MONOGRAFÍAS
DE LA
ACADEMIA
DE
CIENCIAS
Exactas
Físicas
Químicas y
Naturales
DE
ZARAGOZA

Nº 20

Multivariate Approximation and Interpolation with Applications

Mariano GASCA (Editor)

Académico Numerario

2002

An International Workshop on Multivariate Approximation and Interpolation with Applications (MAIA2001) was held in Almuñécar (Spain) during the period of September 10-14, 2001. It was the sixth conference of a series which started in 1986 in Santiago de Chile and was continued in Duisburg, Germany, 1989, Santiago de Chile again in 1992, Montecatini, Italy, in 1995, and Eilat, Israel, in 1998.

Four Spanish Universities organized jointly this Conference: Almería, Granada, Jaén and Zaragoza. The Organizing Committee was formed by R. Carreño and A. Martínez-Finkelstein from Almería, M. Pasadas and V. Ramírez from Granada, D. Cárdenas and M. Muñoz from Jaén and J. Carnicer and M. Gasca from Zaragoza. The Committee wants to remark the great and hard work of M. Pasadas and V. Ramírez in the local arrangements.

The city of Almuñécar, in the region of Costa del Sol close to Granada provided a quiet and fruitful atmosphere to the Workshop, complemented by the pleasant proximity to the beach for the hours after sessions. The terrible events of September 11th in New York City, happened in the middle of the Workshop, gave rise of a strong feeling of solidarity.

There were 60 participants from around the world, invited by the organizers, giving talks in the topics covered by the Conference or contributing with discussions to the scientific success of it. Many of these participants had attended several of the previous Conferences in the series. Two volumes are being published as results of the talks. The Real Academia de Ciencias de Zaragoza agreed to publish the present issue, as one of its Monographs, containing some of those talks. The organizers are indebted to this institution for its offer. Another volume is being published as a special issue of the journal Advances in Computational Mathematics of Kluwer Academic Publishers. Both volumes have been carefully reviewed through a complete refereeing process. J.M. Carnicer provided unselfish help in many technical questions of both editorial processes.

Funding for the Conference was provided by the Universities of Almería, Granada, Jaén and Zaragoza, Ministerio de Educación y Ciencia of Spain, the USA European Office of Aerospace R. and D. (EOARD), Sociedad Española de Matemática Aplicada, Real Academia de Ciencias de Zaragoza, Ibercaja and Banco Santander Central Hispano. The City Hall of Almuñécar provided a welcome party in La Najarra Gardens. The organizers appreciate very much the cooperation of these institutions.

En los días 10 al 14 de setiembre de 2001 tuvo lugar en Almuñécar, Granada, España, un Congreso Internacional sobre Aproximación e Interpolación en Varias Variables y sus Aplicaciones (MAIA2001). Fue el sexto de una serie que comenzó en 1986 en Santiago de Chile y que continuó en 1989 en Duisburg (Alemania), en 1992 en Santiago de Chile de nuevo, en 1995 en Montecatini (Italia) y en 1998 en Eilat (Israel).

Cuatro Universidades españolas, Almería, Granada, Jaén y Zaragoza, organizaron conjuntamente este congreso. El Comité Organizador estuvo formado por R. Carreño y A. Martínez-Finkelstein de Almería, M. Pasadas y V. Ramírez de Granada, D. Cárdenas y M. Muñoz de Jaén and J. Carnicer y M. Gasca de Zaragoza. Debe destacarse el gran trabajo de M. Pasadas y V. Ramírez en el aspecto logístico.

La ciudad de Almuñécar, en la Costa del Sol granadina, proporcionó un tranquilo y fructífero ambiente de trabajo, complementada por la relajante proximidad de la playa de Almuñécar para las horas de descanso. Desgraciadamente los participantes se vieron sorprendidos por los terribles acontecimientos del 11 de setiembre en Nueva York en medio del Congreso, produciéndose un inmediato y fuerte sentido de solidaridad con los damnificados.

Participaron 60 especialistas de todo el mundo, invitados por los organizadores como es habitual en esta serie de congresos, que impartieron charlas sobre los distintos temas objeto de la conferencia o contribuyeron con sus discusiones al éxito científico de ella. Muchos de estos especialistas han participado en varias de las conferencias de la serie

Como resultado de las charlas van a ser publicados dos volúmenes. El que ahora presentamos aquí es uno de ellos. La Real Academia de Ciencias de Zaragoza aceptó publicarlo como el número 20 de su serie de Monografías. Los organizadores agradecen este ofrecimiento. Otro volúmen será publicado como un número especial de la revista Advances in Computational Mathematics por Kluwer Academic Publishers. Ambos volúmenes han sufrido sus correspondientes procesos de selección. Se agradece aquí la generosa ayuda de J. M. Carnicer en ambos procesos editoriales.

Colaboraron en la financiación del congreso las Universidades de Almería, Granada, Jaén y Zaragoza, el Ministerio de Educación y Ciencia español, la USA European Office of Aerospace R. and D. (EOARD), la Sociedad Española de Matemática Aplicada, la Real Academia de Ciencias de Zaragoza y las entidades financieras Ibercaja y Banco Santander Central Hispano. El Ayuntamiento de Almuñécar ofreció una recepción de bienvenida en los preciosos Jardines de La Najarra. Los organizadores agradecen la ayuda de todas estas instituciones.

Contributions

Monografías de la Academia de Ciencias de Zaragoza. **20:** 1–8, (2002).

Una nota sobre aproximación simultánea de funciones*

J. M. Almira, N. Del Toro

Depto. de Matemáticas. Universidad de Jaén. E.U.P. Linares. 23700 Linares (Jaén), Spain

email: `jmalmira@ujaen.es; ndeltoro@ujaen.es`

A. J. López-Moreno

Depto. de Matemáticas. Universidad de Jaén. Las Lagunillas. 23071 Jaén. Spain

email: `ajlopez@ujaen.es`

Abstract

The main purpose of this note is to prove Müntz's type results on simultaneous approximation of functions with polynomials and also with polynomials with integral coefficients. To achieve our goal, we use some variations of Bernstein's polynomials.

Resumen

El objetivo principal de esta nota es demostrar algunos resultados tipo Müntz sobre aproximación simultánea de funciones con polinomios y también con polinomios de coeficientes enteros. Para ello, utilizamos ciertas variaciones de los polinomios de Bernstein.

Clasificación AMS: 41A10, 41A28, 41A29

1 Introducción y primeros resultados

Dada una sucesión de números naturales $\{m(n)\}_{n=1}^{\infty}$, $m(n) \leq n$ $(n \in \mathbb{N})$, queremos estudiar la convergencia simultánea de las sucesiones de polinomios definidas por

$$P_n(f) = \sum_{k=m(n)}^{n} f(k/n) \binom{n}{k} x^k (1-x)^{n-k} \tag{1}$$

---

*Este trabajo ha sido parcialmente financiado por la Junta de Andalucía, Grupo de Investigación "Aproximación y Métodos Numéricos", FQM-0178

y

$$\widetilde{P}_n(f) = \sum_{k=m(n)}^{n} \left[f(k/n)\binom{n}{k} \right] x^k(1-x)^{n-k} \tag{2}$$

a la función $f \in C^{(s)}[0,1]$ y sus derivadas. Veremos que bajo condiciones razonables sobre la sucesión de números naturales $\{m(n)\}_{n=0}^{\infty}$, dicha convergencia se produce sobre ciertos compactos $K \subset [0,1]$. Para motivar este problema, quisiéramos exponer antes algunos resultados que han sido demostrados recientemente por Almira y Luther en [1]. En particular, el Teorema 1 y el Corolario 3 de esta nota aparecen en [1], y el Teorema 2 es una versión más fuerte de otro resultado del mismo artículo. Incluimos sus demostraciones por completitud. Empezamos con un resultado que explica por qué necesitamos imponer ciertas restricciones a la sucesión $\{m(n)\}_{n=1}^{\infty}$, no ya para la aproximación simultánea sino simplemente para la aproximación uniforme de funciones con polinomios $p_n \in \mathbf{span}\{x^{m(n)}, \ldots, x^n\}$.

Teorema 1. *Sea $[a,b]$ un intervalo tal que $0 \notin [a,b]$ y supongamos que para toda función $f \in \mathbf{C}[a,b]$ existe una sucesión de polinomios $\{p_n\}_{n=1}^{\infty}$ tales que $p_n \in \mathbf{span}\{x^{m(n)}, ..., x^n\}$, $(n \geq 1)$ y $\lim_{n\to\infty} \|f - p_n\|_{[a,b]} = 0$. Entonces $\limsup_{n\to\infty} \frac{m(n)}{n} < 1$.*

Demostración. Evidentemente, $\limsup_{n\to\infty} \frac{m(n)}{n} \leq 1$, pues $m(n) \leq n$ para todo n. Si el límite superior fuese igual a 1, entonces existirían números naturales $n_1 < n_2 < \cdots$ tales que $\log \frac{n_k}{m(n_k)-1} \leq 2^{-k}$ para todo $k \geq 1$. Ahora bien, se tiene que la desigualdad

$$\sum_{j=m(n_k)}^{n_k} \frac{1}{j} \leq \sum_{j=m(n_k)}^{n_k} \int_{j-1}^{j} \frac{dx}{x} = \int_{m(n_k)-1}^{n_k} \frac{dx}{x} = \log \frac{n_k}{m(n_k)-1}$$

se satisface para todo $k \geq 1$, de modo que

$$\sum_{k=1}^{\infty} \sum_{j=m(n_k)}^{n_k} \frac{1}{j} \leq \sum_{k=1}^{\infty} \log \frac{n_k}{m(n_k)-1} \leq \sum_{k=1}^{\infty} 2^{-k} < \infty.$$

Se sigue entonces del teorema de Müntz clásico que el espacio vectorial

$$\mathbf{H}_{\{m(n_k)\}} = \mathbf{span}\{x^h : h \in \bigcup_{k=1}^{\infty} \{m(n_k), ..., n_k\}\}$$

no es denso en $\mathbf{C}[a,b]$. Ahora, tomemos una función $f \in \mathbf{C}[a,b]$ que no sea uniformemente aproximable por polinomios de $\mathbf{H}_{\{m(n_k)\}}$. Si $\{p_n\}_{n=1}^{\infty}$ es una sucesión de polinomios tal que $p_n \in \mathbf{span}\{x^{m(n)}, ..., x^n\}$ $(n \geq 1)$ y $\lim_{n\to\infty} ||f - p_n||_{[a,b]} = 0$, entonces $p_{n_k} \in \mathbf{H}_{\{m(n_k)\}}$ para todo k, y $\lim_{n\to\infty} ||f - p_{n_k}||_{[a,b]} = 0$, lo que entra en contradicción con nuestras hipótesis sobre la función f. □

Evidentemente, los polinomios $P_n(f)$ y $\widetilde{P}_n(f)$ definidos por (1) y (2), satisfacen las relaciones $P_n(f), \widetilde{P}_n(f) \in \mathbf{span}\{x^{m(n)}, \dots, x^n\}$. Por tanto, para obtener resultados sobre aproximación uniforme y/o simultánea para dichos polinomios es necesario suponer que $\limsup_{n\to\infty} \frac{m(n)}{n} = c$ para cierta constante $c < 1$. De hecho, podemos utilizar una ligera variación de los polinomios $P_n(f)$ para demostrar que la condición $\limsup_{n\to\infty} \frac{m(n)}{n} = c < \frac{\min\{|a|,|b|\}}{\max\{|a|,|b|\}}$ es suficiente en todos los casos.

Teorema 2. *Supongamos que* $0 < a < b$ *y* $\limsup \frac{m(n)}{n} = c < \frac{a}{b}$. *Entonces toda función* $f \in \mathbf{C}[a,b]$ *se puede aproximar uniformemente en* $[a,b]$ *con una sucesión de polinomios* $\{p_n\}_{n=1}^{\infty}$*, tales que* $p_n \in \mathbf{span}\{x^{m(n)}, \dots, x^n\}$ *para todo* $n \geq 1$*.Además, si* $f \in \mathbf{C}^{(s)}[a,b]$*, el mismo resultado se satisface para la aproximación simultánea de* f *y sus derivadas hasta orden* s *en el intervalo* $[a,b]$*. Finalmente, si* $a < b < 0$ *entonces se tienen resultados análogos si cambiamos* $c < \frac{a}{b}$ *por* $c < \frac{b}{a}$*.*

Demostración. Suponemos, sin pérdida de generalidad, que $0 < a < b$ y $c < \frac{a}{b}$. Dada la función $f \in \mathbf{C}[a,b]$, tomamos $\overline{f}$ una extension de f tal que $\overline{f} \in \mathbf{C}[0,b]$ y $\overline{f}_{|[0,(c+\varepsilon)b]} = 0$ para cierto $\varepsilon > 0$ Evidentemente, esto es posible porque $c < a/b$ implica que $cb < a$. Además, si f es de clase $\mathbf{C}^{(s)}$, entonces podemos suponer también que la extensión $\overline{f}$ es de clase $\mathbf{C}^{(s)}$. Se sigue que la sucesión de polinomios de Bernstein de $\overline{f}$,

$$B_n(\overline{f})_{[0,b]} = \sum_{k=0}^{n} \overline{f}(kb/n)\binom{n}{k}\frac{x^k}{b^k}(1-\frac{x}{b})^{n-k},\ n = 1, 2, \dots$$

converge a $\overline{f}$ (y a sus derivadas de orden $\leq s$, en el caso de que $\overline{f} \in \mathbf{C}^{(s)}[0,b]$) uniformemente en $[0,b]$. Por tanto, se tiene la convergencia (convergencia simultánea, resp.) de $B_n(\overline{f})_{[0,b]}$ a f en $[a,b]$. Ahora bien, como $\limsup \frac{m(n)}{n} = c$ y $\varepsilon > 0$, se tiene que existe $n_0 \in \mathbb{N}$ tal que, si $n > n_0$ entonces $\frac{m(n)b}{n} < (c+\varepsilon)b$. Por tanto, fijado $n > n_0$, se tiene que si $k < m(n)$ entonces $kb/n < (c+\varepsilon)b$, y $\overline{f}(kb/n) = 0$. Esto significa que podemos reescribir los polinomios de Bernstein de $\overline{f}$ como

$$B_n(\overline{f})_{[0,b]} = \sum_{k=m(n)}^{n} \overline{f}(kb/n)\binom{n}{k}\frac{x^k}{b^k}(1-\frac{x}{b})^{n-k},\ n > n_0$$

y, por tanto, $B_n(\overline{f})_{[0,b]} \in \mathbf{span}\{x^{m(n)}, \dots, x^n\}$ para todo $n > n_0$. □

Corolario 3. *Si mantenemos la notación utilizada en la demostración del teorema anterior, y suponemos que* $0 < a < b = 1$*,* $c < a$*,* $f(1) \in \mathbb{Z}$ *entonces la sucesión de polinomios*

$$\widetilde{Q}_n(f) = \sum_{k=m(n)}^{n} \left[\overline{f}(k/n)\binom{n}{k}\right] x^k(1-x)^{n-k} \tag{3}$$

converge uniformemente a la función f *en* $[a,1]$*.*

Demostración. Es conocido que si $\overline{f} \in \mathbf{C}_0[0,1] = \{h \in \mathbf{C}[0,1] : h(0), h(1) \in \mathbb{Z}\}$ entonces los polinomios de Bernstein modificados,

$$\widetilde{B}_n \overline{f}(x) = \sum_{k=0}^{n} \left[\overline{f}(k/n) \binom{n}{k} \right] x^k (1-x)^{n-k}$$

(introducidos por Kantorovich en 1931) convergen a $\overline{f}$ uniformemente en $[0,1]$. Ahora bien, la función $\overline{f}$ introducida en la demostración del Teorema 2 satisface la identidad $\widetilde{Q}_n(f) = \widetilde{B}_n \overline{f}$ para todo $n \geq n_0$. Esto concluye la prueba. $\square$

El interés principal de los resultados anteriores radica en el hecho de que permiten demostrar de forma sencilla un resultado tipo Müntz para la aproximación con polinomios de coeficientes enteros (también llamada aproximación diofántica). A saber: Si $\{k_i\}_{i=0}^{\infty}$ es una sucesión de números naturales que contiene un conjunto de la forma

$$\bigcup_{j=1}^{\infty} \{m(n_j), m(n_j)+1, \ldots, n_j\}$$

para ciertas sucesiones de números naturales $\{n_j\} \to \infty$ y $\{m(n)\}_{n=1}^{\infty} \subset \mathbb{N}$ tal que $\limsup_{n\to\infty} \frac{m(n)}{n} = c < a < 1$, entonces $\mathbb{Z}[x] \cap \mathbf{span}\{x^{k_i}\}_{i=0}^{\infty}$ es denso en $C[a,1]$. Este resultado es un teorema tipo Müntz para aproximación diofántica en $[0,1]$. En realidad, no es el mejor resultado posible, ya que en 1976 Ferguson y Golitscheck [3] demostraron que existe un análogo completo del teorema de Múntz para aproximación con polinomios de $\mathbb{Z}[x]$ en $\mathbf{C}[0,1]$, si los exponentes se toman en $\mathbb{N}$. Sin embargo, el teorema de Golitschek y Ferguson es muy difícil de probar (en contraposición a los que aparecen en este trabajo).

La desventaja, desde nuestro punto de vista, es que para construir los aproximantes de una función $f \in \mathbf{C}[a,b]$, necesitamos antes construir cierta extensión $\overline{f}$ de f, y esto no nos parece natural. Esta es la razón por la que en esta nota nos interesamos por la convergencia de las sucesiones definidas en (1) y (2). La pregunta natural, que pretendemos resolver, es: Si $f \in \mathbf{C}^{(s)}[0,1]$, ¿En qué subconjuntos de $[0,1]$ podemos garantizar la convergencia simultánea a f y sus derivadas de las sucesiones de polinomios $P_n(f)$ y $\widetilde{P}_n(f)$?

