CONSTRUCTIVE MATHEMATICAL ANALYSIS

Volume V
Issue I

CONSTRUCTIVE MATHEMATICAL ANALYSIS

ISSN 2651-2939

https://dergipark.org.tr/en/pub/cma

VOLUME V ISSUE I
ISSN 2651-2939

March 2022
https://dergipark.org.tr/en/pub/cma

CONSTRUCTIVE MATHEMATICAL ANALYSIS

CONSTRUCTIVE MATHEMATICAL ANALYSIS

Contents

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 1, pp. 1-6
http://dergipark.gov.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Norm attaining multilinear forms on the spaces c_0 or l_1

SUNG GUEN KIM*

ABSTRACT. $T \in \mathcal{L}(^nE)$ is called a *norming attaining* if there are $x_1, \ldots, x_n \in E$ such that $\|x_1\| = \cdots = \|x_n\| = 1$ and $|T(x_1, \ldots, x_n)| = \|T\|$, where $\mathcal{L}(^nE)$ denotes the space of all continuous n-linear forms on E. We investigate norm attaining multilinear forms on c_0 or l_1.

Keywords: Norming attaining multilinear forms, norming points, norming sets.

2020 Mathematics Subject Classification: 46A22.

1. INTRODUCTION

Let us sketch a brief history of norm attaining multilinear mappings and polynomials on Banach spaces. In 1961, Bishop and Phelps [3] initiated and showed that the set of norm attaining functionals on a Banach space is dense in the dual space. Shortly after, attention was paid to possible extensions of this result to more general settings, specially bounded linear operators between Banach spaces. The problem of denseness of norm attaining functions has moved to other types of mappings like multilinear forms or polynomials. The first result about norm attaining multilinear forms appeared in a joint work of Aron, Finet and Werner [2], where they showed that the Radon-Nikodym property is sufficient for the denseness of norm attaining multilinear forms. Choi and Kim [4] showed that the Radon-Nikodym property is also sufficient for the denseness of norm attaining polynomials. Jimenez-Sevilla and Paya [7] studied the denseness of norm attaining multilinear forms and polynomials on preduals of Lorentz sequence spaces. Acosta and Dávila [1] characterized real Banach spaces Y such that the pair (l_∞^n, Y) has the Bishop-Phelps-Bollobás property for operators. Recently, Dantas et al. [5] introduced and studied a concept of norm-attainment in the space of nuclear operators and in the projective tensor product space of given two Banach spaces.

Let $n \in \mathbb{N}$. We write B_E and S_E for the unit ball and sphere of a Banach space E. We denote by $\mathcal{L}(^nE)$ the Banach space of all continuous n-linear forms on E endowed with the norm $\|T\| = \sup_{(x_1,\cdots,x_n)\in S_E\times\cdots\times S_E} |T(x_1, \cdots, x_n)|$. $\mathcal{L}_s(^nE)$ denotes the closed subspace of all continuous symmetric n-linear forms on E. An element $(x_1, \ldots, x_n) \in E^n$ is called a *norming point* of T if $\|x_1\| = \cdots = \|x_n\| = 1$ and $|T(x_1, \ldots, x_n)| = \|T\|$. For $T \in \mathcal{L}(^nE)$, we define

$$\text{Norm}(T) = \{(x_1, \ldots, x_n) \in E^n : (x_1, \ldots, x_n) \text{ is a norming point of } T\}.$$

Norm(T) is called the *norming set* of T. Notice that $(x_1, \ldots, x_n) \in \text{Norm}(T)$ if and only if $(\epsilon_1 x_1, \ldots, \epsilon_n x_n) \in \text{Norm}(T)$ for some $\epsilon_k = \pm 1$ $(k = 1, \ldots, n)$. Indeed, if $(x_1, \ldots, x_n) \in \text{Norm}(T)$

Received: 12.08.2021; Accepted: 08.02.2022; Published Online: 11.02.2022
*Corresponding author: Sung Guen Kim; sgk317@knu.ac.kr
DOI: 10.33205/cma.981877

then
$$|T(\epsilon_1 x_1, \ldots, \epsilon_n x_n)| = |\epsilon_1 \cdots \epsilon_n T(x_1, \ldots, x_n)| = |T(x_1, \ldots, x_n)| = \|T\|$$
which shows that $(\epsilon_1 x_1, \ldots, \epsilon_n x_n) \in \mathrm{Norm}(T)$. If $(\epsilon_1 x_1, \ldots, \epsilon_n x_n) \in \mathrm{Norm}(T)$ for some $\epsilon_k = \pm 1$ $(k = 1, \ldots, n)$, then
$$(x_1, \ldots, x_n) = (\epsilon_1(\epsilon_1 x_1), \ldots, \epsilon_n(\epsilon_n x_n)) \in \mathrm{Norm}(T).$$
For $m \in \mathbb{N}$, let $l_\infty^m := \mathbb{R}^m$ with the supremum norm. Notice that for every $T \in \mathcal{L}(^n l_\infty^m)$, $\mathrm{Norm}(T) \neq \emptyset$ since $S_{l_\infty^m}$ is compact. Kim [10] classified $\mathrm{Norm}(T)$ for every $T \in \mathcal{L}_s(^2 l_\infty^2)$. If $\mathrm{Norm}(T) \neq \emptyset$, $T \in \mathcal{L}(^n E)$ is called ([2, 4]) a *norm attaining* n-linear form and we denote by
$$\mathrm{NA}(\mathcal{L}(^n E)) = \{T \in \mathcal{L}(^n E) : T \text{ is norm attaining }\}.$$
If S_E is compact, then $\mathrm{NA}(\mathcal{L}(^n E)) = \mathcal{L}(^n E)$. Notice that if $T \in \mathrm{NA}(\mathcal{L}(^n E))$, then $\lambda T \in \mathrm{NA}(\mathcal{L}(^n E))$ for every $\lambda \in \mathbb{R}$. A mapping $P : E \to \mathbb{R}$ is a continuous n-homogeneous polynomial if there exists a continuous n-linear form L on the product $E \times \cdots \times E$ such that $P(x) = L(x, \ldots, x)$ for every $x \in E$. We denote by $\mathcal{P}(^n E)$ the Banach space of all continuous n-homogeneous polynomials from E into $\mathbb{R}$ endowed with the norm $\|P\| = \sup_{\|x\|=1} |P(x)|$.

An element $x \in E$ is called a *norming point* of $P \in \mathcal{P}(^n E)$ if $\|x\| = 1$ and $|P(x)| = \|P\|$. For $P \in \mathcal{P}(^n E)$, we define
$$\mathrm{Norm}(P) = \{x \in E : x \text{ is a norming point of } P\}.$$
$\mathrm{Norm}(P)$ is called the *norming set* of P. Notice that $\mathrm{Norm}(P) = \emptyset$ or a finite set or an infinite set. Kim [9] classify $\mathrm{Norm}(P)$ for every $P \in \mathcal{P}(^2 l_\infty^2)$. If $\mathrm{Norm}(P) \neq \emptyset$, $P \in \mathcal{P}(^n E)$ is called [4] a *norm attaining* n-homogeneous polynomial.

For more details about the theory of multilinear mappings and polynomials on a Banach space, we refer to [6].

It seems to be natural and interesting to study about $\mathrm{NA}(\mathcal{L}(^n E))$. In this paper, we investigate $\mathrm{NA}(\mathcal{L}(^n E))$ for $E = c_0$ or l_1, where
$$c_0 = \{(x_j)_{j \in \mathbb{N}} : x_j \in \mathbb{R}, \lim_{j \to \infty} x_j = 0\},$$
$$l_1 = \{(x_j)_{j \in \mathbb{N}} : x_j \in \mathbb{R}, \sum_{j=1}^{\infty} |x_j| < \infty\}.$$

2. Results

Throughout the paper, we let $n \in \mathbb{N}, n \geq 2$. For a real sequence $(x_j)_{j \in \mathbb{N}}$, we denote by $\mathrm{supp}((x_j)_{j \in \mathbb{N}}) = \{j \in \mathbb{N} : x_j \neq 0\}$. For $T \in \mathcal{L}(^n c_0)$ or $\mathcal{L}(^n l_1)$ with
$$T((x_j^{(1)})_{j \in \mathbb{N}}, \ldots, (x_j^{(n)})_{j \in \mathbb{N}}) = \sum_{(j_1, \ldots, j_n) \in \mathbb{N}^n} a_{j_1 \cdots j_n} \, x_{j_1}^{(1)} \cdots x_{j_n}^{(n)}$$
for some $a_{j_1 \cdots j_n} \in \mathbb{R}$, we denote by $\mathrm{supp}(T) = \{(j_1, \ldots, j_n) \in \mathbb{N}^n : a_{i_1 \cdots i_n} \neq 0\}$. Notice that if $\mathrm{supp}(T)$ is finite, then T is norm attaining. Without loss of generality, we may restrict T such that $\mathrm{supp}(T)$ is infinite.

The following theorem presents a sufficient condition that the norm of $T \in \mathcal{L}(^n c_0)$ is less than of the sum of the absolute values of its coefficients.

Theorem 2.1. *Let $T \in \mathcal{L}(^n c_0)$ be such that*
$$T((x_j^{(1)})_{j \in \mathbb{N}}, \ldots, (x_j^{(n)})_{j \in \mathbb{N}}) = \sum_{(j_1, \ldots, j_n) \in \mathbb{N}^n} a_{j_1 \cdots j_n} \, x_{j_1}^{(1)} \cdots x_{j_n}^{(n)}$$
for some $a_{j_1 \cdots j_n} \in \mathbb{R}$. If $T \in NA(\mathcal{L}(^n c_0))$ and $\mathrm{supp}(T)$ is infinite, then $\|T\| < \sum_{(j_1, \ldots, j_n) \in A} |a_{j_1 \cdots j_n}|$.

Proof. Assume the contrary. Let $((x_j^{(1)})_{j\in\mathbb{N}},\ldots,(x_j^{(n)})_{j\in\mathbb{N}}) \in \mathrm{Norm}(T)$. Let $A = \mathrm{supp}(T)$ and $A_l := \{i_l \in \mathbb{N} : (i_1,\ldots,i_l,\ldots,i_n) \in A\}$ for $l = 1,\ldots,n$. There is $1 \le l \le n$ such that $\mathrm{supp}((x_j^{(l)})_{j\in\mathbb{N}}) \cap A_l$ is infinite. Without loss of generality, we may assume that $\mathrm{supp}((x_j^{(1)})_{j\in\mathbb{N}}) \cap A_1$ is infinite. Choose $i_1^{'} \in A_1$ such that $|x_{i_1^{'}}^{(1)}| < \frac{1}{2}$. Let $(i_1^{'},\ldots,i_n^{'}) \in A$. It follows that

$$\begin{aligned}
\|T\| &= \Big|T((x_j^{(1)})_{j\in\mathbb{N}},\ldots,(x_j^{(n)})_{j\in\mathbb{N}})\Big| \\
&= \Big|\sum_{(j_1,\ldots,j_n)\in A} a_{j_1\cdots j_n}\, x_{j_1}^{(1)}\cdots x_{j_n}^{(n)}\Big| \\
&\le \sum_{(j_1,\ldots,j_n)\in A} |a_{j_1\cdots j_n}|\, |x_{j_1}^{(1)}|\cdots|x_{j_n}^{(n)}| \\
&= \sum_{(j_1,\ldots,j_n)\in A\setminus\{(i_1^{'},\ldots,i_n^{'})\}} |a_{j_1\cdots j_n}|\, |x_{j_1}^{(1)}|\cdots|x_{j_n}^{(n)}| + |a_{i_1^{'}\cdots i_n^{'}}|\ |x_{i_1^{'}}^{(1)}|\cdots|x_{i_n^{'}}^{(n)}| \\
&\le \sum_{(j_1,\ldots,j_n)\in A\setminus\{(i_1^{'},\ldots,i_n^{'})\}} |a_{j_1\cdots j_n}| + \frac{1}{2}|a_{i_1^{'}\cdots i_n^{'}}| \\
&< \sum_{(j_1,\ldots,j_n)\in A} |a_{j_1\cdots j_n}| \le \|T\|
\end{aligned}$$

which is a contradiction. Therefore, $\|T\| < \sum_{(j_1,\ldots,j_n)\in A} |a_{j_1\cdots j_n}|$. □

Remark 2.1. *The converse of Theorem 2.1 is not true in general.*

In fact, let

$$T((x_j)_{j\in\mathbb{N}},(y_j)_{j\in\mathbb{N}}) = \frac{1}{2}(x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1) + \sum_{k=3}^{\infty}\frac{1}{2^{k-1}}x_ky_k \in \mathcal{L}(^2c_0).$$

Obviously, supp$(T) = \{(k,k),(1,2),(2,1) : k \in \mathbb{N}\}$. *Let* $A =$ *supp*(T).

Claim 1. $1 = \|T\| < \sum_{(i,j)\in A}|a_{ij}| = \frac{5}{2}$.

We may consider the bilinear form $x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1$ *as an element of* $\mathcal{L}(^2l_\infty^2)$. *It was shown* [8] *that for* $T((x_1,x_2),(y_1,y_2)) = ax_1y_1 + bx_2y_2 + cx_1y_2 + dx_2y_1 \in \mathcal{L}(^2l_\infty^2)$,

$$\|T\| = \max\{|a+b| + |c+d|, |a-b| + |c-d|\}. \tag{2.1}$$

By (2.1), $\|x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1\| = 1$. *It follows that*

$$\begin{aligned}
\|T\| &\le \frac{1}{2}\Big\|x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1\Big\| + \sum_{k=3}^{\infty}\Big\|\frac{1}{2^{k-1}}x_ky_k\Big\| \\
&= \frac{1}{2} + \frac{1}{2} = 1.
\end{aligned}$$

For $n \in \mathbb{N}$,

$$\|T\| \ge |T(e_1 + \sum_{k=3}^{n+2} e_k,\ e_1 + \sum_{k=3}^{n+2} e_k)| = 1 - \frac{1}{2^{n+1}} \to 1$$

as $n \to \infty$. *Hence,* $\|T\| = 1$. *Obviously,* $\sum_{(i,j)\in A}|a_{ij}| = \frac{5}{2}$.

Claim 2. $T \notin NA(\mathcal{L}(^2c_0))$.

Assume the contrary. Let $((x_j)_{j\in\mathbb{N}},(y_j)_{j\in\mathbb{N}}) \in \mathrm{Norm}(T)$. *Notice that*

$$S := supp((x_j)_{j\in\mathbb{N}}) \cap supp((y_j)_{j\in\mathbb{N}})$$

is infinite because if S is finite, then $\|T\| < 1$ by the above argument. Choose $i_0 \in S\backslash\{1,2\}$ such that $|x_{i_0}| < \frac{1}{2}$. It follows that

$$\begin{aligned}
1 = \|T\| &= \Big|\frac{1}{2}(x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1) + \sum_{k\in S\backslash\{1,2\}} \frac{1}{2^{k-1}} x_k y_k\Big| \\
&\leq \frac{1}{2}\Big|x_1y_1 - x_2y_2 + x_1y_2 + x_2y_1\Big| + \sum_{k\in S\backslash\{1,2\}} \Big|\frac{1}{2^{k-1}} x_k y_k\Big| \\
&\leq \frac{1}{2} + \sum_{k\in S\backslash\{1,2,i_0\}} \frac{1}{2^{k-1}}|x_k|\,|y_k| + \frac{1}{2^{i_0-1}}|x_{i_0}|\,|y_{i_0}| \ (\textit{by } (2.1)) \\
&< \frac{1}{2} + \sum_{k\in S\backslash\{1,2,i_0\}} \frac{1}{2^{k-1}} + \frac{1}{2^{i_0}} < 1
\end{aligned}$$

which is a contradiction. Hence, $T \notin NA(\mathcal{L}(^2c_0))$.

Lemma 2.1. *Let $T \in NA(\mathcal{L}(^2c_0))$ and $(x_1, x_2) \in \mathrm{Norm}(T)$ with $x_k = (x_j^{(k)})_{j\in\mathbb{N}}$ for $k = 1, 2$. Then, there is $N \in \mathbb{N}$ such that*
(1) if $n \geq N$ and $|x_j^{(1)}| < 1$ for some $j \in \mathbb{N}$, then $T(e_j, e_n) = 0$,
(2) if $n \geq N$ and $|x_j^{(2)}| < 1$ for some $j \in \mathbb{N}$, then $T(e_n, e_j) = 0$.

Proof. (1) Since $x_1, x_2 \in S_{c_0}$, there are $N \in \mathbb{N}$ and $0 < \delta < \frac{1}{2}$ such that if $n \geq N$, then $|x_n^{(k)}| < \delta$ for $k = 1, 2$. It follows that for $0 < \lambda < 1 - |x_j^{(1)}|$ and $0 < \beta < 1 - \delta$,

$$\begin{aligned}
\|T\| &\geq \max\{|T(x_1 \pm \lambda e_j,\ x_2 \pm \beta e_n)|\} \\
&= \max\{|T(x_1,\ x_2) \pm \beta T(x_1, e_n) \pm \lambda T(e_j, x_2) \pm \lambda\beta T(e_j, e_n)|\} \\
&= |T(x_1,\ x_2)| + \beta|T(x_1, e_n)| + \lambda|T(e_j, x_2)| + \lambda\beta|T(e_j, e_n)| \\
&= \|T\| + \beta|T(x_1, e_n)| + \lambda|T(e_j, x_2)| + \lambda\beta|T(e_j, e_n)|
\end{aligned}$$

which shows that $|T(x_1, e_n)| = |T(e_j, x_2)| = |T(e_j, e_n)| = 0$.
(2) follows by the similar argument as in the proof of (1). □

The following theorem presents a sufficient condition that $T \in \mathrm{NA}(\mathcal{L}(^2c_0))$ is a finite-type bilinear form.

Theorem 2.2. *Let $T \in NA(\mathcal{L}(^2c_0))$ and $(x_1, x_2) \in \mathrm{Norm}(T)$ with $x_k = (x_j^{(k)})_{j\in\mathbb{N}}$ for $k = 1, 2$. Suppose that $|\{j \in \mathbb{N} : |x_j^{(k)}| = 1\}| = 1$ for $k = 1, 2$. Then $T((x_j)_{j\in\mathbb{N}}, (y_j)_{j\in\mathbb{N}}) = \sum_{1\leq i,j\leq N} a_{ij}x_iy_j$ for some $a_{ij} \in \mathbb{R}$ and $N \in \mathbb{N}$. Hence, $\mathrm{supp}(T)$ is finite.*

Proof. Let $N \in \mathbb{N}$ be the number in the proof of Lemma 2.1. Let $j_1, j_2 \in \mathbb{N}$ be such that $|x_{j_k}^{(k)}| = 1$ and $|x_j^{(k)}| < 1$ for all $j \neq j_k$. By the proof of Lemma 2.1, $T(x_1, e_n) = T(e_j, e_n) = 0$ for every $j \neq j_1$ and $n \geq N$. It follows that

$$\begin{aligned}
0 = T(x_1, e_n) &= T(\sum_{1\leq k\leq N} x_k^{(1)} e_k,\ e_n) \\
&= \sum_{1\leq k\leq N} x_k^{(1)}\ T(e_k, e_n) = x_{j_1}^{(1)}\ T(e_{j_1}, e_n)
\end{aligned}$$

which implies that $T(e_{j_1}, e_n) = 0$. Hence, $T(e_j, e_n) = 0$ for all $j \in \mathbb{N}$ and $n \geq N$. By the proof of Lemma 2.1, $T(e_n, x_2) = T(e_n, e_j) = 0$ for every $j \neq j_2$ and $n \geq N$. It follows that

$$\begin{aligned} 0 = T(e_n, x_2) &= T(e_n, \sum_{1 \leq k \leq N} x_k^{(2)} e_k) \\ &= \sum_{1 \leq k \leq N} x_k^{(2)}\, T(e_n, e_k) = x_{j_2}^{(2)}\, T(e_n, e_{j_2}) \end{aligned}$$

which implies that $T(e_n, e_{j_2}) = 0$. Hence, $T(e_n, e_j) = 0$ for all $j \in \mathbb{N}$ and $n \geq N$. Therefore, $T((x_j)_{j\in\mathbb{N}}, (y_j)_{j\in\mathbb{N}}) = \sum_{1 \leq i,j \leq N} a_{ij} x_i y_j$ for some $a_{ij} \in \mathbb{R}$. □

Motivated by Theorem 2.2, we propose some question.
Question. *Is it true that $NA(\mathcal{L}(^2 c_0)) = \{T \in \mathcal{L}(^2 c_0) : \text{supp}(T)$ is finite$\}$?*
The following theorem characterizes $\text{NA}(\mathcal{L}(^n l_1))$.

Theorem 2.3. *Let $T \in \mathcal{L}(^n l_1)$ be such that*

$$T((x_j^{(1)})_{j\in\mathbb{N}}, \ldots, (x_j^{(n)})_{j\in\mathbb{N}}) = \sum_{(j_1,\ldots,j_n)\in\mathbb{N}^n} a_{j_1\cdots j_n}\, x_{j_1}^{(1)} \cdots x_{j_n}^{(n)}$$

for some $a_{j_1\cdots j_n} \in \mathbb{R}$. Then $T \in NA(\mathcal{L}(^n l_1))$ if and only if there are $j_1', \ldots, j_n' \in \mathbb{N}$ such that $\|T\| = \left|a_{j_1'\cdots j_n'}\right|$.

Proof. Without loss of generality, we may assume that $T \neq 0$.
($\Rightarrow$) Assume the contrary. Let $((x_j^{(1)})_{j\in\mathbb{N}}, \ldots, (x_j^{(n)})_{j\in\mathbb{N}}) \in \text{Norm}(T)$. Let $B = \text{supp}(T)$. We claim that B is infinite. Assume that B is finite. Let $\delta := \max\{|a_{j_1\cdots j_n}| : (j_1, \ldots, j_n) \in B\} < \|T\|$. It follows that

$$\begin{aligned} \|T\| &= \left|T((x_j^{(1)})_{j\in\mathbb{N}}, \ldots, (x_j^{(n)})_{j\in\mathbb{N}})\right| = \Big|\sum_{(j_1,\ldots,j_n)\in B} a_{j_1\cdots j_n}\, x_{j_1}^{(1)} \cdots x_{j_n}^{(n)}\Big| \\ &\leq \sum_{(j_1,\ldots,j_n)\in B} |a_{j_1\cdots j_n}|\, |x_{j_1}^{(1)}| \cdots |x_{j_n}^{(n)}| \leq \delta \sum_{(j_1,\ldots,j_n)\in\mathbb{N}^n} |x_{j_1}^{(1)}| \cdots |x_{j_n}^{(n)}| = \delta < \|T\| \end{aligned}$$

which is a contradiction. Hence, B is infinite. Since $T \neq 0$, there are $(j_1', \ldots, j_n') \in B$ such that $j_k' \in \text{supp}((x_j^{(k)})_{j\in\mathbb{N}})$ for $k = 1, \ldots, n$. Then

$$\begin{aligned} \|T\| &= \left|T((x_j^{(1)})_{j\in\mathbb{N}}, \ldots, (x_j^{(n)})_{j\in\mathbb{N}})\right| \\ &= \Big|\sum_{(j_1,\ldots,j_n)\in B} a_{j_1\cdots j_n}\, x_{j_1}^{(1)} \cdots x_{j_n}^{(n)}\Big| \\ &\leq \sum_{(j_1,\ldots,j_n)\in B} |a_{j_1\cdots j_n}|\, |x_{j_1}^{(1)}| \cdots |x_{j_n}^{(n)}| \\ &= |a_{j_1'\cdots j_n'}|\, |x_{j_1'}^{(1)}| \cdots |x_{j_n'}^{(n)}| + \sum_{(j_1,\ldots,j_n)\in B\setminus\{(j_1',\ldots,j_n')\}} |a_{j_1\cdots j_n}|\, |x_{j_1}^{(1)}| \cdots |x_{j_n}^{(n)}| \end{aligned}$$

$$
\begin{aligned}
&< \|T\|\, |x^{(1)}_{j_1'}| \cdots |x^{(n)}_{j_n'}| + \sum_{(j_1,\ldots,j_n)\in B\setminus\{(j_1',\ldots,j_n')\}} |a_{j_1\cdots j_n}|\, |x^{(1)}_{j_1}| \cdots |x^{(n)}_{j_n}| \\
&\leq \|T\| \sum_{(j_1,\ldots,j_n)\in\mathbb{N}^n} |x^{(1)}_{j_1}| \cdots |x^{(n)}_{j_n}| \\
&= \|T\| (\sum_{j_1\in\mathbb{N}} |x^{(1)}_{j_1}|) \cdots (\sum_{j_n\in\mathbb{N}} |x^{(n)}_{j_n}|) = \|T\|
\end{aligned}
$$

which is a contradiction.

