STUDIA MATHEMATICA 214 (3) (2013)

Explicit representation of compact linear operators in Banach spaces via polar sets

by

David E. Edmunds (Brighton) and Jan Lang (Columbus, OH)

Abstract. We consider a compact linear map T acting between Banach spaces both of which are uniformly convex and uniformly smooth; it is supposed that T has trivial kernel and range dense in the target space. It is shown that if the Gelfand numbers of T decay sufficiently quickly, then the action of T is given by a series with calculable coefficients. This provides a Banach space version of the well-known Hilbert space result of E. Schmidt.

1. Introduction. Apart from its intrinsic interest and wide applicability, the famous Schmidt decomposition of compact linear maps acting between Hilbert spaces may also be given credit for stimulating work on those compact linear maps from one Banach spaces to another, such as nuclear maps, that may also be represented in series form. Of course, it is not to be expected that an exact analogue of Schmidt's theorem will be generally true outside a Hilbert space setting. Our main result is that such a series representation with coefficients that are recursively calculable is possible for maps the Gelfand numbers of which decrease sufficiently rapidly.

More precisely, let X and Y be real Banach spaces, both of which are uniformly convex and uniformly smooth, and let $T : X \to Y$ be linear and compact, with trivial kernel and range dense in Y. We show that if the Gelfand numbers $c_n(T)$ of T decay sufficiently quickly as $n \to \infty$, then the action of T is given by a series:

$$Tx = \sum_n \alpha_n^Y(x) y_n \quad (x \in X), \tag{1.1}$$

where $\alpha_n^Y \in X^*$ and $y_n \in Y$, for each n. The terms in this series originate in the construction of a decreasing sequence of linear subspaces X_n of X with

2010 *Mathematics Subject Classification*: Primary 46B50, 47A75; Secondary 47A80, 47A58.

Key words and phrases: eigenvalues, Banach spaces, compact operators, nuclear maps, Gelfand numbers.

DOI: 10.4064/sm214-3-5

finite codimension and trivial intersection; the norm λ_n of the restriction of T to X_n is attained at x_n and $y_n = Tx_n/\lambda_n$; the coefficients $\alpha_n^Y(x)$ are recursively calculable by using projections and polar sets.

Banach space versions of the Schmidt result are also given in the recent papers [3]–[5], and in particular it has been shown that, under an additional assumption, there are representations of the form

$$x = \sum_n \xi_n^X(x)x_n, \quad Tx = \sum_n \lambda_n \xi_n^X(x)y_n, \quad y_n = Tx_n/\|Tx_n\| \quad (x \in X).$$

The additional assumption is that for each $x \in X$, the elements $S_n^{X,T}x := \sum_{j=1}^{n-1} \xi_j^X(x)x_j$ should have uniformly bounded norms:

$$\sup_n \|S_n^{X,T}x\|_X < \infty.$$

This condition is automatically satisfied for every compact map T if X is a Hilbert space, no matter what Y is (within the class of spaces studied), and also in certain other special cases, but it remains unclear exactly how wide is the class of spaces and operators for which it holds. Here we cast some light on this question by showing that the Y-analogue of this assumption holds if the Gelfand number $c_n(T)$ of T is bounded above by $2^{-n+1}(2^n-1)^{-1}$ for all $n \in \mathbb{N}$, and that consequently we have the series representation of T given in (1.1). Under these conditions on the Gelfand numbers the map T is nuclear, so that the existence of a series representation is guaranteed.

The main point of our work is to establish the particular decomposition given by (1.1), in which the coefficients are recursively calculable by procedures analogous to those used in the Hilbert space case. Note also that the requirements of uniform convexity and uniform smoothness imposed on X and Y are met by such commonly used spaces as L_p $(1 < p < \infty)$ and the associated Sobolev spaces.

A crucial part of the proof is the study of the projections $P_n^Y : Y \to Y_n := T(X_n)$ $(n \in \mathbb{N})$ that take each $y \in Y$ to the nearest point in Y_n. In general these are nonlinear, but we show that because of the special properties of the Y_n they have certain linearity properties that we are able to exploit to obtain the result.

2. Preliminaries. Throughout the paper we shall suppose that X and Y are real, reflexive, infinite-dimensional Banach spaces with norms $\|\cdot\|_X, \|\cdot\|_Y$ and duals X^*, Y^* that are strictly convex; the closed unit ball in X is denoted by B_X and the family of all bounded linear maps from X to Y by $B(X,Y)$ ($B(X)$ if $X = Y$); T will be a compact linear map from X to Y. We denote the value of $x^* \in X^*$ at $x \in X$ by $\langle x, x^*\rangle_X$, and given any closed linear subspaces M, N of X, X^* respectively, their polar sets are

$$M^0 = \{x^* \in X^* : \langle x, x^* \rangle_X = 0 \text{ for all } x \in M\},$$
$$^0N = \{x \in X : \langle x, x^* \rangle_X = 0 \text{ for all } x^* \in N\}.$$

