Operator Theory
Advances and Applications
238

David E. Edmunds
W. Desmond Evans

Representations of Linear Operators Between Banach Spaces

Birkhäuser

Birkhäuser

Operator Theory: Advances and Applications

Volume 238

Founded in 1979 by Israel Gohberg

David E. Edmunds
W. Desmond Evans

Representations of Linear Operators Between Banach Spaces

Birkhäuser

David E. Edmunds
Department of Mathematics
University of Sussex
Brighton
United Kingdom

W. Desmond Evans
School of Mathematics
University of Cardiff
Cardiff
United Kingdom

ISSN 0255-0156 ISSN 2296-4878 (electronic)
ISBN 978-3-0348-0641-1 ISBN 978-3-0348-0642-8 (eBook)
DOI 10.1007/978-3-0348-0642-8
Springer Basel Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013948130

Mathematics Subject Classifications: 47A75, 47B06, 47B40, 35P30

Printed on acid-free paper

Springer Basel is part of Springer Science+Business Media (www.birkhauser-science.com)

Contents

Preface

The main object of this book is to give a self-contained account of recent work ([34], [35], [36], [40], [41]) concerning the representation in series form of compact linear maps $T : X \to Y$, where X and Y are reflexive Banach spaces with strictly convex duals. When X and Y are Hilbert spaces such a result is classical and is due to the work by Hilbert, F. Riesz and E. Schmidt in the early years of the twentieth century; it is an apotheosis of early operator theory. Outside Hilbert spaces the problems encountered by the lack of orthogonality are so severe that it is hardly surprising that some restrictions on X and Y are needed to make any progress. By use of polar sets and projections (in general, nonlinear) as substitutes for orthogonal complements the following are constructed:

(i) a decreasing sequence (X_n) of linear subspaces of X with finite codimension and trivial intersection if T has a trivial kernel;

(ii) a sequence (x_n) of points in the unit sphere of X such that the norm λ_n of the restriction of T to X_n is attained at x_n, and with a semi-orthogonality property in a Banach space sense;

(iii) a recursively calculable sequence of scalars $(\xi_n(x))$ analogous to the sequence of Fourier coefficients of x that appears in the Hilbert space case.

Despite the obvious parallels between the recursive process used in the Hilbert space case and that adopted here, there is in fact a significant difference: the eigenvalues of T or $|T|$ (the positive square root of T^*T) that play such an important part when X and Y are Hilbert spaces no longer appear. The numbers λ_n that replace them have dual characteristics, for they are both

(a) eigenvalues of a **nonlinear** operator,

and (which emphasises their **linear** origins)

(b) norms of restrictions of T to the subspaces X_n of X with finite codimension.

This tends to suggest that when leaving the safe and comfortable framework of Hilbert spaces, the classical spectrum of the linear operator T may lose some of its overwhelming importance insofar as the representation of T is concerned. Of course, addiction to classical eigenvalues is firmly rooted in mathematicians' psyche, but nevertheless we think that such a point of view deserves serious consideration even though it may serve as grist to the mill of those who consider Banach spaces that are not Hilbert spaces to be strange and exotic objects.

The λ_n and x_n correspond to an "eigenvalue" and "eigenvector" respectively of a nonlinear operator equation involving a duality map that becomes the identity map in the Hilbert spaces case. In fact, when X and Y are Hilbert spaces, the λ_n are eigenvalues of $|T|$ and (x_n) is orthonormal. The most attractive form of the results presented here, when T has trivial kernel, is that under an additional assumption,

$$x = \sum_n \xi_n(x) x_n \quad \text{and} \quad Tx = \sum_n \lambda_n \xi_n(x) y_n$$

for all $x \in X$; here $y_n = Tx_n / \|Tx_n\|$. Thus the x_n form a basis of X and the representation of T is directly comparable with that known for the Hilbert space situation. The additional hypothesis is that, with $S_n x := \sum_{j=1}^{n-1} \xi_j(x) x_j \ (x \in X)$,

$$\sup_n \|S_n\| < \infty. \tag{H}$$

This condition is satisfied whenever X is a Hilbert space, no matter what Y is (within the class considered), and also in certain other cases. We do not know the extent of spaces and operators for which (H) holds, but if X is so unpleasant that it does not have the approximation property, there can be no compact linear map $T : X \to Y$ (whatever Y is, within the class considered) with trivial kernel for which (H) is satisfied, for otherwise X would have a basis, in contradiction to the lack of the approximation property.