($\Leftarrow$) Since $\|T\| = |T(e_{j_1'}, \ldots, e_{j_n'})|$ for some $(j_1^{'}, \ldots, j_n^{'}) \in \mathbb{N}^n$, $(e_{j_1'}, \ldots, e_{j_n'}) \in \mathrm{Norm}(T)$ and $T \in \mathrm{NA}(\mathcal{L}(^n l_1))$. We complete the proof. $\square$

References

[1] M. D. Acosta, J. L. Dávila: *A basis of $\mathbb{R}^n$ with good isometric properties and some applications to denseness of norm attaining operators*, J. Funct. Anal., **279** (6) (2020), 108602, 26 pp.

[2] R. M. Aron, C. Finet and E. Werner: *Some remarks on norm-attaining n-linear forms*, Function spaces (Edwardsville, IL, 1994), 19–28, Lecture Notes in Pure and Appl. Math., **172**, Dekker, New York, (1995).

[3] E. Bishop, R. Phelps: *A proof that every Banach space is subreflexive*, *Bull. Amer. Math. Soc.*, **67** (1961), 97–98.

[4] Y. S. Choi, S. G. Kim: *Norm or numerical radius attaining multilinear mappings and polynomials*, J. London Math. Soc., **54** (1) (1996), 135–147.

[5] S. Dantas, M. Jung, O. Roldán and A. R. Zoca: *Norm-attaining tensors and nuclear operators*, to appear in Mediterr. J. Math. (2022). DOI: https://doi.org/10.1007/s00009-021-01949-5

[6] S. Dineen: *Complex Analysis on Infinite Dimensional Spaces*, Springer-Verlag, London, (1999).

[7] M. Jimenez Sevilla, R. Paya: *Norm attaining multilinear forms and polynomials on preduals of Lorentz sequence spaces*, Studia Math., **127** (1998), 99–112.

[8] S. G. Kim: *The geometry of $\mathcal{L}(^2 l^2_\infty)$*, Kyungpook Math. J., **58** (2018), 47–54.

[9] S. G. Kim: *The norming set of a polynomial in $\mathcal{P}(^2 l^2_\infty)$*, Honam Math. J., **42** (3) (2020), 569-576.

[10] S. G. Kim: *The norming set of a symmetric bilinear form on the plane with the supremum norm*, Mat. Stud., **55** (2) (2021), 171–180.

Sung Guen Kim
Kyungpook National University
Department of Mathematics
Daegu 702-701, Republic of Korea
ORCID: 0000-0001-8957-3881
E-mail address: sgk317@knu.ac.kr

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 1, pp. 7-13
http://dergipark.gov.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Oscillation of noncanonical second-order advanced differential equations via canonical transform

MARTIN BOHNER*, KUMAR S. VIDHYAA, AND ETHIRAJU THANDAPANI

ABSTRACT. In this paper, we develop a new technique to deduce oscillation of a second-order noncanonical advanced differential equation by using established criteria for second-order canonical advanced differential equations. We illustrate our results by presenting two examples.

Keywords: Advanced differential equation, canonical transform, second-order, oscillation.

2020 Mathematics Subject Classification: 34C10, 34K11.

1. INTRODUCTION

Consider the second-order noncanonical advanced differential equation

$$(\mu_1\eta')'(t) + f_1(t)\eta(\sigma(t)) = 0, \quad t \geq t_0 \tag{1.1}$$

subject to

(P$_1$) $\mu_1, f_1 \in \mathrm{C}([t_0,\infty),(0,\infty))$,
(P$_2$) $\sigma \in \mathrm{C}^1([t_0,\infty),\mathbb{R})$, $\sigma'(t) > 0$, $\sigma(t) \geq t$ for all $t \geq t_0$,
(P$_3$) Equation (1.1) is in noncanonical form, that is,

$$\Omega(t_0) := \int_{t_0}^{\infty} \frac{\mathrm{d}t}{\mu_1(t)} < \infty.$$

If (P$_3$) does not hold, then we say that (1.1) is in canonical form.

In recent years, there are many results dealing with the oscillation of (1.1) and its modifications for the delay case, that is, $\sigma(t) \leq t$, see for example [2, 4, 7, 11, 12, 14], and few results in the case of $\sigma(t) \geq t$, see [1, 3, 5, 6, 8–10, 13–17, 19, 20, 23]. Many authors paid attention to a comparison technique, which is a powerful tool in the theory of oscillation, see, for instance, the papers [11, 19, 21, 24] for more details. Further, many authors used the Riccati transformation method to obtain oscillation criteria for delay equations. For the mixed case, that is, $\sigma(t) \leq t$ and $\sigma(t) \geq t$, the author in [22] discussed the oscillatory and nonoscillatory behavior of systems of differential equations based on the analysis of the corresponding characteristic equations. On the other hand in [10], Jozef Džurina already obtained oscillation criteria for the canonical second-order advanced differential equation

$$(ru')'(t) + p(t)u(\sigma(t)) = 0$$

Received: 09.01.2022; Accepted: 13.02.2022; Published Online: 16.02.2022
*Corresponding author: Martin Bohner; bohner@mst.edu
DOI: 10.33205/cma.1055356

from those of a related ordinary differential equation

$$(ru')'(t) + q(t)u(t) = 0.$$

In this paper, we will rewrite (1.1) in noncanonical form equivalently as an equation in canonical form, then apply the results established by Jozef Džurina in [10] to the obtained equation in canonical form, thus establishing new results for our equation (1.1) in noncanonical form.

Section 2 contains some preliminary results, the main results are presented in Section 3, and two illustrative examples are offered in Section 4.

2. Preliminary Results

Throughout, without loss of generality, considering nonoscillatory solutions of (1.1), we restrict our attention to the positive case, since the negative case is similar.

Lemma 2.1. *We have*

$$(\mu_1\eta')' = \frac{1}{\Omega}\left(\mu_1\Omega^2\left(\frac{\eta}{\Omega}\right)'\right)'. \tag{2.1}$$

Proof. A straightforward calculation shows that

$$\begin{aligned}
\frac{1}{\Omega}\left(\mu_1\Omega^2\left(\frac{\eta}{\Omega}\right)'\right)' &= \frac{1}{\Omega}\left(\mu_1\Omega^2\frac{\eta'\Omega-\eta\Omega'}{\Omega^2}\right)' \\
&= \frac{1}{\Omega}\left(\mu_1\left(\eta'\Omega-\eta\left(-\frac{1}{\mu_1}\right)\right)\right)' \\
&= \frac{1}{\Omega}\left(\mu_1\eta'\Omega+\eta\right)' \\
&= \frac{1}{\Omega}\left(\Omega(\mu_1\eta')'+\mu_1\eta'\Omega'+\eta'\right) \\
&= (\mu_1\eta')' + \frac{1}{\Omega}\left(\mu_1\eta'\left(-\frac{1}{\mu_1}\right)+\eta'\right) \\
&= (\mu_1\eta')',
\end{aligned}$$

completing the proof. □

Lemma 2.2. *Equation* (1.1) *can be written in the equivalent canonical form as*

$$(\mu z')'(t) + f(t)z(\sigma(t)) = 0 \tag{2.2}$$

where

$$\mu = \mu_1\Omega^2, \quad z = \frac{\eta}{\Omega}, \quad \textit{and} \quad f = \Omega(\Omega\circ\sigma)f_1.$$

Proof. The equivalence of (1.1) and (2.2) follows from Lemma 2.1. Moreover, since

$$\int_{t_0}^{\infty}\frac{\mathrm{d}t}{\mu_1(t)\Omega^2(t)} = \lim_{t\to\infty}\frac{1}{\Omega(t)} - \frac{1}{\Omega(t_0)} = \infty,$$

(2.2) is in canonical form. □

Corollary 2.1. *The noncanonical differential equation* (1.1) *has an eventually positive solution if and only if the canonical equation* (2.2) *has an eventually positive solution.*

From Corollary 2.1, it is clear that the investigation of oscillation of (1.1) is reduced to that of (2.2), and therefore, we deal with only one class of an eventually positive solution, namely,

$$z(t) > 0, \quad \mu(t)z'(t) > 0 \quad \text{and} \quad (\mu(t)z'(t))' < 0 \tag{2.3}$$

for $t \geq t_1 \geq t_0$, see [10, Lemma 2.1]. Define

$$w(t) = \int_{t_0}^{t} \frac{\mathrm{d}s}{\mu(s)}.$$

Now, we state a basic oscillation result given in [10, 18], which will be improved in the next section.

Theorem 2.1. *Assume that there exists a constant δ such that*

$$w(t) \int_{t}^{\infty} f(s)\mathrm{d}s \geq \delta > \frac{1}{4} \tag{2.4}$$

eventually. Then (2.2) *is oscillatory.*

3. Oscillation Results

In this section, we obtain results for (1.1) by applying results from [10] to the equivalent equation (2.2). If the condition (2.4) fails to hold ($\delta \leq 1/4$), then we can derive a new oscillation criterion using the constant δ.

Theorem 3.2. *Let η be a positive solution of* (1.1) *and suppose*

$$w(t) \int_{t}^{\infty} f(s)ds \geq \delta > 0 \tag{3.1}$$

eventually. Then

$$\frac{\eta(t)}{\Omega(t)w^{\delta}(t)}$$

is increasing eventually.

Proof. Let $\eta > 0$ be a solution of (1.1). By Lemma 2.2, $z > 0$ is a solution of (2.2) satisfying (2.3). Hence, the assumption [10, (3.1) of Theorem 3.1] is satisfied, and therefore the conclusion of [10, Theorem 3.1] holds, which says that z/w^{δ} is strictly increasing, completing the proof. □

Next, we present a new comparison result.

Theorem 3.3. *Let* (3.1) *hold. If the differential equation*

$$(\mu z')'(t) + \left(\frac{w(\sigma(t))}{w(t)}\right)^{\delta} f(t)z(t) = 0 \tag{3.2}$$

is oscillatory, then so is (1.1).

Proof. Since [10, assumption (E_2) of Theorem 3.3] is satisfied, (2.2) is oscillatory, and then so is (1.1). □

Using any criterion for the oscillation of (3.2), we immediately obtain an oscillation result for (1.1).

Theorem 3.4. *Let* (3.1) *hold. If there exists a constant δ_1 such that*

$$w(t) \int_{t}^{\infty} \left(\frac{w(\sigma(s))}{w(s)}\right)^{\delta} f(s)\mathrm{d}s \geq \delta_1 > \frac{1}{4} \tag{3.3}$$

eventually, then (1.1) *is oscillatory.*

Proof. Use [10, Theorem 3.4] to complete the proof. □

If the condition (3.3) fails to hold ($\delta_1 \le 1/4$), then we can derive a new oscillation criterion using the constant δ_1.

Theorem 3.5. *Let* (3.1) *hold. Assume that* η *is a positive solution of* (1.1) *and*

$$w(t)\int_t^\infty \left(\frac{w(\sigma(s))}{w(s)}\right)^\delta f(s)\mathrm{d}s \ge \delta_1 > 0$$

eventually. Then

$$\frac{\eta(t)}{\Omega(t)w^{\delta_1}(t)}$$

is increasing eventually.

Proof. Use [10, Theorem 3.8] to complete the proof. □

Theorem 3.6. *Let* (3.1) *and* (3.3) *hold. If the differential equation*

$$(\mu z')'(t) + \left(\frac{w(\sigma(t))}{w(t)}\right)^{\delta_1} f(t)z(t) = 0 \tag{3.4}$$

is oscillatory, then so is (1.1).

Theorem 3.7. *Let* (3.1) *and* (3.3) *hold. If there exists a constant* δ_2 *such that*

$$w(t)\int_t^\infty \left(\frac{w(\sigma(s))}{w(s)}\right)^{\delta_1} f(s)\mathrm{d}s \ge \delta_2 > \frac{1}{4} \tag{3.5}$$

eventually, then (1.1) *is oscillatory.*

The proofs of Theorems 3.6 and 3.7 follow from [10, Theorems 3.9 and 3.10].

For convenience, let us use the additional condition that there is a positive constant β such that

$$\frac{w(\sigma(t))}{w(t)} \ge \beta > 1 \tag{3.6}$$

eventually. Thus, in view of (3.1), conditions (3.3) and (3.5) can be written in simpler forms as

$$\begin{aligned}\delta_1 =& \beta^\delta \delta > \frac{1}{4},\\ \delta_2 =& \beta^{\delta_1} \delta > \frac{1}{4},\end{aligned}$$

respectively. Repeating the above process, we have the increasing sequence $\{\delta_n\}$ defined by

$$\begin{aligned}\delta_0 =& \delta,\\ \delta_{n+1} =& \beta^{\delta_n}\delta.\end{aligned}$$

Now as in [10, Theorem 3.12], one can generalize the oscillation criteria obtained in Theorems 3.4 and 3.7.

Theorem 3.8. *Let* (3.1) *and* (3.6) *hold. If there exists* $n \in \mathbb{N}$ *such that* $\delta_j \le 1/4$ *for* $j = 0, 1, 2, \ldots, n-1$ *and*

$$\delta_n > \frac{1}{4},$$

then (1.1) *is oscillatory.*

4. Examples

We support the obtained results with some examples.

Example 4.1. *Consider the second-order advanced differential equation*

$$(t^2\eta'(t))' + a\lambda\eta(\lambda t) = 0, \quad t \geq 1, \tag{4.1}$$

where $a > 0$*. Here* $\mu_1(t) = t^2$*,* $f_1(t) = a\lambda$*,* $\sigma(t) = \lambda t$*,* $\lambda > 1$*. A simple calculation shows that*

$$\Omega(t) = \frac{1}{t}, \quad \mu(t) = 1, \quad w(t) = t, \quad f(t) = \frac{a}{t^2}.$$

The transformed canonical equation is

$$z''(t) + \frac{a}{t^2} z(\lambda t) = 0.$$

Condition (3.1) *clearly holds, and* (3.3) *becomes*

$$a\lambda^{\delta} > \frac{1}{4}.$$

Now $\delta = a$*, and by Theorem 3.4,* (4.1) *is oscillatory provided*

$$a\lambda^{a} > \frac{1}{4}.$$

For example, if $a = \frac{1}{5}$*, then we see that* $\lambda \geq 3.052$*, and for* $\lambda = 1.8$*, we need* $a \geq 0.22$*.*

Example 4.2. *Consider the second-order advanced differential equation*

$$(t^2\eta'(t))' + 0.35742\eta(1.61t) = 0. \tag{4.2}$$

The transformed canonical equation is

$$z''(t) + \frac{0.222}{t^2} z(1.61t) = 0.$$

For (4.2)*,* $\delta_0 = 0.222$ *and* $\lambda = 1.61$*. A simple calculation shows that*

$$\delta_1 = 0.2468 \quad \textit{and} \quad \delta_2 = 0.24968.$$

Therefore, Theorems 3.4 and 3.7 fail for (4.2)*. But*

$$\delta_3 = 0.25003 > \frac{1}{4}$$

and Theorem 3.8 implies the oscillation of (4.2)*. However, it is easy to see that [5, Theorems 3, 5, 6], [8, Theorems 3.3, 3.4, and Corollary 4.4] and [4, Theorem 2] do not get oscillation of* (4.2)*. Thus, our result improve these results.*

5. Conclusion

In this paper, we derive oscillation criteria for the noncanonical equation (1.1) by transforming it to the canonical equation (2.2), and then we use the comparison technique available for the canonical equation (2.2) to get new oscillation criteria for the studied equation (1.1). Our oscillation criteria improve [5, Theorems 3, 5, 6], [8, Theorems 3.3, 3.4 and Corollary 4.4] and [4, Theorem 2] for the special case $\alpha = \beta = 1$. Finally, the results obtained in [10] cannot be applied to (4.1) and (4.2) since they are of noncanonical type.

Acknowledgements

The authors would like to thank both anonymous referees as well as the handling Editor Professor Tuncer Acar, for pointing out several shortcomings in the submitted version, which have been fixed in the final version of this paper.

References

[1] R. P. Agarwal, M. Bohner and W. T. Li: *Nonoscillation and oscillation: theory for functional differential equations*, volume 267 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, Inc., New York, 2004.

[2] R. P. Agarwal, S. R. Grace and D. O'Regan: *Oscillation theory for second order linear, half-linear, superlinear and sublinear dynamic equations*, Kluwer Academic Publishers, Dordrecht, 2002.

[3] R. P. Agarwal, C. Zhang and T. Li: *New Kamenev-type oscillation criteria for second-order nonlinear advanced dynamic equations*, Appl. Math. Comput., **225** (2013), 822–828.

[4] B. Baculíková: *Oscillation of second-order nonlinear noncanonical differential equations with deviating argument*, Appl. Math. Lett., **91** (2019), 68–75.

[5] G. E. Chatzarakis, J. Džurina and I. Jadlovská: *New oscillation criteria for second-order half-linear advanced differential equations*, Appl. Math. Comput., **347** (2019), 404–416.

[6] G. E. Chatzarakis, S. R. Grace and I. Jadlovská: *A sharp oscillation criterion for second-order half-linear advanced differential equations*, Acta Math. Hungar., **163** (2) (2021), 552–562.

[7] G. E. Chatzarakis, I. Jadlovská: *Improved oscillation results for second-order half-linear delay differential equations*, Hacet. J. Math. Stat., **48** (1) (2019) 170–179.

[8] G. Chatzarakis, O. Moaaz, T. Li and B. Qaraad: *Some oscillation theorems for nonlinear second-order differential equations with an advanced argument*, Adv. Difference Equ., Paper No. 160 (2020), 17 pages.

[9] J. Džurina: *Oscillation of second order differential equations with advanced argument*, Math. Slovaca, **45** (3) (1995), 263–268.

[10] J. Džurina: *Oscillation of second order advanced differential equations*, Electron. J. Qual. Theory Differ. Equ., Paper No. 20 (2018), 9 pages.

[11] J. Džurina, S. R. Grace, I. Jadlovská and T. Li: *Oscillation criteria for second-order Emden-Fowler delay differential equations with a sublinear neutral term*, Math. Nachr., **293** (5) (2020), 910–922.

[12] J. Džurina, I. Jadlovská: *A sharp oscillation result for second-order half-linear noncanonical delay differential equations* Electron. J. Qual. Theory Differ. Equ., Paper No. 46 (2020), 14 pages.

[13] J. Džurina, I. P. Stavroulakis: *Oscillation criteria for second-order delay differential equations*, Appl. Math. Comput., **140** (2-3) (2003), 445–453.

[14] J. R. Graef: *Oscillation of higher order functional differential equations with an advanced argument*, Appl. Math. Lett., **111** (2021), Paper No. 106685, 6.

[15] I. Győri, G. Ladas: *Oscillation theory of delay differential equations*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1991. With applications, Oxford Science Publications.

[16] I. Jadlovská: *Oscillation criteria of Kneser-type for second-order half-linear advanced differential equations*. Appl. Math. Lett., **106** (2020), 106354, 8.

[17] N. Kılıç, Ö. Öcalan, and U. M. Özkan: *Oscillation tests for nonlinear differential equations with several nonmonotone advanced arguments*, Appl. Math. E-Notes, **21** (2021), 253–262.

[18] T. Kusano, M. Naito: *Comparison theorems for functional-differential equations with deviating arguments*. J. Math. Soc. Japan, **33** (3) (1981), 509–532.

[19] T. Li, Y. V. Rogovchenko: *Oscillation of second-order neutral differential equations*, Math. Nachr., **288** (10) (2015), 1150–1162.

[20] T. Li, Y. V. Rogovchenko: *Oscillation criteria for second-order superlinear Emden-Fowler neutral differential equations*, Monatsh. Math., **184** (3) (2017), 489–500.

[21] T. Li, Y. V. Rogovchenko: *On the asymptotic behavior of solutions to a class of third-order nonlinear neutral differential equations*, Appl. Math. Lett., **105** (2020), 106293, 7.

[22] A. M. Pedro: *Oscillatory behavior of linear mixed-type systems*, Rend. Circ. Mat. Palermo (2), 2021. doi: 10.1007/s12215-021-00658-y

[23] S. Tang, T. Li, R. P. Agarwal and Martin Bohner: *Oscillation of odd-order half-linear advanced differential equations*, Commun. Appl. Anal., **16** (3) (2012), 349–357.

[24] C. Vetro, D. Wardowski: *Asymptotics for third-order nonlinear differential equations: Nonoscillatory and oscillatory cases*, Asymptot. Anal., (2021), 1–19. doi:10.3233/ASY-211710

MARTIN BOHNER
MISSOURI S&T
DEPARTMENT OF MATHEMATICS AND STATISTICS
ROLLA, MO 65409, USA
ORCID: 0000-0001-8310-0266
E-mail address: bohner@mst.edu

KUMAR S. VIDHYAA
SRM EASWARI ENGINEERING COLLEGE
DEPARTMENT OF MATHEMATICS
CHENNAI – 600 089, INDIA
ORCID: 0000-0003-2965-4553
E-mail address: vidyacertain@gmail.com

ETHIRAJU THANDAPANI
UNIVERSITY OF MADRAS
RAMANUJAN INSTITUTE FOR ADVANCED STUDY IN MATHEMATICS
CHENNAI – 600 005, INDIA
ORCID: 0000-0001-6801-4191
E-mail address: ethandapani@yahoo.co.in

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 1, pp. 14-36
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Padua points and "fake" nodes for polynomial approximation: old, new and open problems

STEFANO DE MARCHI*

ABSTRACT. Padua points, discovered in 2005 at the University of Padua, are the first set of points on the square $[-1,1]^2$ that are explicitly known, unisolvent for total degree polynomial interpolation and with Lebesgue constant increasing like $\log^2(n)$ of the degree. One of the key features of the *Padua points* is that they lie on a particular *Lissajous curve*. Other important properties of Padua points are

(1) in two dimensions, Padua points are a WAM for interpolation and for extracting approximate Fekete points and discrete Leja sequences.
(2) in three dimensions, Padua points can be used for constructing tensor product WAMs on different compacts.