2 Resultados principales

Las demostraciones en esta sección están basadas en las propiedades básicas de las funciones analíticas en dominios del plano complejo, así como el uso de los polinomios de Bernstein y sus propiedades de convergencia. En particular, hacemos uso de que si $\{f_n\}_{n=0}^{\infty}$ es una sucesión de funciones analíticas que converge uniformemente sobre compactos en un abierto y conexo $\Omega \subset \mathbb{C}$ a una función f, entonces f es analítica en Ω y la sucesión $\{f_n^{(v)}\}_{n=0}^{\infty}$ converge uniformemente sobre compactos de Ω a la función $f^{(v)}$, para todo $v \geq 0$. De hecho, una consecuencia inmediata de este hecho, es el siguiente resultado técnico:

Lema 4. *Sea $f : [0,1] \to \mathbb{C}$ una función arbitraria. Entonces*

$$\lim_{n\to\infty} \|(B_n f)^{(v)} - (\widetilde{B}_n f)^{(v)}\|_K = 0$$

para todo compacto $K \subset \Omega := \{z \in \mathbb{C} : \max\{|z|, |1-z|\} < 1\}$ y todo número natural $v \geq 0$.

Demostración. Como $(B_n f)^{(v)}$ y $(\widetilde{B}_n f)$ son funciones enteras, sea quien sea la función $f : [0,1] \to \mathbb{C}$, el resultado quedará demostrado si probamos que la sucesión $\{(B_n f)^{(v)} - (\widetilde{B}_n f)^{(v)}\}_{n=0}^{\infty}$ converge a cero uniformemente sobre compactos de Ω. Ahora bien, esto es precisamente lo que se prueba con la siguiente acotación:

$$\begin{aligned} \left| B_n f(z) - \widetilde{B}_n f(z) \right| &\leq \varphi_n(z) := \sum_{k=0}^{n} |z|^k |1-z|^{n-k} \\ &= \begin{cases} \frac{|z|^{n+1} - |1-z|^{n+1}}{|z| - |1-z|} & \text{si } |z| \neq |1-z| \\ (n+1)|z|^n & \text{si } |z| = |1-z| \end{cases} . \end{aligned}$$

□

Corolario 5. *Si $f \in \mathbf{C}^{(s)}[0,1]$, entonces*

$$\lim_{n\to\infty} \|f^{(v)} - (\widetilde{B}_n f)^{(v)}\|_{[a,b]} = 0$$

para todo intervalo $[a,b] \subset (0,1)$ y todo $v \in \{0, 1, \ldots, s\}$.

Demostración. Basta tener en cuenta que los polinomios de Bernstein $B_n f$ convergen a f en la norma de $\mathbf{C}^{(s)}[0,1]$, y el Teorema 4. □

Vamos a centrarnos ahora en nuestro problema. Supongamos que hemos fijado la función $f \in \mathbf{C}^{(s)}[0,1]$ y la sucesión de números naturales $\{m(n)\}_{n=1}^{\infty}$, de modo que $\limsup \frac{m(n)}{n} = c < 1$. Es evidente que el estudio del conjunto de puntos del intervalo $(0,1)$ para los que $P_n(f)^{(v)}$ y $\widetilde{P}_n(f)^{(v)}$ convergen a $f^{(v)}$ para $v \leq s$ es equivalente al estudio del conjunto de puntos de $[0,1]$ donde podemos asegurar que las derivadas $\frac{d^v}{dx^v}(\Delta_n f)(x)$ y $\frac{d^v}{dx^v}(\widetilde{\Delta}_n f)(x)$ de las diferencias

$$\Delta_n f(x) = B_n f - P_n(f) = \sum_{k=0}^{m(n)-1} f(k/n) \binom{n}{k} x^k (1-x)^{n-k}$$

y

$$\widetilde{\Delta}_n f(x) = \widetilde{B}_n f - \widetilde{P}_n(f) = \sum_{k=0}^{m(n)-1} \left[f(k/n) \binom{n}{k} \right] x^k (1-x)^{n-k}$$

convergen a cero para $v \leq s$. Podemos ahora demostrar los resultados principales de esta nota.

Teorema 6. *Supongamos que* $f \in \mathbf{C}^{(s)}[0,1]$ *y* $\{m(n)\}_{n=1}^{\infty} \subset \mathbb{N}$ *de manera que* $\limsup_{n\to\infty} \frac{m(n)}{n} = c < 1$*. Entonces las sucesiones de polinomios*

$$P_n(f) = \sum_{k=m(n)}^{n} f(k/n)\binom{n}{k} x^k(1-x)^{n-k}$$

y

$$\widetilde{P}_n(f) = \sum_{k=m(n)}^{n} \left[f(k/n)\binom{n}{k}\right] x^k(1-x)^{n-k},$$

convergen a f *en la norma de* $\mathbf{C}^{(s)}[a,b]$ *siempre que* $1-2^{-1/(1-c)} < a < b < 1$.

Demostración. Sea $a \in (0,1)$ y supongamos que $z \in K_a := \{z \in \mathbb{C} : |z| \leq 1, |1-z| \leq 1-a\}$. Fijemos $c < \alpha < 1$ y sea $n_0(\alpha) \in \mathbb{N}$ tal que $m(n) < \alpha n$ para todo $n \geq n_0(\alpha)$. Si definimos $M = \|f\|_{[0,1]} + 1$, entonces obtenemos que

$$\begin{aligned}
\max\{|\Delta_n f(z)|, |\widetilde{\Delta}_n f(z)|\} &\leq M \sum_{k=0}^{m(n)-1} \binom{n}{k}|z|^k(1-a)^{n-k} \\
&\leq M(1-a)^{n-m(n)+1} \sum_{k=0}^{m(n)-1} \binom{n}{k}|z|^k \\
&\leq M(1-a)^{n-m(n)+1}(1+|z|)^n \\
&\leq M(1-a)^{n(1-\alpha)}2^n \text{ (pues } m(n) < \alpha n)
\end{aligned}$$

para todo $n \geq n_0(\alpha)$. Ahora bien, $((1-a)^{(1-\alpha)}2)^n$ converge a cero siempre que $(1-a)^{(1-\alpha)}2 < 1$, lo que equivale a decir que $a > 1-2^{-1/(1-\alpha)}$. Si $a > 1-2^{-1/(1-c)}$ entonces existe un α lo bastante cerca de c como para que $a > 1-2^{-1/(1-\alpha)}$. Esto demuestra que las sucesiones $\{\Delta_n f(z)\}_{n=0}^{\infty}$ y $\{\widetilde{\Delta}_n f(z)\}_{n=0}^{\infty}$ convergen a cero uniformemente sobre compactos de $\Omega_c = \{z \in \mathbb{C} : |z| < 1, |1-z| < 2^{-1/(1-c)}\}$. Se sigue, pues, la convergencia a cero de las derivadas de dichas diferencias uniformemente sobre compactos de $(1-2^{-1/(1-c)}, 1)$, lo que completa la demostración. □

En realidad, el teorema anterior no permite afirmar nada para intervalos $[a,b]$ con $0 < a \leq 1/2$, pues $1-2^{-1/(1-c)} \geq 1/2$ para todo $c \in [0,1]$. Sin embargo, parece razonable pensar que si $\limsup_{n\to\infty} \frac{m(n)}{n} = 0$, entonces será posible la aproximación simultánea cerca del origen de coordenadas. El siguiente resultado proporciona una respuesta parcial a dicho problema.

Teorema 7. *Si* $\limsup_{n\to\infty} \frac{m(n)\log_2 n}{n} = c < 1$*, entonces las conclusiones del Teorema* 6 *también se satisfacen en los intervalos* $[a,b]$ *tales que* $1-2^{-c} < a < b < 1$*. En particular, si* $\limsup_{n\to\infty} \frac{m(n)\log_2 n}{n} = 0$*, entonces*

$$\lim_{n\to\infty} \|P_n(f) - f\|_{\mathbf{C}^{(s)}[a,b]} = \lim_{n\to\infty} \|\widetilde{P}_n(f) - f\|_{\mathbf{C}^{(s)}[a,b]} = 0$$

para toda función $f \in \mathbf{C}^{(s)}[0,1]$ *y para todo* $0 < a < b < 1$.

Demostración. podemos suponer, sin pérdida de generalidad, que $1-2^{-c} < a < 1/2$, pues el Teorema 6 garantiza, en el caso que estamos estudiando, la convergencia simultánea sobre compactos de $(1/2,1)$ (ya que por hipótesis $\limsup_{n\to\infty} \frac{m(n)}{n} = 0$). Así pues, sólo tenemos que demostrar la convergencia simultánea en $[a, 1/2+\varepsilon]$ para algún $\varepsilon > 0$. Con este objetivo en mente, definimos $\alpha \in (0,1)$ mediante la identidad $a = 1-2^{-\alpha}$ y tomamos $M = ||f||_{[0,1]} + 1$. Entonces, para todo $z \in K_\alpha^* = \{z \in \mathbb{C} : |z| < 2^{-\alpha}, |1-z| < 2^{-\alpha}\}$, podemos estimar las diferencias $\Delta_n f(z)$ y $\widetilde{\Delta}_n f(z)$ como sigue:

$$
\begin{aligned}
&\max\{|\Delta_n f(z)|, |\widetilde{\Delta}_n f(z)|\} \leq M \sum_{k=0}^{m(n)-1} \binom{n}{k} |z|^k |1-z|^{n-k} \\
&\leq 2^{-\alpha n} M \sum_{k=0}^{m(n)-1} \frac{n^k}{k!} \leq 2^{-\alpha n} n^{m(n)-1} M \sum_{k=0}^{m(n)-1} \frac{1}{k!} \leq 2^{-\alpha n} n^{m(n)-1} M \exp(1) \\
&= n^{-1} \left(2^{-\alpha} n^{m(n)/n}\right)^n M \exp(1) \qquad (4)
\end{aligned}
$$

Ahora bien, como $\limsup \frac{m(n)\log_2 n}{n} = c$, entonces $\frac{m(n)\log_2 n}{n} \leq \alpha$, para todo $n \geq n_0(\alpha)$, para cierto $n_0(\alpha) \in \mathbb{N}$. Se tiene, por tanto, que

$$
2^{\frac{m(n)\log_2 n}{n}} = 2^{\log_2 n^{\frac{m(n)}{n}}} = n^{\frac{m(n)}{n}} \leq 2^\alpha
$$

para todo $n \geq n_0(\alpha)$, lo que implica que la cota superior que hemos estimado en (4) converge a cero para $n \to \infty$. Esto demuestra que las sucesiones $\{\Delta_n f(z)\}_{n=0}^\infty$ y $\{\widetilde{\Delta}_n f(z)\}_{n=0}^\infty$ convergen a cero uniformemente sobre compactos de $\Omega_c^* = \{z \in \mathbb{C} : |z| < 2^{-c}, |1-z| < 2^{-c}\}$. Se sigue, pues, la convergencia a cero de las derivadas de dichas diferencias uniformemente sobre compactos de $(1-2^{-c}, 2^{-c})$, lo que completa la demostración, ya que habíamos reducido el problema a estudiar el caso $2^{-c} > 1/2$. □

3 Notas finales y algunos problemas abiertos

Los resultados que se han expuesto en esta nota admiten una generalización obvia para la aproximación con polinomios de varias variables. Más concretamente, si consideramos los operadores de Bernstein en el cubo multidimensional,

$$
\begin{aligned}
&B_{(n_1,\ldots,n_d)} f(x_1,\ldots,x_d) = \\
&= \sum_{i_1,\ldots,i_d=0}^{n_1,\ldots,n_d} f\left(\frac{i_1}{n_1},\ldots,\frac{i_d}{n_d}\right) \prod_{j=1}^d \binom{n_j}{i_j} \prod_{j=1}^d x_j^{i_j}(1-x_j)^{n_j-i_j}
\end{aligned}
$$

y definimos

$$
\begin{aligned}
&\widetilde{B}_{(n_1,\ldots,n_d)} f(x_1,\ldots,x_d) = \\
&= \sum_{i_1,\ldots,i_d=0}^{n_1,\ldots,n_d} \left[f\left(\frac{i_1}{n_1},\ldots,\frac{i_d}{n_d}\right) \prod_{j=1}^d \binom{n_j}{i_j}\right] \prod_{j=1}^d x_j^{i_j}(1-x_j)^{n_j-i_j},
\end{aligned}
$$

entonces es fácil comprobar que

$$\left|B_{(n_1,\dots,n_d)}f(z_1,\dots,z_d) - \widetilde{B}_{(n_1,\dots,n_d)}f(z_1,\dots,z_d)\right| \leq$$

$$\leq \sum_{i_1,\dots,i_d=0}^{n_1,\dots,n_d} \prod_{j=1}^{d} |z_j|^{i_j}(|1-z_j|^{n_j-i_j} = \prod_{j=1}^{d} \left(\sum_{i_j=0}^{n_j} |z_j|^{i_j}|1-z_j|^{n_j-i_j}\right) = \prod_{j=1}^{d} \varphi_{n_j}(z_j)$$

y, por tanto,

$$B_{(n_1,\dots,n_d)}f(z_1,\dots,z_d) - \widetilde{B}_{(n_1,\dots,n_d)}f(z_1,\dots,z_d)$$

converge uniformemente a cero sobre compactos de Ω^d. Ahora bien, es conocido (ver [4, pag. 12]) que las funciones de varias variables complejas también poseen la propiedad de que si una sucesión $\{f_n\}_{n=0}^{\infty}$ de funciones holomorfas en un abierto W de $\mathbb{C}^d$ converge uniformemente sobre compactos de W a cierta función f, entonces f es holomorfa en W y las derivadas parciales $\left\{\frac{\partial^{v_1+\cdots+v_d} f_n}{\partial z_1^{v_1}\dots\partial z_d^{v_d}}\right\}_{n=0}^{\infty}$ convergen a $\frac{\partial^{v_1+\cdots+v_d} f}{\partial z_1^{v_1}\dots\partial z_d^{v_d}}$ uniformemente sobre compactos de W para todo $v_1,\dots,v_d \geq 0$. Teniendo esto en cuenta, es ahora fácil definir aproximantes que generalizan a los P_n, $\widetilde{P}_n$ y para los que se satisfacen resultados análogos a los Teoremas 6 y 7 de este trabajo. En particular, se sigue que se pueden demostrar resultados tipo Müntz para la aproximación diofántica simultánea de funciones de varias variables en compactos del cubo unitario d-dimensional abierto, incluso eliminando un conjunto infinito de monomios de $\mathbb{Z}[x_1,\dots,x_d]$.

Terminamos esta nota estableciendo los siguientes problemas abiertos:

- Estudiar si la condición $\limsup_{n\to\infty} \frac{m(n)}{n} = 0$ es suficiente para garantizar que las sucesiones de polinomios definidas en (1) y (2) convergen simultáneamente sobre compactos de $(0,1)$ a funciones $f \in \mathbf{C}^{(s)}[0,1]$.
- Estudiar si los intervalos de convergencia que aparecen en el Teorema 6 son óptimos.
- Estudiar los dos puntos anteriores en el caso de varias variables.
- Analizar otras versiones de los operadores multidimensionales de Bernstein definidas en otros dominios con objeto de extender los resultados conseguidos en este trabajo.

Referencias

[1] J. M. Almira and U. Luther, “A note on simultaneous diophantine approximation”, *Applied. Math. E-Notes* **2** (2002) 29-35.

[2] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer (1993).

[3] O. Ferguson and M. V. Golitscheck, “Müntz-Szász Theorem with integral coefficients, II”, *Transactions of the AMS* **213** (1975) 115-126.

[4] L. Kaup and B. Kaup, *Holomorphic Functions of Several Variables*, de Gruyter (1983).

Monografías de la Academia de Ciencias de Zaragoza. **20:** 9–58, (2002).

On Variation–diminishing Schoenberg Operators: New Quantitative Statements *

Laura Beutel, Heiner Gonska, Daniela Kacsó

Inst. of Mathematics, Gerhard Mercator University. D–47048 Duisburg. Germany

and

Gancho Tachev

Dept. of Mathematics, Univ. of Architecture, Civil Eng. and Geodesy.

BG–1421 Sofia. Bulgaria

Abstract

We give quantitative results for variation–diminishing splines, focusing on the case of equidistant knots. New direct inequalities are obtained, both in terms of the classical second modulus of continuity and in terms of the second Ditzian–Totik modulus. These new results are based upon a detailed analysis of the second moments and very recent theorems for positive linear operator approximation. The potential for simultaneous approximation is described by means of an estimate involving both the first and the second classical modulus of continuity. The topic of global smoothness preservation is also addressed. Furthermore, we discuss the degree of simultaneous approximation in the multivariate case, namely for Boolean sums and tensor products of Schoenberg splines.

Keywords: Variation–diminishing splines, degree of approximation, simultaneous approximation, global smoothness preservation, Boolean sums, tensor products.
2000 MSC: 41A15, 41A25, 41A28, 41A36, 41A63, 65D07, 65D17.

