Existence of best m-term approximation

P. Wojtaszczyk*
Institut for Applied Mathematics
Warsaw University
02-097 Warszawa ul. Banacha 2

September 17, 2002

Abstract

We discuss the structure of the sets where the best m–term approximation with respect to a natural biorthogonal system is attained.

In this paper we will discuss existence of the best m-term approximation in the framework of natural biorthogonal systems in a Banach space X. Let us recall that a countable system of vectors $\Phi = (x_n, x_n^*)_{n\in A} \subset X \times X^*$ is called a *biorthogonal system* if for $n, m \in A$ we have

$$x_n^*(x_m) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m. \end{cases} \tag{1}$$

Such a system is called natural (c.f. [3]) if

$$0 < \inf_{n\in A} \|x_n\| \leq \sup_{n\in A} \|x_n\| < \infty \tag{2}$$

$$0 < \inf_{n\in A} \|x_n^*\| \leq \sup_{n\in A} \|x_n^*\| < \infty \tag{3}$$

$$\overline{\operatorname{span}\{x_n\}_{n\in A}} = X. \tag{4}$$

Let us introduce some notation: For $B \subset A$ we put $X(B) = \overline{\operatorname{span}\{x_n\}_{n\in B}}$. For $m = 1, 2, \ldots$ we put $\Sigma_m = \bigcup_{|B|=m} X(B)$. For $x \in X$ and $m = 1, 2, \ldots$ we put

$$\sigma_m(x) = \inf\{\|x - y\| \,:\, y \in \Sigma_m\}. \tag{5}$$

and

$$\mathcal{P}_m(x) = \{y \in \Sigma_m \,:\, \|x - y\| = \sigma_m(x)\}. \tag{6}$$

A system such that $\mathcal{P}_m(x) \neq \emptyset$ for all $x \in X$ and $m = 1, 2, \ldots$ we will call an *existence system*.

Let us now discuss the question of existence of best m-term approximation i.e. the question when $\mathcal{P}_m(x) \neq \emptyset$ for all $x \in X$ and $m = 1, 2, \ldots$. This question in a more general context motivated by the case of classical algebraic polynomials was investigated by B. Baishanski in [1]. Some isolated results we obtained in our context in [2] and [4, Prop. 7]. We will show that the arguments of Baishanskyi cover all cases considered so far. Let us present with the proof Theorem 1 from [1] in our context.

*This research was partially supported by KBN grant 5P03A 03620 located at the Institute of Mathematics of the Polish Academy of Sciences.

Theorem 1 ([1]). *Let $(x_n, x_n^*)_{n\in A}$ be a natural biorthogonal system in X. Assume that there exists a subspace $Y \subset X^*$ such that*

1. *Y is norming i.e. for all $x \in X$*
$$\sup\{|y(x)| \,:\, y \in Y \text{ and } \|y\| \le 1\} = \|x\| \tag{7}$$
2. *for every $y \in Y$ we have $\lim_{n\to\infty} y(x_n) = 0$*

Then $\mathcal{P}_m(x) \neq \emptyset$ for each $x \in X$ and $m = 1, 2, \ldots$.

Proof. Let us take $z_n \in X(A_n)$ with $|A_n| = m$ such that $\|x - z_n\| \to \sigma_m(x)$. Passing to a subsequence if necessary we can assume that $z_n = z_n^B + z_n^R$ where $B \subset A$ satisfies $|B| \le m$, $z_n^B \in X(B)$ and z_n^B converges in norm to certain $z^B \in X(B)$; moreover $z_n^R \in X(B_n)$ where B_n's are disjoint and all are disjoint from B. Put $z = x - z^B$, we clearly have $\lim_{n\to\infty} \|z - z_n^R\| = \sigma_m(x)$. Now let us fix $\epsilon > 0$ and $y \in Y$ such that $y(z) \ge \|z\| - \epsilon$ and $\|y\| = 1$. From our assumption we infer that $y(z_n^R) \to 0$ so also $y(z - z_n^R) \to y(z) \ge \|z\| - \epsilon$. From this (since ϵ can be arbitrarily small) we infer that
$$\liminf_{n\to\infty} \|z - z_n^R\| \ge \|z\| \ge \sigma_m(x). \tag{8}$$
On the other hand
$$\begin{aligned} \|z - z_n^R\| &= \|x - z^B - z_n^R\| = \|x - z_n^B - z_n^R + z_n^B - z^B\| \\ &\le \|z - z_n^R\| + \|z_n^B - z^B\| \end{aligned}$$
so
$$\limsup_{n\to\infty} \|z - z_n^R\| \le \sigma_m(x). \tag{9}$$
From (8) and (9) we see that
$$\lim_{n\to\infty} \|z - z_n^R\| = \|z\| = \sigma_m(x)$$
which shows that $z^B \in \mathcal{P}_m(x)$.

