. Mat. Ital. B(6)* 13(1976), 560–573.
15. H. KÖNIG, D.R. LEWIS, P.-K. LIN, Finite dimensional projection constants, *Studia Math.* 75(1983), 341–358.
16. H. KÖNIG, N. TOMCZAK-JAEGERMANN, Norms of minimal projections, *J. Func. Anal.* 119(1994), 253–280.
17. C.A. MICCHELLI, T.J. RIVLIN, eds., *Optimal estimation in approximation theory*, Plenum, New York 1977.
18. C.A. MICCHELLI, T.J. RIVLIN, Lectures on optimal recovery, in *Numerical Analysis, Lancaster 1984*, P.R. Turner ed., Springer-Verlag, New York 1985, pp. 21–93.
19. P.D. MORRIS, E.W. CHENEY, On the existence and characterization of minimal projections, *J. Reine Angew. Math.* 270(1974), 61–76.
20. W. ODYNYEC, G. LEWICKI, *Minimal projections in Banach spaces*, Springer-Verlag, Berlin 1990.
21. K.C. PAN, B. SHEKHTMAN, On minimal interpolating projections and trace duality, *J. Approx. Theory* 65(1991), 216–230.
22. G. PISIER, *Factorization of linear operators and geometry of Banach spaces*, C.B.M.S. Regional Conf. Ser. in Math. Vol. 60(1986), Amer. Math. Soc. .

23. I. SINGER, *Best approximation in normed linear spaces by elements of linear subspaces*, Springer-Verlag, Berlin 1970.
24. N. TOMCZAK-JAEGERMANN, *Banach-Mazur distances and finite-dimensional operator ideals*, John Wiley & Sons, New York 1989.

B.L. CHALMERS and F.T. METCALF
Department of Mathematics
University of California
Riverside, CA 92521
U.S.A.

Received July 22, 1991; revised October 5, 1993.