Unfortunately, their extension to higher dimensions is still the biggest open problem.

The concept of mapped bases has been widely studied (cf. e.g. [35] and references therein), which turns out to be equivalent to map the interpolating nodes and then construct the approximant in the classical form without the need of resampling. The mapping technique is general, in the sense that works with any basis and can be applied to continuous, piecewise or discontinuous functions or even images. All the proposed methods show convergence to the interpolant provided that the function is resampled at the mapped nodes. In applications, this is often physically unfeasible. An effective method for interpolating via mapped bases in the multivariate setting, referred as *Fake Nodes Approach* (FNA), has been presented in [37]. In this paper, some interesting connection of the FNA with Padua points and "families of relatives nodes", that can be used as "fake nodes" for multivariate approximation, are presented and we conclude with some open problems.

Keywords: Padua points, Lissajous curves and points, mapped polynomial basis.

2020 Mathematics Subject Classification: 41A17, 41A63.

1. INTRODUCTION

Let $\mathbb{P}_n(\mathbb{R})$ be the space of the univariate polynomials of total degree $\leq n$ on $\mathbb{R}$ and $C(\mathbb{R})$ the linear space of continuous functions on $\mathbb{R}$. Further, for the basis of monomials $\mathcal{M} = \{1, x, x^2, \ldots, x^n\}$ and a set $X = \{x_0, \ldots, x_n\}$ of $n+1$ distinct points, we denote by

$$Vdm(X; \mathcal{M}) = \prod_{i<j}(x_i - x_j) \tag{1.1}$$

the corresponding Vandermonde determinant which plays an important role for the unisolvency of a given set of points.

The classical univariate interpolation problem of f by polynomials of degree n can be stated as follows.

Problem 1. *Let K be a closed and bounded interval of $\mathbb{R}$. Consider X a set of $n+1$ pairwise distinct points of K, the values $\{f(x_i), i = 0, \ldots, n\}$ and the basis of monomials $\mathcal{M} = \{1, x, \ldots, x^n\}$. Find the*

Received: 08.02.2022; Accepted: 01.03.2022; Published Online: 03.03.2022
*Corresponding author: Stefano De Marchi; stefano.demarchi@unipd.it
DOI: 10.33205/cma.1070020

polynomial $p_n = \sum_{k=0}^{n} a_k x^k$, *so that*

$$p_n(x_i) = f(x_i), \quad i = 0, \dots, n.$$

Being $x_i \neq x_j$, $i \neq j$, p_n is unique because $Vdm(X; \mathcal{M}) \neq 0$. Using the *Lagrange basis* $L = \{l_i, i = 0, ..., n\}$ with

$$l_i(x) = \prod_{i=0, i\neq j}^{n} \frac{x - x_j}{x_i - x_j} = \frac{Vdm(X_i; \mathcal{M})}{Vdm(X; \mathcal{M})},$$

where X_i is the set X in which we substitute x_i with x, we can then write

$$p_n(x) = \sum_{i=0}^{n} l_i(x) f(x_i), \; x \in K. \tag{1.2}$$

This process generates an interpolation error $e_n(x) = |f(x) - p_n(x)|, \; x \in K$ or in norm $E_n = \|f - p_n\|_\infty$. Using the Lagrange form (1.2) of the interpolant, we can bound this error by

$$E_n \leq (1 + \Lambda_n) E_n^* \tag{1.3}$$

with $\Lambda_n = \sup_{x\in K} \sum_{i=0}^{n} |l_i(x)|$ the *Lebesgue constant* which depends on n and on the node set X. As well-known, Λ_n represents the sup-norm of the linear operator (cf. e.g. [26]) $L : C(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$, $Lf = \sum_{i=0}^{n} f(x_i) l_i$, where E_n^* is the error of best-uniform approximation that is $E_n^* := \inf_{p_n \in \mathbb{P}_n(\mathbb{R})} E_n(f)$.

In the one dimensional case we know

- $\Lambda_n \approx 2^n$ when the set X is made of equally spaced points of K (or even worse when X are randomly chosen);
- $\Lambda_n \approx \log(n)$ when X is made of *Chebyshev-like* points of K.

We call *Chebyshev-like* points, those points that have the so-called *arccos-distribution* which characterizes for instance the Chebyshev-Gauss-Lobatto points (or Chebyshev extrema)

$$\left\{ x_k = -\cos\left(\frac{k\pi}{n}\right), \; k = 0, ..., n \right\}$$

and all zeros of orthogonal polynomials on a finite interval with respect to some positive measure. All these points are *near-optimal* in the sense that their Lebesgue constant grows logarithmically with respect to the degree n. Two other important sets of points are *Fekete points* and *Leja sequences* (cf. e.g. [32]) whose definition and properties will be discussed later on in the paper.

Fundamental question: Are there quasi-optimal interpolation nodes explicitly known in the multivariate setting for polynomial interpolation of total degree?

The answer is partially negative, *except for some known cases and in small dimensions* (see also the seminal paper by L. Bos [5]).

The previous question was the spring which pushed us in studying new families of near-optimal points, starting from the square $[-1, 1]^2$, being the square a simple domain, intrinsically tensorial, easy to be mapped to other domains (see [23]).

There are then many other questions and many more open problems, in this paper we present the answers to the following that were the main reasons why we discovered the *Padua points* on the square $\Omega = [-1, 1]^2$.

- We looked for *well-distributed nodes*. We found various nodal sets for polynomial interpolation of *even* degree n in the square Ω, which turned out to be equidistributed with respect to the *Dubiner metric* [45] and which show *near-optimal Lebesgue* constant growth [20].

- We also required *efficient interpolant evaluation*: the interpolant should be constructed without solving the Vandermonde system whose complexity is $\mathcal{O}(N^3)$, for each pointwise evaluation, with $N = \binom{n+2}{2}$ the dimension of the bivariate polynomials of total degree $\le n$. Moreover, we looked for closed formulae.
- We required *efficient cubature formulas*: in particular a fast computation of cubature weights for non-tensorial cubature formulae.

The last two points were inspired by the *rule of* 10 claimed by Nick L. Trefethen in [60] (also in a talk given in 2009 at the Dolomites Workshops in Alba di Canazei): a good implementation should last for 10 seconds, have a 10 digits precision and does not consist of more than 10 lines of executable code.

In section 2, we start by introducing the Dubiner metric and which is the one we used for the square. Then, in section 3 we recall the construction of the Padua points, their properties and outline some open problems. Section 4 is devoted to the description of the problem of approximating discontinuous functions, which was the main reason of studying the “fake" nodes. In Section 5, we then introduce the idea of the “fake” nodes approach and its equivalence with the mapping polynomial basis. Also in this section we outline some open problems and possible future developments. We finally conclude in Section 6.

As a final note, many of the figures are taken from the papers cited in the bibliography of which I am a co-author and that can be reproduced with the Matlab codes freely available online.

2. From Dubiner metric to Padua points

In his seminal paper [45], M. Dubiner introduced what we call the *Dubiner metric* which in $[-1, 1]$ corresponds to the arccosine distance between two points:

$$\mu_{[-1,1]}(x, y) = |\arccos(x) - \arccos(y)|, \ \ \forall x, y \in [-1, 1]. \tag{2.4}$$

By using the *Van der Corput-Schaake inequality* for trigonometric polynomials $T(\theta)$ of degree m and $|T(\theta)| \le 1$, that is,

$$|T'(\theta)| \le m\sqrt{1 - T^2(\theta)} \tag{2.5}$$

we want to show that the Dubiner metric is

$$\mu_{[-1,1]}(x, y) := \sup_{\|P\|_{\infty,[-1,1]}\le 1} \frac{1}{m}|\arccos(P(x)) - \arccos(P(y))| \tag{2.6}$$

with $P \in \mathbb{P}_n([-1, 1])$. Firstly, inequality (2.5) is equivalent to

$$\left|\frac{d}{d\theta}\arccos(T(\theta))\right| \le m. \tag{2.7}$$

The following result then holds.

Lemma 2.1. *Take $x, y \in [-1, 1]$ and $P \in \mathbb{P}_m([-1, 1])$, then*

$$|\arccos(x) - \arccos(y)| = \sup_{\|P\|_{\infty,[-1,1]}\le 1} \frac{1}{m}|\arccos(P(x)) - \arccos(P(y))|.$$

Proof. Letting $T(\theta) = P(\cos(\theta))$ and $x = \cos(\theta_x)$, $y = \cos(\theta_y)$. By using (2.7), we get

$$|\arccos(T(\theta_x)) - \arccos(T(\theta_y))| = \int_{\theta_x}^{\theta_y}\left|\frac{d}{d\theta}\arccos(T(\theta))\right| d\theta \le \int_{\theta_x}^{\theta_y} m d\theta \le m|\theta_x - \theta_y|.$$

But $\arccos(x) = \theta_x$, $\arccos(y) = \theta_y$ giving

$$|\arccos(T(\theta_x)) - \arccos(T(\theta_y))| \le m|\arccos(x) - \arccos(y)|$$

and

$$\sup_{\|P\|_{\infty,[-1,1]}\leq 1} \frac{1}{m} |\arccos(P(x)) - \arccos(P(y))| = |\arccos(x) - \arccos(y)|.$$

This concludes the proof. □

This metric generalizes to compact sets $\Omega \subset \mathbb{R}^d$, $d > 1$ (see e.g. [32]):

$$\mu_\Omega(\mathbf{x}, \mathbf{y}) := \sup_{\|P\|_{\infty,\Omega}\leq 1} \frac{1}{m} |\arccos(P(\mathbf{x})) - \arccos(P(\mathbf{y}))|.$$

This metric is important because there is an interesting unproved conjecture quoted in [20]:

Conjecture 1. *Nearly optimal interpolation points on a compact $\Omega \subset \mathbb{R}^d$ are asymptotically equidistributed with respect to the Dubiner metric on Ω.*

Hence, once we know the Dubiner metric on a compact Ω, we have at least a method for producing "good" interpolation points.

For $d = 2$, let $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$

- Dubiner metric on the square, $S = [-1, 1]^2$:

$$\mu_S(\mathbf{x}, \mathbf{y}) = \max\{|\arccos(x_1) - \arccos(y_1)|, |\arccos(x_2) - \arccos(y_2)|\}. \tag{2.8}$$

- Dubiner metric on the disk, $D = \{|\mathbf{x}| \leq 1\}$:

$$\mu_D(\mathbf{x}, \mathbf{y}) = \left|\arccos\left(x_1 y_1 + x_2 y_2 + \sqrt{1 - x_1^2 - x_2^2}\,\sqrt{1 - y_1^2 - y_2^2}\right)\right|. \tag{2.9}$$

As an example, by using the previous definition of the Dubiner metric on the square, we can extract points from a discretization of the square itself. In Fig. 1, we show 496 Dubiner nodes (corresponding on taking $n = 30$), Random and Euclidean points as well as their Lebesgue constants. Notice that the

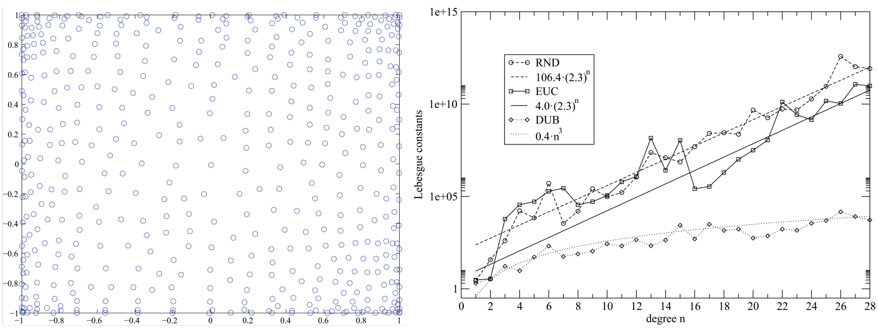


FIGURE 1. Left: Dubiner points. Right: Lebesgue constants growth.

Euclidean points, are *Leja-like* points, given by $\max_{\mathbf{x}\in\Omega} \min_{\mathbf{y}\in X_n} \|\mathbf{x} - \mathbf{y}\|_2$. There is a tight connection with the *Morrow-Patterson* (MP)-points (see [63]) which are a set of $N = \binom{n+2}{2} = \mathtt{dim}(\mathbb{P}_n^2)$ points in the square $[-1, 1]^2$, equidistributed with respect to the Dubiner metric (2.8). To be more precise, let n be a positive even integer, the MP-points are given by the following

$$x_m = \cos\left(\frac{m\pi}{n+2}\right), \quad y_k = \begin{cases} \cos\left(\frac{2k\pi}{n+3}\right), & \text{if } m \text{ odd} \\ \cos\left(\frac{(2k-1)\pi}{n+3}\right), & \text{if } m \text{ even} \end{cases}$$

$1 \le m \le n+1$, $1 \le k \le n/2+1$ and are unisolvent for the total degree interpolation problem.

The interest of these points where noticed by Len Bos who showed, in an unpublished note, that their Lebesgue constant grows polynomially in n and $\Lambda_{MP} = O(n^6)$. Later on, in [39] we showed, by using (the reciprocal of) Christoffel functions for estimating the Lebesgue constant of the hyperinterpolation operator on various 2-dimensional domains, that indeed $\Lambda_{MP} = O(n^3)$. Numerically, we actually found a growth of $O(n^2)$. So this is an open problem to show that the $\Lambda_{MP} = O(n^2)$.

Brutman introduced the so-called *extended* Chebyshev points [17].

$$\tilde{T}_n = \left\{ \tilde{x}_k = -\frac{1}{\gamma_n} \cos\left(\frac{(2k-1)\pi}{2n}\right), k = 1, ..., n \right\},$$

where $\gamma_n = \cos\left(\frac{\pi}{2n}\right)$, that is the set of Chebyshev points stretched to the boundary of the interval.

Similarly, we can define the *Extended Morrow-Patterson* points (EMP) as the points

$$x_m^{EMP} = \frac{1}{\alpha_n} x_m^{MP}, \quad y_k^{EMP} = \frac{1}{\beta_n} y_k^{MP},$$

$\alpha_n = \cos(\pi/(n+2))$, $\beta_n = \cos(\pi/(n+3))$.

Note: Both MP and the EMP points are equally distributed with respect to Dubiner metric on the square $[-1,1]^2$ and unisolvent for polynomial interpolation of degree n on the square $[-1,1]^2$ (see [20]). The *Padua points (PD)* are modified Morrow-Patterson points and were discovered "miraculously" in summer 2003, by Len Bos, Shayne Waldron, Marco Vianello and myself. They are the points in the square $[-1,1]^2$ with coordinates

$$x_m^{PD} = \cos\left(\frac{(m-1)\pi}{n}\right), \quad y_k^{PD} = \begin{cases} \cos\left(\frac{(2k-1)\pi}{n+1}\right), & \text{if } m \text{ odd} \\ \cos\left(\frac{2(k-1)\pi}{n+1}\right), & \text{if } m \text{ even} \end{cases}$$

$1 \le m \le n+1$, $1 \le k \le n/2+1$, $N = \binom{n+2}{2}$.

We recall here some fundamental properties proved in [8].

- The PD points are equispaced with respect to Dubiner metric μ_S on $[-1,1]^2$.
- The interior points are the MP points of degree $n-2$ while the boundary points are "natural" points of the grid. In Fig. 2 to the left, we show the set of Padua points for $n = 8$ as well as the MP and EMP.
- There are *4 families* of PD points obtained by taking rotations of 90 degrees: clockwise for *even* degrees and counterclockwise for *odd* degrees.
- The Lebesgue constant of the Padua points has optimal growth (see Fig. 2, right)

$$\Lambda(PD_n) = O((\log n)^2). \tag{2.10}$$

As a final note, their construction can be obtained in this simple way. Consider the $n+1$ Chebyshev-Lobatto points on $[-1,1]$

$$C_{n+1} = \left\{ z_j^n = \cos\left(\frac{(j-1)\pi}{n}\right), \ j = 1, \ldots, n+1 \right\}$$

and the two subsets of points with O=odd and E=even indexes

$$C_{n+1}^{\mathrm{O}} = \left\{ z_j^n, \ j = 1, \ldots, n+1, \ j \text{ odd} \right\},$$
$$C_{n+1}^{\mathrm{E}} = \left\{ z_j^n, \ j = 1, \ldots, n+1, \ j \text{ even} \right\}.$$

Then, the Padua points of degree n are the set

$$\mathrm{PD}_n = C_{n+1}^{\mathrm{O}} \times C_{n+2}^{\mathrm{E}} \cup C_{n+1}^{\mathrm{E}} \times C_{n+2}^{\mathrm{O}} \subset C_{n+1} \times C_{n+2}.$$

As a nice and interesting observation, the Padua points lie on n concentric squares with sides at the zeros of U_n and U_{n-1} (the inner) except the external and the center [29]. With U_k we indicate the classical orthogonal Chebyshev polynomials of second kind, see also Fig. 3.

3. Padua points: generating curve, WAMs, applications and open problems

There exists an alternative construction consisting of the self-intersections and boundary contacts of the parametric and periodic curve, called *generating curve*:

$$\gamma(t) = (-\underbrace{\cos((n+1)t)}_{T_{n+1}(t)}, -\underbrace{\cos(nt)}_{T_n(t)}), \qquad t \in [0, \pi].$$

For instance, in the figure below we display the curve $\gamma(t)$ for $n = 4$. The generating curve $\gamma(t)$ turns out to be a *Lissajous curve*. In particular, it is an algebraic curve such that $T_{n+1}(x) = T_n(y)$ (for the first family!). Being a Lissajous curve, we recall some important properties of these curves

- Their implicit equations can be found by using Chebyshev polynomials. Chebyshev polynomials are indeed Lissajous curves (cf. [62]).
- Lissajous curves are planar parametric curves studied by the astronomer Nathaniel Bowditch (1815) and later on by the mathematician Jules A. Lissajous (1857). They can be written in a general form as
$$\gamma(t) = (A_x \cos(\omega_x t + \alpha_x), A_y \sin(\omega_y t + \alpha_y)),$$
where A_x, A_y are amplitudes, ω_x, ω_y are pulsations and α_x, α_y are phases.

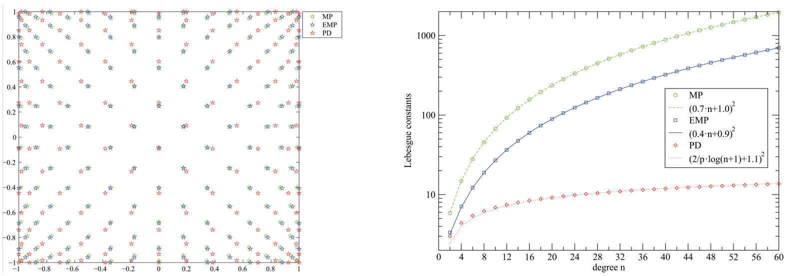


Figure 2. Left: the graphs of MP, EMP, PD for $n = 8$. Right: the growth of the corresponding Lebesgue constants.

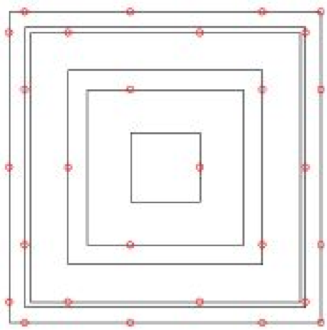


Figure 3. Padua for $n = 6$ are distributed on n concentric squares with sides at the zeros of U_n and U_{n-1} (the inner) except the external and the center (just a dot!).

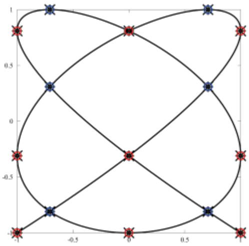


FIGURE 4. PD_4 on the generating curve and the two grids (with different colors).

In two dimensions, there is an interesting general definition described in [46].

Definition 3.1.

$$\gamma^{\boldsymbol{n}}_{\boldsymbol{\kappa},\boldsymbol{u}}(t) = \begin{pmatrix} u_1 \cos(n_2 t - \kappa_1\pi/(2n_1)) \\ u_2 \cos(n_1 t - \kappa_2\pi/(2n_2)) \end{pmatrix}, \ t \in [0, 2\pi],$$

with $\boldsymbol{n} = (n_1, n_2) \in \mathbb{N}^2$, $\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in \mathbb{R}^2$ and $\boldsymbol{u} = (u_1, u_2) \in \{-1, 1\}^2$. The values n_1, n_2 are called frequencies (like for the pendulum) and $\boldsymbol{u}$ reflection parameter.

It is nice and also quite instructive to see how Lissajous curves can be constructed by playing with the *sand pendulum* (see the video `https://www.youtube.com/watch?v=7f16hAs1FB4`).

The construction in the square $[-1, 1]^2$ goes as follows. Let $\boldsymbol{n} = (n_1, n_2)$ with $n_1, n_2 \in \mathbb{N}$ *relatively primes*. Then, we may consider the curves $\gamma^{\boldsymbol{n}}_{\epsilon} : [0, 2\pi] \to [-1, 1]^2$

$$\gamma^{\boldsymbol{n}}_{\epsilon}(t) := \gamma^{\boldsymbol{n}}_{(0,\epsilon-1),\mathbf{1}}(t) = \begin{pmatrix} \cos(n_2 t) \\ \cos(n_1 t + (\epsilon - 1)\pi/(2n_2)) \end{pmatrix} \tag{3.11}$$

with $\epsilon \in \{1, 2\}$ and *fixed* reflection parameter $\mathbf{1} = (1, 1)$.

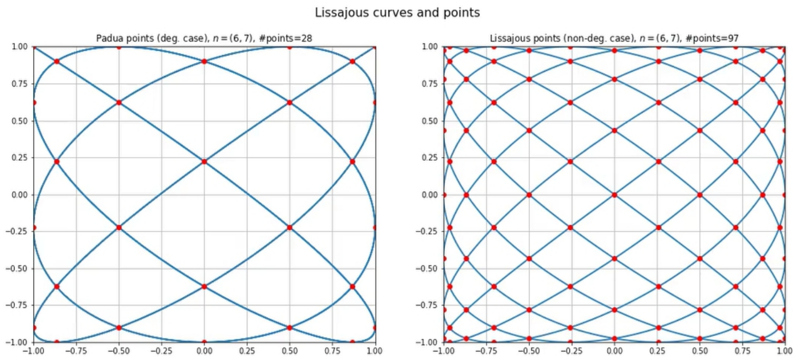


FIGURE 5. Left: Padua points, Right: Lissajous points. Both sets are relative to degree $\boldsymbol{n} = (6, 7)$, as used in (3.11).

Two special cases, whose details are discussed in [46], allow to classify Lissajous curves on the square in two main families.

- For $\epsilon = 1$, that is $\gamma^{\boldsymbol{n}}_1(t)$, is called a *degenerate* curve.
- For $\epsilon = 2$, that is $\gamma^{\boldsymbol{n}}_2(t)$, is called *non-degenerate* curve.