The space X^* is strictly convex if and only if $\|\cdot\|_X$ is Gâteaux differentiable on $X \setminus \{0\}$; the *Gâteaux derivative* $\widetilde{J}_X(x) := \operatorname{grad}\|x\|_X$ of $\|x\|_X$ at $x \in X \setminus \{0\}$ is the unique element of X^* such that

$$\|\widetilde{J}_X(x)\|_{X^*} = 1 \quad \text{and} \quad \langle x, \widetilde{J}_X(x)\rangle_X = \|x\|_X.$$

A *gauge function* is a map $\mu : [0,\infty) \to [0,\infty)$ that is continuous, strictly increasing and such that $\mu(0) = 0$ and $\lim_{t\to\infty} \mu(t) = \infty$; the map $J_X : X \to X^*$ defined by

$$J_X(x) = \mu(\|x\|_X)\widetilde{J}_X(x) \quad (x \in X \setminus \{0\}), \qquad J_X(0) = 0,$$

is called a *duality map* on X with gauge function μ. For all $x \in X$,

$$\langle x, J_X(x)\rangle_X = \|J_X\|_{X^*}\|x\|_X, \qquad \|J_X\|_{X^*} = \mu(\|x\|_X).$$

From now on we suppose that X and Y are equipped with duality maps corresponding to gauge functions μ_X, μ_Y respectively, normalised so that $\mu_X(1) = \mu_Y(1) = 1$. Let M be a closed linear subspace of X. Then if X is strictly convex, so are M and $X \setminus M$; if X^* is strictly convex, so are $(X\setminus M)^*$ and M^0. A semi-inner product is defined on X by

$$(x, h)_X = \|x\|_X \langle h, \widetilde{J}_X x\rangle_X \quad (x \neq 0), \qquad (0, h)_X = 0.$$

Proofs of these assertions and further details of Banach space geometry may be found in [2], [10], [11] and [12].

Next we summarise some of the results of [3], [4] and [5], beginning with the elementary fact that, under the conditions on X, Y and T, there exists $x_1 \in X$, with $\|x_1\|_X = 1$, such that $\|T\| = \|Tx_1\|_Y$, and that x_1 satisfies the equation

$$T^*\widetilde{J}_Y T x_1 = \nu \widetilde{J}_X x_1, \qquad \nu = \|T\|,$$

or equivalently,

$$T^* J_Y T x_1 = \nu_1 J_X x_1, \qquad \nu_1 = \|T\|\mu_Y(\|T\|).$$

Set $X_1 = X$, $M_1 = \operatorname{sp}\{J_X x_1\}$ (where sp denotes the linear span), $X_2 = {}^0M_1$, $N_1 = \operatorname{sp}\{J_Y T x_1\}$, $Y_2 = {}^0N_1$ and $\lambda_1 = \|T\|$. Since X_2 and Y_2 are closed subspaces of reflexive spaces they are reflexive. Since $X_2^* = ({}^0M_1)^*$ is isometrically isomorphic to X_1^*/M_1 it follows that X_2^* is strictly convex; the same argument applies to Y_2^*. Because

$$\langle Tx, J_Y T x_1\rangle_Y = \nu_1 \langle x, J_X x_1\rangle_X \quad \text{for all } x \in X,$$

we see that T maps X_2 to Y_2. The restriction T_2 of T to X_2 is thus a compact linear map from X_2 to Y_2, and if it is not the zero operator we can repeat the above argument: there exists $x_2 \in X_2 \setminus \{0\}$ such that

$$\langle T_2 x, J_{Y_2} T_2 x_2 \rangle_Y = \nu_2 \langle x, J_{X_2} x_2 \rangle_X \quad \text{for all } x \in X_2,$$

where $\nu_2 = \lambda_2 \mu_Y(\lambda_2)$, $\lambda_2 = \|Tx_2\|_Y = \|T_2\|$. Evidently $\lambda_2 \le \lambda_1$ and $\nu_2 \le \nu_1$. In this way we obtain elements $x_1, \dots, x_n$ of X, all with unit norm, subspaces $M_1, \dots, M_n$ of X^* and $N_1, \dots, N_n$ of Y^*, where

$$M_k = \mathrm{sp}\{J_X x_1, \dots, J_X x_k\}, \quad N_k = \mathrm{sp}\{J_Y T x_1, \dots, J_Y T x_k\}, \quad k = 1, \dots, n,$$

and decreasing families $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ of subspaces of X and Y respectively given by

$$X_k = {}^0 M_{k-1}, \quad Y_k = {}^0 N_{k-1}, \quad k = 2, \dots, n.$$

For each $k \in \{1, \dots, n\}$, T maps X_k into Y_k, $x_k \in X_k$ and, with $T_k := T\restriction_{X_k}$, $\lambda_k(T) = \lambda_k = \|T_k\|, \nu_k = \lambda_k \mu(\lambda_k)$, we have

$$\langle T_k x, J_{Y_k} T_k x_k \rangle_{Y_k} = \nu_k \langle x, J_{X_k} x_k \rangle_{X_k} \quad \text{for all } x \in X_k, \tag{2.1}$$

and so

$$T_k^* J_{Y_k} T_k x_k = \nu_k J_{X_k} x_k.$$

In fact, (2.1) is equivalent to

$$\langle Tx, J_Y T x_k \rangle_Y = \nu_k \langle x, J_X x_k \rangle_X \quad \text{for all } x \in X_k.$$