1. Introduction

Consider the knot sequence $\Delta_n = \{x_i\}_{-k}^{n+k}$ $(n > 0,\ k > 0)$, with

$$x_{-k} = x_{-k+1} = \ldots = x_0 = 0 < x_1 < \ldots < x_n = \ldots = x_{n+k} = 1.$$

Dedicated to Prof. D.D. Stancu (February 11, 1927) on the occasion of his 75th birthday

For a function $f \in \mathbb{R}^{[0,1]}$, the variation–diminishing spline of degree k w.r.t. Δ_n is given by

$$\begin{aligned} S_{\Delta_n,k}f(x) &:= \sum_{j=-k}^{n-1} f(\xi_{j,k}) \cdot N_{j,k}(x) \quad \text{for } 0 \le x < 1 \text{ and} \\ S_{\Delta_n,k}f(1) &:= \lim_{\substack{y\to 1 \\ y<1}} S_{\Delta_n,k}f(y), \end{aligned}$$

with the nodes (Greville abscissas) $\xi_{j,k} := \dfrac{x_{j+1} + \ldots + x_{j+k}}{k}$, $-k \le j \le n-1$, and the normalized B–splines as fundamental functions

$$N_{j,k}(x) := (x_{j+k+1} - x_j)[x_j, x_{j+1}, \ldots, x_{j+k+1}](\cdot - x)_+^k.$$

This method of approximation was introduced by Schoenberg [72] in 1965 as a "natural" extension of the classical Bernstein polynomial approximation; an important predecessor is a paper by Curry and Schoenberg [18] written in 1945 and completed by 1947, but "for no good reason" not published until 1966. One further key article on the method is one by Marsden and Schoenberg [58] which appeared in Romania, Schoenberg's native country, in 1966. Due to the early work of Marsden [55], [56] on the subject, Schoenberg's variation–diminishing splines (colloquially just denoted as "Schoenberg splines") became known to the mathematical community in the early 1970's and immediately attracted considerable interest. Before continuing these short historical remarks, we list some of their most important properties:

P1) $S_{\Delta_n,k}$ is a positive linear operator which reproduces linear functions, i.e.,

$$\begin{aligned} \sum_{j=-k}^{n-1} N_{j,k}(x) &= 1, \quad 0 \le x \le 1, \\ \sum_{j=-k}^{n-1} \xi_{j,k} \cdot N_{j,k}(x) &= x, \quad 0 \le x \le 1. \end{aligned}$$

P2) **Theorem 1** *(see [55, Theorem 3]) A necessary and sufficient condition that*

$$\lim S_{\Delta_n,k}f(x) = f(x), \ \textit{uniformly in } [0,1]$$

for every $f \in C[0,1]$*, is that*

$$\lim \frac{\|\Delta_n\|}{k} = 0.$$

P3) $S_{\Delta_n,k}$ is a discretely defined operator, which maps $\mathbb{R}^{[0,1]}$ into that subspace of $C^{k-1}[0,1]$ containing all functions which are on each interval $[x_i, x_{i+1}]$ a polynomial of degree at most k;

- P4) Besides of the Bernstein operators, $S_{\Delta_n,k}$ also generalizes piecewise linear interpolation at the knots of Δ_n;
- P5) $S_{\Delta_n,k}$ has the convex hull property and interpolates at the endpoints;
- P6) $S_{\Delta_n,k}$ has the variation–diminishing property, i.e., $V(S_{\Delta_n,k}f - l) \leq V(f - l)$ on $[0,1]$, for all linear functions l, where $V(g)$ denotes the number of sign changes of the function g.

We said before that this method attracted interest in the mathematical community already in the early 1970's. The reader ought to consult the book of DeVore [21], and papers by Leviatan [53], Meyer and Thomas [60], Scherer [71] (for an L_p modification), and by Coman and Frenţiu [14], [15] (for multivariate approaches) in order to confirm our statement. An important contribution from the period 1970–1975 is due to Munteanu and Schumaker [62]. We will cite their article on several occasions in the sequel.

During the late 1970's, the 80's and the 90's further contributions concerning modifications and generalizations of Schoenberg's original method were given, both for the univariate and multivariate cases. With a few exceptions the results given there were of a positive nature. It should not be overlooked, though, that the behaviour in the vicinities of the endpoints 0 and 1 is somewhat poor due to the coalescence of the knots there. We will also discuss this below. Since the present note is not intended to be a survey paper, we have chosen to add several references to the bibliography which are not explicitely cited in the text, but should provide the reader with an idea of the continuing interest among approximation theorists. We make no claim for completeness.

However, this introduction is not yet finished. Schoenberg's variation–diminishing spline operator is in much use in Computer–Aided Geometric Design and has become an indispensible tool there. In CAGD the method has an early history of its own. In his most interesting thesis Riesenfeld [69] introduced Schoenberg splines to the field, having Gordon as his principal advisor. See [46] and [4] in order to confirm that Gordon was always the driving force behind introducing B–spline methods into CAGD at a very early stage of its development. These historical facts seem to be frequently overlooked (or neglected). For more details in regard to their use in CAGD see the books by Farin [26] and by Hoschek and Lasser [47] where more references can be found.

In the present note we will supplement the quantitative information available on Schoenberg's method. In doing so we will in part follow the organization of the Munteanu and Schumaker paper, but also cover further aspects. We will consequently use second order moduli of various types in our assertions. In the late 1960's and early 70's, that is, at the time of writing of the fundamental papers on the subject, these were quantities not too well understood and hardly ever used. The estimates given here are to the most part based upon very recent general results for positive linear and, more generally, convex operators of various orders. This will enable us to also provide new statements on the degree of simultaneous approximation for first and second order derivatives in both the univariate and certain bivariate cases. There will be an emphasis on small explicit constants.

In the foreground of our considerations will mostly be the mesh gauge (rather than the degree) of the splines. Sometimes we will restrict ourselves to the case of equidistant knots $x_j = \frac{j}{n},\ 0 \le j \le n$, because we have not found corresponding statements for the general case which reduce to the "equidistant ones" we are able to give. Throughout this paper we will always denote the $k-$th degree Schoenberg splines with equidistant knots $x_j = \frac{j}{n},\ 0 \le j \le n$, by $S_{n,k}$.

2. The second moments

As for any positive linear operator, the second moments $(S_{\Delta_n,k}(e_1-x)^2)(x),\ x \in [0,1],\ e_i(t) = t^i$ for $i \ge 0$, play an important role for the quantitative behaviour of $S_{\Delta_n,k}$. It is thus instructive to have an idea of where the graph of the function

$$[0,1] \ni x \mapsto (S_{\Delta_n,k}(e_1 - x)^2)(x) \in \mathbb{R}$$

is located. For $\xi_{j,k} \le x \le \xi_{j+1,k}$ we have

$$0 \le (x - \xi_{j,k})(\xi_{j+1,k} - x) \le (S_{\Delta_n,k}(e_1 - x)^2)(x) \le (B_k(e_1 - x)^2)(x) = \frac{x(1-x)}{k}, \qquad (1)$$

$n, k \ge 1$, where B_k is the $k-$th Bernstein operator given by

$$B_k f(x) = \sum_{i=0}^{k} f\left(\frac{i}{k}\right)\binom{k}{i} x^i (1-x)^{k-i},\ x \in [0,1].$$

The second inequality in (1) follows from the fact that, for x fixed, the graph of $S_{\Delta_n,k}(e_1 - x)^2$ lies in the convex hull of its convex control polygon. The third inequality is a consequence of an observation made by Goodman and Sharma [43, Theorem 1], namely that, for a convex function f, one has

$$S_{\Delta_n,k} f(t) \le B_k f(t),\ t \in [0,1].$$

One further exact representation is

$$(S_{\Delta_n,1}(e_1 - x)^2)(x) = (x - x_j)(x_{j+1} - x),\ x \in [x_j, x_{j+1}],\ 0 \le j \le n-1. \qquad (2)$$

In the equidistant case this reduces to

$$(S_{n,1}(e_1 - x)^2)(x) = \frac{\{nx\}(1 - \{nx\})}{n^2},\ x \in [0,1],$$

where $y = [y] + \{y\}$, i.e.,$\{y\}$ is the fractional part of y (see [54]).

We continue to discuss the general case. As shown by DeVore [21]

$$\begin{aligned} 0 &\leq (S_{\Delta_n,k}(e_1 - x)^2)(x) \\ &= \sum_{j=-k}^{n-1} \frac{1}{k^2} \cdot \frac{1}{k-1} \sum_{1 \leq r < s \leq k} (x_{j+r} - x_{j+s})^2 \cdot N_{j,k}(x) \\ &\leq \alpha^2_{\Delta_{n,k}} := \frac{1}{k} \cdot \max_{-k \leq j \leq n} (x_{j+k} - x_j)^2. \end{aligned}$$

The above equation is not very instructive. It was shown by Marsden [56] that one has

$$0 \leq (S_{\Delta_n,k}(e_1 - x)^2)(x) \leq \min \left\{ \frac{1}{2k}, \frac{(k+1)\|\Delta_n\|^2}{12} \right\},\ 0 \leq x \leq 1, \tag{3}$$

where $\|\Delta_n\| := \max_j (x_{j+1} - x_j)$ is the mesh gauge.

However, the upper bound is not a pointwise one. Such pointwise bound is, for example, needed for expressing the fact that one has interpolation at the endpoints.

For the case of equidistant knots we will give such inequalities in this section. We will restrict ourselves first to a discussion of the cases $k \in \{1, 2, 3\},\ n \geq 2$ and present in detail the case $k = 3$.

To this end, we have to estimate the quantity $\dfrac{(S_{n,3}(e_1 - x)^2)(x)}{x(1-x)}$.

For the case $k = 3$ and equidistant knots we get the Greville abscissas

$$\begin{aligned} &\xi_{-3,3} = 0,\ \xi_{-2,3} = \frac{1}{3n},\ \xi_{-1,3} = \frac{1}{n}, \\ &\xi_{j,3} = x_{j+2} = \frac{j+2}{n},\ j = 0, \ldots, n-4, \\ &\xi_{n-3,3} = \frac{n-1}{n},\ \xi_{n-2,3} = 1 - \frac{1}{3n},\ \xi_{n-1,3} = 1. \end{aligned}$$

For $0 \leq x \leq \frac{1}{n}$ we have $S_{n,3}f(x) = \sum_{j=-3}^{0} f(\xi_{j,3}) \cdot N_{j,3}(x)$. The divided differences which we are interested in are equal to

$$\begin{aligned} \left[0,0,0,0,\frac{1}{n}\right](\cdot - t)^3_+ &= n^4 \cdot \left(\frac{1}{n} - t\right)^3_+, \\ \left[0,0,0,\frac{1}{n},\frac{2}{n}\right](\cdot - t)^3_+ &= \frac{n^4}{8} \cdot \left(\frac{2}{n} - t\right)^3_+ - n^4 \cdot \left(\frac{1}{n} - t\right)^3_+, \\ \left[0,0,\frac{1}{n},\frac{2}{n},\frac{3}{n}\right](\cdot - t)^3_+ &= \frac{n^4}{18} \cdot \left(\frac{3}{n} - t\right)^3_+ - \frac{n^4}{4} \cdot \left(\frac{2}{n} - t\right)^3_+ + \frac{n^4}{2} \cdot \left(\frac{1}{n} - t\right)^3_+, \\ \left[0,\frac{1}{n},\frac{2}{n},\frac{3}{n},\frac{4}{n}\right](\cdot - t)^3_+ &= \frac{n^4}{24} \cdot \left(\frac{4}{n} - t\right)^3_+ - \frac{n^4}{6} \cdot \left(\frac{3}{n} - t\right)^3_+ + \frac{n^4}{4} \cdot \left(\frac{2}{n} - t\right)^3_+ - \frac{n^4}{6} \cdot \left(\frac{1}{n} - t\right)^3_+. \end{aligned}$$

For $x \in [0, \frac{1}{n}]$, the first four B–splines of the basis have the form

$$\begin{aligned} N_{-3,3}(x) &= n^3 \cdot \left(\frac{1}{n^3} - \frac{3x}{n^2} + \frac{3x^2}{n} - x^3 \right), \\ N_{-2,3}(x) &= n^3 \cdot \left(\frac{3x}{n^2} - \frac{9x^2}{2n} + \frac{7x^3}{4} \right), \\ N_{-1,3}(x) &= n^3 \cdot \left(\frac{3x^2}{2n} - \frac{11x^3}{12} \right), \\ N_{0,3}(x) &= n^3 \cdot \frac{x^3}{6}. \end{aligned}$$

For $0 \le x \le \frac{1}{n}$ it follows that

$$(S_{n,3}(e_1 - x)^2)(x) = \frac{x}{3n} - \frac{nx^3}{18},$$

whence

$$\frac{(S_{n,3}(e_1 - x)^2)(x)}{x(1-x)} = \frac{\frac{1}{3n} - \frac{nx^2}{18}}{1-x} \le \frac{\frac{1}{3n} - \frac{nx^2}{18}}{1-\frac{1}{n}} = \frac{6 - n^2x^2}{18(n-1)} \le \frac{1}{3(n-1)}. \tag{4}$$

Analogously, for $0 \le x \le \frac{1}{n}$, one can prove that

$$\frac{(S_{n,2}(e_1 - x)^2)(x)}{x(1-x)} = \frac{\frac{1}{2n} - \frac{x}{4}}{1-x} \le \frac{1}{2n} \quad \text{and} \tag{5}$$

$$\frac{(S_{n,1}(e_1 - x)^2)(x)}{x(1-x)} = \frac{-x + \frac{1}{n}}{1-x} \le \frac{1}{n}. \tag{6}$$

Because of the symmetry of the B–spline basis and the symmetry of the function $(e_1 - x)^2$, we also get the above inequalities for $1 - \frac{1}{n} \le x \le 1$. Furthermore, using (3), for arbitrary k and $\frac{1}{n} \le x \le 1 - \frac{1}{n}$ there holds:

$$\frac{(S_{n,k}(e_1 - x)^2)(x)}{x(1-x)} \le \frac{(k+1) \cdot \|\Delta_n\|^2}{12x(1-x)} \le \frac{k+1}{12} \cdot \frac{1}{n^2} \cdot \max_{\frac{1}{n} \le x \le 1 - \frac{1}{n}} \frac{1}{x(1-x)} = \frac{k+1}{12(n-1)}.$$

For $k \in \{1, 2\}$ and $n \ge 2$, one has $\frac{k+1}{12(n-1)} \le \frac{1}{kn}$, thus

$$\frac{(S_{n,k}(e_1 - x)^2)(x)}{x(1-x)} \le \frac{1}{kn}, \text{ for } x \in [0,1].$$

For $k = 3$ and $n \ge 2$, one can only get

$$\frac{(S_{n,3}(e_1 - x)^2)(x)}{x(1-x)} \le \frac{1}{3(n-1)}, \text{ for } x \in [0,1].$$

We will consider next the case of general k and n and recall that $S_{1,k}f(x) \equiv B_kf(x)$, where B_kf is the Bernstein polynomial of f of degree k. In this case we have

$$(B_k(e_1 - x)^2)(x) = \frac{x(1-x)}{k} = \frac{x(1-x)}{n+k-1} = \frac{1}{2} \cdot \frac{\min\{2x(1-x), \frac{k}{n}\}}{n+k-1}.$$

So it is natural to ask whether it is possible to find a constant $A > 0$ as small as possible such that

$$(S_{n,k}(e_1 - x)^2)(x) \leq \frac{A \cdot \min\{2x(1-x), \frac{k}{n}\}}{n+k-1} \tag{7}$$

holds for all natural n and k. We give a positive answer to this question but do not assert to find the optimal value of A in (7). Moreover, we show that it is not possible to find a positive constant B such that

$$\frac{(S_{n,k}(e_1 - x)^2)(x)}{x(1-x)} \leq \frac{B}{k(n-1)} \tag{8}$$

holds for all k and n, although this is fulfilled for $k \in \{1, 2, 3\}$ as it was proved above.

To show (7) we rely on the technique in [57]. In principle, the results presented there for the second moments are correct, but the main tool – Marsden's function $f_2(x, y)$ in [57, p. 1089] – is not uniquely defined in certain cases. The corrections have recently been made by us together with a new proof which will be presented elsewhere. Here we will restrict ourselves to state the correct definitions. Following [57] we denote

$$E_2(x) := (S_{n,k}(e_1 - x)^2)(x) = \sum_{j=-k}^{n-1} \frac{f_2(\xi_{j,k})}{k-1} \cdot N_{j,k}(x), \; k \geq 2, \tag{9}$$

where

$$\begin{aligned} f_2(\xi_{j,k}) &:= \xi_{j,k}^2 - \eta_{j,k}, \\ \eta_{j,k} &:= \binom{k}{2}^{-1} \cdot \sum_{j<i_1<i_2<j+k+1} x_{i_1} \cdot x_{i_2}, \end{aligned}$$

(see [57, p. 1084], where to our notations $\xi_{j,k}$, $\eta_{j,k}$ and $f_2(\xi_{j,k})$ correspond there ξ_j, $\xi_{j,k}$ and $f_2(x, \xi_j)$, respectively; the function f_2 does not actually depend on x).

The crucial representation is

$$g_2(y) := \frac{f_2(y)}{k-1} \text{ for } y \in [0,1], \; k \geq 2, \tag{10}$$

where the corrected form of g_2 (not to be confused with Marsden's g_2) is

$$g_2(y) = \begin{cases} \frac{1}{k-1} \cdot \left(-y^2 + \frac{1}{3} \cdot y\sqrt{8\frac{k}{n} \cdot y + \frac{1}{n^2}}\right), & \text{for } 0 \leq y \leq \min\left\{\frac{k+1}{2n}, \frac{n-1}{2k}\right\}, \\ \frac{1}{k-1} \cdot \left(y - y^2 - \frac{n^2-1}{6nk}\right), & \text{for } \frac{n-1}{2k} \leq y \leq \frac{1}{2}, \\ \frac{1}{k-1} \cdot \frac{(k+1)(k-1)}{12n^2}, & \text{for } \frac{k+1}{2n} \leq y \leq \frac{1}{2}, \\ g_2(1-y), & \text{for } \frac{1}{2} \leq y \leq 1. \end{cases}$$

The function g_2 is continuous on $[0, 1]$.