The Padua points curve is then a degenerate Lissajous curve, being two points of the curve at two consecutive corners of the square. Moreover, the degenerate Lissajous curve are π-periodic, while the non-degenerate are 2π periodic.

In Figure 5, we have displayed PD_6 and $Lis_{6,7}$. In particular the generating curves and the cardinalities are as follows:

$$\gamma^{PD}_{n,n+1} = (\cos(nt), \cos((n+1)t), \ \#PD_n = (n+2)(n+1)/2$$

$$\gamma^{Lis}_{n_1,n_2} = \left(\cos(n_2 t), \cos(n_1 t + \frac{\pi}{2n_2})\right), \ \#Lis_{n_1,n_2} = 2n_1n_2 + n_1 + n_2.$$

This shows that the Padua points are a unisolvent set for the total degree interpolation problem. While the Lissajous points can be used for polynomial interpolation, not of total degree, and they guarantee stability (slow growth of the Lebesgue constant).

The more general topic of multivariate polynomial approximation on Lissajous Curves turned out to be of interest in the emerging field of Magnetic Particle Imaging (MPI) (see, e.g., some recent publications and the activities of the scientific network MathMPI). Lissajous sampling seems to be relevant also in the field of Atomic Force Microscopy (AFM).

3.1. **Padua points are WAM (Weakly Admissible Meshes).** In the field of multivariate polynomial approximation, the notion of polynomial mesh has recently emerged as a significant concept. Originally introduced in the seminal paper [25], it has been studied in several subsequent papers, from both the theoretical and the computational point of view, interpolation and extracting Fekete points on 2d domains (cf. [14, 9, 12] and references therein). Moreover, approximate Fekete-like points extracted from polynomial meshes have begun to play a role in the framework of high-order methods for PDEs (cf., e.g., [83]).

We simply recall, that a polynomial *Weakly Admissible Mesh* (WAM) is a sequence of discrete subsets $\{A_n\}$ of a *polynomial determining* (i.e. polynomial vanishing there vanish everywhere) compact set $K \subset \mathbb{R}^d$ such that the inequality

$$\|p\|_k \leq C(A_n)\|p\|_{A_n}, \ \forall p \in \mathbb{P}^d_n \tag{3.12}$$

holds, where both the $card(A_n) \geq dim(\mathbb{P}^d_n) = O(n^d)$ and $C(A_n)$ are bounded by n^d. Notice that $\|f\|_X$ is the sup-norm of a function f bounded on the (discrete or continuous) set X. Properties of WAMs and various examples in one and two dimensional domains, are described in [41]. Hence, once we know a WAM, the computation of *discrete estremal sets*, can be done by numerical linear algebra techniques by using *greedy algorithms*. The interested reader can refer to [13, 12].

The following lemma is the fundamental result for the construction of WAMs by using tensor product strategies.

Lemma 3.2. *Let $p \in \mathbb{P}^1_n$ be a univariate algebraic polynomial, and C_n, $\tilde{C}_n$ the Chebyshev and Chebyshev-Lobatto nodal sets, respectively. Let $t \in \mathbb{T}^1_n$ be a univariate trigonometric polynomial, and Θ_n the angular nodal set*

$$\Theta_n(\alpha,\beta) = \phi_\omega(\tilde{C}_{2n}) + \frac{\alpha+\beta}{2} \subset (\alpha,\beta), \ \omega = \frac{\beta-\alpha}{2} \leq \pi,$$

where $\phi_\omega(r) = 2\arcsin(\sin\frac{\omega}{2}r)$, $r \in [-1,1]$. Then, the following polynomial inequalities hold

$$\|p\|_{[a,b]} \leq c_n\|p\|_{C_n} \tag{3.13}$$

$$\|p\|_{[a,b]} \leq c_n\|p\|_{\tilde{C}_n} \tag{3.14}$$

$$\|t\|_{[\alpha,\beta]} \leq c_{2n}\|t\|_{\Theta_n} \tag{3.15}$$

with $c_n = 1 + \frac{2}{\pi}\log(n+1)$.

Padua points can be used in 3-dimensional tensor product WAMs on different domains [43]. Knowing a WAM on a planar compact, say Ω, we can construct 3-dimensional WAMs for cones with base Ω and vertex y, which consists of all the segments connecting y with a point on Ω. Similarly the construction can be done for pyramids (which are cones with polygonal base) and truncated cones. The last is obtained by cutting the cone with a plane parallel to the base. We can also construct 3-dimensional WAMs for solid of rotation with cross section Ω and external axis r. The WAMs is then obtained by rotation of Ω by a given angle $\leq 2\pi$, around a coplanar line r.

For instance in Fig. 6, we show on the left the WAMs for a pyramid obtained by the tensor product of Padua points of degree 10 on the base and Chebyshev-Lobatto points along the z-axis, on the right the WAM on a portion of the torus with circular base. In both sets we have highlighted the approximate Fekete points extracted from the WAM by the greedy algorithm described in [13].

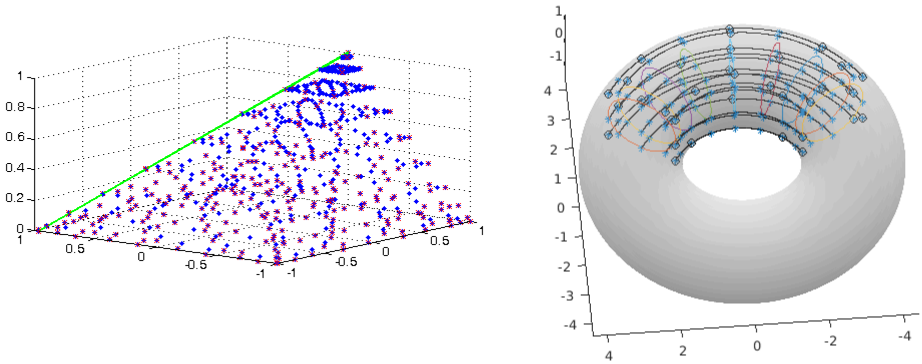


FIGURE 6. 3-dimensional WAMs obtained by using the Padua points.

3.2. **Some recent applications of the Padua points.** Lagrange interpolation at the Padua points has been recently used in several scientific and technological applications.

- Computational Chemistry (the Fun2D subroutine of the CP2K simulation package for Molecular Dynamics, `https://www.cp2k.org/`),
- Image Processing (algorithms for image retrieval by colour indexing),
- Materials Science (Modelling of Composite Layered Materials, [69]),
- Mathematical Statistics (Copula Density Estimation, [67]),
- Quantum Physics (Quantum State Tomography [59]),
- Padua points for solving PDEs with radial basis functions methods [58].

Padua points have been included in the *Chebfun2 package* (whose features have been described in the book [60]). The Padua points can be obtained simply specifying the degree n: `x=paduapts(n)`. For more details, see the web page `http://www.chebfun.org/examples/geom/Lissajous.html`

- **Software:** `www.math.unipd.it/~marcov/CAApadua.html`, J. Burkardt `https://people.sc.fsu.edu/~jburkardt/m_src/padua/padua.html`
- **Scholar citations (to the date): about 7140**.

3.3. **Some open problems.**

(1) We do not know the Padua points on $[-1, 1]^d$, $d \geq 3$.

(2) The Lebesgue function has its maxima in the corners, where there are no Padua points (see Fig. 7 that displays the Lebesgue function and its maximum at the corner points).

(3) The Vandermonde determinant associated to the Padua and Padua-like points has variables that separate. Using a notation similar to (1.1), for a point set $A = \{a_1, ..., a_N\} \in [-1, 1]^2$ and a basis $\mathcal{B} = \{b_1, \ldots, b_N\}$, we may construct the Vandermonde matrix

$$V(A; \mathcal{B}) = (b_i(a_j))_{i,j=1}^N,$$

where the i-th row of V consists of i-th polynomial of the basis $\mathcal{B}$ evaluated at all points. For Padua-like points $N = \binom{n+2}{2}$ and we denote with $Vdm(A; \mathcal{B})$ the corresponding determinant. Using the standard monomial basis of $\mathbb{P}_n(\mathbb{R}^2)$,

$$\mathcal{B}_n = \{x^\alpha y^\beta, \mid \alpha + \beta \le n\},$$

the tensor product basis

$$\mathcal{T}_n = \{x^\alpha y^\beta, \mid \max(\alpha, \beta) \le n\}$$

and the univariate polynomials

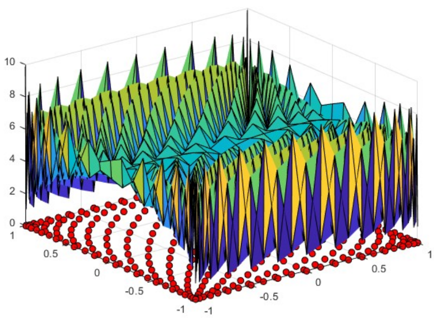
$$a(x) := \prod_{i=0}^{n/2} (x - x_{2i+1})$$

$$b(y) := \prod_{j=0}^{n/2} (y - y_{2j+1}),$$

another basis for $\mathbb{P}_n(\mathbb{R}^2)$ is

$$\mathcal{B}' = a(x)\mathcal{B}_{n/2-1} \cup b(y)\mathcal{B}_{n/2-1} \cup \mathcal{T}_n \tag{3.16}$$

such that $Vdm(A; \mathcal{B}_n) = \pm Vdm(A; \mathcal{B}'_n)$ being the transition matrix diagonal with 1 on the diagonal. This construction allowed to manipulate the Vandermonde matrix splitting it along the even and odd grids of the Padua-like points, providing an unexpected commutative property of the Vandermonde determinant associated to each direction. The claim in [11, Lemma 1] had a "gap". After some years, the Lemma was completely proved [42]. Moreover, we noticed that this "commutative" property of the Vandermonde determinant associated to Padua-like points, holds for general functions and general rectangular grids [31].



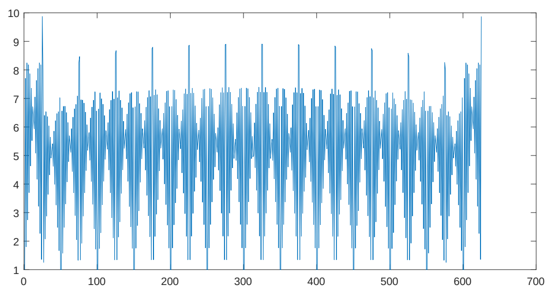


FIGURE 7. Padua points for $n = 25$ and its Lebesgue function. On the right the profile in 1d of the function.

4. Approximation of discontinuous functions

In this section, we deal with an important problem in data analysis, that is the reconstruction of functions with discontinuities or with jumps. The approach we describe is the *mapping bases technique* which turns out to be equivalent to the *"fake" nodes approach* [35, 37]. We recall that general approaches to overtake unavoidable reconstruction instabilities around the discontinuities are based on a clever choice of interpolation points before and after the jumps (cf. e.g. [33]), rational approximation (cf. e.g. [54, 4]), sinc-approx, filtering (cf. e.g. [36]). This list is not complete, but shows the wide interest to the topic. In particular, in image analysis in medicine (Computerized Tomography (CT), Magnetic Resonance (MR), and their variants (SPECT, fMRI)) or the above mentioned Magnetic Particle Imaging (MPI) or in geosciences, where satellite images are used to analyzed ground characteristics (humidity, temperature, water distribution and so on), often the images need to be geometrically aligned, registered or simply reconstructed by sampling them properly. In Figs. 8 and 9, we show some images connected to these applications.

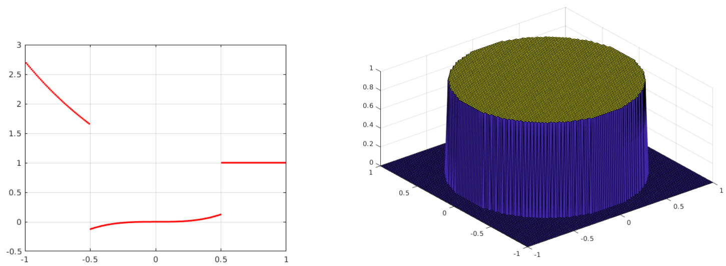


Figure 8. Discontinous functions in 1d and 2d.

- Interpolation by polynomials and rational functions of discontinuous functions is historically well-studied. Two related well-known phenomena are the *Runge and Gibbs* effects [71, 51]. In both cases, unwanted oscillations appears near the boundary of the domain or close to the discontinuities, respectively.

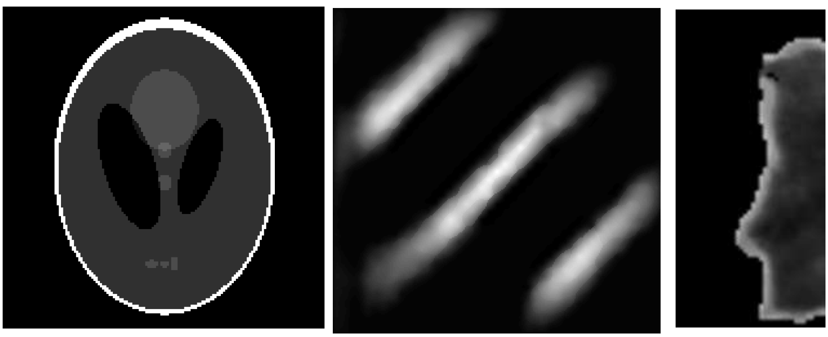


Figure 9. Left: the Shepp-Logan phantom used in medicine for testing. Center: an MPI acquisition reconstructed by Gaussian kernels. Right: RBF reconstruction of the soil of Portugal.

- More recently, interpolation by kernels, mainly radial basis functions has become a powerful tool for high-dimensional scattered data problems [52, 80, 18] and application to the solution of PDES [55], machine learning [72, 48], image registration and many other more.

5. The Fake Nodes Approach (FNA)

We start observing three facts from which "fake" nodes ideas originated.

(1) In applications, samples are given. *Resampling*, which is often necessary, can be done at Chebyshev points, or by extracting *mock Chebyshev* points from the data, or finding good interpolation points depending on applications (like Padua points, approximate Fekete points, discrete Leja sequences, Lissajous points, (P, f, β)-greedy points, minimal energy points and so on). For more details, see [35, 37].

(2) When the function has steep gradients, like $f(x) = \arctan(20x)$, $x \in (-0.22, 0.22)$, its reconstruction gives rise to oscillations nearby the boundaries. This is a well-known fact from the Fourier analysis of the coefficients of the corresponding series known as *Gibbs phenomenon*.

(3) For analytic functions on compact intervals, Adcock and Platte [1] investigated weighted least-squares approximation of mapped polynomial basis via the Kosloff and Tal-Azer map [57]:

$$\kappa_\alpha(x) = \frac{\sin(\alpha\pi x/2)}{\sin(\alpha\pi/2)},\ x \in [-1, 1],\ \alpha \in (0, 1]$$

giving rise to the α-polynomial space

$$\mathbb{P}_n^\alpha = \{p \circ \kappa_\alpha,\ p \in \mathbb{P}_n\},$$

which corresponds to the space of trigonometric polynomials when $\alpha = 1$ and the classical polynomial space when $\alpha = 0$ (which is excluded).

These observations are the main ingredients of the FNA which, as we shall see, is equivalent to a polynomial mapping of the original polynomial space. We need some notations. Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map. The main idea behind the FNA, is that of constructing an interpolant $R_f \in \mathcal{B}_N^S := \mathrm{span}\{B_1^S, \ldots, B_N^S\}$ of the function f, so that

$$R_f(\boldsymbol{x}) = \sum_{i=1}^N \alpha_i^S B_i^S(\boldsymbol{x}) = \sum_{i=1}^N \alpha_i^S B_i(S(\boldsymbol{x})) = P_g(S(\boldsymbol{x})), \forall \boldsymbol{x} \in \Omega. \tag{5.17}$$

The function g has the property that $g_{|S(X_N)} = f_{|X_N}$, that is, it assumes the same values of f at the mapped interpolation points $S(X_N)$. Thus, having the mapped basis $\mathcal{B}_N^S$, the construction of the interpolant R_f is equivalent to build a classical interpolant $P_g \in \mathcal{B}_N$ at the "fake" or mapped nodes $S(X_N)$. In what follows we will use the words *"fake" nodes*, thinking of this mapping process.

Provided we have a unisolvent set of points for the given basis, $X_N = \{x_1, ..., x_N\}$, and the corresponding values $\boldsymbol{f} = \{f(x_1), \ldots, f(x_N)\}$, R_f can be constructed by solving the linear system

$$\mathsf{A}^S \boldsymbol{\alpha}^S = \boldsymbol{f}, \tag{5.18}$$

where $\boldsymbol{\alpha}^S = (\alpha_1^S, \ldots, \alpha_N^S)^\intercal$, and

$$\mathsf{A}^S = \begin{pmatrix} B_1^S(\boldsymbol{x}_1) & \ldots & B_1^S(\boldsymbol{x}_N) \\ \vdots & \ddots & \vdots \\ B_N^S(\boldsymbol{x}_1) & \ldots & B_N^S(\boldsymbol{x}_N) \end{pmatrix}.$$

Concerning the cardinal form of the mapped interpolant, we may state the following proposition.

Proposition 5.1 (Cardinal form)**.** *Let $X_N = \{\boldsymbol{x}_i,\ i = 1, \dots, N\} \subseteq \Omega$ be a set of pairwise distinct data points and let $u_i \in \mathcal{B}_N$, $i = 1, \dots, N$ be the basis functions. Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map. The functions $\{u_1, \dots, u_N\}$ are cardinal on $S(\Omega)$ for the "fake" nodes $S(X_N)$ if and only if the mapped functions $\{u_1 \circ S, \dots, u_N \circ S\}$ are cardinal for the original set of nodes X_N.*

The proof is trivial and comes immediately asking the cardinality property to the functions u_i^S. Hence, we can write the interpolant at the "fake" nodes in *cardinal* form:

$$R_f^S(\boldsymbol{x}) = \boldsymbol{f}^\intercal \boldsymbol{u}^S(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \tag{5.19}$$

where $\boldsymbol{u}^S(\boldsymbol{x}) = (u_1^S(\boldsymbol{x}), \dots, u_N^S(\boldsymbol{x}))^\intercal$.

The Lebesgue constant of the points mapped via R_f^S is equivalent to that of the image Ω though S (see [37] for details).

Proposition 5.2 (Equivalence of the Lebesgue constant)**.** *Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map. Let $X_N \subseteq \Omega$ be a unisolvent set of nodes for the space $\mathcal{B}_N$, and $u_i^S \in \mathcal{B}_N^S$, $i = 1, \dots, N$, be the associated cardinal functions. Then, the Lebesgue constant $\Lambda^S(\Omega)$ associated to the mapped nodes is*

$$\Lambda^S(\Omega) = \Lambda(S(\Omega)).$$

Remark 5.1. *The proposition states that the interpolation at the mapped basis $\mathcal{B}_N^S$ inherits the Lebesgue constant of the "fake" nodes $S(X_N)$ over the 'standard' basis $\mathcal{B}_N$.*

The Lebesgue constant, as well-known, represents the stability constant of the interpolation process. For analyzing the stability, we thus study an interpolant of perturbed data $\tilde{f}(\boldsymbol{x}_i)$ sampled at $\boldsymbol{x}_i$, $i = 1, \dots, N$, i.e. data affected by measurement errors.

Proposition 5.3 (Stability)**.** *Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map and $X_N \subseteq \Omega$ be a unisolvent set of nodes for the space $\mathcal{B}_N$. Let $\boldsymbol{f}$ be the associated vector of function values and $\tilde{\boldsymbol{f}}$ be the vector of perturbed values. Let R_f^S and $R_{\tilde{f}}^S$ be the interpolant of the function values $\boldsymbol{f}$ and $\tilde{\boldsymbol{f}}$, respectively. Then,*

$$\|R_f^S - R_{\tilde{f}}^S\|_{\infty,\Omega} \leq \Lambda^S(\Omega)\, \|\boldsymbol{f} - \tilde{\boldsymbol{f}}\|_{\infty,X_N}.$$

Proof. Taking into account that $g_{|S(X_N)} = f_{|X_N}$ and thus also $\tilde{g}_{|S(X_N)} = \tilde{f}_{|X_N}$, we deduce that

$$\begin{aligned}
\|R_f^S - R_{\tilde{f}}^S\|_{\infty,\Omega} = \|P_g - P_{\tilde{g}}\|_{\infty,S(\Omega)} &= \sup_{x \in S(\Omega)} \left| \sum_{i=1}^N (g_i(\boldsymbol{x}_i) - \tilde{g}_i(\boldsymbol{x}_i))\ u_i(\boldsymbol{x}) \right| \\
&= \sup_{x \in \Omega} \left| \sum_{i=1}^N (g_i(S(\boldsymbol{x}_i)) - \tilde{g}_i(S(\boldsymbol{x}_i)))\ u_i(S(\boldsymbol{x})) \right| \\
&\leq \sup_{x \in \Omega} \sum_{i=1}^N |u_i(S(\boldsymbol{x}))|\ |g_i(S(\boldsymbol{x}_i)) - \tilde{g}_i(S(\boldsymbol{x}_i))| \\
&\leq \sup_{x \in \Omega} \sum_{i=1}^N |u_i(S(\boldsymbol{x}))| \max_{i=1,\dots,N} |g_i(S(\boldsymbol{x}_i)) - \tilde{g}_i(S(\boldsymbol{x}_i))| \\
&= \Lambda(S(\Omega)) \max_{i=1,\dots,N} \left| f(\boldsymbol{x}_i) - \tilde{f}_i(\boldsymbol{x}_i) \right| \\
&= \Lambda^S(\Omega)\, \|\boldsymbol{f} - \tilde{\boldsymbol{f}}\|_{\infty,X_N}.
\end{aligned}$$

This concludes the proof. □

Consistently with Remark 5.1, the FNA approach also inherits the error of the classical approach, as shown in the following proposition.

Proposition 5.4 (Error bound inheritance)**.** *Letting S, X_N, f and R_f^S, as above. Then, for any given function norm, we have*

$$\|R_f^S - f\|_\Omega = \|P_g - g\|_{S(\Omega)},$$

where $g_{|S(X_N)} = f_{|X_N}$.

Proof. From (5.17), we know that $R_f^S = P_g \circ S$. Choosing g such that $g \circ S = f$ on Ω (this g exists being S injective), we get

$$\|R_f^S - f\|_\Omega = \|P_g \circ S - g \circ S\|_\Omega = \|P_g - g\|_{S(\Omega)},$$

which gives the claimed result. □

5.1. **Mapped bases.** As discussed above, let $S : I \to \mathbb{R}$ be a given map. We are interested to the function

$$R_{n,f}^S(x) := P_{n,g}(S(x)) = \sum_{i=0}^{n} c_i S_i(x) \tag{5.20}$$

for some $g : S(I) \to \mathbb{R} \in C^r(I)$ such that

$$g_{|S(X_n)} = f_{|X_n}.$$

$R_{n,f}^S \in \mathsf{span}\{S_i = m_i \circ S,\ i = 0, \dots, n\}$ is the interpolant at (X_n, F_n), that is no resampling is done. This mapping construction is equivalent to the "fake" nodes approach.