Since $Tx_k \in Y_k = {}^0 N_{k-1}$, we have

$$\langle Tx_k, J_Y T x_l \rangle_Y = 0 \quad \text{if } l < k.$$

The process stops with λ_n, x_n and X_{n+1} if and only if the restriction of T to X_{n+1} is the zero operator while $T_n \ne 0$. With respect to the semi-inner product on X the x_n have the semi-orthogonality property

$$(x_r, x_s)_X = \delta_{r,s} \quad \text{if } r \le s,$$

and correspondingly the $y_n := Tx_n / \|Tx_n\|_Y = Tx_n/\lambda_n$ satisfy

$$(y_r, y_s)_Y = \delta_{r,s} \quad \text{if } r \le s,$$

where $\delta_{r,s}$ is the Kronecker delta.

The iterative process just described can be expressed in terms of the notion of orthogonality given by James [9], using the notation of [6] and [7]. We say that an element $x \in X$ is *j-orthogonal* (or *orthogonal in the sense of James*) *to* $y \in X$, and write $x \perp^j y$, if

$$\|x\|_X \le \|x + ty\|_X \quad \text{for all } t \in \mathbb{R}.$$

If x is j-orthogonal to every element of a subset W of X, it is said to be *j-orthogonal to* W, written $x \perp^j W$. A subset W_1 of X is j-orthogonal to $W_2 \subset X$ (written $W_1 \perp^j W_2$) if $x \perp^j y$ for all $x \in W_1$ and all $y \in W_2$.

In general, j-orthogonality is not symmetric, that is, $x \perp^j y$ need not imply $y \perp^j x$; indeed, symmetry would imply that X is a Hilbert space. The connection between j-orthogonality and the semi-inner product defined earlier is given by the following result of [9]: if $x, h \in X$, then $x \perp^j h$ if and only if

$$(x, h)_X := \|x\|_X \langle h, \widetilde{J}_X x \rangle_X = 0.$$

It now follows immediately that

$$x_r \perp^j x_s \quad \text{and} \quad y_r \perp^j y_s \quad \text{if } r < s,$$

and

$$x_r \perp^j X_r \quad \text{and} \quad y_r \perp^j Y_r \quad \text{for all } r.$$

A decomposition of X in terms of James orthogonality was given by Alber [1], who introduced the following terminology: given closed subsets M_1, M_2 of X, the space X is said to be the *James orthogonal direct sum* of M_1 and M_2, and we write $X = M_1 \uplus M_2$, if

(1) for each $x \in X$ there is a unique decomposition $x = m_1 + m_2$, where $m_1 \in M_1$, $m_2 \in M_2$;
(2) $M_2 \perp^j M_1$;
(3) $M_1 \cap M_2 = \{0\}$.

Alber established the following

THEOREM 1. *Let X be uniformly convex and uniformly smooth, and let M be a closed linear subspace of X; let J_X be a duality map that is normalised in the sense that it has gauge function μ with $\mu(t) = t$ for all $t \geq 0$. Then*

$$X = M \uplus J_X^{-1} M^0 \quad \textit{and} \quad X^* = M^0 \uplus J_X M.$$

Returning to the properties of T we note that when the rank of T is infinite,

$$\{\lambda_n\} \text{ is an infinite sequence that converges to } 0;$$

moreover,

$$X_\infty := \bigcap_{n \in \mathbb{N}} X_n = \ker(T).$$

For these results see [5].

Next, we define

$$Z_n^X = \mathrm{sp}\{x_1, \ldots, x_n\}, \qquad Z_n^Y = \mathrm{sp}\{y_1, \ldots, y_n\} \quad (n \in \mathbb{N}),$$

and introduce the family of maps

$$S_k^{X,T} : X \to Z_{k-1}^X \quad (k \geq 2)$$

determined by the condition that $x - S_k^{X,T} x \in X_k$ for all $x \in X$. When the meaning is clear we shall simply write S_k^X instead of $S_k^{X,T}$. It turns out (see [3]) that S_k^X is uniquely given by

$$S_k^X x = \sum_{j=1}^{k-1} \xi_j^X(x) x_j,$$

where

$$\xi_j^X(x) = \Big\langle x - \sum_{i=1}^{j-1} \xi_i^X(x) x_i, J_X x_j \Big\rangle_X \text{ for } j \geq 2,\ \xi_1^X(x) = \langle x, J_X x_1 \rangle_X. \tag{2.2}$$

Let us note that $\xi_j^X(\cdot)$ is a linear functional; this guarantees the linearity of S_k^X. Maps $S_k^Y : Y \to Z_{k-1}^Y$ are defined in an analogous fashion:

$$S_k^Y y = \sum_{j=1}^{k-1} \xi_j^Y(y) y_j,$$

where the $\xi_j^Y(y)$ are given by expressions corresponding to (2.2). From the uniqueness it follows that $(S_k^X)^2 = S_k^X$ and hence $((S_k^X)^*)^2 = (S_k^X)^*$, so that S_k^X and $(S_k^X)^*$ are linear projections of X onto Z_{k-1}^X and X^* onto M_{k-1} respectively.