In order to compute a constant A in (7) we have to consider in the sequel the following three cases (for $n \geq 2$, because for $n = 1$ we get the Bernstein operator, which was already discussed):

Case 1: $k = n - 1$

$$g_2(y) = \begin{cases} \frac{1}{k-1} \cdot \left(-y^2 + \frac{1}{3} \cdot y \sqrt{8\frac{k}{n} \cdot y + \frac{1}{n^2}} \right), & \text{for } 0 \leq y \leq \frac{1}{2}, \\ g_2(1-y), & \text{for } \frac{1}{2} \leq y \leq 1. \end{cases}$$

Case 2: $k < n - 1$

$$g_2(y) = \begin{cases} \frac{1}{k-1} \cdot \left(-y^2 + \frac{1}{3} \cdot y \sqrt{8\frac{k}{n} \cdot y + \frac{1}{n^2}} \right), & \text{for } 0 \leq y \leq \frac{k+1}{2n}, \\ \frac{1}{k-1} \cdot \frac{(k+1)(k-1)}{12n^2}, & \text{for } \frac{k+1}{2n} \leq y \leq \frac{1}{2}, \\ g_2(1-y), & \text{for } \frac{1}{2} \leq y \leq 1. \end{cases}$$

Case 3: $k > n - 1$

$$g_2(y) = \begin{cases} \frac{1}{k-1} \cdot \left(-y^2 + \frac{1}{3} \cdot y \sqrt{8\frac{k}{n} \cdot y + \frac{1}{n^2}} \right), & \text{for } 0 \leq y \leq \frac{n-1}{2k}, \\ \frac{1}{k-1} \cdot \left(y - y^2 - \frac{n^2-1}{6nk} \right), & \text{for } \frac{n-1}{2k} \leq y \leq \frac{1}{2}, \\ g_2(1-y), & \text{for } \frac{1}{2} \leq y \leq 1. \end{cases}$$

We have now the necessary ingredients in order to prove the following

Theorem 2 *For $n \geq 1,\ k \geq 1,\ x \in [0,1]$ we have*

$$(S_{n,k}(e_1 - x)^2)(x) \leq 1 \cdot \frac{\min\{2x(1-x), \frac{k}{n}\}}{n+k-1}.$$

Proof:

For brevity we write again $E_2(x) = (S_{n,k}(e_1 - x)^2)(x)$.

a) $n = 1,\ k \geq 1$. This is the Bernstein operator case in which we have

$$E_2(x) = \frac{1}{2} \cdot \frac{\min\{2x(1-x), \frac{k}{n}\}}{n+k-1}.$$

b) $n \geq 2,\ k = 1$. This is piecewise linear interpolation at $\frac{l}{n},\ 0 \leq l \leq n$. Here (see (2)),

$$E_2(x) = \left(x - \frac{l}{n} \right) \left(\frac{l+1}{n} - x \right) \quad \text{for } x \in \left[\frac{l}{n}, \frac{l+1}{n} \right].$$

For $l = 0$ one has

$$E_2(x) = x \left(\frac{1}{n} - x \right) \leq E_2 \left(\frac{1}{2n} \right) = \frac{1}{4n^2} \leq \frac{1}{2} \cdot \frac{\min\{2x(1-x), \frac{1}{n}\}}{n} \quad \text{for } x \in \left[0, \frac{1}{n} \right].$$

For $1 \le l \le n-2$, i.e., $x \in \left[\frac{1}{n}, \frac{n-1}{n}\right]$, we have

$$E_2(x) \le \frac{1}{4n^2} \le \frac{1}{2n^2} = \frac{1}{2} \cdot \frac{\min\{2x(1-x), \frac{1}{n}\}}{n}.$$

The case $l = n-1$ is symmetric to $l = 0$.

c) $n \ge 2,\ k \in \{2,3\}$. First we observe that

$$\min\left\{2x(1-x), \frac{k}{n}\right\} = \begin{cases} 2x(1-x), & \text{for } 0 \le x \le \frac{1}{n}, \\ 2x(1-x), & \text{for } \frac{1}{n} \le x \le \frac{1}{2} \text{ and } 2 \le n \le 2k, \\ 2x(1-x), & \text{for } \frac{1}{n} \le x \le \frac{k}{n\left(1+\sqrt{1-\frac{2k}{n}}\right)} \text{ and } n \ge 2k+1, \\ \frac{k}{n}, & \text{for } \frac{k}{n\left(1+\sqrt{1-\frac{2k}{n}}\right)} \le x \le \frac{1}{2} \text{ and } n \ge 2k+1. \end{cases}$$

<u>Case 1</u>: $0 \le x \le \frac{1}{n}$

From (4) and (5) it follows

$$\begin{aligned} E_2(x) &\le 2x(1-x) \cdot \frac{1}{kn} \cdot \frac{1}{2(1-x)} \le 2x(1-x) \cdot \frac{1}{kn} \cdot \frac{1}{2(1-\frac{1}{n})} \\ &= \frac{2x(1-x)}{n+k-1} \cdot \frac{n+k-1}{2k(n-1)}. \end{aligned}$$

We need now a constant a such that for all $n \ge 2$ and $k \in \{2,3\}$ there holds

$$\frac{n+k-1}{2k(n-1)} \le a$$

which is equivalent to

$$n \ge \frac{k-1+2ak}{2ak-1}.$$

We impose

$$\frac{k-1+2ak}{2ak-1} = 2,$$

which implies $a = \frac{k+1}{2k}$. Thus we get $a = \frac{3}{4}$ for $k=2$, and $a = \frac{2}{3}$ for $k = 3$, respectively.

<u>Case 2</u>: ($\frac{1}{n} \le x \le \frac{1}{2}$ and $2 \le n \le 2k$) or ($\frac{1}{n} \le x \le \frac{k}{n\left(1+\sqrt{1-\frac{2k}{n}}\right)}$ and $n \ge 2k+1$).

Under these assumptions we can always write

$$\frac{2\frac{1}{n}\left(1-\frac{1}{n}\right)}{n+k-1} \le \frac{2x(1-x)}{n+k-1}.$$

From Marsden's paper [57] we know that

$$E_2(x) \le \frac{k+1}{12n^2}. \tag{11}$$

We need now to find again a constant a such that the inequality

$$\frac{k+1}{12n^2} \leq a \cdot \frac{2(n-1)}{n^2(n+k-1)}$$

holds for all $n \geq 2$ and $k \in \{2,3\}$. We thus get $a = \frac{k^2+2k+1}{24}$, which gives $a = \frac{3}{8}$ for $k = 2$, and $a = \frac{2}{3}$ for $k = 3$.

<u>Case 3</u>: $\frac{k}{n\left(1+\sqrt{1-\frac{2k}{n}}\right)} \leq x \leq \frac{1}{2}$ and $n \geq 2k+1$.
Inequality (11) holds also in this case. We require

$$\frac{k+1}{12n^2} \leq a \cdot \frac{k}{n(n+k-1)}$$

for $n \geq 2$, which leads to $a = \frac{(k+1)^2}{24k}$. Whence we get $a = \frac{3}{16}$ for $k = 2$, and $a = \frac{2}{9}$ for $k = 3$.

Taking the maximum over all the constants a which have been computed, we find that for $n \geq 2$ and $k \in \{2,3\}$

$$E_2(x) \leq \frac{3}{4} \cdot \frac{\min\{2x(1-x), \frac{k}{n}\}}{n+k-1}.$$

d) $n \geq 2,\ k \geq 4$. Here we proceed differently using the function g_2 from above. It is our aim to show that in this case we have

$$g_2(y) \leq h_2(y) := \frac{\min\{2y(1-y), \frac{k}{n}\}}{n+k-1},\ y \in [0,1]. \tag{12}$$

Since $h_2(y)$ is a concave function, one has

$$S_{n,k}(h_2(\cdot);y) \leq h_2(y), \quad y \in [0,1].$$

Due to the positivity of $S_{n,k}$ we also have

$$S_{n,k}(g_2(\cdot);y) \leq S_{n,k}(h_2(\cdot);y), \quad \text{or}$$

$$\sum_{j=-k}^{n-1} \frac{f_2(\xi_{j,k})}{k-1} \cdot N_{j,k}(y) \leq h_2(y) \quad \text{for all } y \in [0,1]. \tag{13}$$

Setting $y = x$ and combining (9), (12) and (13) then shows that for $n \geq 2$ and $k \geq 4$

$$E_2(x) \leq 1 \cdot \frac{\min\{2x(1-x), \frac{k}{n}\}}{n+k-1}.$$

Hence it remains to prove (12).

<u>Case $k = n-1$</u>:

We may consider $n \geq 5$, because $n < 5$ (that is $k \leq 3$) was considered already. Since

$$8\frac{k}{n}y + \frac{1}{n^2} \leq 4\frac{n-1}{n} + \frac{1}{n^2} = \left(2 - \frac{1}{n}\right)^2$$

we obtain

$$g_2(y) \;\leq\; \frac{1}{k-1}\left[-y^2 + \frac{y}{3}\cdot\left(2-\frac{1}{n}\right)\right] \leq \frac{y(1-y)}{3(k-1)}\left(2-\frac{1}{n}\right) = \frac{y(1-y)}{n-1}\cdot a(n),$$

where

$$a(n) := \frac{n-1}{n-2}\cdot\frac{1}{3}\cdot\left(2-\frac{1}{n}\right) \leq a(5) = \frac{4}{5}.$$

Finally it follows

$$g_2(y) \leq \frac{4}{5}\cdot\frac{y(1-y)}{n-1} = \frac{4}{5}\cdot\frac{\min\{2y(1-y),\frac{k}{n}\}}{n+k-1} \tag{14}$$

for all $y \in [0,1]$.

Case $k < n-1$:

We may consider $n \geq 6$, because $k \geq 4$.

For $0 \leq y \leq \frac{k+1}{2n}$ we obtain successively

$$\begin{aligned} g_2(y) \;&\leq\; \frac{1}{k-1}\left(-y^2 + \frac{y}{3}\sqrt{8\frac{k}{n}\cdot\frac{k+1}{2n}+\frac{1}{n^2}}\right) = \frac{1}{k-1}\left(-y^2+\frac{y}{3}\cdot\frac{2k+1}{n}\right) \\ &\leq\; \frac{y(1-y)}{k-1}\cdot\frac{2k+1}{3n} = \frac{2y(1-y)}{n+k-1}\cdot\left[\left(\frac{n}{k-1}+1\right)\cdot\frac{2k+1}{6n}\right] \\ &\leq\; \frac{2y(1-y)}{n+k-1}\cdot\left[\frac{1}{3}+\frac{1}{2(k-1)}+\frac{2n-1}{6n}\right] \leq \frac{29}{36}\cdot\frac{2y(1-y)}{n+k-1}. \end{aligned}$$

For $0 \leq y \leq \frac{k+1}{2n}$ we also have

$$g_2(y) \;\leq\; \frac{y(1-y)}{k-1}\cdot\frac{2k+1}{3n} \leq \frac{35}{64}\cdot\frac{k}{n(n+k-1)}.$$

We conclude now that for $0 \leq y \leq \frac{k+1}{2n}$ one has

$$g_2(y) \leq \max\left\{\frac{29}{36},\frac{35}{64}\right\}\cdot\frac{\min\{2y(1-y),\frac{k}{n}\}}{n+k-1} = \frac{29}{36}\cdot\frac{\min\{2y(1-y),\frac{k}{n}\}}{n+k-1}. \tag{15}$$

Further we consider $y \in [\frac{k+1}{2n},\frac{1}{2}]$ and observe that for the continuous function g_2 we can write

$$g_2(y) \;=\; g_2\left(\frac{k+1}{2n}\right) = \lim_{z\nearrow\frac{k+1}{2n}} g_2(z) \leq \lim_{z\nearrow\frac{k+1}{2n}} \frac{29}{36}\cdot\frac{\min\{2y(1-y),\frac{k}{n}\}}{n+k-1},$$

for all $y \in [\frac{k+1}{2n},\frac{1}{2}]$.

Hence (15) holds for $y \in [0,1]$.

Case $k > n-1$:
It follows immediately that $\min\{2y(1-y), \frac{k}{n}\} = 2y(1-y)$.
For $0 \le y \le \frac{n-1}{2k}$ we obtain successively

$$\begin{aligned} g_2(y) &\le \frac{1}{k-1}\left(-y^2 + \frac{y}{3}\sqrt{8\frac{k}{n}\cdot\frac{n-1}{2k} + \frac{1}{n^2}}\right) \\ &= \frac{1}{k-1}\left[-y^2 + \frac{y}{3}\left(2 - \frac{1}{n}\right)\right] \le \frac{y(1-y)}{k-1}\frac{1}{3}\left(2-\frac{1}{n}\right) \\ &= \frac{2y(1-y)}{n+k-1}\cdot\frac{1}{6}\left(2-\frac{1}{n}\right)\left(1 + \frac{n}{k-1}\right) \\ &\le \frac{7}{9}\cdot\frac{2y(1-y)}{n+k-1}. \end{aligned}$$

If $y \in [\frac{n-1}{2k}, \frac{1}{2}]$ then we have

$$g_2(y) = \frac{1}{k-1}\left(y - y^2 - \frac{n^2-1}{6nk}\right).$$

We have to show that

$$\frac{1}{k-1}\left(y - y^2 - \frac{n^2-1}{6nk}\right) \le \frac{2y(1-y)}{n+k-1},$$

the latter being equivalent to

$$\frac{n-k+1}{n+k-1}\cdot y(1-y) \le \frac{n^2-1}{6nk}.$$

For $y \in [\frac{n-1}{2k}, \frac{1}{2}]$ the left hand side does not exceed

$$\frac{n-k+1}{4(n+k-1)} \le \frac{n^2-1}{6nk} \quad \text{for } n \ge 2 \text{ and } k > n-1.$$

Thus

$$g_2(y) \le 1\cdot\frac{2y(1-y)}{n+k-1} \tag{16}$$

for all $y \in [0,1]$, and the proof of Theorem 2 is complete. □

Remark 3 *We prove here that (8) is not possible. Suppose (8) holds for some $B > 0$. From [57, Theorem 2] we get*

$$\lim_{(n+k+1)\to\infty}(n+k+1)(S_{n,k}(e_1-x)^2)(x) = \frac{1}{12}t(t+1),$$

for $t := \lim_{(n+k+1)\to\infty}\frac{k}{n}$, $0 \le t \le 1$ *and* $\frac{t}{2} \le x \le 1 - \frac{t}{2}$.

For $k = n$ we obtain $t = 1$ and $x = \frac{1}{2}$. In this case one gets

$$\frac{B}{4} \cdot \lim_{(n+k+1)\to\infty} \left[\frac{n+k+1}{k(n-1)}\right] = \frac{B}{4} \cdot \lim_{n\to\infty} \left[\frac{2n+1}{n(n-1)}\right] \geq \frac{1}{12} \cdot 2 = \frac{1}{6}.$$

The latter inequality is not true. $\square$

3. New direct inequalities

In this section we prove direct inequalities for arbitrary functions in $C[0,1]$.

3.1 Uniform estimates in terms of the classical second order modulus of smoothness

While there are many estimates in terms of the first order modulus of smoothness available in the literature – starting with the ones by Marsden and Schoenberg ([72], [58], [55], [56]) and by Munteanu and Schumaker [62] – the first estimates with ω_2 were given by Esser in [25] and later further improved by Gonska [31]. One advantage of the use of ω_2 is the fact that this quantity annihilates linear functions. The desirability to have estimates in terms of such a quantity was already observed at the end of the paper by Marsden and Schoenberg [58] where

$$\omega_1^*(f;\delta) := \inf_{c\in\mathbb{R}} \omega_1(f - ce_1;\delta)$$

was used. As was noted by one of the present authors in [33, p. 17] there is no constant $c > 0$ such that

$$\omega_1^*(f;\delta) \leq c \cdot \omega_2(f;\delta) \text{ for all } f \in C[0,1] \text{ and all } \delta > 0.$$

The following elegant general result of Păltănea is the key for our subsequent applications to Schoenberg splines.

Lemma 4 *(see [68, Corollary 3.1]) Let $K = [a,b]$ be a compact interval of the real axis and K' a compact subinterval of K. If $L : C(K) \to C(K')$ is a positive linear operator, then for $f \in C(K)$, $x \in K'$, and each $0 < h \leq \frac{1}{2}$length(K), the following holds:*

$$\begin{aligned} |(Lf)(x) - f(x)| \;&\leq\; |(Le_0)(x) - 1| \cdot |f(x)| + \frac{1}{h} \cdot |(L(e_1 - x))(x)| \cdot \omega_1(f;h) \qquad (17)\\ &+\; \left[(Le_0)(x) + \frac{1}{2h^2} \cdot (L(e_1 - x)^2)(x)\right] \cdot\; \omega_2(f;h). \end{aligned}$$

Remark 5 *Condition $h \leq \frac{1}{2} \cdot$ length(K) in the above can be eliminated for operators which preserve linear functions.*

Thus we can state

Theorem 6 *For all $f \in C[0,1]$, $x \in [0,1]$ and $h > 0$, there holds*

$$|S_{\Delta_n,k}f(x) - f(x)| \leq \left(1 + \frac{1}{2h^2} \cdot \min\left\{\frac{1}{2k}, \frac{(k+1)\|\Delta_n\|^2}{12}\right\}\right) \cdot \omega_2(f;h). \tag{18}$$

Proof:

Applying (17) and taking into account that $S_{\Delta_n,k}$ reproduces linear functions yields

$$|S_{\Delta_n,k}f(x) - f(x)| \leq \left[1 + \frac{1}{2h^2} \cdot (S_{\Delta_n,k}(e_1 - x)^2)(x)\right] \cdot \omega_2(f;h).$$

Since

$$(S_{\Delta_n,k}(e_1 - x)^2)(x) \leq \min\left\{\frac{1}{2k}, \frac{(k+1)\|\Delta_n\|^2}{12}\right\},$$

the statement of our theorem follows. □

To achieve the goal of this subsection there are two meaningful choices for the parameter h in (18), namely in terms of the degree k and in terms of the mesh gauge $\|\Delta_n\|$. A direct application of Theorem 6 yields in these cases:

Corollary 7 *For all $f \in C[0,1]$, $x \in [0,1]$, one has the following uniform estimates*

$$\|S_{\Delta_n,k}f - f\|_\infty \leq \frac{5}{4} \cdot \omega_2\left(f; \frac{1}{\sqrt{k}}\right), \text{ and} \tag{19}$$

$$\|S_{\Delta_n,k}f - f\|_\infty \leq \left(1 + \frac{k+1}{24}\right) \cdot \omega_2\left(f; \|\Delta_n\|\right). \tag{20}$$

Remark 8 *From (19) and (20), using the properties of the moduli, one gets for $f \in C^1[0,1]$, $x \in [0,1]$, that*

$$\|S_{\Delta_n,k}f - f\|_\infty \leq \frac{5}{4\sqrt{k}} \cdot \omega_1\left(f'; \frac{1}{\sqrt{k}}\right), \text{ and} \tag{21}$$

$$\|S_{\Delta_n,k}f - f\|_\infty \leq \left(1 + \frac{k+1}{24}\right) \cdot \|\Delta_n\| \cdot \omega_1\left(f'; \|\Delta_n\|\right), \tag{22}$$

respectively.