- The mapped bases approach on I ask to *"interpolate f on the set X_n via $R_{n,f}^s$ in the function space S_n."*
- The FNA on $S(I)$ ask to *"interpolate g on the set $S(X_n)$ via $P_{n,g}$ in the polynomial space M_n."*

Remark 5.2. *This approach is rather general, in the sense that we may use any space of linear independent functions (polynomials, rational function, radial basis functions and so on). The only point to clarify is the choice of the map S.*

Problem 2. *How can we find a suitable admissible map S for mitigating the Runge and Gibbs effects?*

The map S should be taken so that the resulting set of "fake" nodes $S(X_n)$ guarantees a stable interpolation process. A "natural" choice for a stable interpolation is to map X_n for example, to the set of Chebyshev-Lobatto (CL) nodes on the interval I.

The following algorithms, S-Runge and S-Gibbs, provide a constructive solution to Problem 2.

Algorithm 1 (S-Runge)**.**

Input: *X_n, C_n. Note: X_n is ordered left-right, C_n are the CL nodes.*

Core

- *If $x \in [x_i, x_{i+1}]$, for $i \in \{0, \dots, n-1\}$, S is the (piecewise) linear map*

$$S(x) = \beta_{1,i}(x - x_i) + \beta_{2,i},$$

where

$$\beta_{1,i} = \frac{c_{i+1} - c_i}{x_{i+1} - x_i}, \quad \beta_{2,i} = c_i.$$

Output: $S(x)$.

For S-Gibbs, we need to identify the set of discontinuities

$$D_m := \{(\xi_i, d_i) \mid \xi_i \in (a, b),\ \xi_i < \xi_{i+1},\ \text{ and } d_i := |f(\xi_i^+) - f(\xi_i^-)|\},\ i = 0, \dots, m$$

by an edge-detection algorithm. This can be done by well-known and stable techniques, such as the the Canny algorithm described in [24] or, for irregularly samples signals and images, in [2]. When Radial basis functions are used, the analysis of the coefficients of the interpolant, can give information on the location of the discontinuities, as described in [70]. Recently, we proposed another approach to extract the location of the discontinuities through a segmentation method based on a classification algorithm from machine learning (see [38]).

Algorithm 2 (S-Gibbs)**.**

Inputs: *X_n, D_m, x and $k \in \mathbb{R}_+$.*
Core
(a) $\alpha_i := kd_i,\ i = 0, \ldots, m.$

(b) Letting $A_i = \sum_{j=0}^{i} \alpha_j$, define S as follows:

$$S(x) = \begin{cases} x, & \textit{for } x \in [a, \xi_0[, \\ x + A_i, & \textit{for } x \in [\xi_i, \xi_{i+1}[,\ 0 \le i < m, \textit{ or } x \in [\xi_m, b]. \end{cases}$$

Output: *$S(x)$.*

Remarks. *Some comments are in order.*

- *Our strategy consists in constructing the map S in such a way that it sufficiently increases the gap between the node right before and the one right after the discontinuities via the real parameters α_i.*
- *About the shifting parameter $k > 0$. We experimentally observed that its selection is not critical. The resulting interpolation process is not sensitive to its choice, provided that it is sufficiently large, i.e. in such a way that in the mapped space the so-constructed function g has no steep gradients.*
- *The "fake" nodes mapping, S-Runge, enables one to obtain an interpolant on equispaced points that may converge efficiently while avoiding Runge phenomenon. The connection worth to be emphasized regards the application of this mapping on a polynomial basis. In particular, if we consider the Chebychev polynomials of the first kind, that is*

  $$T_k(x) = \cos(k \arccos(x)),\ \ for\ x \in [-1; 1],\ k \ge 0,$$

  then, it appears that applying the "fake" nodes mapping to T_k on a general interval $[a, b]$, provides a Fourier basis $\hat{T}_k$:

  $$\hat{T}_k(x) = T_k(\cos(\pi(x - a)/(b - a))) = \cos(k\pi(x - a)/(b - a)).$$

  In other words, interpolating with the "fake" nodes mapping is equivalent to a particular decomposition in Fourier series. It also means that one can make direct connections with several tricks used e.g. by the software `Chebfun` *[64] and easily find the series coefficients via an FFT. An application of this idea has recently been explored in [56].*

In Fig. 10, we plot the cardinal functions on 4 nodes (so cubics), at varying the location of the discontinuity ξ and the shift parameter k. The cardinals become discontinuous at ξ. When ξ is not at the center of the interval, they do not look anymore cubics.

5.2. **Examples.**

5.2.1. *Runge phenomenon.* The first example of the FNA deals with the interpolation of the *Runge function*. We take $I = [-5, 5],\ f_1(x) = 1/(1 + x^2),\ X_n$: equally spaced. As evaluation points we consider a set of 100 equally spaced points.

We computed the Relative Max Approximation Error (RMAE), that is

$$\mathtt{RMAE} = \max_{z \in E} \frac{|R^s_{n,f}(z) - f(z)|}{|f(z)|},$$

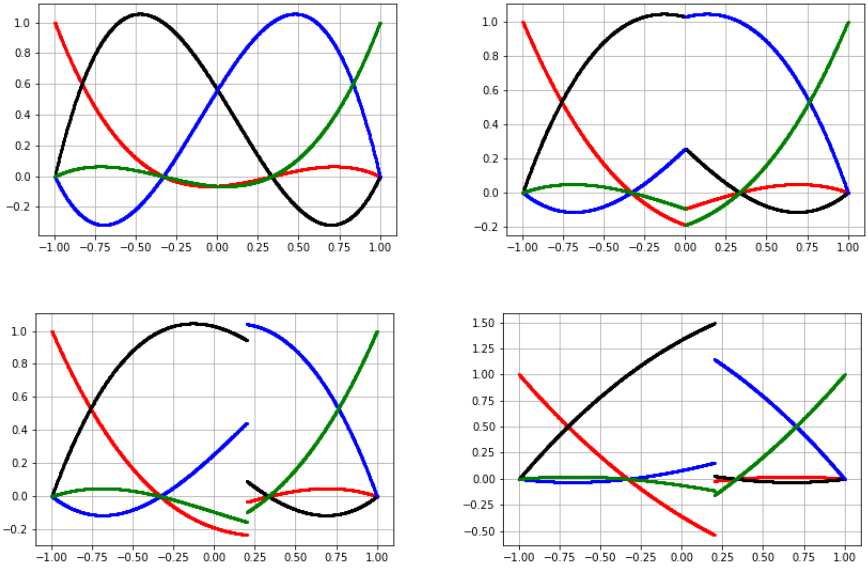


FIGURE 10. Left-right, up-down: the original cardinals on 4 nodes, the cardinals around $\xi = 0, k = 0$ the cardinals around $\xi = 0.2, k = 1$,the cardinals around $\xi = 0, k = 0.5$.

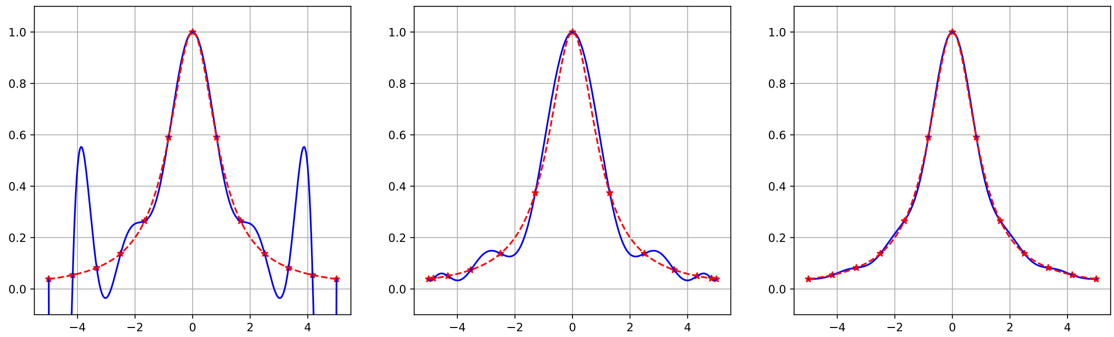


FIGURE 11. Interpolation at 13 points of f_1. Using equispaced (left), CL (center) and "fake" nodes (right). The original and reconstructed functions are plotted with continuous red and dotted blue lines, respectively.

5.2.2. *Gibbs phenomenon.* The second example deals with the Gibbs effect. We consider the discontinuous function below

$$f_2(x) := \begin{cases} \dfrac{x^2}{10}, & -5 \leq x < -\frac{3}{2}, \\ \dfrac{1}{4}x + \dfrac{19}{8}, & -\frac{3}{2} \leq x < \frac{5}{2}, \\ -\dfrac{x^3}{30} + 4, & \frac{5}{2} \leq x \leq 5. \end{cases}$$

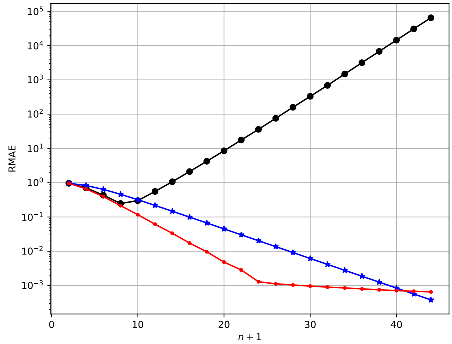


FIGURE 12. The RMAE for the Runge function varying the number of nodes. The results with equispaced, CL and “fake” nodes are represented by black circles, blue stars and red dots, respectively.

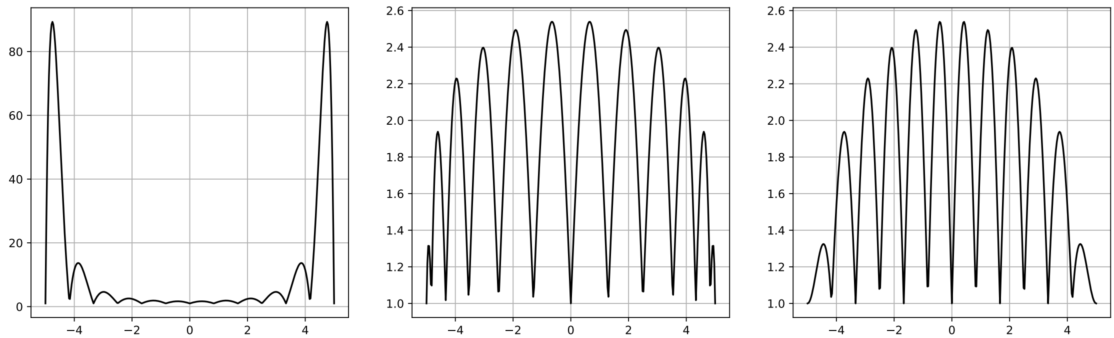


FIGURE 13. Lebesgue functions of equispaced (left), CL (center) and “fake” CL (right) nodes.

In this example $\mathcal{D} = \{(-3/2, 1.775), (5/2, 0.479)\}$. As before, we compare:

a) the interpolating polynomial at equispaced points E_n and associated function values $f_2(E_n)$;
b) the interpolating polynomial at the CL nodes C_n in I and resampled function values $f_2(C_n)$;
c) the approximant built upon the polynomial interpolant at the “fake” nodes, $S(E_n)$, and function values related to the equispaced points $f_2(E_n)$. In this setting, we fix $k = 50$ and the map S of the S-Gibbs algorithm.

5.3. **Extensions.** The mapped basis approach suggested many interesting applications. Here, we enumerate the most interesting ones and the corresponding references in which interested readers can refer to.

- Quadrature weights of the “fake" Chebyshev-Lobatto nodes are those of the composite trapezoidal rule [34].
- In 2d and 3d, as we have already seen, we can extract approximate Fekete points on various domains (disk, sphere, polygons, spherical caps, lunes, etc.). With these points we can apply the mapped basis approach for least-squares approximation [37]. In the 2d case, we have results on the approximation of discontinuous functions on the square, using polynomial approximation at the Padua points or tensor product meshes, see Figs. 17 and 18. It is interesting to see Fig.

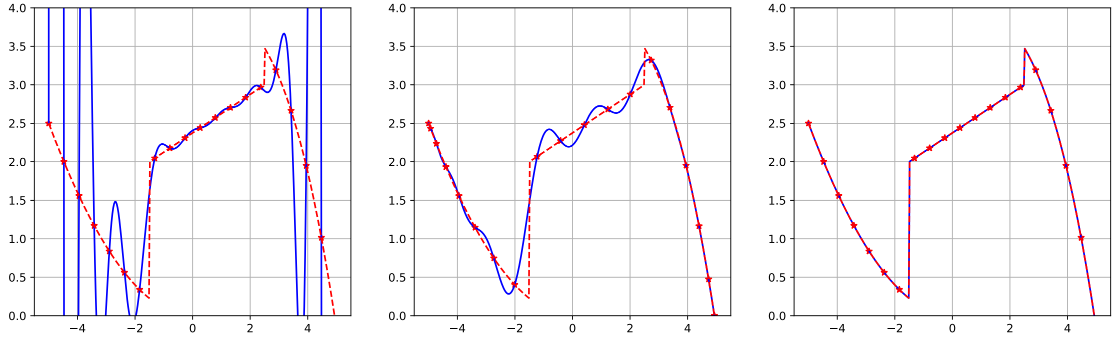


FIGURE 14. Interpolation at 20 points of the function f_2 on $[-5, 5]$, using equispaced (left), CL nodes (center) and the discontinuous map (right). The nodes are represented by stars, the original and reconstructed functions are plotted with continuous red and dotted blue lines, respectively.

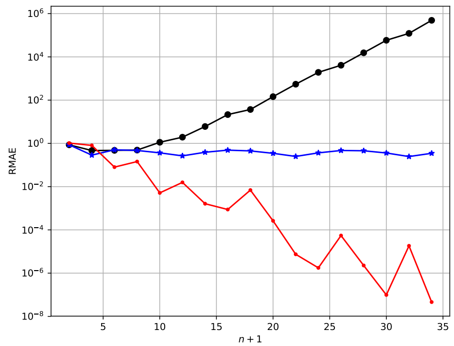


FIGURE 15. The RMAE for the function f_2 varying the number of nodes. The results with equispaced, CL and "fake" nodes are represented by black circles, blue stars and red dots respectively.

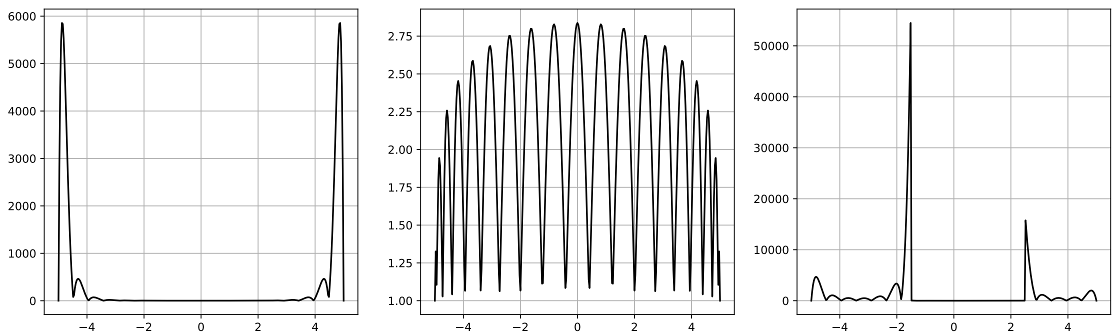


FIGURE 16. Lebesgue functions of equispaced (left), CL (center) and "fake" nodes (right).

18 where we show how to extract and map at the Padua points, *fake Padua*, starting from an original grid.

- In higher dimensions, where Padua points are not known, we may sample the function at the so-called Lissajous points or in the case of scattered data approximate by variably scaled discontinuous kernels [38].
- Extensions to rational interpolation/approximation: Floater-Hormann (FH) and trigonometric FH (for periodic signals) interpolants and the *AAA-approximation* (see [4] and references therein).
- The original proposed S-Gibbs map suffers of a subtle instability when the interpolation is done at equidistant nodes, a consequence of the Runge's phenomenon. A new approach, termed *Gibbs-Runge-Avoiding Stable Polynomial Approximation (GRASPA)* has been introduced in [33], which allows to mitigate both Runge and Gibbs phenomena
- In multimodal medical imaging, it is a common practice to undersample the anatomically-derived segmentation images to measure the mean activity of a co-acquired functional image. This avoids the resampling-related Gibbs effect that would occur in oversampling the functional image. It turns out that the FNA for image resampling it is designed to reduce the Gibbs effect when oversampling the functional image. This has been proved by a tight error analysis in [66].
- Links: `https://en.wikipedia.org/wiki/Runge%27s_phenomenon#S-Runge_algorithm_without_resampling`

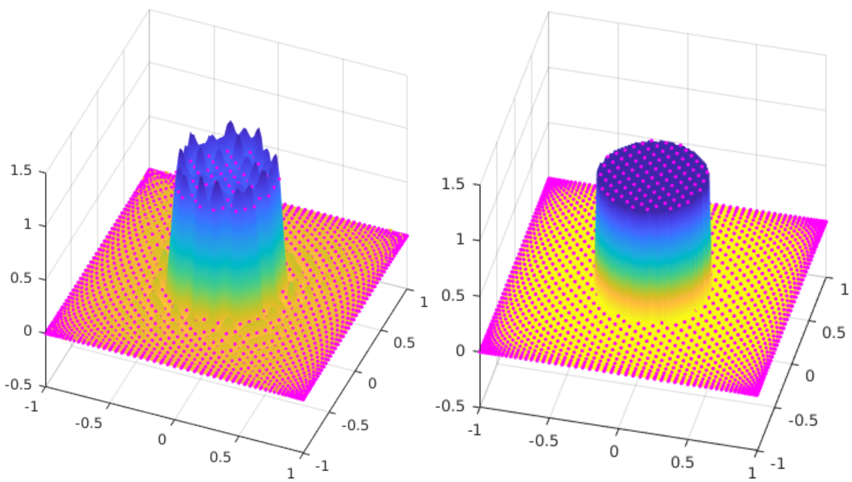


FIGURE 17. Left: interpolation with PD60 of a function with a circular jump. Right: the same by mapping circularly the PD points, and using least-squares fake-Padua.

5.4. **Some open problems.**

- As mentioned above, S-Runge and S-Gibbs have been improved in [33] via the GRASPA approach. Extension, at least to two dimensions, is needed.
- Recently two dimensional mock-Chebyshev points plus regression have been investigated [44]. Is this approach an alternative to the "fake" one?
- Error analysis and tight Lebesgue constant bounds should be investigated.

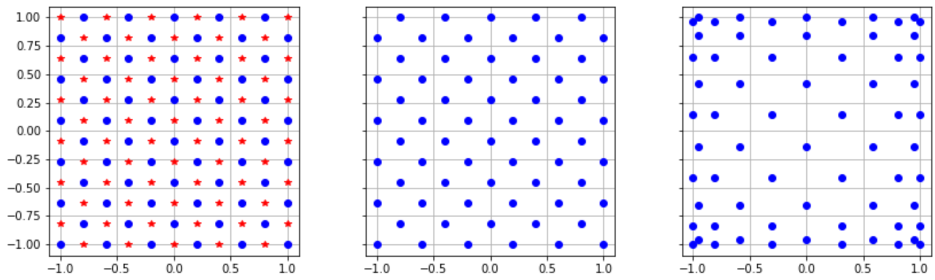


FIGURE 18. Here $n = 10$. On the left the set X_{66} (represented by blue dots) is extracted from a 11×12 equispaced grid (represented by both blue dots and red stars). The set X_{66} (centre) is then mapped on the set of Padua points Pad_{66} via the mapping S (right).

6. CONCLUSIONS

In this paper, we have reviewed the most important facts concerning the Padua points and the mapped bases approach for polynomial approximation of functions and data. We also outlined some open problems with the hope that some researcher can be interested in these topics and can propose a solution.

Acknowledgments. This work has been accomplished within the "Rete Italiana di Approssimazione" (RITA), the thematic group on "Approximation Theory and Applications" of the Unione Matematica Italiana (UMI). The paper in its final form, has been completed during the Erasmus mobility at the University "Lucian Blaga" of Sibiu, invited by Prof A. Acu of the Department of Mathematics and Computer Science.

REFERENCES

[1] B. Adcock, R. B. Platte: *A mapped polynomial method for high-accuracy approximations on arbitrary grids*, SIAM J. Numer. Anal., **54** (2016), 2256–2281.

[2] R. Archibald, A. Gelb and J. Yoon: *Polynomial Fitting for Edge Detection in Irregularly Sampled Signals and Images*, SIAM J. Numer. Analysis, **43** (1) (2005), 259–279.

[3] J. Baglama, D. Calvetti and L. Reichel: *Fast Leja points,* Electron. Trans. Numer. Anal., **7** (1998), 124–140.

[4] J.-P. Berrut, S. De Marchi, G. Elefante and F. Marchetti: *Treating the Gibbs phenomenon in barycentric rational interpolation and approximationvia the S-Gibbs algorithm*, Appl. Math. Letters, **103** (2020), 106196.

[5] L. Bos: *On certain configurations of points in $\mathbb{R}^n$ which are unisolvent for polynomial interpolation*, J. Approx. Theory, **64** (3) (1991), 271–280.

[6] L. Bos: *Multivariate interpolation and polynomial inequalities*, Ph.D. course held at the University of Padua (2001), unpublished.

[7] L. Bos, M. Caliari, S. De Marchi and M. Vianello: *A numerical study of the Xu interpolation formula*, Computing, **76** (3-4) (2006), 311–324.

[8] L. Bos, M. Caliari, S. De Marchi, M. Vianello and Y. Xu: *Bivariate Lagrange interpolation at the Padua points: the generating curve approach*, J. Approx. Theory, **143** (1) (2006), 15–25.

[9] L. Bos, J.-P. Calvi, N. Levenberg, A. Sommariva and M. Vianello: *Geometric Weakly Admissible Meshes, Discrete Least Squares Approximations and Approximate Fekete Points*, Math. Comp., **80** (2011), 1601–1621.

[10] L. Bos, S. De Marchi, M. Vianello and Y. Xu: *Bivariate Lagrange interpolation at the Padua points: the ideal theory approach*, Numer. Math., **108** (1) (2007), 43–57.

[11] L. Bos, S. De Marchi and S. Waldron: *On the Vandermonde Determinant of Padua-like Points (on Open Problems section)*, Dolomites Res. Notes on Approx., **2** (2009), 1–15.

[12] L. Bos, S. De Marchi, A. Sommariva and M. Vianello: *Weakly Admissible Meshes and Discrete Extremal Sets*, Numer. Math. Theory Methods Appl., **4** (2011), 1–12.