Lemma 2. *The spaces X and X^* have the direct sum decompositions*

$$X = X_k \oplus Z_{k-1}^X, \quad X^* = M_{k-1} \oplus (Z_{k-1}^X)^0 \quad \textit{for each } k \geq 2. \tag{2.3}$$

The operators S_k^X, $(S_k^X)^*$ are respectively linear projections of X onto Z_{k-1}^X and X^* onto M_{k-1}. For all $k \in \mathbb{N}$,

$$X_k = X_{k+1} \uplus \operatorname{sp}\{x_k\}.$$

Corresponding statements hold for Y.

Proof. We give the proofs for X only, and for simplicity omit the superscript X. The decomposition for X folllows from $I = (I - S_k) + S_k$, where I is the identity map of X to itself, since $I - S_k$ maps X into X_k by definition, and S_k has range Z_{k-1}. It is unique in view of the construction of the X_k.

From the definition of the S_k and (2.2) we have

$$S_k x = S_{k-1} x + \langle x - S_{k-1} x, J_X x_{k-1} \rangle_X x_{k-1}$$

and so, on setting $E_j = \langle \cdot, J_X x_j \rangle_X x_j$,

$$S_k = S_{k-1} + E_{k-1}(I - S_{k-1}),$$

which gives

$$I - S_k = (I - E_{k-1}) \cdots (I - E_1), \qquad k \geq 2.$$

For $x \in X$ and $x^* \in X^*$,

$$\langle E_{k-1} x, x^* \rangle_X = \langle x, \langle x_k, x^* \rangle_X J_X x_k \rangle_X.$$

Thus the map $E_{k-1} : X \to X$ has adjoint $E^*_{k-1} : X^* \to X^*$ given by $E^*_{k-1} = \langle x_k, \cdot \rangle_X J_X x_k$; and

$$I^* - S^*_k = (I^* - E^*_1) \cdots (I^* - E^*_{k-1}), \quad k \geq 2.$$

It follows by induction that S^*_k and $I^* - S^*_k$ have ranges $M_{k-1}$and Z^0_{k-1} respectively, and hence $X^* = M_{k-1} \oplus (Z_{k-1})^0$.

From the decomposition of X we see that $S_k x = 0$ if $x \in X_k$. Since $S_{k+1}x = S_k x + \xi_k(x)x_k$, it follows that for all $x \in X_k$, $S_{k+1}x = \xi_k(x)x_k$, and as $(I - S_{k+1})x \in X_{k+1} \subset X_k$, we have $X_k = X_{k+1} \oplus \mathrm{sp}\{x_k\}$, from which the final part of the lemma follows. ∎

Now let P^X_n be the projection of X onto X_n ($n \in \mathbb{N} \cup \{\infty\}$), by which we mean that $P^X_n : X \to X_n$ takes $x \in X$ to the point in X_n nearest to x. Projections $P^Y_n : Y \to Y_n$ are defined analogously. In [3] and [4] it is shown that for all $x \in X$,

$$x = \lim_{n\to\infty} (S^X_n x - P^X_n S^X_n x) + P^X_\infty x, \quad P^X_\infty x = \lim_{n\to\infty} P^X_n x$$

in the sense of weak convergence (strong if X is uniformly convex) and

$$Tx = \lim_{n\to\infty} \Big(\sum_{i=1}^{n-1} \lambda_i \xi^X_i(x) y_i - T P^X_n S^X_n x \Big)$$

in the sense of strong convergence, even if X is not uniformly convex. More can be said if the additional assumption of uniform boundedness of the S^X_n is made: it then turns out (see [4, Corrigendum]) that if T has trivial kernel and $(S^X_n)_{n\in\mathbb{N}}$ is bounded, then $(x_n)_{n\in\mathbb{N}}$ is a basis of X and accordingly there is a Schmidt-type representation of T. If the assumption that $(S^X_n)_{n\in\mathbb{N}}$ is bounded is omitted, this result no longer holds. Readers interested in decomposition of Banach spaces should look at [1, Section 3], [3] or [6] for more details.

3. The main results. Here we retain the notation of the last section and strengthen the assumptions (concerning X, Y and T) by requiring that both X and Y should be uniformly convex and uniformly smooth, and that the compact map T should have trivial kernel and range dense in Y. We begin with a lemma connecting P^X_2 and S^X_2.

LEMMA 3. *For all $x \in X$,*

$$x - P^X_2 x = S^X_2 x.$$

Proof. Let $x \in X$. Since $P^X_2 x$ is the nearest element in X_2 to x, it follows that for all $\beta \in \mathbb{R}$ and all $z \in X_2$,

$$\|x - P^X_2 x\|_X \leq \|x - P^X_2 x + \beta z\|_X,$$

which means that $x - P_2^X x \perp^j X_2$. By Theorem 1, applied with $M = X_2$, we see that $X = X_2 \uplus \mathrm{sp}\{x_1\}$, and in view of the uniqueness of decomposition we have $x - P_2^X x = \alpha x_1$ for some $\alpha \in \mathbb{R}$. The uniqueness of $S_2^X x$ now gives the result. ∎