Representations of $x \in X$ and Tx are also derived without hypothesis (H), but they are generally less elegant and involve nonlinear projections. However, additional progress can be made via the Gelfand numbers $c_n(T)$ of T: we recall that these are defined by

$$c_n(T) = \inf \left\{ \left\| T J_M^X \right\| : \operatorname{codim} M < n \right\},$$

where J_M^X is the natural embedding from the closed linear subspace M of X into X. It turns out that if the $c_n(T)$ decay sufficiently quickly (2^{-2n} will do), then a Y-analogue of (H) holds, from which may be obtained a series representation of Tx, though not of x. We emphasise that, solely on account of this rapid decay and now standard theory, the map T is nuclear and hence may be expressed as a series: the point of the result presented here is that the series representation given involves coefficients recursively calculable by procedures broadly similar to those used in the Hilbert space case.

Chapter 1 provides the basic facts concerning the geometry of Banach spaces, orthogonality, bases, the approximation property and p-trigonometric functions that are needed later; the treatment is virtually self-contained as the proofs of most assertions are given in full. It also presents a brief account of s-numbers and entropy numbers in which, principally for reasons of space, detailed proofs are eschewed in favour of precise references. However, particular attention is paid to the Gelfand numbers $c_n(T)$ and the Gelfand widths $\widetilde{c}_n(T)$ of T, defined by

$$\widetilde{c}_n(T) = \inf_{L_n} \sup \left\{ \|Tx\|_Y : \|x\|_X \leq 1, Tx \in L_n \right\},$$

where the infimum is taken over all closed linear subspaces L_n of Y with codimension at most $n-1$. In general these two sets of quantities can be very different, but we show that if both X and Y are uniformly convex and uniformly smooth, and T has trivial kernel and range dense in Y, then $c_n(T) = \widetilde{c}_n(T)$ for all $n \in \mathbb{N}$.

This machinery is deployed in Chapter 2 to give the results mentioned above concerning the representation in series form of compact operators. When (H) holds, not only do the x_n form a basis of X, but (ξ_n) is a basis of X^* and, for each $x \in X$, the sequence $(\xi_n(x))$ belongs to l_q for some q expressible in terms of the basis constant of (ξ_n). This may be thought of as a weak version of the Hausdorff–Young theorem about Fourier coefficients.

The assumption that the Banach spaces X and Y have strictly convex duals means that their norms are Gâteaux-differentiable (except at the origin). This enables duality maps $J_X : X \to X^*$ and $J_Y : Y \to Y^*$ to be defined and these have a prominent role in the theory. Sobolev spaces $\overset{0}{W}{}^k_p(\Omega)$, and Lebesgue spaces $L_p(\Omega)$ defined on an open subset Ω of $\mathbb{R}^n$, and sequence spaces l_p, have strictly convex duals for $1 < p < \infty$ and explicit duality maps can be determined in each case. For instance, for $\overset{0}{W}{}^k_p(\Omega)$, the duality map is the p-Laplacian when $k = 1$ and the p-biharmonic operator when $k = 2$. Moreover, if Ω is bounded, the embedding $\overset{0}{W}{}^k_p(\Omega) \hookrightarrow L_q(\Omega)$ is compact for a well-known range of values of k, p, q, and taking T to be this embedding, the abstract theory yields a rich vein of results. As a sample take $k = 1$, $1 < p < \infty$, $q = p$: the abstract results apply to establish the existence of an infinite sequence of “eigenfunctions” x_n and “eigenvalues” λ_n of the Dirichlet problem on Ω for the p-Laplacian:

$$-\Delta_p u := -\sum_{j=1}^{n} D_j \left(|D_j u|^{p-2} D_j u \right) = \gamma |u|^{p-2} u, \ \text{ on } \ \Omega, \quad u = 0 \ \text{ on } \ \partial\Omega.$$