[13] L. Bos, S. De Marchi, A. Sommariva and M. Vianello: *Computing multivariate Fekete and Leja points by numerical linear algebra*, SIAM J. Num. Anal., **48** (5) (2010), 1984–1999.
[14] L. Bos, N. Levenberg: *On the Approximate Calculation of Fekete Points: the Univariate Case*, Elec. Trans. Numer. Anal., **30** (2008), 377–397.
[15] L. Bos, A. Sommariva and M. Vianello: *Least-squares polynomial approximation on weakly admissible meshes: disk and triangle*, J. Comput. Appl. Math., **235** (2010), 660–668.
[16] L. Bos, M. A. Taylor and B. A. Wingate: *Tensor product Gauss-Lobatto points are Fekete points for the cube*, Math. Comp., **70** (2001), 1543–1547.
[17] L. Brutman: *Lebesgue functions for polynomial interpolation: a survey*, Ann. Numer. Math., **4** (1997), 111–127.
[18] M. D. Buhmann: *Radial Basis Functions: Theory and Implementation*, Cambridge Monogr. Appl. Comput. Math., Vol. 12, Cambridge Univ. Press, Cambridge, (2003).
[19] CAA: Padova-Verona Research Group on Constructive Approximation webpage: `https://sites.google.com/view/caa-padova-verona/home`
[20] M. Caliari, S. De Marchi and M. Vianello: *Bivariate polynomial interpolation on the square at new nodal sets*, Appl. Math. Comput., **165** (2) (2005), 261–274.
[21] M. Caliari, S. De Marchi, M. Vianello: *Algorithm 886: Padua2D: Lagrange Interpolation at Padua Points on Bivariate Domains*, ACM Trans. Math. Software, **35** (3) (2008), 1–11.
[22] M. Caliari, S. De Marchi and M. Vianello: *Bivariate Lagrange interpolation at the Padua points: computational aspects*, J. Comput. Appl. Math., **221** (2008), 284–292.
[23] M. Caliari, S. De Marchi, A. Sommariva and M. Vianello: *Padua2DM: fast interpolation and cubature at Padua points in Matlab/Octave*, Numer. Algorithms, **56** (1) (2011), 45–60.
[24] J. Canny: *A Computational Approach to Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **8** (6) (1986), 679–698.
[25] J. P. Calvi, N. Levenberg: *Uniform approximation by discrete least squares polynomials*, J. Approx. Theory, **152** (2008), 82–100.
[26] W. Cheney, W. Light: *A Course on Approximation Theory*, AMS, Vol. 101, (2009).
[27] K. C. Chung, T. H. Yao: *On lattices adimmitting unique Lagrange interpolations*, SIAM J. Numer. Anal., **14** (1977), 735–743.
[28] P. Davis: *Interpolation and Approximation*, Blaisdell Pub Company, New York, (1963).
[29] A. Cuyt, I. Yaman, B. A. Ibrahimoglu and B. Benouahmane: *Radial orthogonality and Lebesgue constants on the disk*, Numer. Algorithms, **61** (2) (2012), 291–313.
[30] C. de Boor: *A Practical Guide to Splines*, revised edition, Springer, New York, (2001).
[31] A. P. de Camargo, S. De Marchi: *A few remarks on "On certain Vandermonde determinants whose variables separate"*, Dolomites Res. Notes Approx., **8** (2015), 1–11.
[32] S. De Marchi: *On Leja sequences: some results and applications*, Appl. Math. Comput., **152** (3) (2004), 621–647.
[33] S. De Marchi, G. Elefante and F. Marchetti: *Stable discontinuous mapped bases: the Gibbs-Runge-Avoiding Stable Polynomial Approximation (GRASPA) method*, Comput. Appl. Math., **40**:299 (2021).
[34] S. De Marchi, G. Elefante, E. Perracchione and D. Poggiali: *Quadrature at fake nodes*, Dolomites Res. Notes Approx., **14** Special Issue MATA2020 (2021), 39–45.
[35] S. De Marchi, F. Marchetti, E. Perracchione and D. Poggiali: *Polynomial interpolation via mapped bases without resampling*, J. Comput. Appl. Math., **364** (2020), 112347.
[36] S. De Marchi, W. Erb and F. Marchetti: *Lissajous sampling and spectral filtering in MPI applications: the reconstruction algorithm for reducing the Gibbs phenomenon*, 2017 International Conference on Sampling Theory and Applications SampTA (2017), 580–584.
[37] S. De Marchi, F. Marchetti, E. Perracchione and D. Poggiali: *Multivariate approximation at fake nodes*, Appl. Math. Comput., **391** (2021), 125628.
[38] S. De Marchi, W. Erb, F. Marchetti, E. Perracchione and M. Rossini: *Shape-Driven Interpolation with Discontinuous Kernels: Error Analysis, Edge Extraction and Applications in Magnetic Particle Imaging*, SIAM J. Sci. Comput., **42** (2) (2020), B472-B491.
[39] S. De Marchi, A. Sommariva and M. Vianello: *Multivariate Christoffel functions and hyperinterpolation*, Dolomites Res. Notes Approx., **7** (2014), 36–33.
[40] S. De Marchi, R. Schaback and H. Wendland: *Near-Optimal Data-Independent Point Locations for Radial Basis Function Interpolation*, Adv. Comput. Math., **23** (3) (2005), 317–330.
[41] S. De Marchi, F. Piazzon, A. Sommariva and M. Vianello: *Polynomial Meshes: Computation and Approximation*, Proceedings of the 15th International Conference on Computational and Mathematical Methods in Science and Engineering CMMSE (2015), 414–425.
[42] S. De Marchi, K. Usevich: *On certain multivariate Vandermonde determinants whose variables separate*, Linear Alg. Appl., **449** (2014), 17–27.

[43] S. De Marchi, M. Vianello: *Polynomial approximation on pyramids, cones and solids of rotation*, Dolomites Res. Notes Approx., Proceedings DWCAA12, **6** (2013), 20–26.

[44] F. Dell'Accio, F. Di Tommaso and F. Nudo: *Generalizations of the constrained mock-Chebyshev least squares in two variables: Tensor product vs total degree polynomial interpolation*, Appl. Math. Letters, **125** (2022), 107732.

[45] M. Dubiner: *The theory of multi-dimensional polynomial approximation*, J. Anal. Math., **67** (1995), 39–116.

[46] W. Erb, C. Kathner, P. Denker and M. Alhborg: *A survey on bivariate Lagrange interpolation on Lissajous nodes*, Dolomites Res. Notes Approx., **8** (2015), 23-36.

[47] G. E. Fasshauer: *Meshfree Approximation Methods with Matlab*, World Scientific Publishing, Interdisciplinary Mathematical Sciences, Vol. 6, Singapore, (2007).

[48] G. E. Fasshauer, M. J. McCourt: *Kernel-based Approximation Methods Using Matlab*, World Scientific Publishing, Interdisciplinary Mathematical Sciences, Vol. 17, Singapore, (2015).

[49] L. Fernández, T. E. Pérez and M. A. Piñar: *On Koornwinder classical orthogonal polynomials in two variables*, J. Comput. Appl. Math., **236** (2012), 3817–3826.

[50] G. J. Gassner, F. Lörcher, C.-D. Munz and J. S. Hesthaven: *Polymorphic nodal elements and their application in discontinuous Galerkin methods*, J. Comput. Phys., **228** (2009), 1573–1590.

[51] J. W. Gibbs: *Fourier's Series*, Nature, **59** (1898), 200.

[52] R. L. Hardy: *Multiquadric equations of topography and other irregular surfaces*, J. Geophys. Res., **76** (1971), 1905–1915.

[53] N. J. Higham: *The numerical stability of barycentric Lagrange interpolation*, IMA J. Numer. Anal., **24** (2004), 547–556.

[54] K. Hormann, G. Klein and S. De Marchi: *Barycentric rational interpolation at quasi-equidistant nodes*, Dolomites Res. Notes Approx., **5** (2012), 1–6.

[55] E. J. Kansa: *Application of Hardy's multiquadric interpolation to hydrodynamics*, Proceeding Multiconference on Computer Simulation: Aerospace, San Diego (1986), 111–117.

[56] M. Krebsbach, B. Trauzette and A. Calzona: *Optimization of Richardson extrapolation for quantum error mitigation*, preprint on ResearchGate (21 January 2022).

[57] D. Kosloff, H. Tal-Ezer: *A modified Chebyshev pseudospectral method with an $O(N^{-1})$ time step restriction*, J. Comput. Phys., **104** (1993), 457–469.

[58] M. Koushki, E. Jabbari and M. Ahmadinia: *Evaluating RBF methods for solving PDEs using Padua points distribution*, Alexandria Eng. J., **59** (5) (2020), 2999–3018.

[59] O. Landon-Cardinal, L. C. G. Govia and A. A. Clerk: *Quantitative Tomography for Continuous Variable Quantum Systems*, Phys. Rev. Lett., **120** (9) (2018), 090501.

[60] N. T. Lloyd: *Approximation Theory and Approximation Practice*, SIAM, (2013).

[61] G. Mastroianni, D. Occorsio: *Optimal systems of nodes for Lagrange interpolation on bounded intervals. A survey.*, J. Comput. Appl. Math., **134** (1-2) (2001), 325–341.

[62] J. C. Merino: *Lissajous Figures and Chebyshev Polynomials*, College Math. J., **32** (2) (2003), 122–127.

[63] C. R. Morrow, T. N. L. Patterson: *Construction of Algebraic Cubature Rules Using Polynomial Ideal Theory*, SIAM J. Numer. Anal., **15** (5) (1978), 953–976.

[64] Numerical computing with functions: Chebfun. `www.chebfun.org`

[65] R. Pachón, L. N. Trefethen: *Barycentric-Remez algorithms for best polynomial approximation in the chebfun system*, BIT Numer. Math., **49** (2009), 721–741.

[66] D. Poggiali, D. Cecchin, C. Campi and S. De Marchi: *Oversampling errors in multimodal medical imaging are due to the Gibbs effect*, Mathematics, **9** (12) (2021), 1348.

[67] L. Qu: *Copula density estimation by Lagrange interpolation at the Padua points*, Conference on *Data Science, Statistics & Visualization 2017*, Book of abstacts p. 67.

[68] T. Rivlin: *An Introduction to the Approximation of Functions*, Dover Pub. Inc, (1969).

[69] G. Rodeghiero, Y. Zhong et al.: *An efficient interpolation for calculation of the response of composite layered material and its implementation in MUSIC imaging*, Proceedings 19th Conference on the Computation of Electromagnetic Fields COMPUMAG, Budapest (Hungary) (2013).

[70] L. Romani, M. Rossini and D. Schenone: *Edge detection methods based on RBF interpolation*, J. Comput. Applied Math., **349** (2019), 532–547.

[71] C. Runge: *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten*, Zeit. Math. Phys., **46** (1901), 224–243.

[72] B. Schölkopf, A. J. Smola: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Cambridge, MA, USA, (2002).

[73] I. J. Schoenberg: *Metric spaces and completely monotone functions*, Ann. of Math., **39** (1938), 811–841.

[74] L. L. Schumaker: *Spline Functions - Basic Theory*, Wiley-Interscience, New York, (1981).

[75] A. Sommariva, M. Vianello: *Computing approximate Fekete points by QR factorizations of Vandermonde matrices*, Comp. Math. App., **57** (2009), 1324–1336.

[76] A. Sommariva, M. Vianello and R. Zanovello: *Nontensorial Clenshaw–Curtis cubature*, Numer. Algorithms, **49** (2008), 409–427.
[77] I. H. Sloan, R. S. Womersley: *Extremal systems of points and numerical integration on the sphere*. Adv. Comput. Math., **21** (2004), 107–125.
[78] M. A. Taylor, B. A. Wingate and R. E. Vincent: *An algorithm for computing Fekete points in the triangle*, SIAM J. Numer. Anal., **38** (5) (2000), 1707–1720.
[79] P. Vértesi: *On the Lebesgue function and Lebesgue constant: a tribute to Paul Erd́os*, Bolyai Society of Mathematical Studies, Vol. 11, Budapest, Janos Bolyai Math. Soc., (2002), 705–728.
[80] H. Wendland: *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge Univ. Press, (2005).
[81] Wikipedia: Padua points `https://en.wikipedia.org/wiki/Padua_points`
[82] Y. Xu: *Christoffel functions and Fourier series for multivariate orthogonal polynomials*, J. Approx. Theory, **82** (1995), 205–239.
[83] P. Zitnan: *The collocation solution of Poisson problems based on approximate Fekete points*, Eng. Anal. Bound. Elem., **35** (2011) 594–599.

Stefano De Marchi
University of Padova
Department of Mathematics "Tullio Levi-Civita"
35122, Padova PD, Italy
ORCID: 0000-0002-2832-8476
E-mail address: `stefano.demarchi@unipd.it`

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 1, pp. 37-45
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Parameters in Banach spaces and orthogonality

MARCO BARONTI AND PIER LUIGI PAPINI*

ABSTRACT. In Banach spaces, plenty of parameters have been considered: they are often defined by using pairs of vectors. Rarely, they are defined by considering pairs of vectors which are orthogonal in the sense of Birkhoff and James; in that case the study is often not easy. In fact, it can be difficult to identify pairs of orthogonal vectors; so to calculate the value of these parameters, to compare them with the other parameters, to see if they have some stability with respect to changes of the norm. In this paper, we shall do this for a couple of new parameters.

Keywords: Orthogonal vectors, James constant, uniform non squareness.

2020 Mathematics Subject Classification: 46B20, 46B99, 46C15.

1. INTRODUCTION

Let X be a real Banach space; we shall denote by S_X (or simply by S, if no confusion can arise) its unit sphere. As known, it is possible to consider in X several different notions of orthogonality. The most popular and used seems to be the one suggested by Birkhoff and James, that we shall consider here. We say that x is orthogonal to y, and we write $x \perp y$, if $||x|| \leq ||x + ty||$ for every $t \in \mathbb{R}$.
Considering parameters defined by using orthogonal pairs of vectors is not usual (and simple). Among the few attempts done in this direction, we recall that for example I. Serb considered the "orthogonal version" of a modulus of smoothness, indicating only "weak" results (see [6] and the references therein). We give a simple example showing that parameters defined by using orthogonal pairs can hardly be "stable".
Consider X as the space $\mathbb{R}^2$ with the maximum norm: for the vectors $x = (1, 1)$ and $y = (-1, 0)$ we have $x \perp y$. Now, we "slightly" change the norm: for $x = (x_1, x_2)$ we set $||x|| = (|x_1|^p + |x_2|^p)^{1/p}$, with p "large"; then $||x||_p$ is near to 1 but on S_X only $y' = (1, -1)$ is such that $x \perp \pm y'$ and only $x' = (0, 1)$ $(\in S_X)$ is such that $\pm x' \perp y$.
In this paper, we consider "orthogonal versions" of two deeply studied parameters. We prove several facts, giving also new characterizations of uniformly nonsquare spaces; we show by examples that our parameters have different behaviors with respect to the classical ones. We underline how much the new parameters can differ from the corresponding classical ones; everything is explained also by means of numerous examples.

2. OLD AND NEW PARAMETERS

The following parameters received much attention during the last decades (see for example [3]) and are still studied in deep:

Received: 03.02.2022; Accepted: 07.03.2022; Published Online: 09.03.2022
*Corresponding author: Pier Luigi Papini; pierluigi.papini@unibo.it
DOI: 10.33205/cma.1067323

$$J(X) = \sup\{\min\{||x-y||, ||x+y||\} : x, y \in S_X\} \text{ (James constant);}$$

$$g(X) = \inf\{\max\{||x-y||, ||x+y||\} : x, y \in S_X\} \text{ (Schäffer constant).}$$

As known, we always have: $1 \le g(X) \le \sqrt{2} \le J(X) \le 2\,;\, g(X)J(X) = 2$. Also $g(X) = J(X) = \sqrt{2}$ in Hilbert spaces. Recall the definition of uniformly non square spaces, (UNS) for short. The space X is (UNS) when there exists $\epsilon > 0$ such that for $x, y \in S_X$ either $||x-y|| < 2-\epsilon$ or $||x+y|| < 2-\epsilon$. Clearly:

$$X \text{ is } (UNS) \iff J(X) = 2 \iff g(X) = 1. \tag{2.1}$$

We define now:

$$J_\perp(X) = \sup\{\min\{||x-y||, ||x+y||\} : x, y \in S_X,\ x \perp y\};$$

$$g_\perp(X) = \inf\{\max\{||x-y||, ||x+y||\} : x, y \in S_X,\ x \perp y\}.$$

Of course $g(X) \le g_\perp(X)$ and $J_\perp(X) \le J(X)$ always hold. Sometimes we shall simply write J, $J_\perp$, g, $g_\perp$ when it is clear which is the underlying space.
Note that $J(X) = \sup\{J(Y) : Y \subset X,\ dim(Y) = 2\}$; and a similar remark applies for $J_\perp$, g, $g_\perp$. This indicates that studying these parameters for 2-dimensional spaces (where some specific pathologies can also arise) is essentially studying them in general.

3. Studying $g_\perp(X)$

An equivalent formulation for (UNS) is the following: there exists $\epsilon > 0$ such that for all $x, y \in S_X$ either $||x-y|| > 1-\epsilon$ or $||x+y|| > 1-\epsilon$.
Also: reading the proof of [2, Theorem 3.2], we see that the following fact (based on orthogonal vectors) is true:

X is not (UNS) if and only if there exist $x, y \in S_X$, $x \perp y$, such that $||x \pm \lambda y|| \approx 1$ with $\lambda \approx 2$.

Next result gives a sharper result concerning orthogonal pairs.

Theorem 3.1. *Let X be a real Banach space; assume that S_X contains two points x, y such that*

$$||x \pm y|| \le 1+\epsilon,\ \epsilon \in [0,1). \tag{3.2}$$

Then S_X contains y' such that: $x \perp y'$ and $||x \pm y'|| \le \dfrac{1-\epsilon^2+2\epsilon}{1-\epsilon}$.

Proof. Let x, y as indicated; assume that x is not orthogonal to y, thus $\epsilon > 0$ (otherwise there is nothing to prove). Take a norm-one functional f_x such that $f_x(x) = 1$ and $f_x(y) \neq 0$. Eventually exchanging y and $-y$, we can assume that $f_x(y) > 0$. Let $y' = \alpha x + \beta y \in S_X$ be such that $f_x(y') = 0$ (so $x \perp y'$). We have $\beta \neq 0$ (otherwise also $\alpha = 0$ against $y' \in S_X$). Again, we can assume $\beta > 0$ (eventually exchanging y' and $-y'$). Then: $f_x(y') = \alpha + \beta f_x(y) = 0$, so $\alpha = -\beta f_x(y) < 0$. Also, $1 + f_x(y) = f_x(x+y) \le ||x+y|| \le 1+\epsilon$, so $0 \le f_x(y) \le \epsilon$. This implies $|\alpha| \le \beta\epsilon$. We have

$$1 = ||y'|| \ge \beta - |\alpha| \ge \beta(1-\epsilon)$$

thus

$$\beta \le \frac{1}{1-\epsilon};$$

$$1 = ||y'|| \le |\alpha| + \beta \le \beta(1+\epsilon)$$

so we have

$$\frac{1}{1+\epsilon} \le \beta \le \frac{1}{1-\epsilon}.$$

Therefore

$$||y - y'|| = ||y - \alpha x - \beta y|| = ||\alpha x + (\beta - 1)y|| \leq |\alpha| + |\beta - 1|.$$

If $\beta \geq 1$, then

$$||y - y'|| \leq |\alpha| + \beta - 1 \leq \beta(\epsilon + 1) - 1 \leq \frac{2\epsilon}{1 - \epsilon};$$

if $\beta \leq 1$, then

$$||y - y'|| \leq |\alpha| + 1 - \beta \leq \beta(\epsilon - 1) + 1 \leq \frac{\epsilon - 1}{1 + \epsilon} + 1 = \frac{2\epsilon}{1 + \epsilon} < \frac{2\epsilon}{1 - \epsilon}.$$

So we obtain

$$||x - y'|| = ||x - y - y' + y|| \leq ||x - y|| + ||y - y'|| \leq 1 + \epsilon + \frac{2\epsilon}{1 - \epsilon};$$

$$||x + y'|| = ||x + y + y' - y|| \leq ||x + y|| + ||y - y'|| \leq 1 + \epsilon + \frac{2\epsilon}{1 - \epsilon}$$

and so

$$||x \pm y'|| \leq \frac{1 - \epsilon^2 + 2\epsilon}{1 - \epsilon}.$$

Note that the last function of $\epsilon \in [0, 1)$ is increasing. □

By using Theorem 3.1, we can prove the following result:

Theorem 3.2. *Fon any space X, we have*

$$g_\perp(X) \leq \frac{-g^2(X) + 4g(X) - 2}{2 - g(X)}. \tag{3.3}$$

In particular, $g_\perp(X) = 1$ characterizes non (UNS) spaces (in fact $g_\perp(X) = 1$ if and only if $g(X) = 1$ since $1 \leq g(X) \leq g_\perp(X)$ always).

Proof. Take $\epsilon > g(X) - 1$: thus S_X contains pair x, y satisfying (3.2). According to Theorem 3.1 we have

$$g_\perp(X) \leq \frac{1 - \epsilon^2 + 2\epsilon}{1 - \epsilon}.$$

Since this is true for all $\epsilon > g(X) - 1$, we obtain

$$g_\perp(X) \leq \frac{1 - (g(X) - 1)^2 + 2(g(X) - 1)}{1 - (g(X) - 1)} = \frac{-g^2(X) + 4g(X) - 2}{2 - g(X)}.$$

□

In the last theorem, the majorizing function (of $g(X)$) is increasing.

We note that Theorem 3.2 gives an estimate that is not very sharp in general; for example if X is a Hilbert space then $g_\perp(X) = g(X) = \sqrt{2}$ but that estimate gives $g_\perp(X) \leq 2\sqrt{2}$; on the contrary that estimate is "fine" if $g(X) \approx 1$.

4. STUDYING $J_\perp(X)$

We start with a remark concerning $g(X)$ and $J(X)$.

Remark 4.1. *It is not difficult to see that*

$$J(X) = \sup\{\min\{||x-y||, ||x+y||\} : x, y \in S_X;\ ||x-y|| = ||x+y||\};$$

$$g(X) = \inf\{\max\{||x-y||, ||x+y||\} : x, y \in S_X;\ ||x-y|| = ||x+y||\}\,.$$

Proof. : See for example [5] for this and a general discussion of this. □

We prove now a result related to $J(X)$ and $J_\perp(X)$.

Theorem 4.3. *In any space X, we have*

$$J_\perp(X) \geq 2J(X) - 2\,. \tag{4.4}$$

In particular $J(X) = 2$ implies (so it is equivalent to) $J_\perp(X) = 2$, and this condition is equivalent to X being not (UNS).