Repetition of this argument, with X, X_2 replaced by X_2, X_3 and use of Theorem 1 to give $X_2 = X_3 \uplus \mathrm{sp}\{x_2\}$, shows that for all $x \in X$,

$$P_2^X x - S_3^X P_2^X x = P_3^X P_2^X x \in X_3.$$

Hence

$$x - S_2^X x - S_3^X P_2^X x = P_3^X P_2^X x.$$

Since

$$S_2^X x + S_3^X P_2^X x = S_2^X x + S_3^X (x - S_2^X x) = S_3^X x,$$

it follows that

$$x - S_3^X x = P_3^X P_2^X x.$$

An obvious extension of this argument leads to

LEMMA 4. *Let $n \in \mathbb{N} \setminus \{1\}$, $x \in X$ and $y \in Y$. Then*

$$x - S_n^X x = P_n^X P_{n-1}^X \cdots P_2^X x \quad \text{and} \quad y - S_n^Y y = P_n^Y P_{n-1}^Y \cdots P_2^Y y.$$

Note that even though the projections P_k^X and P_k^Y are nonlinear, in general, nevertheless the linearity of S_n^X and S_n^Y means that the products $P_n^X P_{n-1}^X \cdots P_2^X$ and $P_n^Y P_{n-1}^Y \cdots P_2^Y y$ are linear.

LEMMA 5. *Let $n \in \mathbb{N}$ and let K_n be the convex hull of $x_1, \ldots, x_n$, where the x_i are defined in Section 2. Then*

$$(2^n - 1)^{-1} \le \inf\{\|x\| : x \in K_n\} \le 1.$$

The same holds with each x_i replaced by y_i.

Proof. Let $x \in K_n$. Then $x = \sum_{i=1}^n \alpha_i x_i$ for some nonnegative α_i such that $\sum_{i=1}^n \alpha_i = 1$. Since $\sum_{i=2}^n \alpha_i x_i \in X_2$ and $\alpha_1 x_1 \perp^j X_2$,

$$\alpha_1 = \|\alpha_1 x_1\|_X \le \Big\|\alpha_1 x_1 + \sum_{i=2}^n \alpha_i x_i\Big\|_X = \|x\|_X.$$

As $\sum_{i=3}^n \alpha_i x_i \in X_3$ and $\alpha_2 x_2 \perp^j X_3$,

$$\begin{aligned}\alpha_2 - \alpha_1 = \|\alpha_2 x_2\|_X - \|\alpha_1 x_1\|_X &\le \Big\|\alpha_2 x_2 + \sum_{i=3}^n \alpha_i x_i\Big\|_X - \|\alpha_1 x_1\|_X \\ &\le \Big\|\sum_{i=1}^n \alpha_i x_i\Big\|_X = \|x\|_X.\end{aligned}$$

In the same way it follows that for all $l \in \mathbb{N}$,

$$\alpha_{l+1} - \alpha_l \le \Big\| \sum_{i=l}^{n} \alpha_i x_i \Big\|_X.$$

Thus

$$\alpha_3 - \alpha_2 - \alpha_1 \le \Big\| \sum_{i=2}^{n} \alpha_i x_i \Big\|_X - \|\alpha_1 x_1\|_X \le \Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\|_X = \|x\|_X,$$

and

$$\begin{aligned} \alpha_4 - \alpha_3 - \alpha_2 - \alpha_1 &\le \Big\| \sum_{i=3}^{n} \alpha_i x_i \Big\|_X - \|\alpha_2 x_2\|_X - \|\alpha_1 x_1\|_X \\ &\le \Big\| \sum_{i=2}^{n} \alpha_i x_i \Big\|_X - \|\alpha_1 x_1\|_X \le \Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\|_X = \|x\|_X. \end{aligned}$$

More generally we have, for each $l \in \mathbb{N}$ with $l \le n$,

$$\alpha_l - \sum_{i=1}^{l-1} \alpha_i \le \|x\|_X.$$

Thus

$$\inf \max(\alpha_1, \alpha_2 - \alpha_1, \alpha_3 - \alpha_2 - \alpha_1, \ldots, \alpha_n - \alpha_{n-1} - \cdots - \alpha_1) \le \inf_{x \in K_n} \|x\|_X,$$

where the infimum on the left-hand side is taken over all $\alpha_i \ge 0$ such that $\sum_{i=1}^{n} \alpha_i = 1$. When $n = 2$ we have to find $\inf_{0 \le t \le 1} \max(t, 1-2t)$. Since

$$\max(t, 1-2t) = \begin{cases} 1-2t, & 0 \le t \le 1/3, \\ t, & 1/3 \le t \le 1, \end{cases}$$

we see that the infimum is attained when the two entries t and $1-2t$ are equal, that is, when $t = 1/3$, as otherwise one of the two entries would be greater than this value. For a general value of n, we need to find

$$\begin{aligned} \inf_{\sum_{i=1}^{n} \alpha_i = 1,\, \alpha_i \ge 0} &\max(\alpha_1, \alpha_2 - \alpha_1, \alpha_3 - \alpha_2 - \alpha_1, \ldots, \alpha_n - \alpha_{n-1} - \cdots - \alpha_1) \\ &= \inf \max(\alpha_1, \alpha_2 - \alpha_1, \ldots, 1 - 2(\alpha_1 + \cdots + \alpha_{n-1})). \end{aligned}$$