□

From this Theorem we easily obtain some useful corollaries.

Corollary 1. *Every natural biorthogonal system in a reflexive space is an existence system.*

Proof. Take $Y = X^*$. □

Corollary 2. *Let (x_n, x_n^*) be a natural system. Suppose that there exists a sequence of norm 1 linear operators $T_n : X \to X$ such that*
$$T_n^*(X^*) \quad \subset \quad \overline{\operatorname{span}\{x_n^*\}_{n\in A}} \tag{10}$$
$$T_n(x) \to x \qquad \textit{for all } x \in X. \tag{11}$$
In particular we can assume that (x_n) is a monotone basis. Then $(x_n)_{n\in A}$ is the existence system.

Proof. It suffices to show that $Y = \overline{\operatorname{span}\{x_n\}_{n\in A}}$ is norming. Given $x \in X$ take $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$. Given $\epsilon > 0$ using (11) we find n such that $\|T_n(x) - x\| \leq \epsilon$ so $x^*(T_n x) = T_n^*(x^*)(x) \geq \|x\| - \epsilon$. But now we see that $\|T_n^*(x^*)\| \leq 1$ and from (10) we infer that $T_n^*(x^*) \in Y$ so Y is norming. □

REMARK: Corollary 1 shows in particular that all natural biorthogonal systems in L_p with $1 < p < \infty$ are existence systems. We can also apply Theorem 1 to many concrete systems in L_1 or $C(K)$ type spaces. Taking as $Y = C[0,1]$ we see that Haar, Franklin, Walsh, trigonometric and many other properly normalized orthogonal systems are existence systems in $L_1[0,1]$. Analogously taking $Y = C_0(\mathbb{R})$ we get that good wavelet bases are existence systems in $L_1(\mathbb{R})$. Also taking $Y = L_1[0,1]$ we see that trigonometric system and Franklin system are existence systems in $C[0,1]$ while Haar and Walsh systems are existence systems in their sup closures.

EXAMPLE: The above observations and results may suggest that Theorem 1 gives an answer in all possible cases. This is not so. Let us consider the summing basis in c_0 i.e. the system of vectors $v_n = \sum_{j=1}^n e_j$ where $(e_j)_{j=1}^\infty$ are unite vectors in c_0. As is well known this is a basis in c_0 and biorthogonal functionals are given as $v_n^* = e_n^* - e_{n+1}^*$ for $n = 1, 2, \ldots$. If a functional $\phi \in \ell_1 = c_0^*$ satisfies condition 2. of Theorem 1 then clearly

$$\sum_{j=1}^{\infty} \phi_j = 0 \tag{12}$$

but the subspace $Y \subset \ell_1$ of all $\phi \in \ell_1$ satisfying (12) is not norming. Moreover it has codimension 1 so the only bigger subspace is the whole ℓ_1 which is clearly norming but does not satisfy 2. So we cannot apply Theorem 1 to the summing basis. On the other hand we can show directly that *the summing basis in c_0 is the existence system.* To see it observe that in this case Σ_m consists of null sequences which have at most m jumps. Assume now that for some $x \in c_0$ we have a sequence $z_n \in \Sigma_m$ such that $\|x - z_n\| \to \sigma_m(x)$. Passing to a subsequence we may assume that z_n converges coordinatewise to a bounded sequence z. Clearly $\|x - z\|_\infty = \sigma_m(x)$ and if $z \in c_0$ then it can have at most m jumps and we see that $\sigma_m(x)$ is attained. If however $z \notin c_0$ then it must be eventually constant and have at most $m-1$ jumps. So let us fix M so big that $z_j = \alpha \neq 0$ and $|x_j| < \frac{1}{10}\max(\sigma_m(x), |\alpha|)$ for $j \geq M$. We define $z' = (z'_j)$ by the conditions $z'_j = z_j$ for $j \leq M$ and $z'_j = 0$ for $j > M$. Then z' has at most m jumps so is in Σ_m and one easily checks that $\|x - z'\| = \|x - z\| = \sigma_m(x)$.