This is satisfied in a weak sense relative to a decreasing family of subspaces X_n of $\overset{0}{W}{}^1_p(\Omega)$, and the “eigenfunctions” and “eigenvalues” are respectively the critical points and critical levels of an associated functional. The x_n satisfy a semi-orthogonality condition investigated by James in [52] and are referred to as j-eigenfunctions; the λ_n are called j-eigenvalues. Similar implications of the abstract results follow for other values of k and q. In the one-dimensional case of the p-Laplacian the actual eigenfunctions and eigenvalues are known and these coincide with those obtained by the Lusternik–Schnirelmann procedure. This is also shown to be so more generally for a Sturm–Liouville operator which is similar in form to the p-Laplacian. However, these eigenfunctions do not have the semi-orthogonality property of James and the eigenvalues are not equal to the j-eigenvalues. The eigenvalue problem for a compact Hardy-type operator T on $L_p(0,\infty)$ reduces to a Sturm–Liouville eigenvalue problem, and in this case, a Prüfer-type transformation is on hand to help prove that the eigenvalues coincide with the approximation numbers of T.

The final chapter is concerned with the problems encountered when seeking to extend the earlier results to non-compact maps. Integral representations are obtained for any $x \in X$ and Tx in terms of a right-continuous, non-decreasing family of projection operators $(P_\lambda)_{\lambda>0}$ (a resolution of the identity), and, in particular, if $\ker(T) = \{0\}$ and X is uniformly convex,

$$x = \int_{(0,\|T\|]} dP_\lambda x, \quad Tx = \int_{(0,\|T\|]} dTP_\lambda x.$$

When X is a Hilbert space, we derive an integral representation

$$\langle Tx, J_X Tx\rangle_Y = \int_{(0,\|T\|]} f(\lambda; x) d(P_\lambda x, x)_X$$

for each $x \in X$, where $\langle y, y^*\rangle_Y$ denotes the value of $y^* \in Y^*$ at $y \in Y$ and $f(\cdot; x) \in L_1((0,\|T\|])$. The difficulties faced in trying to find more specific information about $f(\cdot; x)$ and to establish such a result for a general X are underlined by the fact that even in the Hilbert space case, such an integral representation is available only for normal operators, and as Dunford and Schwarz point out in [31], the theory of normal operators is not an infallible guide to the theory for more general operators, even in finite-dimensional spaces. However, in line with our remarks above it seems possible to us that the problems in that situation are compounded by an insistence that the decomposition be expressed in spectral terms. Such a view appears to be in line with that of Davies in [24], who remarks that for non-self-adjoint problems “one must give up any hope that theorems about self-adjoint operators will provide useful signposts: they regularly lead in the wrong direction”; and that the significance of eigenvalues becomes a moot point. In our Banach space environment such comments have added force. Here we report on preliminary work in the hope of stimulating further attacks on this challenging problem.

This book has its roots in various papers written by the authors together with Desmond Harris; the incisive contributions of Jan Lang have also been of great importance. It is a pleasure to acknowledge here the many stimulating discussions we have had with both of them over the years.

Basic Notation

$\mathbb{R}$: real numbers.
$\mathbb{R}^n$: n-dimensional Euclidean space.
$\mathbb{C}$: complex numbers.
$\mathbb{Z}$: integers.
$\mathbb{N}$: natural numbers.
$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
$\mathbb{Q}$: rational numbers.
$\delta_{i,j}$: Kronecker delta.
sgn: signum function: $\operatorname{sgn} 0 = 0$, $\operatorname{sgn} z = z/|z|$ if $z \neq 0$.
χ_E: characteristic function of E.
$\ker(T)$: kernel of a map T.
$\upharpoonright_E$: restriction to E.
$A \subset B$: A contained in B, or possibly equal to B.
Ω: open subset of $\mathbb{R}^n$.
$\partial\Omega$: boundary of Ω.
$\overline{\Omega}$: closure of Ω.
$\|\cdot\|_X$ or $\|\cdot|X\|$: norm on X.
$X,\ Y$: Banach spaces with duals $X^*,\ Y^*$.
B_X: closed unit ball in X.
S_X: unit sphere in X.
$\operatorname{sp} S$: linear span of $S \subset X$.
$\overline{\operatorname{sp} S}$: closed linear span of S.
$B(X,Y)$ ($B(X)$ when $Y = X$): space of bounded linear operators from X to Y.
$\langle x, x^* \rangle_X$: value of $x^* \in X^*$ at $x \in X$.
$\rightharpoonup$: weak convergence.
$\rightharpoonup^*$: weak* convergence.
$X \hookrightarrow Y$: X is continuously embedded in Y.
$L_p(\Omega),\ 1 < p < \infty$: Lebesgue space of functions f with $|f|^p$ integrable on Ω.
$\|\cdot\|_p$ or $\|\cdot\|_{p,I}$: norm on $L_p(I)$.
$l_p,\ 1 < p < \infty$: space of sequences $(x_n)_{n\in\mathbb{N}}$ such that $\sum_{n=1}^{\infty} |x_n|^p < \infty$.
$W_p^r(\Omega)$: Sobolev spaces.
$C_0^\infty(\Omega)$: infinitely differentiable functions with compact supports in Ω.
$\overset{0}{W}{}_p^r(\Omega)$: closure of $C_0^\infty(\Omega)$ in $W_p^r(\Omega)$.