We listed inequalities (21) and (22) here, because they improve the corresponding ones by Munteanu and Schumaker [62, (2.19) and (2.18), respectively] (the second one, however, only for $k \geq 2$).

The inequality

$$(S_{\Delta_n,k}(e_1 - x)^2)(x) \leq \min\left\{\frac{1}{2k}, \frac{(k+1)\|\Delta_n\|^2}{12}\right\}$$

is not quite satisfactory because it does not reflect the fact that

$$(S_{\Delta_n,k}(e_1 - x)^2)(x) = 0 \text{ for } x \in \{0,1\}.$$

For the case of equidistant knots the situation is different as we showed in Section 2. The pointwise inequalities from there will be employed in Sections 3.2 and 3.3.

3.2 Uniform estimates in terms of the Ditzian–Totik modulus of smoothness

In the sequel we use the following particular case of a very recent result by Gonska and Păltănea:

Lemma 9 *(see [37]) If $L : C[0,1] \to C[0,1]$ is a linear positive operator reproducing linear functions, then we have*

$$|L(f,x) - f(x)| \leq \left[1 + \frac{7}{4} \cdot \frac{(L(e_1 - x)^2)(x)}{(h\varphi(x))^2}\right] \omega_2^\varphi(f;h), \tag{23}$$

for all $f \in C[0,1]$, $x \in (0,1)$ and $h \in (0,1]$.

Here

$$\omega_2^\varphi(f;h) = \sup\{|\Delta^2_{\rho\varphi(x)} f(x)|, \ \ x \pm \rho\varphi(x) \in [0,1], \ \ 0 < \rho \leq h\}$$

is the second order Ditzian–Totik modulus, with $\varphi(x) = \sqrt{x(1-x)}$ and $\Delta^2_\eta f(y) = f(y - \eta) - 2f(y) + f(y+\eta)$, if $\eta > 0$, $y \pm \eta \in [0,1]$, $f \in \mathbb{R}^{[0,1]}$.

Applying Lemma 9 we first consider the three cases in which we have an exact representation of $(S_{n,k}(e_1 - x)^2)(x)$ close to the endpoints.

Theorem 10 *For all $f \in C[0,1]$, $x \in [0,1]$, $h \in (0,1]$ and $n \geq 2$, one has*

$$\begin{aligned} |S_{n,3}f(x) - f(x)| &\leq \left[1 + \frac{7}{4} \cdot \frac{1}{h^2} \cdot \frac{1}{3(n-1)}\right] \omega_2^\varphi(f;h), \text{ and} \\ |S_{n,k}f(x) - f(x)| &\leq \left[1 + \frac{7}{4} \cdot \frac{1}{h^2} \cdot \frac{1}{kn}\right] \omega_2^\varphi(f;h), \text{ for } k \in \{1,2\}. \end{aligned}$$

Thus it follows immediately

Corollary 11

$$\begin{aligned} \|S_{n,3}f - f\|_\infty &\leq \frac{19}{12} \cdot \omega_2^\varphi\left(f; \frac{1}{\sqrt{n-1}}\right), \\ \|S_{n,2}f - f\|_\infty &\leq \frac{15}{8} \cdot \omega_2^\varphi\left(f; \frac{1}{\sqrt{n}}\right), \text{ and} \\ \|S_{n,1}f - f\|_\infty &\leq \frac{11}{4} \cdot \omega_2^\varphi\left(f; \frac{1}{\sqrt{n}}\right). \end{aligned}$$

In the general case we apply again Lemma 9 and use Theorem 2.

Theorem 12 *For all $f \in C[0,1]$, $x \in [0,1]$, $h \in (0,1]$ and $n, k \geq 1$ one has:*

i) If $\frac{k}{n} \geq \frac{1}{2}$, then

$$|S_{n,k}f(x) - f(x)| \leq \left[1 + \frac{7}{2} \cdot \frac{1}{h^2(n+k-1)}\right] \cdot \omega_2^\varphi(f;h).$$

ii) If $\frac{k}{n} < \frac{1}{2}$, then

$$|S_{n,k}f(x) - f(x)| \leq \left[1 + \frac{7}{4} \cdot \frac{k}{h^2 \cdot n(n+k-1)}\right] \cdot \omega_2^{\varphi}(f; h).$$

In particular we get

Corollary 13 *i) If $\frac{k}{n} \geq \frac{1}{2}$, then*

$$\begin{aligned}
\|S_{n,k}f - f\|_\infty &\leq \left[1 + \frac{7}{2} \cdot \frac{n}{n+k-1}\right] \cdot \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{n}}\right), \\
\|S_{n,k}f - f\|_\infty &\leq \left[1 + \frac{7}{2} \cdot \frac{k}{n+k-1}\right] \cdot \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{k}}\right), \\
\|S_{n,k}f - f\|_\infty &\leq \frac{9}{2} \cdot \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{n+k-1}}\right).
\end{aligned}$$

ii) If $\frac{k}{n} < \frac{1}{2}$, then

$$\begin{aligned}
\|S_{n,k}f - f\|_\infty &\leq \left[1 + \frac{7}{4} \cdot \frac{k}{n+k-1}\right] \cdot \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{n}}\right), \\
\|S_{n,k}f - f\|_\infty &\leq \left[1 + \frac{7}{4} \cdot \frac{k^2}{n(n+k-1)}\right] \cdot \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{k}}\right), \\
\|S_{n,k}f - f\|_\infty &\leq \frac{11}{4} \cdot \omega_2^{\varphi}\left(f; \sqrt{\frac{k}{n(n+k-1)}}\right).
\end{aligned}$$

Remark 14 *We recall here that Schoenberg's original intention was to introduce a natural "spline extension" of the Bernstein polynomials. This was definitely achieved. Since then impressive progress was made in the investigation of Bernstein operators. One particular highlight is the result of Knoop and Zhou [51]. They showed that for the second order Ditzian–Totik modulus one has*

$$\|B_k f - f\|_\infty \approx \omega_2^{\varphi}\left(f; \frac{1}{\sqrt{k}}\right), \ k \to \infty.$$

The authors are not aware of any corresponding result for the $S_{\Delta_{n,k}}$'s which generalizes the assertion of Knoop and Zhou. We feel that the proof of a strong converse inequality (in what form soever) would be a significant and most valuable contribution to both Approximation Theory and CAGD.

3.3 Pointwise inequalities

For the case of equidistant knots Lemma 4 can also be used to give pointwise inequalities. We have

Theorem 15 *For all* $f \in C[0,1]$, $x \in [0,1]$, $h \in (0,1]$ *and* $n, k \geq 1$ *one has:*

$$|S_{n,k}f(x) - f(x)| \leq \left[1 + \frac{1}{2h^2} \cdot \frac{\min\{2x(1-x); \frac{k}{n}\}}{n+k-1}\right] \cdot \omega_2(f; h). \tag{24}$$

In particular, one has

$$|S_{n,k}f(x) - f(x)| \leq \frac{3}{2} \cdot \omega_2\left(f; \sqrt{\frac{2x(1-x)}{n+k-1}}\right). \tag{25}$$

Proof:

The first inequality is an immediate consequence of Theorem 2 and Lemma 4; for the second one we consider two cases:

Case 1: $\frac{k}{n} \geq \frac{1}{2}$. In this case we have for $x \in [0,1]$

$$\min\left\{2x(1-x); \frac{k}{n}\right\} = 2x(1-x).$$

Putting $h = \sqrt{\frac{2x(1-x)}{n+k-1}} \leq 1$ yields

$$|S_{n,k}f(x) - f(x)| \leq \frac{3}{2} \cdot \omega_2\left(f; \sqrt{\frac{2x(1-x)}{n+k-1}}\right). \tag{26}$$

Case 2: $\frac{k}{n} < \frac{1}{2}$.

Depending on the position of x we have two possibilities.

For $x \in \left[0, \frac{1}{2}\left(1 - \sqrt{1 - \frac{2k}{n}}\right)\right] \cup \left[\frac{1}{2}\left(1 + \sqrt{1 - \frac{2k}{n}}\right), 1\right]$ it follows $\min\left\{2x(1-x); \frac{k}{n}\right\} = 2x(1-x)$, thus (26) also holds in this case.

For $x \in \left(\frac{1}{2}\left(1 - \sqrt{1 - \frac{2k}{n}}\right), \frac{1}{2}\left(1 + \sqrt{1 - \frac{2k}{n}}\right)\right)$ it follows $\min\left\{2x(1-x); \frac{k}{n}\right\} = \frac{k}{n}$. In this case Theorem 15 implies

$$|S_{n,k}f(x) - f(x)| \leq \left[1 + \frac{1}{2h^2} \cdot \frac{k}{n(n+k-1)}\right] \cdot \omega_2(f; h). \tag{27}$$

Setting $h = \sqrt{\frac{k}{n(n+k-1)}}$ we obtain

$$|S_{n,k}f(x) - f(x)| \leq \frac{3}{2} \cdot \omega_2\left(f; \sqrt{\frac{k}{n(n+k-1)}}\right) \leq \frac{3}{2} \cdot \omega_2\left(f; \sqrt{\frac{2x(1-x)}{n+k-1}}\right). \tag{28}$$

This concludes the proof. □

Remark 16 *(i) Case 1 in the proof of Theorem 15, namely* $\frac{k}{n} \geq \frac{1}{2}$, *is the one similar to that of the Bernstein operators* B_k. *For these we obtain*

$$|B_kf(x) - f(x)| \leq \frac{3}{2} \cdot \omega_2\left(f; \sqrt{\frac{2x(1-x)}{k}}\right).$$

(ii) *The case* $\frac{k}{n} < \frac{1}{2}$ *(where the piecewise polynomial degree* k *is small in comparison to the number* $n-1$ *of interior knots) is more similar to "true spline interpolation". In the middle of the interval* $[0,1]$ *we used the inequality* $\frac{k}{n} \le 2x(1-x)$ *in order to arrive at (25), thus loosing one power of* $\frac{1}{n}$ *under the square root in the special situation* $\frac{k}{n} = \frac{k_n}{n} = \mathcal{O}(\frac{1}{n}),\ n \to \infty$.

There is a second possibility to prove pointwise inequalities. This bridges the gap between pointwise ones in terms of the classical second modulus and uniform estimates using the Ditzian–Totik modulus. In order to indicate what can be done in this direction, we give without proof the following

Theorem 17 *Under the conditions of Theorem 15 one has*

$$|S_{n,k}f(x) - f(x)| \le 2 \cdot c\left(\lambda, \left(\frac{1}{2}\right)^{1-\lambda}\right) \cdot \omega_2^{\varphi^\lambda}\left(f; \frac{\varphi(x)^{1-\lambda}}{\sqrt{n+k-1}}\right).$$

Here, $\varphi(x) = \sqrt{x(1-x)},\ 0 \le \lambda \le 1$*, the constant* $c(\lambda, t_0)$ *is chosen such that* $K_2^{\varphi^\lambda}(f;t^2) \le c(\lambda,t_0) \cdot \omega_2^{\varphi^\lambda}(f;t)$ *for* $0 \le t \le t_0$; $K_2^{\varphi^\lambda}(f;t^2) := \inf\{\|f-g\|_\infty + t^2 \cdot \|\varphi^{2\lambda} \cdot g''\|_\infty\},\ t \ge 0$*, where the infimum is taken over all* g *such that* $g' \in AC_{loc}[0,1]$ *and* $\|\varphi^{2\lambda} \cdot g''\|_\infty < \infty$*, and* $\omega_2^{\varphi^\lambda}(f;t) := \sup\limits_{0 \le h \le t} \|\Delta^2_{h\varphi^\lambda} f\|_\infty$ *with*

$$\Delta^2_{h\varphi^\lambda} f(x) := \begin{cases} f(x - h\varphi^\lambda(x)) - 2f(x) + f(x + h\varphi^\lambda(x)), & \textit{if } [x - h\varphi^\lambda(x), x + h\varphi^\lambda(x)] \subseteq [0,1]; \\ 0, & \textit{otherwise.} \end{cases}$$

For details on the technique employed here see [38, Theorem 4.1] or [22, 27, 28]. Refinements of Theorem 17 are possible.

4. Approximation of derivatives

While, for functions $f \in C^r[0,1]$, the rth derivatives of the corresponding Bernstein polynomials converge uniformly to the rth derivative of the function f on $[0,1]$, this fact does not hold for variation–diminishing spline approximations in general, except for $r=1$ (see, for example, [55, Theorem 9]).

For the special case of equidistant knots,

$$n > 1 \text{ and } x_j = \frac{j}{n},\ 0 \le j \le n,$$

Marsden [55, Section 10] noted that the pth derivative $(1 < p \le r)$ of the spline approximation of degree k to $f(x)$ converges to $f^{(p)}(x)$ as

$$\frac{\|\Delta_n\|}{k} \to 0 \text{ or, equivalently } k + n \to \infty$$

if and only if $0 < x < 1$. The convergence is uniform on compact subintervals of $(0, 1)$.

Of course, the latter statement assumes that all quantities in question are defined. For example, one needs $k - 1 \geq p$ here in order to have sufficiently many derivatives of the splines available.

In the sequel we present quantitative estimates concerning the degree of simultaneous approximation for the first and second derivative. In order to do so, we need some general settings.

Again $K = [a, b]$ is a compact interval of the real axis and $K' \subset K$. We consider the Banach space $X = C^r(K)$ endowed with the norm $\|g\|_X := \max\limits_{0 \leq j \leq r}(\|D^j g\|_K)$. Here $\|\cdot\|_K$ denotes the Chebyshev norm in $C(K) := C^0(K)$ and D^j is the $j-$th differential operator.

Let $\mathcal{K}_K^i := \{f \in C(K) : [x_0, \ldots, x_i; f] \geq 0 \text{ for any } x_0 < \ldots < x_i \in K\}$, where $[x_0, \ldots, x_i; f]$ is an $i-$th order divided difference of f. Note that $\mathcal{K}_K^0$ is the set of all positive functions on K, $\mathcal{K}_K^1$ is the set of non–decreasing functions, and $\mathcal{K}_K^2$ represents the usual convex functions on the same interval.

Knoop and Pottinger [50] generalized the convexity notion for operators as follows: an operator $L : V \to C(K')$ is called almost convex of order $r - 1$ $(r \geq 0)$ if there exist $p \geq 0$ integers i_j, $1 \leq j \leq p$, satisfying $0 \leq i_1 < \ldots < i_p < r$ such that

$$f \in \left(\bigcap_{j=1}^{p} \mathcal{K}_K^{i_j}\right) \cap \mathcal{K}_K^r \cap V \text{ implies } Lf \in \mathcal{K}_{K'}^r.$$

Here, the empty intersection $\left(\bigcap_{j=1}^{0} \ldots\right)$ is taken by definition to be the entire subspace V.

The main result that we use in the sequel is the following quantitative Korovkin–type theorem on simultaneous approximation given by Kacsó (see [48], [49]), improving an earlier similar result of Gonska [32]:

Theorem 18 *Let $r \in \mathbb{N}_0$ and the operator $L : C^r(K) \to C^r(K')$ be almost convex of order $r - 1$. If $L(\Pi_{r-1}) \subseteq \Pi_{r-1}$, then for all $f \in C^r(K)$, $x \in K'$ and $0 < h \leq \frac{1}{2}$length(K) there holds:*

$$\begin{aligned} |D^r Lf(x) - D^r f(x)| \;&\leq\; \left|\frac{1}{r!} D^r L e_r(x) - 1\right| \cdot |D^r f(x)| + \frac{1}{h} \cdot |\gamma_L(x)| \cdot \omega_1(D^r f; h) \\ &+\; \left[D^r L\left(\frac{1}{r!} e_r\right)(x) + \frac{1}{2h^2} \cdot \beta_L(x)\right] \cdot \omega_2(D^r f; h), \end{aligned}$$

where

$$\gamma_L(x) := D^r L\left(\frac{1}{(r+1)!}\, e_{r+1} - \frac{1}{r!}\, x \cdot e_r\right)(x), \tag{29}$$

$$\beta_L(x) := D^r L\left(\frac{2}{(r+2)!}\, e_{r+2} - \frac{2}{(r+1)!}\, x \cdot e_{r+1} + \frac{1}{r!}\, x^2 \cdot e_r\right)(x). \tag{30}$$

In regard to the representation and the behaviour of the spline functions, Marsden proved the following results:

Lemma 19 *(see [55, Lemmas 1,2])*

(i) Let $f \in C^1[0,1]$ *and* $k > 1$*. Then*

$$\begin{aligned} DS_{\Delta_n,k}f(x) &= \sum_{j=1-k}^{n-1} \frac{f(\xi_{j,k}) - f(\xi_{j-1,k})}{\xi_{j,k} - \xi_{j-1,k}} \cdot N_{j,k-1}(x) \qquad (31) \\ &= \sum_{j=1-k}^{n-1} Df(\theta_{j,k}) \cdot N_{j,k-1}(x),\ \xi_{j-1,k} < \theta_{j,k} < \xi_{j,k}. \end{aligned}$$

(ii) Let $f \in C^2[0,1]$ *and* $k > 1$*. Then*

$$D^2 S_{\Delta_n,k}f(x) = \sum_{j=2-k}^{n-1} D^2 f(\eta_{j,k}) \cdot \frac{\xi_{j,k} - \xi_{j-2,k}}{2(\xi_{j,k-1} - \xi_{j-1,k-1})} \cdot N_{j,k-2}(x),\ \xi_{j-2,k} < \eta_{j,k} < \xi_{j,k}. \quad (32)$$

Lemma 20 *(see [55, Theorem 10]) Let* $f \in C^3[0,1]$ *and* $k > 2$*. Then*

(i) If $Df(x) \geq 0$ *on* $[0,1]$*, then* $DS_{\Delta_n,k}f(x) \geq 0$ *on* $[0,1]$*.*

(ii) If $D^2f(x) \geq 0$ *on* $[0,1]$*, then* $D^2S_{\Delta_n,k}f(x) \geq 0$ *on* $[0,1]$*.*

However,

(iii) If $D^3f(x) \geq 0$ *on* $[0,1]$*,* $D^3S_{\Delta_n,k}f(x)$ *need not be nonnegative.*

Lemma 21 *(see [55, Theorem 11]) Let* $f \in C^2[0,1]$*, and let* $x_j = \frac{j}{n}$*,* $0 < j < n$*, be the interior knots of* Δ_n*. Let* $k + n \to \infty$*,* $\liminf n > 1$*, and* $\liminf k > 1$*. If*

$$\lim \left(\frac{k-1}{k} \right) = R$$

exists, then

$$\begin{aligned} \lim D^2 S_{n,k} f(0) &= \frac{3R}{2} D^2 f(0), \\ \lim D^2 S_{n,k} f(1) &= \frac{3R}{2} D^2 f(1), \\ \lim D^2 S_{n,k} f(x) &= D^2 f(x), \quad \text{for } 0 < x < 1. \end{aligned}$$

The convergence is uniform on compact subsets of $(0,1)$*.*

The most elegant case in the above is attained for $\frac{3R}{2} = 1$, that is for $k = 3$. This is why for the second derivative cubic splines with equidistant knots play a special role.