If the infimum is attained at $\alpha_i - \alpha_{i-1} - \cdots - \alpha_1$, we can decrease α_i by a sufficiently small $\varepsilon > 0$ while increasing each α_j with $j \in \{i+1, \ldots, n\}$ by $\varepsilon/(n-i)$: this leads to a decrease of $\alpha_i - \alpha_{i-1} - \cdots - \alpha_1$ by ε and an increase of every term $\alpha_j - \alpha_{j-1} - \cdots - \alpha_1$ with $j > i$. Considerations like this show that as before the infimum is attained when the entries coincide: $\alpha_1 = \alpha_2 - \alpha_1$, so that $\alpha_2 = 2\alpha_1$, $\alpha_3 = 2\alpha_2 = 2^2\alpha_1$ and so on up to $\alpha_{n-1} = 2^{n-2}\alpha_1$. Moreover,

$$\alpha_1 = 1 - 2\alpha_1(1 + 2 + \cdots + 2^{n-2}),$$

which gives

$$\alpha_1 = 1/(2^n - 1)$$

and establishes the claimed lower bound of the lemma. As the upper bound is obvious, the proof is complete. The y-version follows in exactly the same way. ∎

We note that since the lower estimate in this lemma was obtained by quite crude means, there is a distinct possibility of improvement in particular spaces, such as L_p $(1 < p < \infty)$.

Next we show how an estimate involving the maps S_n^Y can be obtained when the λ_k decay sufficiently rapidly.

LEMMA 6. *Suppose that $\lambda_k \leq 2^{-k+1}$ for all $k \in \mathbb{N}$. Then for all $y \in T(B_X)$ and all $n \in \mathbb{N}$,*

$$\|y - S_n^Y y\|_Y \leq 1.$$

Proof. Let $y \in T(B_X)$. By Lemma 4, $y = S_2^Y(y) + P_2^Y(y)$; moreover, $y_0 := -S_2^Y(y) = -T(\xi_1^Y(y)x_1)/\|T\| \in T(B_X)$ since $|\xi_1^Y(y)| = |\langle y, J_Y y_1\rangle_Y| \leq \|y\|_Y \leq \|T\|$. Hence $y_1 := (y + y_0)/2 \in T(B_X)$. As $y_1 = \frac{1}{2}P_2^Y(y)$, we see that $\frac{1}{2}P_2^Y(y) \in Y_2 \cap T(B_X)$. Thus

$$\frac{1}{2}P_2^Y T(B_X) \subset Y_2 \cap T(B_X). \tag{3.1}$$

We claim that

$$\sup\{\|y\|_Y : y \in Y_2 \cap T(B_X)\} = \lambda_2.$$

For if $y \in Y_2 \cap T(B_X)$, then $y = Tx$ for some $x \in B_X$, and in view of the decomposition (2.3), $x = u + v$ for some $u \in X_2$ and $v \in Z_1^X$, $v = \alpha x_1$, say. Thus $y = Tu + \alpha T x_1 = Tu + \alpha\lambda_1 y_1$, so that $\alpha\lambda_1 y_1 = y - Tu \in Y_2$. By the Y-form of (2.3) it follows that $\alpha = 0$ and $x \in X_2$, proving the claim. Thus if $\lambda_2 \leq 2^{-1}$, then

$$\sup\{\|y\|_Y : y \in P_2^Y T(B_X)\} \leq 1.$$

Further use of Lemma 4 shows, by similar techniques, that

$$\frac{1}{2}P_3^Y(T(B_X) \cap Y_2) \subset T(B_X) \cap Y_2 \cap Y_3,$$

which together with (3.1) gives

$$\frac{1}{2}P_3^Y\big(\frac{1}{2}P_2^Y T(B_X)\big) \subset T(B_X) \cap Y_3.$$

Hence the condition $\lambda_3 \leq 2^{-2}$ implies that

$$\sup\{\|y\|_Y : y \in P_3^Y P_2^Y T(B_X)\} \leq 1.$$

More generally, the same procedure shows that for any $n \in \mathbb{N} \setminus \{1\}$, if $\lambda_n \leq 2^{-n+1}$, then

$$\sup\{\|y\|_Y : y \in P_n^Y P_{n-1}^Y \cdots P_2^Y T(B_X)\} \leq 1.$$

Together with Lemma 4 this finishes the proof. ∎

From this it follows that the S_n^Y, regarded as maps from the closure in Y of $T(X)$ to itself, have uniformly bounded norms. However, all this is based on the assumption that the λ_n decay quickly enough, and it would clearly be desirable if the result were to hinge instead on the decay of more familiar objects associated with T. The Gelfand numbers come immediately to mind (see, for example, [14]): recall that the nth *Gelfand number* of T is

$$c_n(T) := \inf\{\|TJ_M^X\| : \operatorname{codim} M < n\} \quad (n \in \mathbb{N}),$$

where J_M^X is the natural embedding from the closed linear subspace M of X into X. It is plain that

$$c_n(T) = \inf_{x_1^*,\dots,x_{n-1}^* \in X^*} \{\sup\{\|Tx\|_Y : \|x\|_X = 1,\ \langle x, x_k^*\rangle = 0 \text{ for } k < n\}\}.$$