Chapter 1

Preliminaries

In this chapter we collect information about the geometry of Banach spaces, Schauder bases and p-trigonometric functions that will be useful later on. For the convenience of the reader, proofs of most assertions are given, precise references being provided for the remainder.

As a matter of notation throughout the book, the norm on a normed linear space space X will usually be written as $\|\cdot \mid X\|$ or $\|\cdot\|_X$, depending on the size of the expression X; if no ambiguity is likely we shall simply write $\|\cdot\|$, and often $\|\cdot\|_p$ or $\|\cdot\|_{p,I}$ will be used to represent the norm on $L_p(I)$; by $\dim X$ we mean the dimension of X; and B_X (resp. S_X) will stand for the closed unit ball (resp. the unit sphere) in X. Given any set $S \subset X$, the linear span of S, $\operatorname{sp} S$, is the smallest linear subspace of X containing S; the closed linear span of S, $\overline{\operatorname{sp}}\, S$, has the obvious analogous definition. The space of all bounded linear maps from a normed linear space X to another such space Y is denoted by $B(X, Y)$, or $B(X)$ if $X = Y$, and we write $B(X, \mathbb{K}) = X^*$ (the dual of X) when $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$) is the space of scalars associated with X. If X and Y are Banach spaces, $B(X, Y)$ is a Banach space with respect to the norm $\|\cdot\|$ defined by $\|T\| = \sup\{\|Tx\|_Y : \|x\|_X \leq 1\}$. The value of $x^* \in X^*$ at $x \in X$ is denoted by $\langle x, x^* \rangle_X$ or $\langle x, x^* \rangle$. Weak convergence in X is represented by a half arrow $\rightharpoonup$, weak* convergence in X^* by $\overset{*}{\rightharpoonup}$. Unless special mention is made to the contrary, all statements and proofs hold for both the real and the complex scalar fields. A linear operator $T : X \to Y$, acting between Banach spaces X, Y, is defined to be compact if, for any bounded set B in X, $T(B)$ is precompact, i.e., its closure $\overline{T(B)}$ is compact in Y. Equivalently, given a bounded sequence (x_n) in X, the sequence (Tx_n) contains a subsequence which converges in Y. A compact linear operator is necessarily bounded, and the family of compact linear maps from X to Y is closed in $B(X, Y)$.

D.E. Edmunds and W.D. Evans, *Representations of Linear Operators Between Banach Spaces*,
Operator Theory: Advances and Applications 238, DOI 10.1007/978-3-0348-0642-8_1,

1.1 The geometry of Banach spaces

Definition 1.1.1. A Banach space X is said to be **strictly convex** if whenever $x, y \in X$ are such that $x \neq y$ and $\|x\| = \|y\| = 1$, and $\lambda \in (0,1)$, then $\|\lambda x + (1-\lambda)y\| < 1$.

This means that no sphere in X contains a line segment. Use of the conditions for equality in Minkowski's inequality shows that the spaces l_p and L_p are strictly convex if $1 < p < \infty$ and do not have this property if p is 1 or ∞. Evidently every linear subspace of a strictly convex space is strictly convex with respect to the inherited norm. The next result gives a simple condition equivalent to strict convexity.