Proof. According to Remark 4.1, given $\epsilon > 0$ there exist $x, y \in S_X$ such that $||x-y|| = ||x+y|| = \beta$ for some $\beta \in (J(X) - \epsilon, J(X)\,)$. Set $f(t) = ||x + ty||$: this is a convex, 1-Lipschitz function and $f(0) = 1 < \beta = f(1) = f(-1)$. Let $t_0 \in (-1, 1)$ a point, where the function f attains its minimum $\alpha \in (0, 1]$. This means that $x + t_0 y \perp y$. We can suppose $t_0 > 0$ (eventually exchanging y and $-y$); if $t_0 = 0$ then there is nothing to prove. Also, by considering the slope of f in $[0, 1]$ and the fact that f is 1-Lipschitz, we have $\beta - \alpha + 1 - \alpha = f(1) - f(t_o) + f(0) - f(t_0) \leq 1$, so $\alpha \geq \beta/2$. Set $z = (x + t_0 y)/\alpha$ (so $z \in S_X;\ ||z - (x + t_o y)|| = 1 - \alpha;\ z \perp y$). We have

$$||z+y|| = ||z-(x+t_0y)+(x+t_0y)+y|| \geq ||x+(t_0+1)y||-(1-\alpha) = f(t_0+1)-1+\alpha \geq \frac{\beta - \alpha t_0}{1-t_0} - 1 + \alpha.$$

Hence, $||z + y|| \geq \dfrac{\beta - 1 + \alpha + t_0(1 - 2\alpha)}{1 - t_0}$. Since $J(X) \geq \sqrt{2}$ we can assume $\beta > 4/3$; $\alpha \geq \beta/2$ implies $\beta - 1 + \alpha \geq \frac{3}{2}\beta - 1$; $1 - 2\alpha \geq -1$ implies

$$||z + y|| \geq \frac{(3/2)\beta - 1 - t_0}{1 - t_0} > \frac{3}{2}\beta - 1\,.$$

Considering the average slope of f in $[t_0, 1]$, we have $\frac{\beta - \alpha}{1 - t_0} \leq 1$, so $t_0 \leq 1 + \alpha - \beta$. Then, we have $f(t_0 - 1) = ||x + t_0 y - y|| \geq \beta - t_0 \geq 2\beta - \alpha - 1$. Therefore

$$||z - y|| = ||z - (x + t_0 y) + (x + t_0 y) - y|| \geq f(t_0 - 1) - 1 + \alpha \geq 2(\beta - 1).$$

Since $2(\beta - 1) \leq \frac{3}{2}\beta - 1$ (in fact $\beta \leq 2$), we obtain

$$J_\perp(X) \geq \min\{2(\beta - 1);\ \frac{3}{2}\beta - 1\} = 2(\beta - 1).$$

But we can choose $\epsilon > 0$ arbitrarily small, so β can be arbitrarily near to $J(X)$. Then, we obtain the result. □

We note that the inequality (4.4) is "fine" if $J(X) \approx 2$, but it is not sharp in general: for example in Hilbert spaces it only gives $J_\perp(X) \geq 2(\sqrt{2} - 1)$; it gives $J_\perp(X) \geq \sqrt{2}$ if $J(X) \geq 1 + 1/\sqrt{2}$. Of course $g(X) = g_\perp(X)$ and/or $J(X) = J_\perp(X)$ when $g(X)$ and/or $J(X)$ is realized by orthogonal pairs $x, y \in S_X$.

5. Examples

In this section, we collect several examples of 2-dimensional spaces, where we compute the values of our parameters. We recall (see [3]) that the value of $J(X)$ depends on the modulus of convexity, defined for $\epsilon \in [0,2]$ in this way:

$$\delta_X(\epsilon) = \inf\{1 - \frac{||x+y||}{2} : x, y \in S_X; ||x-y|| \geq \epsilon\}.$$

Namely, we have

$$J(X) = \sup\{\epsilon > 0 : \epsilon \leq 2 - \delta_X(\epsilon)\}.$$

Thus if $J(X) < 2$, we have $J(X) = 2 - 2\delta_X(J(X))$. Computing the values of $g_\perp(X)$ and $J_\perp(X)$ is often not very simple: in many cases the calculation is not difficult but rather tedious; due also to this we shall not give all details here. We shall use these examples later, to clearify the behaviour of our parameters and to see which properties of $g(X)$ and $J(X)$ remain true for $g_\perp(X)$ and $J_\perp(X)$.

Example 5.1. *Consider $X = \mathbb{R}^2$ with the norm determined by a regular hexagon whose vertices are $(\pm 1, 0)$; $(\pm 1, \pm 1)$; $(0, \pm 1)$. In other words the norm in X is given by*

$$||(x,y)|| = \begin{cases} \max\{|x|, |y|\}, & xy \geq 0 \\ |x| + |y|, & xy < 0. \end{cases}$$

As known, for this space we have $\delta(\epsilon) = \max\{0, (\epsilon - 1)/2\}$. So, $J(X) = 3/2$ and $g(X) = 4/3$. We can see that $g_\perp(X) = 3/2$ (achieved when $x = (0,1)$; $y = (1, 1/2)$). Therefore $J_\perp(X) \leq J(X) = 3/2$; for $x = (1/2, -1/2)$ and $y = (1,1)$, $x, y \in S_X$, $x \perp y$ we obtain $J_\perp(X) = 3/2$.

Example 5.2. *Let $X = \mathbb{R}^2$ endowed with the norm determined by a different hexagon whose vertices are $(\pm 1, 0)$; $(\mp 1, \pm 1)$; $(\pm 1/2, \pm 1)$. Concerning the modulus of convexity in this space, we have*

$$\delta_X(\epsilon) = \begin{cases} 0, & \epsilon \leq 3/2 \\ (1/2)\epsilon - (3/4), & 3/2 < \epsilon \leq 2. \end{cases}$$

This implies $J(X) = 7/4$ so, $g(X) = 8/7$. Concerning our parameters, we have $J_\perp(X) = 5/3$ (achieved, for example, for $x = (-1/4, 1), y = (1,0)$); $g_\perp(X) = 5/4$ (achieved, for example, for $x = (-1,1)$, $y = (2/3, 2/3)$).

Example 5.3. *Let $X = \mathbb{R}^2$ endowed with the norm determined by a regular octagon whose vertices are $(\pm(\sqrt{2}-1), \pm 1)$, $(\pm(1-\sqrt{2}), \pm 1)$, $(\pm 1, \pm(\sqrt{2}-1))$, $(\pm 1, \pm(1-\sqrt{2}))$. Thus,*

$$||(x,y)|| = \min\left\{ \max\{|x|, |y|\}, \frac{|x|+|y|}{\sqrt{2}} \right\}.$$

As known, in this case we have $g(X) = J(X) = \sqrt{2}$ (consider for example $(1/\sqrt{2}, 1/\sqrt{2})$ and $(-1/\sqrt{2}, 1/\sqrt{2})$; but we observe that $(1/\sqrt{2}, 1/\sqrt{2}) \perp (-1/\sqrt{2}, 1/\sqrt{2})$ and so we have also $g_\perp(X) = J_\perp(X) = \sqrt{2}$.

Example 5.4. *Let $X = \mathbb{R}^2$ endowed with the norm defined by*

$$||(x,y)|| = \begin{cases} \sqrt{x^2+y^2}, & xy \geq 0 \\ |x| + |y|, & xy < 0. \end{cases}$$

We have (see for example [4, p. 280]) $J(X) = \sqrt{8/3}$. Therefore, $g(X) = \sqrt{3/2}$. Concerning $J_\perp(X)$ (in this case the calculation is non trivial), it is slightly smaller than $J(X)$.
In fact, $J_\perp(X) \approx 1,626$ achieved by $(1,0)$ and $(a, \sqrt{1-a^2})$ with $a \approx 0,321$; or by $(0,1)$ and $(a, \sqrt{1-a^2})$ with $a \approx 0.948$.

Example 5.5. *Let $X = \mathbb{R}^2$ endowed with the norm defined by*

$$||(x,y)|| = \begin{cases} \sqrt{x^2+y^2}, & xy \geq 0 \\ \max\{|x|,|y|\}, & xy < 0. \end{cases}$$

As known, $J(X) = 1 + \sqrt{2}/2$; $g(X) = 4 - 2\sqrt{2}$. $J(X)$ is achieved by $(-1,1)$, $(1/\sqrt{2}, 1/\sqrt{2})$, and $(-1,1) \perp (1/\sqrt{2}, 1/\sqrt{2})$. So, $J_\perp(X) = J(X)$. $g(X)$ is not achieved by an orthogonal pair: we should take $(1,-\alpha)$, $(\alpha,1)$ with $\alpha = 3 - \sqrt{8}$ and so $||x \pm y|| = 1 + \alpha$. We obtain $g_\perp(X) = 5/4$ with the orthogonal pair $(-1/2, 1)$ and $(1,0)$. We note that $g_\perp(X)\, J_\perp(X) > 2$.

Example 5.6. *Let $X = \mathbb{R}^2$ endowed with the norm l^p*

$$||(x,y)|| = \begin{cases} (|x|^p + |y|^p)^{1/p}, & p \geq 1 \\ \\ max\{|x|,|y|\}, & p = +\infty. \end{cases}$$

For $p \in \{1, +\infty\}$, X is not (UNS). So, $J(X) = J_\perp(X) = 2$, $g(X) = g_\perp(X) = 1$. Otherwise (see [3]) by using for example the modulus of convexity, we obtain $J(X) = \max\{2^{1/p}, 2^{1-1/p}\}$ and $g(X) = \min\{2^{1/p}, 2^{1-1/p}\}$. Directly these values can be obtained by using Clarkson's inequality and (respectively) the following pairs of orthogonal vectors: $(0,1)$, $(1,0)$ and $(1/2^{1/p}, 1/2^{1/p})$, $(-1/2^{1/p}, 1/2^{1/p})$. Thus, we have $J(X) = J_\perp(X)$ and $g(X) = g_\perp(X)$.

Example 5.7. *Let $X = \mathbb{R}^2$ endowed with the norm defined by*

$$||(x,y)|| = \begin{cases} (|x|^3 + |y|^3)^{1/3}, & xy \geq 0 \\ |x| + |y|, & xy < 0. \end{cases}$$

According to [7], we have $J(X) \approx 1.5573$ and $g(X) = 2/J(X) \approx 1.2843$. For our parameters, we have $J_\perp(X) = J(X)$; $g_\perp(X) \approx 1.2987 > g(X)$. The calculations are not simple. We only indicate here how the modulus of continuity behaves. Clearly $\delta_X(\epsilon) = 0$ if $\epsilon \leq 2^{1/3}$. For $\epsilon > 2^{1/3}$, the graph of δ_X is formed by two segments intersecting (approximately) at $(1.55,\ 0.23)$ (the other extremes being $(1.26, 0)$ and $(2,\ 0.23)$).

6. Comparison of our parameters with the old ones

We collect a few properties concerning with the parameters $g(X)$ and $J(X)$. We always have

(6.5) $$1 \leq g(X) \leq J(X) \leq 2,$$

(6.6) $$g(X)\, J(X) = 2,$$

(6.7) $$g(X) = \sqrt{2} \iff J(X) = \sqrt{2} \iff g(X) = J(X) \iff g(X) = J(X) = \sqrt{2},$$

(6.8) $$X\ is\ (UNS) \iff J(X) < 2 \iff g(X) > 1,$$

(6.9) $$g(X) = J(X) = \sqrt{2} \text{ if X is Hilbert;}$$

the converse of the last statement is true if dim(X) >2, but not in general: see Example 5.3.

(6.10) It may happen that $J(X) \neq J(X^*)$, $g(X) \neq g(X^*)$.

Concerning (6.10), we can consider Example 5.4 and Example 5.5, namely $X = \mathbb{R}^2$ endowed with the norm defined by

$$||(x,y)|| = \begin{cases} \sqrt{x^2+y^2}, & xy \geq 0 \\ |x| + |y|, & xy < 0 \end{cases}$$

and then $X^* = \mathbb{R}^2$ endowed with the norm defined by

$$||(x,y)|| = \begin{cases} \sqrt{x^2+y^2}, & xy \geq 0 \\ \max\{|x|,|y|\}, & xy < 0. \end{cases}$$

Since X is reflexive, we see that passing to the dual the value of these parameters can both increase or decrease.

The examples we have described in the previous section show that some of these properties fail for $g_\perp(X)$ and $J_\perp(X)$. Now, we list the situation with some details.

$$1 \leq g(X) \leq g_\perp(X) \leq J_\perp(X) \leq J(X) \leq 2. \tag{6.11}$$

This chain of inequalities strengthens (6.5); the only non trivial is the central one.
The proof is based on the following result (see [1, Theorem 6.6]).

Theorem 6.4. *In every 2-dimensional normed plane, there exist $x, y \in S_X$ such that $x \perp y$ and $||x-y|| = ||x+y||$.*

Theorem 6.5. *In any space X, we have $g_\perp(X) \leq J_\perp(X)$.*

Proof. It is enough to prove this for 2-dimensional X. Set for $x \in S_X$

$$\alpha_\perp(x) = \inf\left\{\max\{||x \pm y||\} : x \perp y;\, y \in S_X\right\}$$

and

$$\beta_\perp(x) = \sup\{\min\{||x \pm y||\} : x \perp y;\, y \in S_X\}.$$

Of course $g_\perp(X) = \inf\{\alpha_\perp(x) : x \in S_X\}$ and $J_\perp(X) = \sup\{\beta_\perp(x) : x \in S_X\}$.
According to the Theorem 6.4, if dim(X) = 2, then there is a pair $x_0, y_0 \in S_X$ such that $x_0 \perp y_0$ and $||x_0 - y_0|| = ||x_0 + y_0|| = k$. So, we have $\alpha_\perp(x_0) \leq k \leq \beta_\perp(x_0)$. Thus,

$$g_\perp(X) \leq \alpha_\perp(x_0) \leq k \leq \beta_\perp(x_0) \leq J_\perp(X).$$

□

We note that given two different spaces X, Y we always have $g(X) \leq J(Y)$, but our examples show that instead we can have $g_\perp(X) > J_\perp(Y)$.

(6.6): The analogue of (6.6) is not true for our parameters: for example, we have $g_\perp(X)J_\perp(X) > 2$ in Example 5.1.

(6.7): Example 5.2 shows that both $g(X) \neq g_\perp(X)$ and $J(X) \neq J_\perp(X)$ are possible.

(6.8) According to Theorem 3.2 and Theorem 4.3, we see that this result extends to $g_\perp(X)$ and $J_\perp(X)$ giving new characterizations of (UNS) spaces.

(6.9) $g(X) = J(X) = \sqrt{2}$ implies $g_\perp(X) = J_\perp(X) = \sqrt{2}$, so this does not imply that X is Hilbertian (see Example 5.3).

(6.10) We already noticed that the same results hold for our parameters.

7. Bounds concerning the new parameters

We know (see Example 5.2) that we can have $g_\perp(X) = 3/2 > \sqrt{2}$. We can ask how large $g_\perp(X)$ can be in general. Note that $g(X) \leq g_\perp(X) \leq J_\perp(X) \leq J(X) = 2/g(X)$, thus we have

$$g(X)\, g_\perp(X) \leq 2\,, \quad J(X)\, J_\perp(X) \geq 2. \tag{7.12}$$

Consider the first inequality; in general it only gives $g(X) \leq \sqrt{2}$; it says that $g(X) = \sqrt{2}$ implies $g_\perp(X) = \sqrt{2}$. So, the equality (we already noticed this) holds. We know (Theorem 3.2) that

$$g_\perp(X) \leq \frac{-g^2(X) + 4g(X) - 2}{2 - g(X)}\,.$$

The function on the right side increases with $g \in [1, \sqrt{2}]$. Since also $g_\perp(X) \leq 2/g(X)$ (the majorizing function is decreasing with g), we compute when we have

$$\frac{-g^2(X) + 4g(X) - 2}{2 - g(X)} = 2/g(X)\,.$$

This happens for $g(X) \approx 1.194$ and from this, we obtain $g_\perp(X) \leq a \approx 1.675$. We can also estimate

$$g_\perp(X) - g(X) \leq \min\left\{\frac{2g(X) - 2}{2 - g(X)};\ \frac{2}{g(X)} - g(X)\right\};$$

again we have

$$\frac{2g(X) - 2}{2 - g(X)} = \frac{2}{g(X)} - g(X)$$

if and only if $g(X) \approx 1.194$. So, $g_\perp(X) - g(X) \leq b \approx 0.481$.

Consider now $J_\perp(X)$. According to Theorem 4.3, we have

$$J_\perp(X) \geq 2J(X) - 2\,;$$

but also

$$J_\perp(X) \geq \frac{2}{J(X)}\,.$$

The first minorizing function is increasing and the second is decreasing. Moreover,

$$2J(X) - 2 = \frac{2}{J(X)}$$

exactly for

$$J(X) = \frac{1 + \sqrt{5}}{2}\,;$$

so we have

$$J_\perp(X) \geq \sqrt{5} - 1.$$

Also,

$$J(X) - J_\perp(X) \leq \min\left\{J(X) - (2J(X) - 2)\,;\ J(X) - \frac{2}{J(X)}\right\}$$

and since the two majorizing functions coincide for $J(X) = \frac{1+\sqrt{5}}{2}$, we obtain

$$J(X) - J_\perp(X) \leq \frac{3 - \sqrt{5}}{2} \approx 0.382.$$

Also the estimates given in this section seem to be not so sharp; in fact for example $J_\perp(X) \geq 2J(X) - 2$ implies $J_\perp(X) \geq \sqrt{2}$ if $J(X) \geq 1 + 1/\sqrt{2}$, but in our examples we have always $J_\perp(X) \geq \sqrt{2}$.

References

[1] J. Alonso, H. Martini and S. Wu: *On Birkhoff orthogonality and isosceles orthogonality in normed linear spaces*, Aequationes Math., **83** (1-2) (2012), 153–189.

[2] M. Baronti, E. Casini and P. L. Papini: *Revisiting the rectangular constant in Banach spaces*, Bull. Aust. Math. Soc., **105** (1) (2022), 124–133.

[3] E. Casini: *About some parameters of normed linear spaces*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., **80** (1986), 11–15.

[4] M. Kato, L. Maligranda and Y. Takahashi: *On James and Jordan-von Neumann constants and the normal structure coefficient of Banach spaces*, Studia Math., **144** (3) (2001), 275–295.

[5] P. L. Papini: *Constants and symmetries in Banach spaces*, Ann. Univ. Mariae Curie-Skłodowska Sect. A, **56** (2002), 65–76.

[6] I. Şerb: *On a modified modulus of smoothness of a Banach space,* Rev. Anal. Numér. Théor. Approx., **22** (2) (1993), 217–224.

[7] C. S. Yang, H. Y. Li: *The James constant for the ℓ_3-ℓ_1 space,* Acta Math. Sin. (Engl. Ser.), **32** (9) (2016), 1075–1079.

MARCO BARONTI
UNIVERSITÀ DI GENOVA
DEPARTMENT OF MATHEMATICS
VIA DODECANESO 35, 16100 GENOVA, ITALY
ORCID: 0000-0001-8827-4855
E-mail address: baronti@dime.unige.it

PIER LUIGI PAPINI
UNIVERSITY OF BOLOGNA
DEPARTMENT OF MATHEMATICS
VIA MARTUCCI 19, 40136 BOLOGNA, ITALY
ORCID: 0000-0002-2337-7906
E-mail address: pierluigi.papini@unibo.it

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 1, pp. 46-53
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

On matching distance between eigenvalues of unbounded operators

MICHAEL GIL*

ABSTRACT. Let A and $\tilde{A}$ be linear operators on a Banach space having compact resolvents, and let $\lambda_k(A)$ and $\lambda_k(\tilde{A})$ $(k = 1, 2, ...)$ be the eigenvalues taken with their algebraic multiplicities of A and $\tilde{A}$, respectively. Under some conditions, we derive a bound for the quantity

$$\operatorname{md}(A, \tilde{A}) := \inf_{\pi} \sup_{k=1,2,\dots} |\lambda_{\pi(k)}(\tilde{A}) - \lambda_k(A)|,$$

where π is taken over all permutations of the set of all positive integers. That quantity is called the matching optimal distance between the eigenvalues of A and $\tilde{A}$. Applications of the obtained bound to matrix differential operators are also discussed.

Keywords: Banach space, perturbations of eigenvalues, matching distance, differential operator, tensor product of Hilbert spaces.

2020 Mathematics Subject Classification: 47A10, 47A55, 47B10.

1. INTRODUCTION

Let $\mathcal{X}$ be a Banach space with the unit operator $I = I_{\mathcal{X}}$ and norm $\|.\|$. For a linear operator B, $\sigma(B)$ denotes the spectrum, B^{-1} is the inverse operator, $R_z(B) = (B - zI)^{-1}$ $(z \notin \sigma(B))$ is the resolvent, $\|B\|$ is the operator norm, if B is bounded; B^* is the adjoint operator, $D(B)$ is the domain and

$$d(B, z) := \inf_{s \in \sigma(B)} |s - z|, \quad z \in \mathbb{C}.$$

Throughout this paper, A and $\tilde{A}$ are linear operators on $\mathcal{X}$ having compact resolvents. So A and $\tilde{A}$ can have root vectors and all their eigenspaces are finite dimensional.

Let $\lambda_k(A)$ and $\lambda_k(\tilde{A})$ $(k = 1, 2, ...)$ be the eigenvalues of A and $\tilde{A}$, respectively, enumerated with their algebraic multiplicities taken into account. Introduce the following quantity (called the matching optimal distance between the eigenvalues of A and $\tilde{A}$):

$$\operatorname{md}(A, \tilde{A}) := \inf_{\pi} \sup_{i=1,2,\dots} |\lambda_{\pi(i)}(\tilde{A}) - \lambda_i(A)|,$$

where π is taken over all permutations of the set of all positive integers.

Our definition of $\operatorname{md}(A, \tilde{A})$ is a natural generalization of the well-known definition from the perturbation theory of finite matrices [19, p. 167].

The present paper is devoted to estimating $\operatorname{md}(A, \tilde{A})$. The perturbation theory of operators is very rich. The classical results are presented in the book [15], the recent results can be

Received: 20.01.2022; Accepted: 10.03.2022; Published Online: 11.03.2022
*Corresponding author: Michael Gil; gilmi@bezeqint.net
DOI: 10.33205/cma.1060718

found in [1]-[5], [7, 8, 9], [11], [12, 13, 14, 18] and references, which are given therein, but to the best of our knowledge, the matching optimal distance for infinite dimensional operators was not investigated in the available literature although it is important for the localization of the spectrum.

Below we suggest a bound for $\mathrm{md}(A, \tilde{A})$ assuming that

$$D(A) = D(\tilde{A}) \text{ and } q := \|A - \tilde{A}\| < \infty. \tag{1.1}$$

As a particular case, we consider a class of operators on the tensor product of a Hilbert space and a finite dimensional one. We also discuss applications of our results to matrix differential operators.

2. Operators on a Banach space

In the sequel, by $\hat{\lambda}_k(A)$ $(k = 1, 2, ...)$, we denote the distinct eigenvalues of A and assume that

$$r_0(A) := \inf_{j\neq k; j,k=1,2,...} \frac{|\hat{\lambda}_k(A) - \hat{\lambda}_j(A)|}{2} > 0.$$

Since A has a compact resolvent, if $\sigma(A)$ does not contain limit points, one can wait that this condition holds. If $\sigma(A)$ contains limit points, then $r_0(A) = 0$.

Put $r_j := \inf_{k\neq j} \frac{|\hat{\lambda}_k(A)-\hat{\lambda}_j(A)|}{2}$, $j = 1, 2, ...$,

$$\Omega(c, r) := \{z \in \mathbb{C} : |z - c| < r\}, \quad c \in \mathbb{C}, r > 0$$

and

$$C(c, r) := \{z \in \mathbb{C} : |z - c| = r\}, \quad c \in \mathbb{C}, r > 0.$$

By $\nu_k(A)$, we denote the algebraic multiplicity of each $\hat{\lambda}_k(A)$.