Companion to these numbers are the *Gelfand widths* $\widetilde{c}_n(T)$ defined by

$$\widetilde{c}_n(T) = \inf_{L_n} \sup_{\|x\|_X = 1,\, Tx \in L_n} \|Tx\|_Y,$$

where the infimum is taken over all closed linear subspaces L_n of Y with codimension at most $n-1$; equivalently,

$$\widetilde{c}_n(T) = \inf_{y_1^*,\dots,y_{n-1}^* \in Y^*} \{\sup\{\|Tx\|_Y : \|x\|_X = 1,\ \langle Tx, y_k^*\rangle = 0 \text{ for } k < n\}\}.$$

Clearly $c_n(T) \le \lambda_n$ and $c_n(T) \le \widetilde{c}_n(T)$ for every n. However, to be able to use Lemma 6 we need inequalities in the reverse direction, and to establish these we proceed as follows.

Lemma 7. *For every $n \in \mathbb{N}$,*

$$(2^n - 1)^{-1}\lambda_n \le \widetilde{c}_n(T) \le \lambda_n.$$

Proof. Let $n \in \mathbb{N}$. Then $T(B_X)$ contains the convex hull H_n of $\pm\lambda_n y_1, \dots$ $\dots, \pm\lambda_n y_n$. Let $\varepsilon > 0$. We claim that

$$T(B_X) \cap Z_n^Y \not\supseteq \{y \in Y : \|y\|_Y \le \widetilde{c}_n(T) + \varepsilon\} \cap Z_n^Y.$$

For if K_n is a linear subspace of Y with codimension $n-1$, then $\dim(K_n \cap Z_n^Y) \ge 1$, so that if the claim were false, the definition of $\widetilde{c}_n(T)$ would be contradicted and then $\hat{c}_n(T) \le \lambda_n$. By Lemma 5 applied to the y_i,

$$H_n \supset \{y \in Y : \|y\|_Y \le \lambda_n(2^n - 1)^{-1}\} \cap Z_n^Y,$$

so that

$$T(B_X) \cap Z_n^Y \supset \{y \in Y : \|y\|_Y \le \lambda_n(2^n - 1)^{-1}\} \cap Z_n^Y.$$

It follows that $\widetilde{c}_n(T) + \varepsilon > \lambda_n(2^n - 1)^{-1}$. ∎

We remind the reader of our standing assumption in this section that both X and Y are uniformly convex and uniformly smooth, and that the compact map T should have trivial kernel and range dense in Y. For the Gelfand numbers we need the following result established in [8]:

Theorem 8. *For all $n \in \mathbb{N}$,*

$$c_n(T) = \widetilde{c}_n(T).$$

We emphasise that this result is not generally true for arbitrary compact maps between arbitrary Banach spaces: the proof uses in a crucial way the uniform convexity and smoothness of the spaces together with the assumptions that $T(X)$ is dense in Y and the kernel of T is trivial.

From this and Lemma 6 we immediately have

LEMMA 9. *If* $c_n(T) \le 2^{-n+1}(2^n - 1)^{-1}$ *for all* $n \in \mathbb{N}$, *then* $\|y - S_n^Y y\|_Y \le 1$ *for all* $y \in T(B_X)$ *and all* $n \in \mathbb{N}$.

As mentioned in the Introduction, the rapid decay imposed on the Gelfand numbers of T in Lemma 9 implies that T is nuclear. Indeed, since the approximation numbers $a_n(T)$ of T satisfy $a_n(T) \le 2\sqrt{n}\, c_n(T)$ (see Pietsch [14, Section 6.2.3.14]), it follows that $(a_n(T))_{n\in\mathbb{N}} \in l_1$; the nuclearity of T now results from [15].

PROPOSITION 10. *The subspaces* Y_n *defined by* (2.2) *have trivial intersection:*

$$\bigcap_{n=1}^{\infty} Y_n = \{0\}.$$

Proof. First we show that

$$\Big(\bigcap_{n=1}^{\infty} Y_n\Big) \cap T(X) = \{0\}. \tag{3.2}$$

To begin with, we claim that if $x \in X$ and $Tx \in Y_{n+1}$, then $x \in X_{n+1}$. For by (2.3), $x = x_Z + x_X$, where $x_Z \in Z_n^X$ and $x_X \in X_{n+1}$. Since $T(Z_n^X) \subset Z_n^Y$ and $T(X_{n+1}) \subset Y_{n+1}$, we see that if $x \notin X_{n+1}$, then $x_Z \ne 0$ and $y := Tx = T(x_Z) + T(x_X)$, where $T(x_Z) \ne 0$ as $\ker(T) = \{0\}$. Since $T(x_Z) \in Z_n^Y$ and $T(x_X) \in Y_{n+1}$, application of the decomposition of Y corresponding to (2.3) shows that $Tx \notin Y_{n+1}$, and we have a contradiction.