Proposition 1.1.2. *A Banach space X is strictly convex if and only if it is the case that whenever $x, y \in X$ are such that $\|x + y\| = \|x\| + \|y\|$ then either $y = 0$ or $x = \lambda y$ for some $\lambda \geq 0$.*

Proof. Suppose that X is strictly convex and $x, y \in X \backslash \{0\}$ are such that $x \neq y$ and $\|x + y\| = \|x\| + \|y\|$. Thus $\|x\| \neq \|y\|$, for otherwise $\left\|\frac{x+y}{2}\right\| = \|x\|$, contradicting the strict convexity of X. However, if $\|y\| < \|x\|$, then with $\lambda = \|y\| / \|x\|$ we have

$$1 \geq \left\| \frac{x}{\|x\|} + \lambda \left(\frac{y}{\|y\|} - \frac{x}{\|x\|} \right) \right\| \geq \frac{\|x + y\| - \lambda \|x\|}{\|x\|} = \frac{\|x\| + \|y\| - \lambda \|x\|}{\|x\|} = 1,$$

which means that $x = y/\lambda$. The converse is obvious. $\square$

At a deeper level we have

Proposition 1.1.3. *Let X be a Banach space. Then X is strictly convex if and only if given any $x^* \in X^* \backslash \{0\}$, there exists at most one $x \in X$ such that $\|x\| = 1$ and $\langle x, x^* \rangle = \|x^*\|$; such an x exists if X is reflexive.*

Proof. Let X be strictly convex and $x^* \in X^* \backslash \{0\}$. Suppose there are two distinct such points x, say x_1 and x_2. Then if $0 < \lambda < 1$,

$$\begin{aligned} \|x^*\| &= \lambda \langle x_1, x^* \rangle + (1 - \lambda) \langle x_2, x^* \rangle = \langle \lambda x_1 + (1 - \lambda) x_2, x^* \rangle \\ &\leq \|x^*\| \, \|\lambda x_1 + (1 - \lambda) x_2\| < \|x^*\|, \end{aligned}$$

which is absurd. Conversely, suppose that $\|x + \lambda(y - x)\| = 1$ for some $x, y \in X$ with $\|x\| = \|y\| = 1$ and some $\lambda \in (0, 1)$. By the Hahn–Banach theorem, there exists $x^* \in X^*$ such that $\langle x + \lambda(y - x), x^* \rangle = 1$ and $\|x^*\| = 1$. Then $(1 - \lambda) \langle x, x^* \rangle + \lambda \langle y, x^* \rangle = 1$, and since $|\langle x, x^* \rangle|$, $|\langle y, x^* \rangle| \leq 1$ we must have $\langle x, x^* \rangle = \langle y, x^* \rangle = 1$. By hypothesis, this implies that $x = y$, and so X is strictly convex.

To prove that such an x exists when X is reflexive, let (x_k) be a sequence in X such that $\|x_k\| = 1$ for all $k \in \mathbb{N}$ and $\|x^*\| = \lim_{k \to \infty} \langle x_k, x^* \rangle$. Then there is a weakly convergent subsequence of (x_k), still denoted by (x_k) for simplicity, with weak limit x, say. Hence $\|x\| \leq 1$ and $\langle x, x^* \rangle = \lim_{k \to \infty} \langle x_k, x^* \rangle = \|x^*\|$. $\square$

The existence of a unique element of minimal norm in a convex, closed, non-empty subset of a reflexive, strictly convex space follows quickly.

Proposition 1.1.4. *Let K be a closed, convex, non-empty subset of a strictly convex Banach space X. Then there is at most one element $x \in K$ such that*

$$\|x\| = \inf\{\|y\| : y \in K\}.$$

If in addition X is reflexive, such an x exists.

Proof. Suppose there exist $x, y \in K$ with $\|x\| = \|y\| = \inf\{\|z\| : z \in K\}$, $x \neq y$. Let $0 < \lambda < 1$: then $\lambda x + (1-\lambda)y \in K$, $\|\lambda x + (1-\lambda)y\| < \|x\|$ and we have a contradiction.