Lemma 2.1. *Let conditions (1.1) hold and for an integer j and a positive number $\hat{r}_j \le r_j$, let*

$$q \sup_{z\in C(\hat{\lambda}_j(A),\hat{r}_j)} \|R_z(A)\| < 1. \tag{2.2}$$

Then, $\tilde{A}$ has in $\Omega(\hat{\lambda}_j(A), r_j)$ eigenvalues whose total algebraic multiplicity is equal to $\nu_j(A)$.

Proof. This result is a particular case of the well-known one [15, Theorem IV.3.18]. □

Assume that

$$\|R_\lambda(A)\| \le \phi(1/d(A, \lambda)), \quad \text{for all } \lambda \notin \sigma(A), \tag{2.3}$$

where $\phi(x)$ is a continuous monotonically increasing non-negative function of a non-negative variable x, such that $\phi(0) = 0$ and $\phi(\infty) = \infty$.

Let conditions (1.1) and (2.3) hold, and let there be a positive number $\hat{r}_0 \le r_0(A)$, such that

$$q\phi(1/\hat{r}_0) < 1. \tag{2.4}$$

Then, $\sigma(\tilde{A})$ lies in the set $\cup_{j=1}^{\infty}\Omega(\hat{\lambda}_j(A), \hat{r}_0)$. Indeed, assume that an eigenvalue $\tilde{\lambda}$ of $\tilde{A}$ does not belong to this set. Then for the eigenvalue $\hat{\lambda}_j(A)$ of A nearest to $\tilde{\lambda}$, we have $t = |\tilde{\lambda} - \hat{\lambda}_j(A)| \ge \hat{r}_j$. Thus

$$q\|R_{\tilde{\lambda}}(A)\| \le q\phi(1/t) \le q\phi(1/\hat{r}_0) < 1.$$

According to [15, Theorem IV.1.16], $\tilde{\lambda} \notin \sigma(\tilde{A})$. Hence, due to Lemma 2.1, we arrive at the following result.

Corollary 2.1. *Let conditions* (1.1) *and* (2.3) *hold, and let there be a positive number* $\hat{r}_0 \leq r_0(A)$*, such that inequality* (2.4) *is fulfilled. Then,* $\sigma(\tilde{A})$ *lies in the set* $\cup_{j=1}^{\infty}\Omega(\hat{\lambda}_j(A), \hat{r}_0)$*. Moreover, in each* $\Omega(\hat{\lambda}_j(A), \hat{r}_0)$ $(j = 1, 2, ...)$ *operator* $\tilde{A}$ *has the eigenvalues, whose total algebraic multiplicity is equal to* $\nu_j(A)$*, and therefore* $\mathrm{md}(A, \tilde{A}) \leq \hat{r}_0$*.*

Denote by $x(q)$ the unique positive root of the equation

$$q\phi(1/z) = 1. \tag{2.5}$$

Theorem 2.1. *Let conditions* (1.1) *and* (2.3) *hold, and let* $x(q) < r_0(A)$*. Then*

$$\sigma(\tilde{A}) \subset \cup_{j=1}^{\infty}\Omega(\hat{\lambda}_j(A), x(q)).$$

Moreover, the total algebraic multiplicity of the eigenvalues of $\tilde{A}$*, lying in each* $\Omega(\hat{\lambda}_j(A), x(q))$ $(j = 1, 2, ...)$ *is equal to the algebraic multiplicity* $\nu_j(A)$ *of* $\hat{\lambda}_j(A)$*, and consequently* $\mathrm{md}(A, \tilde{A}) \leq x(q)$*.*

Proof. Since ϕ is an increasing function, for any $\hat{r}_0 \in (x(q), r_0(A))$, we have

$$q\phi(1/\hat{r}_0) < q\phi(1/x(q)) = 1.$$

So, inequality (2.4) is fulfilled. Now, making use of Corollary 2.1, we arrive at the required result. □

3. Operators on the tensor product of a Hilbert space and a finite dimensional one

Throughout this section, $\mathcal{E}$ is a separable Hilbert space with a scalar product $\langle ., . \rangle_{\mathcal{E}}$ and the norm $\|.\|_{\mathcal{E}} = \sqrt{\langle ., . \rangle_{\mathcal{E}}}$, $\mathbb{C}^n$ is the n-dimensional complex Euclidean space with a scalar product $\langle ., . \rangle_n$ and the Euclidean norm $\|.\|_n = \sqrt{\langle ., . \rangle_n}$. Recall the definition of the tensor product $\mathcal{H} = \mathcal{E} \otimes \mathbb{C}^n$ of $\mathcal{E}$ and $\mathbb{C}^n$. To this end, consider the collection of all formal finite sums of the form

$$u = \sum_j y_j \otimes h_j \ \ (y_j \in \mathcal{E}, h_j \in \mathbb{C}^n)$$

with the understanding that

$$\lambda(y \otimes h) = (\lambda y) \otimes h = y \otimes (\lambda h), \ \ (y + y_1) \otimes h = y \otimes h + y_1 \otimes h,$$

$$y \otimes (h + h_1) = y \otimes h + y \otimes h_1, \ \ y, y_1 \in \mathcal{E};\ h, h_1 \in \mathbb{C}^n;\ \lambda \in \mathbb{C}.$$

On that collection define the scalar product as

$$\langle h \otimes y, h_1 \otimes y_1 \rangle_{\mathcal{H}} = \langle y, y_1 \rangle_{\mathcal{E}} \ \langle h, h_1 \rangle_n, \ \ y, y_1 \in \mathcal{E}; h, h_1 \in \mathbb{C}^n$$

and the cross norm is defined by $\|.\|_{\mathcal{H}} = \sqrt{\langle ., . \rangle_{\mathcal{H}}}$. Then, $\mathcal{H}$ is the completion of the considered collection in the norm $\|.\|_{\mathcal{H}}$. Besides, $I_{\mathcal{H}}, I_{\mathcal{E}}$ and I_n are the unit operators in $\mathcal{H}, \mathcal{E}$ and $\mathbb{C}^n$, respectively. From the theory of tensor products, we need only elementary facts which can be found in [6].

Note that the class of operators with compact resolvents is closed under taking the tensor product.

Everywhere below M is an $n \times n$-matrix and S is a normal operator on $\mathcal{E}$ with a compact resolvent. We will consider perturbations of the operator

$$A = S \otimes I_n + I_{\mathcal{E}} \otimes M. \tag{3.6}$$

Let $\hat\lambda_k(M)$ $(k=1,...,m\le n)$ be the distinct eigenvalues of M with the algebraic multiplicities $\nu_k(M)$: $\hat\lambda_k(M)\ne\hat\lambda_j(M)$ $(j\ne k)$ and $\hat\lambda_j(S)$ $(j=1,2,...)$ be the distinct eigenvalues of S with multiplicities $\nu_j(S)$:

$$S=\sum_{j=1}^{\infty}\hat\lambda_j(S)P_j,$$

where P_j are the (mutually orthogonal and finite dimensional) eigen-projections of S. Since

$$I_{\mathcal{E}}=\sum_{k=1}^{\infty}P_k,$$

we have

$$A=\sum_{k=1}^{\infty}\hat\lambda_k(S)P_k\otimes I_n+M\otimes I_{\mathcal{E}}=\sum_{k=1}^{\infty}P_k\otimes(\hat\lambda_k(S)I_n+M).$$

Hence

$$(A-zI_{\mathcal{H}})^{-1}=\sum_{k=1}^{\infty}P_k\otimes((\hat\lambda_k(S)-z)I_n+M)^{-1}$$

and therefore,

$$\|(A-zI_{\mathcal{H}})^{-1}\|=\sup_k\|((\hat\lambda_k(S)-z)I_n+M)^{-1}\|_n. \tag{3.7}$$

Here and below, $\|C\|_n$ means the spectral matrix norm (the operator norm with respect to the Euclidean vector norm) of a matrix C.

Any eigenvalue of A can be written as

$$\hat\lambda_{jk}(A)=\hat\lambda_j(S)+\hat\lambda_k(M),\quad j=1,2...;k=1,...,m.$$

Assume that

$$\begin{aligned}r_0(A)=\inf\{|\hat\lambda_j(S)+\hat\lambda_k(M)-\hat\lambda_{j_1}(S)-\hat\lambda_{k_1}(M)|:\\ j\ne j_1,k\ne k_1;j,j_1=1,2,...;k_1,k=1,...,m\}>0.\end{aligned} \tag{3.8}$$

Denote by $\|M\|_F$ the Frobenius norm of M : $\|M\|_F:=(\text{trace } M^*M)^{1/2}$. The following quantity plays an essential role hereafter:

$$g(M):=[\|M\|_F^2-\sum_{k=1}^{m}\nu_k(M)|\hat\lambda_k(M)|^2]^{1/2}.$$

The following properties of $g(M)$ are checked in [10, Section 3.1]. If M is normal, then $g(M)=0$. In addition,

$$g(e^{it}M+zI_n)=g(M),\quad t\in\mathbb{R};z\in\mathbb{C} \tag{3.9}$$

and

$$g^2(M)\le 2\|M_I\|_F^2\ \ (M_I=(M-M^*)/2i),\text{ and } g^2(M)\le\|M\|_F^2-|\text{trace } M^2|.$$

Due to [10, Theorem 3.2], for any $n\times n$-matrix M, one has

$$\|R_\lambda(M)\|_n\le\sum_{k=0}^{n-1}\frac{g^k(M)}{\sqrt{k!}d^{k+1}(M,\lambda)},\quad \lambda\notin\sigma(M). \tag{3.10}$$

This inequality is sharp: if M is normal, then $g(M)=0$ and with $0^0=1$ (3.10) is attained: $\|R_\lambda(M)\|_n=\frac{1}{d(M,\lambda)}$.

According to (3.7) and (3.10),

$$\|(A - zI_{\mathcal{H}})^{-1}\|_{\mathcal{H}} \le \sup_j \|((\hat\lambda_j(S) - z)I_n + M)^{-1}\|_n$$

$$\le \sup_j \sum_{k=0}^{n-1} \frac{g^k(M)}{\sqrt{k!}d^{k+1}(M, z - \hat\lambda_j(S))}, \quad z - \hat\lambda_j(S) \notin \sigma(M).$$

But

$$d(M, z - \hat\lambda_j(S)) \ge \inf_{j,k} |z - \hat\lambda_{jk}(A)| = d(z, A).$$

Thus

$$\|(A - zI_{\mathcal{H}})^{-1}\|_{\mathcal{H}} \le \sum_{k=0}^{n-1} \frac{g^k(M)}{\sqrt{k!}d^{k+1}(A, z)}, \quad z \notin \sigma(A). \tag{3.11}$$

So, we can take

$$\phi(x) = \sum_{k=0}^{n-1} \frac{g^k(M)x^{k+1}}{\sqrt{k!}}.$$

Besides, equation (2.5) has the form

$$q \sum_{k=0}^{n-1} \frac{g^k(M)}{z^{k+1}\sqrt{k!}} = 1.$$

This equation is equivalent to the equation

$$z^n = q \sum_{k=0}^{n-1} \frac{g^k(A)}{\sqrt{k!}} z^{n-k-1}. \tag{3.12}$$

Now, Theorem 2.1 implies

Theorem 3.2. *Let A be defined by (3.6), condition (3.8) hold and $\tilde A$ be a closed operator on $\mathcal{H}$ satisfying conditions (1.1). Let the unique positive root $y(M, q)$ of (3.12) satisfy the inequality $y(M, q) < r_0(A)$, where $r_0(A)$ is defined by (3.8). Then, $\mathrm{md}(A, \tilde A) \le y(A, q)$.*

If M is normal, then $g(M) = 0$ and with $0^0 = 1$, we have $y(M, q) = q$. Theorem 2.1 gives us the inequality $\mathrm{md}(A, \tilde A) \le q$, provided $q < r_0(A)$.

Now, let M be non-normal: $g(M) \ne 0$. Substitute $z = g(M)w$ into (3.12). We obtain the equation

$$w^n = \frac{q}{g(A)} \sum_{k=0}^{n-1} \frac{1}{\sqrt{k!}} w^{n-k-1}. \tag{3.13}$$

Put

$$p_n = \sum_{j=0}^{n-1} \frac{1}{\sqrt{k!}}.$$

Due to [10, Lemma 3.17], the unique positive root w_0 of equation (3.13) satisfies the inequality

$$w_0 \le \begin{cases} \frac{qp_n}{g(A)} & \text{if } qp_n > g(A), \\ (qp_n/g(A))^{1/n} & \text{if } qp_n \le g(A). \end{cases}$$

But $y(A, q) = w_0 g(A)$. This implies $y(M, q) \le \eta(M, q)$, where

$$\eta(M, q) = \begin{cases} qp_n & \text{if } qp_n > g(M), \\ g^{1-1/n}(M)(qp_n)^{1/n} & \text{if } qp_n \le g(M). \end{cases}$$

Now, Theorem 3.2 yields

Corollary 3.2. *Let A be defined by (3.6), condition (3.8) hold and $\tilde{A}$ be a closed operator on $\mathcal{H}$, such that (1.1) holds. If, in addition, $\eta(M,q) < r_0(A)$, then $\mathrm{md}(A,\tilde{A}) \le \eta(M,q)$.*

Theorem 3.2 is based on the estimate (3.10). If M is diagonalizable, i.e. there is a nonsingular matrix W, such that $W^{-1}MW$ is a normal matrix, then

$$\|R_\lambda(M)\| \le \frac{\kappa}{d(M,\lambda)},$$

where

$$\kappa = \|W^{-1}\|_n \|W\|_n, \quad \lambda \notin \sigma(M).$$

According to (3.7), we obtain

$$\|R_\lambda(A)\| \le \frac{\kappa}{d(A,\lambda)}, \quad \lambda \notin \sigma(A).$$

Equation (2.5) in the considered case takes the form $q\kappa/z = 1$ and thus $x(q) = q\kappa$. So, if M is diagonalizable, then Theorem 2.1 implies

$$\mathrm{md}(A,\tilde{A}) \le q\kappa \text{ provided } q\kappa < r_0(A). \tag{3.14}$$

Some bounds for κ can be found, in particular, in [10, p.105].

4. Differential operators with matrix coefficients

Let $L^2_n = L^2([0,1],\mathbb{C}^n)$ be the space of functions defined on $[0,1]$, with values in $\mathbb{C}^n$ and the scalar product

$$\langle f,h\rangle_{L^2_n} = \int_0^1 \langle f(x),h(x)\rangle_n dx, \quad f,h \in L^2_n.$$

Let $C(x)$ be an $n\times n$-matrix continuously dependent on x. Consider the operators

$$\tilde{A} = -\frac{d^2}{dx^2} + C(x) \tag{4.15}$$

and

$$A = -\frac{d^2}{dx^2} + M, \quad x \in (0,1) \tag{4.16}$$

with a constant $n\times n$-matrix M and the domain

$$D(A) = D(\tilde{A}) = \{u \in L^2_n : u'' \in L^2_n : \ u(0) = u(1) = 0\}.$$

For instance, one can take $M = C(0)$ or $M = \int_0^1 C(x)dx$. Clearly,

$$q = \|A - \tilde{A}\|_{L^2_n} \le \sup_x \|C(x) - M\|_n.$$

Here, $\|A-\tilde{A}\|_{L^2_n}$ is the operator norm in L^2_n of $A-\tilde{A}$.

We have $L^2_n = L^2(0,1)\otimes\mathbb{C}^n$, where $L^2(0,1)$ is the standard complex space of scalar functions. On $D(S) = H^2_0(0,1)$, i.e. on

$$D(S) = \{u \in L^2(0,1) : u'' \in L^2(0,1) : \ u(0) = u(1) = 0\},$$

put $S := -\frac{d^2}{dx^2}$. Since $\hat{\lambda}_j(S) = \pi^2 j^2$ $(j = 1,2,\ldots)$ with $\nu_j(S) = 1$, $\sigma(A)$ consists of the eigenvalues $\lambda_{jk}(A) = \pi^2 j^2 + \hat{\lambda}_k(M)$ $(j = 1,2,\ldots\,; k = 1,\ldots,m)$, and the algebraic multiplicity of $\hat{\lambda}_{jk}(A)$ is equal to $\nu_k(M)$. Let

$$\delta(M) := \inf\{|\pi^2(j^2 - j_1^2) + \hat{\lambda}_k(M) - \hat{\lambda}_{k_1}(M)| :$$

$$j \neq j_1, k_1 \neq k; j, j_1 = 1, 2, ...; k, k_1 = 1, ..., m\} > 0.$$

Then, $r_0(A) = \delta(M) > 0$. Now, Corollary 3.2 yields

Corollary 4.3. *Let $\tilde{A}$ and A be defined by (4.16) and (4.15), $\delta(M) > 0$ and $\eta(M, q) < \delta(M)$. Then,* $\mathrm{md}(A, \tilde{A}) \leq \eta(M, q)$.

In particular, from this corollary, it follows that

$$\sigma(\tilde{A}) \subset \bigcup_{j=1,2,...;\ k=1,...,m} \Omega(\pi^2 j^2 + \hat{\lambda}_k(M), \eta(M, q)),$$

provided $\eta(M, q) < \delta(M)$. If M is diagonalizable, then one can apply inequality (3.14).

For the recent results on the spectra of differential operators see, for instance, the works [16, 17, 20] and the references given therein.

5. Elliptic operators

Let $\omega = [0, 1]^2$ and $L^2(\omega)$ be the space of complex-valued functions defined on ω, with the scalar product

$$\langle f, h \rangle_{L^2(\omega)} = \int_0^1 \int_0^1 f(x, y)\overline{h}(x, y)dxdy, \ \ f, h \in L^2(\omega).$$

Let $c(x, y)$ be a complex continuous function and

$$R := \frac{\partial^2}{\partial x^2} + a\frac{\partial^2}{\partial y^2}, \ \ 0 \leq x, y \leq 1, a \in \mathbb{C}.$$

Consider the operators A and $\tilde{A}$ defined by

$$(\tilde{A}f)(x, y) = (Rf)(x, y) + c(x, y)f(x, y) \tag{5.17}$$

and

$$(Af)(x, y) = (Rf)(x, y) + c_0 f(x, y), \ \ x, y \in (0, 1), f \in D(A) \tag{5.18}$$

with a constant $c_0 \in \mathbb{C}$ and the domain

$$D(A) = \{u \in L^2(\omega) : Ru \in L^2(\omega) : \ u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0; \ 0 \leq x, y \leq 1\}.$$

Clearly,

$$q = \|A - \tilde{A}\|_{L^2(\omega)} \leq \sup_{x,y} |c(x, y) - c_0|.$$

Here, $\|A - \tilde{A}\|_{L^2(\omega)}$ is the operator norm in $L^2(\omega)$ of $A - \tilde{A}$. The eigenfunctions of A are $\sin(\pi jx)\sin(\pi ky)$ and $\sigma(A)$ consists of the simple eigenvalues $\lambda_{jk}(A) = \pi^2(j^2 + ak^2) + c_0$ $(j, k = 1, 2, ...)$. Assume that

$$\delta(R) := \inf\{|\pi^2(j^2 + ak^2 - j_1^2 - ak_1^2)| : j_1 \neq j, k_1 \neq k; j, j_1, k, k_1 = 1, 2, ...\} > 0.$$

Then, $r_0(A) = \delta(R) > 0$. For example, if a is imaginary, then $\delta(R) \geq 3\pi^2(1 + |a|)$. Omitting simple calculations, under consideration, we obtain

$$\|(A - \lambda I)^{-1}\|_{L^2(\omega)} \leq \frac{1}{d(A, \lambda)}.$$

Now, Theorem 2.1 yields

Corollary 5.4. *Let $\tilde{A}$ and A be defined by (5.17) and (5.18), and $\delta(R)>q$. Then,* $\mathrm{md}(A, \tilde{A}) \leq q$.

Similarly, making use of Corollary 3.2, one can consider elliptic operators with matrix coefficients.

REFERENCES

[1] B. Abdelmoumen, A. Jeribi and M. Mnif: *Invariance of the Schechter essential spectrum under polynomially compact operator perturbation*, Extracta Math., **26** (1) (2011), 61–73.

[2] P. Aiena, S. Triolo: *Some perturbation results through localized SVEP*, Acta Sci. Math. (Szeged), **82** (1–2) (2016), 205–219.

[3] A. D. Baranov, D. V. Yakubovich: *Completeness of rank one perturbations of normal operators with lacunary spectrum*, J. Spectr. Theory, **8** (1) (2018), 1–32.

[4] S. Buterin, S.V. Vasiliev: *On uniqueness of recovering the convolution integro-differential operator from the spectrum of its non-smooth one-dimensional perturbation*, Bound. Value Probl., (2018), Paper No. 55, 12 pp.

[5] W. Chaker, A. Jeribi and B. Krichen: *Demicompact linear operators, essential spectrum and some perturbation results*, Math. Nachr., **288** (13) (2015), 1476–1486.

[6] N. Dunford, J.T. Schwartz: *Linear Operators, part I. General Theory*, Wiley Interscience publishers, New York (1966).

[7] M. I. Gil: *Perturbations of operators on tensor products and spectrum localization of matrix differential operators*, J. Appl. Funct. Anal., **3** (3) (2008), 315–332.

[8] M. I. Gil: *Spectral approximations of unbounded non-selfadjoint operators*, Analysis and Mathem. Physics, **3** (1) (2013), 37–44.

[9] M. I. Gil: *Spectral approximations of unbounded operators of the type "normal plus compact"*, Funct. Approximatio. Comment. Math., **51** (1) (2014), 133–140.

[10] M. I. Gil: *Operator Functions and Operator Equations.* World Scientific, New Jersey (2018).

[11] M. I. Gil: *Norm estimates for resolvents of linear operators in a Banach space and spectral variations*, Advances in Operator Theory, **4** (1) (2019), 113–139.

[12] A. Jeribi: *Spectral Theory and Applications of Linear Operators and Block Operator Matrices*, Springer-Verlag, New-York (2015).

[13] A. Jeribi: *Linear Operators and Their Essential Pseudospectra*, CRC Press, Boca Raton (2018).

[14] A. Jeribi: *Perturbation Theory for Linear Operators: Denseness and Bases with Applications*, Springer-Verlag, Singapore (2021).

[15] T. Kato: *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin (1980).

[16] R. Killip: *Perturbations of one-dimensional Schrodinger operators preserving the absolutely continuous spectrum*, Int. Math. Res. Not., **38** (2002), 2029-2061.

[17] R. Ma, H. Wang and M. Elsanosi: *Spectrum of a linear fourth-order differential operator and its applications*, Math. Nachr., **286** (17-18) (2013), 1805-1819.

[18] M. L. Sahari, A. K. Taha and L. Randriamihamison: *A note on the spectrum of diagonal perturbation of weighted shift operator*, Matematiche (Catania), **74** (1) (2019), 35-47.

[19] G. W. Stewart, Ji-guang Sun: *Matrix Perturbation Theory*, Academic Press, New York (1990).

[20] M. Zhang, J. Sun and J. Ao: *The discreteness of spectrum for higher-order differential operators in weighted function spaces*. Bull. Aust. Math. Soc., **86** (3) (2012), 370-376.

MICHAEL GIL
BEN GURION UNIVERSITY OF THE NEGEV
DEPARTMENT OF MATHEMATICS
P.0. BOX 653, BEER-SHEVA 84105, ISRAEL
ORCID: 0000-0002-6404-9618
E-mail address: gilmi@bezeqint.net