Now suppose that there exists $z \in (\bigcap_{n=1}^{\infty} Y_n) \cap T(X)$, $z \ne 0$. Then $z = Tx$ for some $x \ne 0$. Since the X_n are decreasing and have trivial intersection, there exists $N \in \mathbb{N}$ such that for all $n > N$, $x \notin X_N$. But $Tx = z \in Y_n$ for all $n \in \mathbb{N}$, so that by our claim, $x \in X_n$ for all $n \in \mathbb{N}$. This contradiction establishes (3.2).

To finish, simply note that

$$\begin{aligned} Y^* = \{0\}^0 = \Big\{\Big(\bigcap_{n=1}^{\infty} Y_n\Big) \cap T(X)\Big\}^0 &= \overline{\Big(\bigcup_{n=1}^{\infty} Y_n^0\Big) \cup (T(X))^0} \\ &= \overline{\bigcup_{n=1}^{\infty} Y_n^0} = \Big(\bigcap_{n=1}^{\infty} Y_n\Big)^0, \end{aligned}$$

from which the result follows directly. ∎

We can now give the main result:

THEOREM 11. *If* $c_n(T) \leq 2^{-n+1}(2^n-1)^{-1}$ *for all* $n \in \mathbb{N}$, *then for all* $x \in X$,

$$Tx = \sum_n \alpha_n^Y(x) y_n,$$

where $\alpha_n^Y(\cdot) = \xi_n^Y(T\cdot) \in X^*$.

Proof. Let $y \in Y$. Since $S_n^Y y - y \in Y_n \subset Y_k = {}^0N_{k-1}$ if $n > k$, it follows from the definition of a polar set that for all $z \in \bigcup_{k\in\mathbb{N}} N_k$,

$$\langle S_n^Y y - y, z\rangle_Y \to 0 \quad \text{as } n \to \infty.$$

Since by Lemma 9, (S_n^Y) is bounded, this limit also holds for all $z \in \overline{\bigcup_{k\in\mathbb{N}} N_k} = Y^*$; hence $S_n^Y y \rightharpoonup y$. Thus

$$y = \sum_{j=1}^{\infty} \xi_j^Y(y) y_j$$

in the sense of weak convergence in Y. The uniqueness of this weak representation follows from the semi-orthogonality property $(y_l, y_k)_Y = 0$ if $l < k$, and so $(y_n)_{n\in\mathbb{N}}$ is a weak basis of Y. By Banach's weak basis theorem (see, for example, Theorem 5.3 of [13]), it is a basis of Y and the proof is complete. ∎

References

[1] Ya. I. Alber, *James orthogonality and orthogonal decompositions in Banach spaces*, J. Math. Anal. Appl. 312 (2005), 330–342.

[2] F. E. Browder, *Nonlinear operators and nonlinear equations of evolution in Banach spaces*, Proc. Sympos. Pure Math. 18, Part 2, Amer. Math. Soc., Providence, RI, 1976.

[3] D. E. Edmunds, W. D. Evans and D. J. Harris, *Representations of compact linear operators in Banach spaces and nonlinear eigenvalue problems*, J. London Math. Soc. (2) 78 (2008), 65–84.

[4] D. E. Edmunds, W. D. Evans and D. J. Harris, *A spectral analysis of compact linear operators in Banach spaces*, Bull. London Math. Soc. 42 (2010), 726–734; Corrigendum, ibid. 44 (2012), 1079–1081.

[5] D. E. Edmunds, W. D. Evans and D. J. Harris, *The structure of compact linear operators in Banach spaces*, Rev. Mat. Complut., to appear.

[6] D. E. Edmunds and J. Lang, *The j-eigenfunctions and s-numbers*, Math. Nachr. 283 (2010), 463–477.

[7] D. E. Edmunds and J. Lang, *Eigenvalues, Embeddings and Generalised Trigonometric Functions*, Springer, Berlin, 2011.

[8] D. E. Edmunds and J. Lang, *Gelfand numbers and widths*, J. Approx. Theory 166 (2013), 78–84.

[9] R. C. James, *Orthogonality and linear functionals in normed linear spaces*, Trans. Amer. Math. Soc. 61 (1947), 265–292.

[10] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I, II*, Springer, Berlin, 1977, 1979.
[11] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris, 1969.
[12] R. H. Martin, *Nonlinear Operators and Differential Equations in Banach Spaces*, Wiley, New York, 1976.
[13] T. J. Morrison, *Functional Analysis. An Introduction to Banach Space Theory*, Wiley, New York, 2001.
[14] A. Pietsch, *History of Banach Spaces and Linear Operators*, Birkhäuser, Basel, 2007.
[15] A. Pietsch, *Einige neue Klassen von kompakten linearen Abbildungen*, Rev. Math. Pures Appl. (Bucarest) 8 (1963), 427–447.

David E. Edmunds
Department of Mathematics
University of Sussex
Pevensey I
Brighton, BN1 9QH, United Kingdom
E-mail: davideedmunds@aol.com

Jan Lang
Department of Mathematics
The Ohio State University
100 Math Tower
231 West 18th Avenue
Columbus, OH 43210-1174, U.S.A.
E-mail: lang@math.ohio-state.edu

Received October 11, 2012
Revised version February 17, 2013 (7650)