Now let X be reflexive and assume that (x_k) is a sequence in K such that $\lim_{k\to\infty} \|x_k\| = l := \inf\{\|y\| : y \in K\}$. By reflexivity, this sequence has a subsequence, still denoted by (x_k) for convenience, such that $x_k \rightharpoonup x$ for some $x \in X$; in fact, $x \in K$ since K is convex and closed, and hence weakly closed. Moreover, $\|x\| \leq \lim_{k\to\infty} \|x_k\| = l$. □

Next we introduce an important class of strictly convex spaces.

Definition 1.1.5. The **modulus of convexity** $\delta_X : (0,2] \to [0,1]$ of a Banach space X with $\dim X \geq 2$ is defined by

$$\delta_X(\varepsilon) = \inf\left\{1 - \frac{1}{2}\|x+y\| : x, y \in B_X, \|x-y\| \geq \varepsilon\right\}.$$

The space X is called **uniformly convex** if $\delta_X(\varepsilon) > 0$ for all $\varepsilon \in (0,2]$.

Remark 1.1.6.

(i) Clearly

$$\delta_X(\varepsilon) = \inf\left\{\delta_Y(\varepsilon) : Y \text{ is a 2-dimensional subspace of } X\right\}.$$

Moreover, δ_X is increasing on $[0,2]$, continuous on $[0,2)$ (not necessarily at 2) and $\delta_X(0) = 0$; in addition, $\delta_X(2) = 1$ if and only if X is strictly convex.

(ii) The modulus of convexity of X is also given by the formula

$$\delta_X(\varepsilon) = \inf\left\{1 - \frac{1}{2}\|x+y\| : x, y \in X, \|x\| = \|y\| = 1, \|x-y\| = \varepsilon\right\}.$$

To establish this, note that if $x, y \in X$ are such that $\|x\| \leq 1, \|y\| \leq 1$ and $\|x-y\| \geq \varepsilon$, then there are points x_1, y_1 on the line segment joining x to y such that $(x_1+y_1)/2 = (x+y)/2$ and $\|x_1 - y_1\| = \varepsilon$. In the computation of $\delta_X(\varepsilon)$ it is therefore enough to consider only those points $x, y \in B_X$ such that $\|x-y\| = \varepsilon$. All that is left is to show that

$$K_B := \sup\left\{\|x+y\| : x, y \in B_X, \|x-y\| = \varepsilon\right\}$$

coincides with

$$K_S := \sup\{\|x+y\| : x, y \in S_X, \|x-y\| = \varepsilon\}.$$

In view of (i) it is sufficient to deal with the case when X is two-dimensional. Then the suprema are attained; suppose that K_B is attained at $x_0, y_0 \in B_X$. We wish to prove that $x_0, y_0 \in S_X$. Assume that $\|y_0\| < 1$, put $A = \{z \in B_X : \|z - x_0\| = \varepsilon\}$ and let $x^* \in X^*$ be such that $\|x^*\| = 1$ and $\langle x_0 + y_0, x^* \rangle_X = \|x_0 + y_0\|$. Then for all $z \in A$,

$$\operatorname{re}\langle x_0 + z, x^*\rangle_X \leq \|x_0 + z\| \leq \|x_0 + y_0\| = \langle x_0 + y_0, x^*\rangle_X .$$

Hence $\operatorname{re} x^*$ attains its supremum on A at y_0, so that by translation,

$$\operatorname{re}\langle y_0 - x_0, x^*\rangle_X = \|y_0 - x_0\| = \varepsilon.$$

(Note that it is easy to see that for all $f \in X^*$,

$$\sup_{x\in B_X} |\langle x, f\rangle_X| = \sup_{x\in B_X} |\operatorname{re}\langle x, f\rangle_X| .)$$

Thus

$$\begin{aligned}
\|x_0\| &\leq \frac{1}{2}(\|x_0 + y_0\| + \|y_0 - x_0\|) \\
&= \frac{1}{2}(\operatorname{re}\langle x_0 + y_0, x^*\rangle_X + \operatorname{re}\langle y_0 - x_0, x^*\rangle_X) \\
&= \operatorname{re}\langle y_0, x^*\rangle_X < 1.
\end{aligned}$$

However, with $\eta := \frac{1}{2}\min\{1 - \|x_0\|, 1 - \|y_0\|\} > 0$ we have

$$x_1 := x_0 + \eta(x_0 + y_0), y_1 := y_0 + \eta(x_0 + y_0) \in B_X, \|x_1 - y_1\| = \varepsilon$$

and

$$\|x_1 + y_1\| = (1 + 2\eta)\|x_0 + y_0\| > \|x_0 + y_0\|,$$

which contradicts the maximality of $\|x_0 + y_0\|$. It follows that $\|y_0\| = 1$; a similar argument gives $\|x_0\| = 1$.