Our next observations deal with the structure of the sets $\mathcal{P}_m(x)$.

Proposition 1. *If $(x_n, x_n^*)_{n\in A}$ is a natural biorthogonal system then for each $x \in X$ and $m = 1, 2, \ldots$ the set $\mathcal{P}_m(x)$ is closed. The set $\mathcal{P}_m(x)$ is finite for each $x \in X$ if and only if each subspace $X(B)$ for $|B| = m$ is a Chebyshev subspace of X.*

Proof. To show that $\mathcal{P}_m(x)$ is closed let us take $z_n \in \mathcal{P}_m(x)$ convergent to z. Obviously $\|x - z\| = \sigma_m(x)$, so we have to show that $z \in \Sigma_m$. We can repeat the argument in the first paragraph of the proof of Theorem 1 and assume that $z_n = z_n^B + z_n^R$ where all $z_n^B \in X(B)$ for a fixed B with $|B| \leq m$ and z_n^B converge to $z^B \in X(B)$. From this we infer that z_n^R is also a convergent sequence. Since they belong to $X(B_n)$ for disjoint B_n's and the system is natural we infer that z_n^R converges to 0. This gives that $z = z^B$.

Now assume that all spaces $X(B)$ with $|B| = m$ are Chebyshev subspaces. To see that $\mathcal{P}_m(x)$ is finite we assume to the contrary that there are distinct points $z_n \in \mathcal{P}_m(x)$ and repeat the proof of Theorem 1 (clearly $\|x - z_n\| = \sigma_m(x)$ for all n so we can do it). To reach the contradiction we observe that $z^B \in X(B) \subset X(B \cup B_n) \subset \Sigma_m$ and $z_n \in X(B \cup B_n)$, thus *both* z^B and z_n are best approximations to x in a Chebyshev subspace $X(B \cup B_n)$.

Now if $X(B_0)$ with $|B_0| = m$ is not Chebyshev then we can find $x_0 \in X$ such that

$$V =: \{z \in X(B_0) \ : \|x_0 - z\| = \mathrm{dist}(x_0, X(B_0)\}$$

is infinite. We consider points $x_\lambda = x_0 + \lambda \sum_{j \in B_0} x_j$ for $\lambda \in \mathbb{R}$. Clearly $\mathrm{dist}(x_\lambda, X(B_0)) = \mathrm{dist}(x_0, X(B_0))$ for all $\lambda \in \mathbb{R}$ and the set

$$\{z \in X(B_0) \ : \ \|z - x_\lambda\| = \mathrm{dist}(x_\lambda, X(B_0))\}$$

is infinite. But for any $B' \neq B_0$ a subset of A with $|B'| = m$ we may fix $n \in B_0 \setminus B'$ and get

$$\begin{aligned} \mathrm{dist}(x_\lambda, X(B')) &\geq \|x_n^*\|^{-1} |x_n^*(x_\lambda)| \\ &\geq \|x_n^*\|^{-1} \Big(|\lambda| - \max_{j \in B_0} |x_j^*(x_0)| \Big). \end{aligned}$$

Since our biorthogonal system is natural we infer that for sufficiently big λ we have $\mathrm{dist}(x_\lambda, X(B')) > 2\mathrm{dist}(x_\lambda, X(B_0))$ for all $B' \subset A$ with $|B'| = m$. Thus for such λ we have

$$\mathcal{P}_m(x_\lambda) = \{z \in X(B_0) \ : \ \|x_\lambda - z\| = \mathrm{dist}(x_\lambda, X(B_0))\}$$

and is infinite. $\square$

Now we want to present an example of an unconditional basis which is not an existence system.