For the first order derivatives we can prove the following

Theorem 22 *Let $f \in C^1[0,1]$, $x \in [0,1]$ and $h > 0$. Then, for $n \geq 1$, $k \geq 2$, the following estimate holds:*

$$\begin{aligned} & |DS_{\Delta_n,k}f(x) - Df(x)| \qquad (33) \\ & \leq \frac{1}{h} \cdot \|\Delta_n\| \cdot \omega_1(Df;h) + \left[1 + \frac{1}{2h^2}\left(1 + \sqrt{\frac{k}{12}}\right)^2 \cdot \|\Delta_n\|^2\right] \omega_2(Df;h). \end{aligned}$$

Proof:

The above statement will be derived using the result in Theorem 18; to that end we need upper bounds for the quantities appearing there (with $r = 1$).

Using (31) we get immediately

$$DS_{\Delta_n,k}e_1(x) = \sum_{j=1-k}^{n-1} De_1(\theta_{j,k}) \cdot N_{j,k-1}(x) = \sum_{j=1-k}^{n-1} N_{j,k-1}(x) = 1,\ \xi_{j-1,k} < \theta_{j,k} < \xi_{j,k},$$

thus

$$|DS_{\Delta_n,k}e_1(x) - 1| = |DS_{n,k}e_1(x) - 1| = 0.$$

For $\gamma_{S_{\Delta_n,k}}(x)$ we obtain successively

$$\begin{aligned} |\gamma_{S_{\Delta_n,k}}(x)| &:= \left|DS_{\Delta_n,k}\left(\frac{e_2}{2} - xe_1\right)(x)\right| = \left|\frac{1}{2}DS_{\Delta_n,k}e_2(x) - xDS_{\Delta_n,k}e_1(x)\right| \\ &= \left|\frac{1}{2} \cdot 2\sum_{j=1-k}^{n-1} N_{j,k-1}(x) \cdot \theta_{j,k} - \sum_{j=1-k}^{n-1} N_{j,k-1}(x) \cdot \xi_{j,k-1}\right| \\ &= |\sum_{j=1-k}^{n-1} N_{j,k-1}(x)(\theta_{j,k} - \xi_{j,k-1})| \\ &\leq \sum_{j=1-k}^{n-1} N_{j,k-1}(x)|\theta_{j,k} - \xi_{j,k-1}|. \end{aligned}$$

Since

$$\begin{aligned} & \xi_{j-1,k} \leq \xi_{j,k-1} \leq \xi_{j,k},\ -k < j < n \text{ and} \\ & \xi_{j-1,k} < \theta_{j,k} < \xi_{j,k}, \end{aligned}$$

it follows that

$$|\theta_{j,k} - \xi_{j,k-1}| \leq \xi_{j,k} - \xi_{j-1,k} = \frac{x_{j+k} - x_j}{k} \leq \|\Delta_n\|.$$

Substituting this in the above yields

$$|\gamma_{S_{\Delta_n,k}}(x)| \leq \|\Delta_n\| \sum_{j=1-k}^{n-1} N_{j,k-1}(x) = \|\Delta_n\|.$$

In order to obtain an upper bound for $\beta_{S_{\Delta_n,k}}(x)$ we apply formula (31) for the function $f = \frac{e_3}{3} - xe_2 + x^2 e_1$ and write successively

$$\begin{aligned}
\beta_{S_{\Delta_n,k}}(x) &:= DS_{\Delta_n,k}\left(\frac{e_3}{3} - xe_2 + x^2 e_1\right)(x) \\
&= \sum_{j=1-k}^{n-1} (\theta_{j,k}^2 - 2x\theta_{j,k} + x^2) \cdot N_{j,k-1}(x) = \sum_{j=1-k}^{n-1} (\theta_{j,k} - x)^2 \cdot N_{j,k-1}(x) \\
&= \sum_{j=1-k}^{n-1} (\theta_{j,k} - \xi_{j,k-1} + \xi_{j,k-1} - x)^2 \cdot N_{j,k-1}(x) \\
&\leq \sum_{j=1-k}^{n-1} (\theta_{j,k} - \xi_{j,k-1})^2 \cdot N_{j,k-1}(x) + \sum_{j=1-k}^{n-1} (\xi_{j,k-1} - x)^2 \cdot N_{j,k-1}(x) \\
&\qquad +2 \sum_{j=1-k}^{n-1} |\theta_{j,k} - \xi_{j,k-1}| \cdot |\xi_{j,k-1} - x| \cdot N_{j,k-1}(x) \\
&\leq \|\Delta_n\|^2 + S_{\Delta_n,k-1}((e_1 - x)^2; x) + 2\|\Delta_n\| \cdot S_{\Delta_n,k-1}(|e_1 - x|; x) \\
&\leq \|\Delta_n\|^2 + S_{\Delta_n,k-1}((e_1 - x)^2; x) + 2\|\Delta_n\| \cdot \sqrt{S_{\Delta_n,k-1}((e_1 - x)^2; x)} \\
&\leq \|\Delta_n\|^2 + \frac{k\|\Delta_n\|^2}{12} + 2\sqrt{\frac{k}{12}} \cdot \|\Delta_n\|^2 \\
&= \left(1 + \sqrt{\frac{k}{12}}\right)^2 \cdot \|\Delta_n\|^2.
\end{aligned}$$

In the above we used the Cauchy inequality and (3).

Replacing the above quantities into the general estimate of Theorem 18, we obtain the statement of our theorem. □

Taking $h = \|\Delta_n\|$ in Theorem 22 yields

Corollary 23 *Let $f \in C^1[0,1]$, $x \in [0,1]$. Then, for $n \geq 1$, $k \geq 2$, one has*

$$|DS_{\Delta_n,k}f(x) - Df(x)| \leq \omega_1\left(Df; \|\Delta_n\|\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k}{12}}\right)^2 \cdot \omega_2\left(Df; \|\Delta_n\|\right). \quad (34)$$

Remark 24 *For the Schoenberg splines $S_{n,k}$ (with equidistant knots), inequality (33) can be given only in terms of the second order modulus of smoothness if $x \in \left[\frac{k-1}{n}, 1 - \frac{k-1}{n}\right]$ instead of $x \in [0,1]$ since, on this smaller interval, $\gamma_{S_{n,k}}(x) = 0$. Thus we get*

$$|DS_{n,k}f(x) - Df(x)| \leq \left[1 + \frac{1}{2h^2} \cdot \frac{1}{n^2}\left(1 + \sqrt{\frac{k}{12}}\right)^2\right] \omega_2(Df; h), \quad (35)$$

for $2 \leq k \leq \frac{n}{2}+1$ *(the latter inequality following from the requirement* $\frac{k-1}{n} \leq 1-\frac{k-1}{n}$*). In particular, for* $h=\dfrac{1}{n}$*, the latter estimate becomes*

$$|DS_{n,k}f(x)-Df(x)| \leq \frac{3}{2}\left(1+\sqrt{\frac{k}{12}}\right)^2 \cdot \omega_2\left(Df;\frac{1}{n}\right), \tag{36}$$

for $2 \leq k \leq \frac{n}{2}+1$.

For splines with $x_j=\frac{j}{n}$, $0 \leq j \leq n$, and second order derivatives one has uniform convergence on compact subsets of $(0,1)$ only. In this case we can state the following

Theorem 25 *Let* $f \in C^2[0,1]$, $x \in \left[\dfrac{k-1}{n}, 1-\dfrac{k-1}{n}\right]$ *and* $h>0$*. Then there holds:*

$$\begin{aligned} &|D^2S_{n,k}f(x)-D^2f(x)| \qquad (37)\\ &\leq \frac{1}{h}\cdot\frac{1}{n}\cdot\omega_1(D^2f;h)+\left[1+\frac{1}{2h^2}\cdot\frac{1}{n^2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2\right]\cdot\omega_2(D^2f;h), \end{aligned}$$

for $3 \leq k \leq \frac{n}{2}+1$.

Proof:

Putting

$$B_{j,k} := \frac{\xi_{j,k}-\xi_{j-2,k}}{2(\xi_{j,k-1}-\xi_{j-1,k-1})}, \tag{38}$$

formula (32) becomes

$$D^2S_{n,k}f(x)=\sum_{j=2-k}^{n-1} D^2f(\eta_{j,k})\cdot B_{j,k}\cdot N_{j,k-2}(x),\ \xi_{j-2,k}<\eta_{j,k}<\xi_{j,k}. \tag{39}$$

One has

$$\begin{aligned} B_{j,k} &= \frac{1}{2}\cdot\frac{k-1}{k}\cdot\left(1+\frac{x_{j+k}-x_{j-1}}{x_{j+k-1}-x_j}\right)\\ &= \frac{1}{2}\cdot\frac{k-1}{k}\cdot\left(1+\frac{k+1}{k-1}\right),\ \text{for } 1 \leq j \leq n-k\\ &= 1,\ \text{for } 1 \leq j \leq n-k. \end{aligned}$$

Thus, for $x \in \left[\dfrac{k-1}{n}, 1-\dfrac{k-1}{n}\right]$, we obtain

$$\left|\frac{1}{2}D^2S_{n,k}e_2(x)-1\right| = \left|\frac{1}{2}\sum_{j=1}^{n-k}2\cdot B_{j,k}\cdot N_{j,k-2}(x)-1\right| = \left|\sum_{j=1}^{n-k}N_{j,k-2}(x)-1\right| = 0.$$

Here we have taken into account that on $\left[0, \frac{k-1}{n}\right]$ the splines $N_{j,k-2},\ 2-k \le j \le 0$, are zero, and the same is true for $N_{j,k-2}$ on $\left[1-\frac{k-1}{n}, 1\right]$ for $n-k+1 \le j \le n-1$.

Furthermore,

$$\begin{aligned}
|\gamma_{S_{n,k}}(x)| &:= \left| D^2 S_{n,k}\left(\frac{e_3}{3!} - x\frac{e_2}{2}\right)(x)\right| = \left|\frac{1}{3!}D^2 S_{n,k} e_3(x) - x\right| \\
&= \left|\frac{1}{3!}\cdot 6\sum_{j=1}^{n-k}\eta_{j,k}\cdot B_{j,k}\cdot N_{j,k-2}(x) - \sum_{j=1}^{n-k}\xi_{j,k-2}\cdot N_{j,k-2}(x)\right| \quad (\xi_{j-2,k} < \eta_{j,k} < \xi_{j,k}) \\
&= \left|\sum_{j=1}^{n-k} N_{j,k-2}(x)(\eta_{j,k}\cdot B_{j,k} - \xi_{j,k-2})\right| \\
&\le \sum_{j=1}^{n-k} N_{j,k-2}(x)|\eta_{j,k} - \xi_{j,k-2}| \le \sum_{j=1}^{n-k} N_{j,k-2}(x)\cdot\frac{1}{n} = \frac{1}{n}.
\end{aligned}$$

In the above we used the fact that

$$|\eta_{j,k} - \xi_{j,k-2}| \ \le \ \begin{cases} \xi_{j,k} - \xi_{j,k-2} = \frac{1}{n}, & \text{if } \eta_{j,k} \ge \xi_{j,k-2}, \\ \xi_{j,k-2} - \xi_{j-2,k} = \frac{1}{n}, & \text{if } \eta_{j,k} < \xi_{j,k-2}. \end{cases}$$

For $\beta_{S_{n,k}}(x),\ x \in \left[\frac{k-1}{n}, 1-\frac{k-1}{n}\right]$, we use formula (32) for the function $f = \frac{2}{4!}e_4 - \frac{2}{3!}xe_3 + \frac{1}{2!}x^2e_2$ and write successively

$$\begin{aligned}
0 \le \beta_{S_{n,k}}(x) &:= D^2 S_{n,k}\left(\frac{2}{4!}e_4 - \frac{2}{3!}xe_3 + \frac{1}{2!}x^2e_2\right)(x) \\
&= \sum_{j=1}^{n-k} N_{j,k-2}(x)(\eta_{j,k}^2 - 2x\eta_{j,k} + x^2)\cdot B_{j,k} \quad (\xi_{j-2,k} < \eta_{j,k} < \xi_{j,k}) \\
&= \sum_{j=1}^{n-k} N_{j,k-2}(x)(\eta_{j,k} - x)^2 \\
&= \sum_{j=1}^{n-k} N_{j,k-2}(x)(\eta_{j,k} - \xi_{j,k-2} + \xi_{j,k-2} - x)^2 \\
&\le \sum_{j=1}^{n-k} N_{j,k-2}(x)(\eta_{j,k} - \xi_{j,k-2})^2 + \sum_{j=1}^{n-k} N_{j,k-2}(x)(\xi_{j,k-2} - x)^2 \\
&\qquad +2\sum_{j=1}^{n-k}|\eta_{j,k} - \xi_{j,k-2}|\cdot|\xi_{j,k-2} - x|\cdot N_{j,k-2}(x) \\
&\le \frac{1}{n^2} + \frac{k-1}{12n^2} + 2\cdot\frac{1}{n}\cdot\sqrt{\frac{k-1}{12n^2}}
\end{aligned}$$

$$= \frac{1}{n^2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2.$$

In the above we used the Cauchy inequality and (3). An application of Theorem 18 yields the statement of our theorem. □

In particular, for $h = \frac{1}{n}$, we get

Corollary 26 *Let* $f \in C^2[0,1]$, $x \in \left[\frac{k-1}{n}, 1-\frac{k-1}{n}\right]$ *and* $3 \leq k \leq \frac{n}{2}+1$. *Then there holds:*

$$|D^2 S_{n,k} f(x) - D^2 f(x)| \leq \omega_1\left(D^2 f; \frac{1}{n}\right) + \frac{3}{2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2 \cdot \omega_2\left(D^2 f; \frac{1}{n}\right) \quad (40)$$

As was mentioned earlier in this note, close to the endpoints there are problems with second order derivatives. We illustrate this for a simple case in the following

Example 27 *Consider* $S_{n,3}(e_2; x)$ *for* $0 \leq x \leq \frac{1}{n}$. *From the representations for* $N_{j,3}(x)$, $-3 \leq j \leq 0$, *given above it can be derived that on* $[0, \frac{1}{n}]$ *one has*

$$D^2 S_{n,3} e_2(x) = 2 - \frac{1}{3} x n,$$

that is

$$D^2 S_{n,3}(e_2; 0) = 2 = D^2 e_2(0),$$

but, independent of n,

$$D^2 S_{n,3}\left(e_2; \frac{1}{n}\right) = \frac{5}{3} < 2 = D^2 e_2\left(\frac{1}{n}\right).$$

This is why it is impossible to prove uniform convergence for the second derivatives on the whole interval $[0,1]$ *as* $\frac{1}{n} \to 0$.

At the left endpoint $\frac{k-1}{n} = \frac{2}{n}$ *of the interval on which we proved uniform convergence we have*

$$D^2 S_{n,3}\left(e_2; \frac{2}{n}\right) = 2 = D^2 e_2\left(\frac{2}{n}\right), \ 4 \leq n,$$

again due to the general statement. In fact,

$$D^2 S_{n,3}(e_2; x) = D^2 e_2(x)$$

even for all $x \in [\frac{2}{n}, 1-\frac{2}{n}]$.

It is also interesting to note that, while $D^3 e_2(x) = 0$ *for all* $x \in [0,1]$, *the second derivative of* $S_{n,3} e_2$ *strictly* <u>*decreases*</u> *on* $[0, \frac{1}{n}]$.

5. Global smoothness preservation

Over the recent years there has been considerable interest in the preservation of global smoothness in various contexts. This intensive research culminated in the recent book by Anastassiou and Gal [3]. Already in the very first article [2] treating this phenomenon under a systematic point of view, global smoothness preservation by Schoenberg operators $S_{\Delta_n,k}$ with respect to the first order modulus ω_1 was investigated. It was shown there, among other things, that

$$\omega_1(S_{\Delta_n,k}f;t) \leq 2 \cdot \omega_1(f;t), \quad f \in C[0,1], \ t \geq 0.$$

In this section we present an analogous result for a certain "second" $K-$ functional and the classical second order modulus. To that end we use the following tool given earlier by Cottin and Gonska.