(iii) Each of the following two conditions is equivalent to the hypothesis of uniform convexity of X:

(a) If (x_n), (y_n) are sequences in X such that (x_n) is bounded and

$$\lim_{n\to\infty}\left(2\|x_n\|^2 + 2\|y_n\|^2 - \|x_n + y_n\|^2\right) = 0,$$

then $\lim_{n\to\infty}\|x_n - y_n\| = 0$.

(b) If for all $n \in \mathbb{N}$ there are points $x_n, y_n \in B_X$ such that $\lim_{n\to\infty}\|x_n + y_n\| = 2$, then $\lim_{n\to\infty}\|x_n - y_n\| = 0$.

We prove only the implication (b) $\Longrightarrow$ (a) as the remaining parts are fairly routine. Suppose that $(x_n), (y_n)$ are as in (a). Since

$$\begin{aligned} 2\|x_n\|^2 + 2\|y_n\|^2 - \|x_n + y_n\|^2 &\geq 2\|x_n\|^2 + 2\|y_n\|^2 - (\|x_n\| + \|y_n\|)^2 \\ &= (\|x_n\| - \|y_n\|)^2 \geq 0, \end{aligned}$$

we see that $\lim_{n\to\infty} (\|x_n\| - \|y_n\|) = 0$: the sequence (y_n) must be bounded. Passage to a subsequence, if necessary, enables us to assume that $\lim_{n\to\infty} \|x_n\| = \lim_{n\to\infty} \|y_n\| = a$, say. There is nothing further to prove if $a = 0$; suppose that $a > 0$. Then $\|x_n + y_n\| \to 2a$; and $x/\|x_n\|$, $y/\|y_n\| \in B_X$ for all large enough n, while

$$\left\| \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} \right\| \to 2.$$

If (b) holds, then

$$\left\| \frac{x_n}{\|x_n\|} - \frac{y_n}{\|y_n\|} \right\| \to 0,$$

so that $\|x_n - y_n\| \to 0$ and (a) is satisfied.

(iv) Every closed linear subspace of a uniformly convex space is uniformly convex when given the inherited norm. While every uniformly convex space is strictly convex, the converse is false, in general (see [25]). However, if X is strictly convex and $\dim X < \infty$, then X is uniformly convex, for the fact that $1 - \frac{1}{2}\|x+y\| > 0$ whenever $\|x\| = \|y\| = 1$ and $\|x - y\| = \varepsilon > 0$ implies that $\delta_X(\varepsilon) > 0$ since closed bounded sets in X are compact.

(v) Every Hilbert space H is uniformly convex as, with the use of the parallelogram law, we have

$$\begin{aligned} \delta_H(\varepsilon) &= \inf\left\{ 1 - \left(\frac{\|x\|^2}{2} + \frac{\|y\|^2}{2} - \left\| \frac{x-y}{2} \right\|^2 \right)^{1/2} : \|x\| = \|y\| = 1, \|x-y\| = \varepsilon \right\} \\ &= 1 - (1 - \varepsilon^2/4)^{1/2} > 0 \text{ for all } \varepsilon \in (0, 2]. \end{aligned}$$

Note that since

$$(1 - x^q)^{1/q} \leq 1 - x^q/q \ \ (0 \leq x \leq 1, 1 < q < \infty),$$

we have

$$\delta_H(\varepsilon) \geq \varepsilon^2/8.$$

In the opposite direction, for every Banach space X of dimension at least 2, it is known that

$$\delta_X(\varepsilon) \leq \delta_H(\varepsilon) = 1 - (1 - \varepsilon^2/4)^{1/2} \leq C\varepsilon^2;$$

see [62], Vol. II, p. 63. In this sense, Hilbert spaces are the 'most' uniformly convex spaces.