Theorem 2. *There exists an equivalent norm $|||.|||$ on ℓ_1 such that the unit vector basis is not an existence system in $(\ell_1, |||.|||)$.*

Proof. The unit vector basis in ℓ_1 will be as usual denoted by $(e_n)_{n=1}^\infty$. Let us fix two parameters $\alpha, \beta > 0$ such that

$$\tfrac{1}{3} < \alpha - \beta < 1 \text{ and } \alpha > 1 \tag{13}$$

and define the closed convex body $B_\alpha^\beta \subset \ell_1$ as a closed convex hull of the set

$$\mathfrak{B} = \big\{ \pm (e_1 + e_2), \pm(e_1 - e_2), \pm e_n \text{ for } n \geq 3, \pm y_n \text{ for } n \geq 3 \big\} \tag{14}$$

where

$$y_n = (\alpha - \tfrac{1}{n})(e_1 + e_2) + \beta e_n. \tag{15}$$

Clearly $B_{\ell_1} \subset B_\alpha^\beta \subset \max(2, 2\alpha + \beta) B_{\ell_1}$. We define $|||.|||$ as a Minkowski functional of B_α^β; it is an equivalent norm. Explicitly

$$|||x||| = \inf\{\sum |\gamma_z| \ : \ x = \sum_{z \in \mathfrak{B}} \gamma_z z\}. \tag{16}$$

First let us observe that

$$|||ae_1 + be_2||| = \max(|a|, |b|). \tag{17}$$

Writing $ae_1 + be_2$ as a linear combination of $e_1 + e_2$ and $e_1 - e_2$ we see that $|||ae_1 + be_2||| \leq \max(|a|, |b|)$. To see the other inequality let us consider the functional $\varphi_1 = e_1^* - e_3^* - e_4^* - \dots$ which is clearly continuous on ℓ_1 Since

$$\begin{aligned} \varphi_1(e_1 + e_2) &= \varphi_1(e_1 - e_2) = 1 \\ \varphi_1(e_n) &= 1 \text{ for } n \geq 3 \\ |\varphi_1(y_n)| &= |(\alpha - \tfrac{1}{n}) - \beta| < 1 \end{aligned}$$

We see that $|||\varphi_1||| = 1$. But $\varphi_1(ae_1+be_2) = a$ so $|||ae_1+be_2||| \geq |a|$. Considering analogously the functional $\varphi_2 = e_2^* - e_2^* - e_4^* - \dots$ we get $|||ae_1 + be_2||| \geq |b|$.

Now we will calculate $\sigma_1(e_1 + e_2)$ and show that it is not attained. From (17) we infer that

$$\inf_\lambda |||(e_1 + e_2) - \lambda e_1||| = \inf_\lambda |||(e_1 + e_2) - \lambda e_2||| = 1. \tag{18}$$

Now we look at $|||(e_1 + e_2) - \lambda e_n|||$ for $n \geq 3$. Writing

$$(e_1 + e_2) - \lambda e_n = \tfrac{1}{\alpha - \frac{1}{n}} y_n + \big(\lambda + \tfrac{\beta}{\alpha - \frac{1}{n}}\big) e_n$$

we get

$$|||(e_1 + e_2) - \lambda e_n||| \leq \tfrac{1}{|\alpha - \frac{1}{n}|} + \big|\lambda + \tfrac{\beta}{\alpha - \frac{1}{n}}\big|.$$

Taking the infimum over $\lambda \in \mathbb{R}$ we get

$$\inf_\lambda |||(e_1 + e_2) - \lambda e_n||| \leq \tfrac{1}{\alpha - \frac{1}{n}}. \tag{19}$$

Actually we want to show

$$\inf_\lambda |||(e_1 + e_2) - \lambda e_n||| = \tfrac{1}{\alpha - \frac{1}{n}}. \tag{20}$$

To see it we write and analyze an arbitrary combination

$$(e_1 + e_2) - \lambda e_n = A(e_1 + e_2) + B(e_1 - e_2) + \sum_{k \geq 3} C_k e_k + \sum_{k \geq 3} D_k y_k. \tag{21}$$

First we note that $B = 0$ since the rest of the combination gives equal coefficients at e_1 and e_2 and the left hand side also has those coefficients equal.