Lemma 28 *(see Theorem 2.2 in [17]) Let $r \geq 0$ and $s \geq 1$ be integers, and let K and K' be given as above. Furthermore, let $L : C^r(K) \to C^r(K')$ be a linear operator having the following properties:*

- *(i) L is almost convex of orders $r-1$ and $r+s-1$,*
- *(ii) L maps $C^{r+s}(K)$ into $C^{r+s}(K')$,*
- *(iii) $L(\Pi_{r-1}) \subseteq \Pi_{r-1}$ and $L(\Pi_{r+s-1}) \subseteq \Pi_{r+s-1}$*
- *(iv) $L(C^r(K)) \not\subseteq \Pi_{r-1}$.*

Then for all $f \in C^r(K)$ and all $\delta \geq 0$ we have

$$K_s(D^rLf;\delta)_{K'} \leq \frac{1}{r!} \cdot \|D^rLe_r\| \cdot K_s\left(f^{(r)}; \frac{1}{(r+s)_s} \cdot \frac{\|D^{r+s}Le_{r+s}\|}{\|D^rLe_r\|} \cdot \delta\right)_K. \tag{41}$$

In the above, K_s is the Peetre K–functional of order s, $s \geq 1$, given by

$$K_s(f;\delta) := K(f;\delta;C[0,1],C^s[0,1]) := \inf\{\|f-g\| + \delta \cdot \|g^{(s)}\| : \ g \in C^s[0,1]\},$$

$(a)_b$ denotes the Pochhammer symbol defined by

$$(a)_0 := 1, \ \ (a)_b := \prod_{k=0}^{b-1}(a-k), \ \ a \in \mathbb{R}, \ b \in \mathbb{N},$$

and $\Pi_{-1} := \{0\}$.

Now we can state

Theorem 29 *For all $f \in C[0,1]$ and all $\delta \geq 0$, the variation–diminishing splines $S_{\Delta_n,k}$ of degree $k \geq 3$ with $n \geq 2$ satisfy the following estimates:*

$$K_2(S_{\Delta_n,k}f;\delta) \;\leq\; K_2\left(f;\frac{k-1}{k}\cdot\left(\frac{1}{2}+\rho(\Delta_n)\right)\cdot\delta\right), \text{ and} \tag{42}$$

$$\omega_2(S_{\Delta_n,k}f;\delta) \;\leq\; 3\cdot\left[1+\frac{k-1}{4k}\cdot(1+2\cdot\rho(\Delta_n))\right]\cdot\omega_2(f;\delta), \tag{43}$$

where $\rho(\Delta_n) := \dfrac{\|\Delta_n\|}{\min\limits_{0\leq i\leq n-1}(x_{i+1}-x_i)}$ *is the mesh ratio.*

Proof:

It can be easily verified that, for $r = 0$ and $s = 2$, the assumptions of Lemma 28 are satisfied by $S_{\Delta_n,k}$ with $k \geq 3$. Hence (41) reads now as follows:

$$\begin{aligned} K_2(S_{\Delta_n,k}f;\delta) &\leq \|S_{\Delta_n,k}e_0\|\cdot K_2\left(f;\frac{1}{2}\cdot\frac{\|D^2S_{\Delta_n,k}e_2\|}{\|S_{\Delta_n,k}e_0\|}\cdot\delta\right) \\ &= K_2\left(f;\frac{1}{2}\cdot\|D^2S_{\Delta_n,k}e_2\|\cdot\delta\right), \end{aligned} \tag{44}$$

since $\|S_{\Delta_n,k}e_0\| = 1$.

Furthermore, for $k \geq 3$,

$$\begin{aligned} B_{j,k} &= \frac{1}{2}\cdot\frac{k-1}{k}\cdot\left(1+\frac{x_{j+k}-x_{j-1}}{x_{j+k-1}-x_j}\right) \\ &\leq \frac{1}{2}\cdot\frac{k-1}{k}\cdot\left(1+\max_{-k+2\leq j\leq n-1}\frac{x_{j+k}-x_{j-1}}{x_{j+k-1}-x_j}\right) \\ &\leq \frac{1}{2}\cdot\frac{k-1}{k}\cdot\left(1+\max\left\{2,\frac{k+1}{k-1}\right\}\cdot\rho(\Delta_n)\right) \\ &= \frac{1}{2}\cdot\frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n)). \end{aligned}$$

The 2 appearing in $\max\left\{2,\frac{k+1}{k-1}\right\}$ in the above is due to certain special cases when considering equidistant knots.

Thus

$$\begin{aligned} |D^2S_{\Delta_n,k}f(x)| &\leq \|f''\|\cdot\frac{1}{2}\cdot\frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n))\sum_{j=2-k}^{n-1}N_{j,k-2}(x) \\ &= \|f''\|\cdot\frac{1}{2}\cdot\frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n)), \end{aligned}$$

and, in particular,

$$\|D^2S_{\Delta_n,k}e_2\| \leq \frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n)).$$

Substituting this upper bound into (44) yields (42).

For the second statement of our theorem we employ the function $Z_\delta(f)$ from Žuk's paper [81] (see Lemma 1 there), also observing the fact that

$$K_2(f;\delta) = K(f;\delta;C[0,1],C^2[0,1]) = K(f;\delta;C[0,1],W_{2,\infty}[0,1]).$$

Here,

$$W_{2,\infty}[0,1] := \{f \in C[0,1] : f' \text{ absolutely continuous, } \|f''\|_{L_\infty} < \infty\},$$

where

$$\|f''\|_{L_\infty} = \text{vrai} \sup_{x\in[0,1]} |f''(x)|.$$

Let now $f \in C[0,1]$, $0 < \delta \leq \frac{1}{2}$ be arbitrarily given, and let $|h| \leq \delta$. Then for a typical difference figuring in the definition of $\omega_2(S_{\Delta_n,k}f;\delta)$ we have

$$\begin{aligned}
&|S_{\Delta_n,k}f(x-h) - 2S_{\Delta_n,k}f(x) + S_{\Delta_n,k}f(x+h)|\\
&= |\{S_{\Delta_n,k}(f-g;x-h) - 2S_{\Delta_n,k}(f-g;x) + S_{\Delta_n,k}(f-g;x+h)\}\\
&\qquad +\{S_{\Delta_n,k}(g;x-h) - 2S_{\Delta_n,k}(g;x) + S_{\Delta_n,k}(g;x+h)\}|,
\end{aligned}$$

where $g \in W_{2,\infty}[0,1]$ may be arbitrarily chosen.

The absolute value of the first term in curly parentheses can be estimated from above by

$$4\|S_{\Delta_n,k}(f-g)\|_\infty \leq 4\|f-g\|_\infty.$$

For the modulus of the second expression in curly brackets we have

$$\begin{aligned}
&|S_{\Delta_n,k}(g;x-h) - 2S_{\Delta_n,k}(g;x) + S_{\Delta_n,k}(g;x+h)|\\
&= |D^2S_{\Delta_n,k}(g;\xi)| \cdot h^2 \text{ (for some } \xi \text{ between } x-h \text{ and } x+h)\\
&\leq \|D^2S_{\Delta_n,k}g\| \cdot h^2 \leq \frac{1}{2}\cdot\frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n))\cdot h^2\cdot\|g''\|_{L_\infty}.
\end{aligned}$$

We now substitute the function $g \in W_{2,\infty}[0,1]$ by $Z_h(f)$ from Žuk's paper [81], satisfying for $0 < h \leq \frac{1}{2}$ the inequalities

$$\begin{aligned}
\|f - Z_h(f)\| &\leq \frac{3}{4}\cdot\omega_2(f;h), \text{ and}\\
\|Z_h''(f)\|_{L_\infty} &\leq \frac{3}{2}\cdot\frac{1}{h^2}\cdot\omega_2(f;h).
\end{aligned}$$

Combining these estimates leads to

$$\begin{aligned}
\omega_2(S_{\Delta_n,k}f;\delta) &\leq 3\cdot\omega_2(f;\delta) + \frac{3}{4}\cdot\frac{k-1}{k}\cdot(1+2\cdot\rho(\Delta_n))\cdot\omega_2(f;\delta)\\
&= 3\cdot\left[1+\frac{k-1}{4k}\cdot(1+2\cdot\rho(\Delta_n))\right]\cdot\omega_2(f;\delta),
\end{aligned}$$

which completes the proof. □

Since, for equidistant knots $x_j = \frac{j}{n}$, $0 \leq j \leq n$, one has $\rho(\Delta_n) = 1$, it follows immediately

Corollary 30 *For all $f \in C[0,1]$ and all $\delta \geq 0$, the variation–diminishing splines $S_{n,k}$ of degree $k \geq 3$ with $n \geq 2$ satisfy the following estimates:*

$$\begin{aligned} K_2(S_{n,k}f;\delta) &\leq K_2\left(f; \frac{3}{2} \cdot \frac{k-1}{k} \cdot \delta\right), \text{ and} \\ \omega_2(S_{n,k}f;\delta) &\leq \left(3 + \frac{9}{4} \cdot \frac{k-1}{k}\right) \cdot \omega_2\left(f;\delta\right). \end{aligned}$$

6. Multivariate approaches

In the sequel we present statements on the degrees of approximation and simultaneous approximation for first and second derivatives in certain bivariate cases. We restrict ourselves to state inheritance principles (Theorems 31, 37) in terms of the classical second order modulus of smoothness, but similar statements can be formulated also for ω_2^{φ} and $\omega_2^{\varphi^\lambda}$ (see, e.g., [24, 16]).

All our results below should be compared with corresponding ones by Munteanu and Schumaker [62]. Due to the consequent use of ω_2 all our estimates will be of at least the orders given by Munteanu and Schumaker or improve them. Furthermore, again thanks to ω_2, we are able to better exploit smoothness properties of a given function f than they were able to do. This is true in particular in those cases in which f has two continuous partials in either x, y, or both. Note that we are also able to give quantitative information for more partials than they did.

Several results of this section will be given in terms of so–called partial moduli of smoothness of order r, given for the compact intervals $I, J \subset \mathbb{R}$, for $f \in C(I \times J)$, $r \in \mathbb{N}_0$ and $\delta \in \mathbb{R}_+$ by

$$\omega_r(f;\delta,0) := \sup\left\{\left|\sum_{\nu=0}^{r}(-1)^{r-\nu}\binom{r}{\nu} \cdot f(x+\nu h, y)\right| \,:\, (x,y),(x+rh,y) \in I \times J,\ |h| \leq \delta\right\}$$

and symmetrically by

$$\omega_r(f;0,\delta) := \sup\left\{\left|\sum_{\nu=0}^{r}(-1)^{r-\nu}\binom{r}{\nu} \cdot f(x, y+\nu h)\right| \,:\, (x,y),(x,y+rh) \in I \times J,\ |h| \leq \delta\right\}.$$

Occasionally we will use total moduli of smoothness of order r, defined by

$$\begin{aligned} \omega_r(f;\delta_1,\delta_2) := \sup\Bigg\{\left|\sum_{\nu=0}^{r}(-1)^{r-\nu}\binom{r}{\nu} \cdot f(x+\nu h_1, y+\nu h_2)\right| \,: \\ (x,y),(x+rh_1,y+rh_2) \in I \times J,\ |h_1| \leq \delta_1,\ |h_2| \leq \delta_2\Bigg\}, \end{aligned}$$

for the compact intervals $I, J \subset \mathbb{R}$, for $f \in C(I \times J)$, $r \in I\!N_0$ and $\delta_1, \delta_2 \in \mathbb{R}_+$.

The third type of moduli figuring in this section will be the mixed moduli of smoothness, given for $r, s \in I\!N_0$ by

$$\omega_{r,s}(f;\delta_1,\delta_2) := \sup\left\{\left|\sum_{\nu=0}^{r}\sum_{\mu=0}^{s}(-1)^{r+s-\nu-\mu}\binom{r}{\nu}\binom{s}{\mu}\cdot f(x+\nu h_1, y+\mu h_2)\right| : \right.$$

$$\left. (x,y), (x+rh_1, y+sh_2) \in I \times J,\ |h_i| \le \delta_i,\ i=1,2\right\}.$$

Several properties of these moduli can be found in Schumaker's book [73] and in [34].

6.1 Boolean sums

In order to cover Boolean sums of Schoenberg spline operators we will use the inheritance principle in the theorem below. The theorem is in analogy to two previous versions given in [35], [36], but it is adapted here to the situation we are dealing with.

In particular the operators L and M will be discretely defined, i.e., for finitely many, mutually distinct points $x_e,\ e \in E$ (E a suitable index set) of the compact interval I and fundamental functions A_e the operator L will be of the form

$$L(g;x) = \sum_{e \in E} g(x_e) \cdot A_e(x).$$

If $A_e \in C^p(I'),\ p \ge 0,\ I' \subseteq I$, then $L : C^p(I) \to C^p(I')$.

Likewise M will be of the form

$$M(h;y) = \sum_{f \in F} h(y_f) \cdot B_f(y)$$

and under analogous assumptions will map $C^q(J)$ into $C^q(J')$.

If L is of the form given above, then its parametric extension to $C^{p,q}(I \times J)$ is given by

$${}_xL(F;x,y) = L(F_y;x) = \sum_{e \in E} F_y(x_e) \cdot A_e(x) = \sum_{e \in E} F(x_e,y) \cdot A_e(x).$$

If we apply the partial differential operator $\frac{\partial^q}{\partial y^q} = D^{(0,q)}$ to this function we get

$$\begin{aligned} \left(D^{(0,q)} \circ {}_xL\right)(F;x,y) &= \frac{\partial^q}{\partial y^q}\sum_{e \in E} F(x_e,y)A_e(x) \\ &= \sum_{e \in E}\frac{\partial^q}{\partial y^q}F(x_e,y)\cdot A_e(x) = \sum_{e \in E}\left(F^{(0,q)}\right)_y(x_e)\cdot A_e(x) \\ &= \left({}_xL \circ D^{(0,q)}\right)(F;x,y), \end{aligned}$$

that is, $D^{(0,q)}$ and ${}_xL$ commute on $C^{p,q}(I \times J)$.

Analogously

$$D^{(p,0)} \circ {}_yM = {}_yM \circ D^{(p,0)}.$$

These commutativities will be crucial for the proof of the inheritance principle given in the theorem below.

One additional difference in comparison to earlier statements consists in our introducing certain intervals I', J', with $I' \subseteq I,\ J' \subseteq J$. This is due to the fact that, already for the second derivative, Schoenberg splines show a certain deficiency close to the endpoints. This was also observed by Marsden [55].

In other words, we will thus be able to give better estimates on $I' \times J'$ in the particular case where $I' \subsetneq I$ and $J' \subsetneq J$. Details will become clear in the applications.

Theorem 31 *Let I, I', J, J' be non–trivial compact intervals of the real axis $\mathbb{R}$, such that $I' \subseteq I$ and $J' \subseteq J$. For $(0,0) \leq (p',q') \leq (p,q)$ let discretely defined operators $L : C^p(I) \to C^{p'}(I')$ and $M : C^q(J) \to C^{q'}(J')$ be given such that for fixed $r, s \in I\!N_0$*

$$|(g - Lg)^{(p)}(x)| \leq \sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x) \cdot \omega_\rho(g^{(p)}; \Lambda_{\rho,p,L}(x)), \quad x \in I', g \in C^p(I), \tag{45}$$

and

$$|(h - Mh)^{(q)}(y)| \leq \sum_{\sigma=0}^{s} \Gamma_{\sigma,q,M}(y) \cdot \omega_\sigma(h^{(q)}; \Lambda_{\sigma,q,M}(y)), \quad y \in J', h \in C^q(J). \tag{46}$$

Here, Γ and Λ are positive, bounded functions.

Then we have for any $(x,y) \in I' \times J'$ and for all $f \in C^{p,q}(I \times J)$

$$\begin{aligned} &\left|(f - ({}_xL \oplus {}_yM)f)^{(p,q)}(x,y)\right| \\ &\qquad \leq \sum_{\rho=0}^{r} \sum_{\sigma=0}^{s} \Gamma_{\rho,p,L}(x) \cdot \Gamma_{\sigma,q,M}(y) \cdot \omega_{\rho,\sigma}\left(f^{(p,q)}; \Lambda_{\rho,p,L}(x), \Lambda_{\sigma,q,M}(y)\right)\right). \end{aligned}$$

Proof:

We want to estimate

$$\begin{aligned} &\left|D^{(p,q)} \circ [Id - ({}_xL \oplus {}_yM)]\,(f;x,y)\right| \\ &\quad = \left|D^{(p,0)} \circ D^{(0,q)} \circ [Id - ({}_xL \oplus {}_yM)]\,(f;x,y)\right| \\ &\quad = \left|D^{(p,0)}\left[D^{(0,q)}((Id - {}_yM)(f) - {}_xL \circ (Id - {}_yM)(f))\right](x,y)\right| \\ &\quad = \left|D^{(p,0)}\left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f) - \left(D^{(0,q)} \circ {}_xL \circ (Id - {}_yM)\right)(f)\right](x,y)\right|. \end{aligned}$$

Using now the commutativity $D^{(0,q)} \circ {}_xL = {}_xL \circ D^{(0,q)}$, $x \in I'$, for the discretely defined operator L we get

$$\begin{aligned} &\left|D^{(p,0)}\left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f) - \left(D^{(0,q)} \circ {}_xL \circ (Id - {}_yM)\right)(f)\right](x,y)\right| \\ &\qquad = \left|D^{(p,0)}\left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f) - {}_xL \circ \left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right](x,y)\right|. \end{aligned}$$

Now the assumption on the quantitative behaviour of the univariate operator L may be used since the function in [...] can also be written as a univariate function of x with parameter y, namely as

$$I' \ni x \mapsto \left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y (x) - L\left(\left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y ; x\right) \in \mathbb{R}.$$

Applying $D^{(p,0)}$ to the function in [...] is the same as differentiating the latter univariate function with respect to x. Hence, by assumption (45), the quantity which we are interested in is bounded from above by

$$\sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x) \cdot \omega_\rho \left(\left(\frac{d}{dx}\right)^p \left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y ; \Lambda_{\rho,p,L}(x)\right).$$

The ρ–th modulus of smoothness can be replaced by

$$\left| {}_x\Delta^\rho_{\delta^*}\left[\left(\frac{d}{dx}\right)^p \left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y\right](x^*)\right|$$

for some $x^* \in I'$ and $|\delta^*| \le \Lambda_{\rho,p,L}(x)$.