If for $k \neq n$ we have $C_k \neq 0$ then $D_k = -\frac{1}{\beta} C_k$ because those are the only places where e_k appears. Now we replace the combination (21) by the combination where both C_k and D_k are zero. This diminishes the sum of absolute values of coefficients by $|C_k| + |D_k| = (1 + \frac{1}{\beta})|C_k|$. However to preserve the equality we have to add $D_k(\alpha - \frac{1}{k})$ to A. The resulting combination will have smaller sum of absolute values of coefficients because

$$|D_k(\alpha - \tfrac{1}{k})| < \alpha |D_k| = \tfrac{\alpha}{\beta} |C_k| < (1 + \tfrac{1}{\beta}) |C_k|.$$

So we infer that the optimal combination has to have the form

$$\begin{aligned}(e_1+e_2)-\lambda e_n &= A(e_1+e_2)+Ce_n+Dy_n && (22)\\ &= A(e_2+e_2)+D(\alpha-\tfrac{1}{n})(e_1+e_2)+Ce_n+D\beta e_n && (23)\end{aligned}$$

so

$$\begin{aligned}A+D(\alpha-\tfrac{1}{n}) &= 1 && (24)\\ C+D\beta &= -\lambda && (25)\end{aligned}$$

To minimize $|A|+|B|+|C|$ we use (24) and (25) and write

$$|A|+|B|+|C|=\phi(D)=|D|+|1-D(\alpha-\tfrac{1}{n})|+|\lambda+\beta D|.$$

Clearly $\phi(D)$ is a positive, continuous piecewise linear function with nodes at $D=0$, $D=-\frac{\lambda}{\beta}$, $D=\frac{1}{\alpha-\frac{1}{n}}$, so its infimum is attained at nodes. Thus we have to look at

$$\inf_{\lambda}\min(1+|\lambda|),\big(\big|\tfrac{\lambda}{\beta}\big|+\big|1+\tfrac{\lambda}{\beta}(\alpha-\tfrac{1}{n})\big|\big),\big(\tfrac{1}{|\alpha-\frac{1}{n}|}+\big|\lambda+\tfrac{\beta}{\alpha-\frac{1}{n}}\big|\big). \qquad (26)$$

We change the order of inf and min and we look at three infima of positive, continuous piecewise linear functions of λ. We get

$$\begin{aligned}\inf_{\lambda}\big(1+|\lambda|\big) &= 1\\ \inf_{\lambda}\big(\big|\tfrac{\lambda}{\beta}\big|+\big|1+\tfrac{\lambda}{\beta}(\alpha-\tfrac{1}{n})\big|\big) &= \min\big(1,\tfrac{1}{|\alpha-\frac{1}{n}|}\big)\\ \inf_{\lambda}\Big(\tfrac{1}{|\alpha-\frac{1}{n}|}+\big|\lambda+\tfrac{\beta}{\alpha-\frac{1}{n}}\big|\Big) &= \tfrac{1}{|\alpha-\frac{1}{n}|}\end{aligned}$$

so clearly the inf in (26) equals $|\alpha-\frac{1}{n}|^{-1}$ and we have (20). Since $\sigma_1(e_1+e_2)=\inf_n\inf_\lambda |||(e_1+e_2)-\lambda e_n|||$ we see from (18) and (20) that $\sigma_1(e_1+e_2)=\frac{1}{\alpha}$ and is not attained.

□

References

[1] B. Baishanski, *Approximation by polynomials of given length,* Illinois J. Math. 27(1983) pp.449-458

[2] V.N. Temlyakov, *Unpublished lecture notes*, (2001)

[3] P. Wojtaszczyk, *Greedy type bases in Banach spaces*, in CONSTRUCTIVE FUNCTIONS THEORY, Varna 2002: (B. Bojanov, Ed.), DARBA, Sofia, 2002,

[4] P. Wojtaszczyk, *Greedy Algorithm for General Biorthogonal Systems*, J. Approx. Theory, 107(2000) 293-314