Next we investigate the latter quantity by using the information available on M. The absolute value of the ρ–th order difference is equal to

$$\left|\sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left[\left(\frac{d}{dx}\right)^p \left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y\right](x^* + i\delta^*)\right|.$$

As in the above for L, we use now the commutativity for M, namely $D^{(p,0)} \circ {}_yM = {}_yM \circ D^{(p,0)}$, $y \in J'$. Since

$$\begin{aligned} \left(\frac{d}{dx}\right)^p \left[\left(D^{(0,q)} \circ (Id - {}_yM)\right)(f)\right]_y (x) &= \left(D^{(p,0)} \circ D^{(0,q)} \circ (Id - {}_yM)\right)(f; x, y) \\ &= \left(D^{(0,q)} \circ D^{(p,0)} \circ (Id - {}_yM)\right)(f; x, y) \\ &= \left(D^{(0,q)} \circ (Id - {}_yM) \circ D^{(p,0)}\right)(f; x, y), \end{aligned}$$

it follows that the ρ–th order difference can be written as

$$\begin{aligned} &\left|\sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(0,q)} \circ (Id - {}_yM)\right)\left(D^{(p,0)} f; x^* + i\delta^*, y\right)\right| \\ &\quad = \left|\sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left\{\left(\frac{d}{dy}\right)^q \left(D^{(p,0)} f\right)_{x^*+i\delta^*}(y) - \left(\frac{d}{dy}\right)^q \left({}_yM \circ D^{(p,0)}(f)\right)_{x^*+i\delta^*}(y)\right\}\right| \end{aligned}$$

$$= \left| \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left\{ \left(\frac{d}{dy}\right)^q \left(D^{(p,0)} f\right)_{x^*+i\delta^*}(y) - \left(\frac{d}{dy}\right)^q M\left(\left(D^{(p,0)} f\right)_{x^*+i\delta^*}; y\right) \right\} \right|$$

$$= \left| \left[\left(\frac{d}{dy}\right)^q - \left(\frac{d}{dy}\right)^q \circ M \right] \left(\sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(p,0)} f\right)_{x^*+i\delta^*}; y \right) \right| .$$

This difference may now be evaluated using assumption (46) on M. Hence, its absolute value is less than or equal to

$$\sum_{\sigma=0}^{s} \Gamma_{\sigma,q,M}(y) \cdot \omega_\sigma \left(\left(\frac{d}{dy}\right)^q \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(p,0)} f\right)_{x^*+i\delta^*}; \Lambda_{\sigma,q,M}(y) \right) .$$

The σ–th order modulus can be written as

$$\left| {}_y\Delta_{\eta^*}^{\sigma} \left[\left(\frac{d}{dy}\right)^q \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(p,0)} f\right)_{x^*+i\delta^*} \right] (y^*) \right|$$

for some $y^* \in J'$ and a suitable η^* such that $|\eta^*| \leq \Lambda_{\sigma,q,M}(y)$. More explicitly, the latter quantity is equal to

$$\begin{aligned}
&\left| {}_y\Delta_{\eta^*}^{\sigma} \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(\frac{d}{dy}\right)^q \left(D^{(p,0)} f\right)_{x^*+i\delta^*}(y^*) \right| \\
&\quad = \left| {}_y\Delta_{\eta^*}^{\sigma} \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(p,q)} f\right)_{x^*+i\delta^*}(y^*) \right| \\
&\quad = \left| \sum_{j=0}^{\sigma} (-1)^j \binom{\sigma}{j} \sum_{i=0}^{\rho} (-1)^i \binom{\rho}{i} \left(D^{(p,q)} f\right)(x^*+i\delta^*, y^*+j\eta^*) \right| \\
&\quad = \left| \sum_{j=0}^{\sigma} \sum_{i=0}^{\rho} (-1)^{j+i} \binom{\sigma}{j} \binom{\rho}{i} \left(D^{(p,q)} f\right)(x^*+i\delta^*, y^*+j\eta^*) \right| \\
&\quad \leq \omega_{\rho,\sigma}\left(f^{(p,q)};\ \Lambda_{\rho,p,L}(x),\ \Lambda_{\sigma,q,M}(y)\right) .
\end{aligned}$$

Combining the latter inequality with the observations made earlier in this proof shows the validity of the statement of Theorem 31. □

We will now give a number of applications for Boolean sums of Schoenberg splines. In doing so we will not strive to be as general as possible, but restrict ourselves to some cases of special interest, namely to estimates involving only the mesh gauge of the splines, but other direct inequalities from Section 3 can be used as well. The results given here should be also compared to the early paper [14] by Coman.

Theorem 32 *We consider the operators $S_{\Delta_n,k} : C[0,1] \to C^{k-1}[0,1]$ and $S_{\Delta_m,l} : C[0,1] \to C^{l-1}[0,1]$ for $n, m \geq 1$ and $k, l \geq 1$. For their Boolean sums we have*

$$\|f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f\|_\infty \leq \left(1 + \frac{k+1}{24}\right) \cdot \left(1 + \frac{l+1}{24}\right) \cdot \omega_{2,2}\left(f; \|\Delta_n\|, \|\Delta_m\|\right) .$$

The proof is immediate: put $r = s = 2$ and $\Gamma_{0,p,L} = \Gamma_{1,p,L} = \Gamma_{0,q,M} = \Gamma_{1,q,M} = 0$ in the general theorem for Boolean sums and use (20) twice.

We now turn to statements concerning also the approximation of derivatives. However, the upper bounds which are derived from Theorem 31 are quite complex. We thus focus in the following results on certain smooth functions to derive more instructive upper bounds. Nonetheless the following proof will provide the reader with an idea of what can be stated for less smooth functions.

Theorem 33 *For the operators* $S_{\Delta_n,k} : C^1[0,1] \to C^{k-1}[0,1]$ *and* $S_{\Delta_m,l} : C^1[0,1] \to C^{l-1}[0,1]$ *for* $n, m \geq 1$ *and* $k, l \geq 2$ *the following inequalities hold for any function* $f \in C^{2,2}[0,1]^2$*:*

$$\begin{aligned}
(i) &\quad \|f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f\|_\infty = \mathcal{O}\left(\|\Delta_n\|^2 \cdot \|\Delta_m\|^2\right);\\
(ii) &\quad \|(f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f)^{(1,0)}\|_\infty = \mathcal{O}\left(\|\Delta_n\| \cdot \|\Delta_m\|^2\right);\\
(iii) &\quad \|(f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f)^{(0,1)}\|_\infty = \mathcal{O}\left(\|\Delta_n\|^2 \cdot \|\Delta_m\|\right);\\
(iv) &\quad \|(f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f)^{(1,1)}\|_\infty = \mathcal{O}\left(\|\Delta_n\| \cdot \|\Delta_m\|\right).
\end{aligned}$$

In all four cases $\mathcal{O}$ *depends on* k *and* l*.*

Proof:

(i) is an immediate consequence of Theorem 32. It is only necessary to observe that

$$\omega_{2,2}\left(f; \|\Delta_n\|, \|\Delta_m\|\right) \leq \|\Delta_n\|^2 \cdot \|\Delta_m\|^2 \cdot \|f^{(2,2)}\|_\infty.$$

(ii) We apply Theorem 31 (for $p = 1$, $q = 0$) with $r = s = 2$, $\Gamma_{0,0,S_{\Delta_n,k}} = \Gamma_{0,1,S_{\Delta_n,k}} = 0$ and collect the others Γ's and the Λ's from the univariate case, that is

$$\begin{aligned}
&\Gamma_{1,0,S_{\Delta_n,k}}(x) = 0, && \\
&\Gamma_{2,0,S_{\Delta_n,k}}(x) = 1 + \tfrac{k+1}{24}, && \Lambda_{2,0,S_{\Delta_n,k}}(x) = \|\Delta_n\|;\\
&\Gamma_{1,1,S_{\Delta_n,k}}(x) = 1, && \Lambda_{1,1,S_{\Delta_n,k}}(x) = \|\Delta_n\|;\\
&\Gamma_{2,1,S_{\Delta_n,k}}(x) = \tfrac{3}{2}\left(1 + \sqrt{\tfrac{k}{12}}\right)^2, && \Lambda_{2,1,S_{\Delta_n,k}}(x) = \|\Delta_n\|.
\end{aligned}$$

The Γ's and the Λ's with respect to $S_{\Delta_m,l}$ are to be chosen analogously.

For brevity we write in the sequel $\Gamma_{1,1}(x)$ instead of $\Gamma_{1,1,S_{\Delta_n,k}}(x)$, etc. The upper bound which is derived from Theorem 31 is then as follows:

$$\begin{aligned}
&\|(f - ({}_xS_{\Delta_n,k} \oplus {}_yS_{\Delta_m,l})f)^{(1,0)}\|_\infty\\
&\quad \leq \Gamma_{1,1}(x) \cdot [\Gamma_{1,0}(y) \cdot \omega_{1,1}(f^{(1,0)}; \Lambda_{1,1}(x), \Lambda_{1,0}(y)) + \Gamma_{2,0}(y) \cdot \omega_{1,2}(f^{(1,0)}; \Lambda_{1,1}(x), \Lambda_{2,0}(y))]
\end{aligned}$$

$$+\Gamma_{2,1}(x)\cdot[\Gamma_{1,0}(y)\cdot\omega_{2,1}(f^{(1,0)};\Lambda_{2,1}(x),\Lambda_{1,0}(y))+\Gamma_{2,0}(y)\cdot\omega_{2,2}(f^{(1,0)};\Lambda_{2,1}(x),\Lambda_{2,0}(y))]$$

$$=\left(1+\frac{l+1}{24}\right)\cdot\omega_{1,2}\left(f^{(1,0)};\|\Delta_n\|,\|\Delta_m\|\right)$$

$$+\frac{3}{2}\left(1+\sqrt{\frac{k}{12}}\right)^2\cdot\left(1+\frac{l+1}{24}\right)\cdot\omega_{2,2}\left(f^{(1,0)};\|\Delta_n\|,\|\Delta_m\|\right)$$

$$=\mathcal{O}\left(\omega_{1,2}\left(f^{(1,0)};\|\Delta_n\|,\|\Delta_m\|\right)+\omega_{2,2}\left(f^{(1,0)};\|\Delta_n\|,\|\Delta_m\|\right)\right)$$

$$=\mathcal{O}\left(\|\Delta_n\|\cdot\|\Delta_m\|^2\cdot\|f^{(2,2)}\|_\infty\right)$$

$$=\mathcal{O}\left(\|\Delta_n\|\cdot\|\Delta_m\|^2\right)\text{ for }f\in C^{2,2}[0,1]^2.$$

(iii) This is analogous to case (ii).

(iv) The functions Γ and Λ are the same as in case (ii); they just appear in different combinations now:

$$\|(f-({}_xS_{\Delta_n,k}\oplus{}_yS_{\Delta_m,l})f)^{(1,1)}\|_\infty$$

$$\leq\Gamma_{1,1}(x)\cdot[\Gamma_{1,1}(y)\cdot\omega_{1,1}(f^{(1,1)};\Lambda_{1,1}(x),\Lambda_{1,1}(y))+\Gamma_{2,1}(y)\cdot\omega_{1,2}(f^{(1,1)};\Lambda_{1,1}(x),\Lambda_{2,1}(y))]$$

$$+\Gamma_{2,1}(x)\cdot[\Gamma_{1,1}(y)\cdot\omega_{2,1}(f^{(1,1)};\Lambda_{2,1}(x),\Lambda_{1,1}(y))+\Gamma_{2,1}(y)\cdot\omega_{2,2}(f^{(1,1)};\Lambda_{2,1}(x),\Lambda_{2,1}(y))]$$

$$=\omega_{1,1}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)+\frac{3}{2}\left(1+\sqrt{\frac{l}{12}}\right)^2\cdot\omega_{1,2}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)$$

$$+\frac{3}{2}\left(1+\sqrt{\frac{k}{12}}\right)^2\cdot\left[\,\omega_{2,1}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)\right.$$

$$\left.+\frac{3}{2}\left(1+\sqrt{\frac{l}{12}}\right)^2\cdot\omega_{2,2}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)\,\right]$$

$$=\mathcal{O}\left(\omega_{1,1}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)+\omega_{1,2}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)\right.$$

$$\left.+\,\omega_{2,1}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)+\omega_{2,2}\left(f^{(1,1)};\|\Delta_n\|,\|\Delta_m\|\right)\right)$$

$$=\mathcal{O}\left(\|\Delta_n\|\cdot\|\Delta_m\|\cdot\|f^{(2,2)}\|_\infty\right)$$

$$=\mathcal{O}\left(\|\Delta_n\|\cdot\|\Delta_m\|\right)\text{ for }f\in C^{2,2}[0,1]^2.\ \square$$

Remark 34 *If in Theorem 33 we take the* sup *norms over* $\left[\frac{k-1}{n},1-\frac{k-1}{n}\right]\times\left[\frac{l-1}{n},1-\frac{l-1}{n}\right]$ *only and* $f\in C^{3,3}[0,1]^2$*, we get* $\mathcal{O}\left(\|\Delta_n\|^2\cdot\|\Delta_m\|^2\right)$ *as an upper bound for all the quantities from (i) to (iv) there.*

In order to give inequalities for the partial derivatives of orders up to $(2,2)$ which are not covered by the previous theorem we restrict our attention to $\left[\frac{k-1}{n},1-\frac{k-1}{n}\right]\times\left[\frac{l-1}{n},1-\frac{l-1}{n}\right]$ and to the case of equidistant knots.

Theorem 35 *For* $S_{n,k} : C^2[0,1] \to C^{k-1}[0,1]$ *and* $S_{m,l} : C^2[0,1] \to C^{l-1}[0,1]$, $3 \leq k \leq \frac{n}{2}+1$, $3 \leq l \leq \frac{m}{2}+1$, *the following inequalities are true for* $f \in C^{3,3}[0,1]^2$*:*

$$(i) \quad \|(f - ({}_xS_{n,k} \oplus {}_yS_{m,l})f)^{(2,0)}\|_\infty = \mathcal{O}\left(\frac{1}{n} \cdot \frac{1}{m^2}\right);$$

$$(ii) \quad \|(f - ({}_xS_{n,k} \oplus {}_yS_{m,l})f)^{(2,1)}\|_\infty = \mathcal{O}\left(\frac{1}{n} \cdot \frac{1}{m^2}\right);$$

$$(iii) \quad \|(f - ({}_xS_{n,k} \oplus {}_yS_{m,l})f)^{(2,2)}\|_\infty = \mathcal{O}\left(\frac{1}{n} \cdot \frac{1}{m}\right).$$

Analogous statements hold for the partial derivatives of orders $(0,2)$ *and* $(1,2)$*. The* $\mathcal{O}$*'s depend on* k *and* l *and the* sup *norms are to be taken over* $\left[\frac{k-1}{n}, 1-\frac{k-1}{n}\right] \times \left[\frac{l-1}{n}, 1-\frac{l-1}{n}\right]$.

Proof:

The functions needed now are $\Gamma_{0,0} = \Gamma_{0,1} = \Gamma_{0,2} = 0$ and

$$\begin{aligned}
&\Gamma_{1,0}(x) = 0, \\
&\Gamma_{2,0}(x) = 1 + \tfrac{k+1}{24}, && \Lambda_{2,0}(x) = \tfrac{1}{n}; \\
&\Gamma_{1,1}(x) = 0, \\
&\Gamma_{2,1}(x) = \tfrac{3}{2}\left(1+\sqrt{\tfrac{k}{12}}\right)^2, && \Lambda_{2,1}(x) = \tfrac{1}{n}; \\
&\Gamma_{1,2}(x) = 1, && \Lambda_{1,2}(x) = \tfrac{1}{n}; \\
&\Gamma_{2,2}(x) = \tfrac{3}{2}\left(1+\sqrt{\tfrac{k-1}{12}}\right)^2, && \Lambda_{2,2}(x) = \tfrac{1}{n}.
\end{aligned}$$

Again, the Γ's and the Λ's with respect to $S_{m,l}$ are analogous.

(i) From the general theorem we obtain

$$\begin{aligned}
&\|(f - ({}_xS_{n,k} \oplus {}_yS_{m,l})f)^{(2,0)}\|_\infty \\
&\leq \Gamma_{1,2}(x) \cdot \left[\Gamma_{1,0}(y) \cdot \omega_{1,1}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right) + \Gamma_{2,0}(y) \cdot \omega_{1,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right)\right] \\
&\qquad + \Gamma_{2,2}(x) \cdot \left[\Gamma_{1,0}(y) \cdot \omega_{2,1}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right) + \Gamma_{2,0}(y) \cdot \omega_{2,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right)\right] \\
&= \left(1 + \frac{l+1}{24}\right) \cdot \omega_{1,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right) \\
&\qquad + \frac{3}{2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2 \cdot \left(1 + \frac{l+1}{24}\right) \cdot \omega_{2,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right) \\
&= \mathcal{O}\left(\omega_{1,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right) + \omega_{2,2}\left(f^{(2,0)}; \frac{1}{n}, \frac{1}{m}\right)\right) \\
&= \mathcal{O}\left(\frac{1}{n} \cdot \frac{1}{m^2}\right) \text{ for } f \in C^{3,3}[0,1]^2.
\end{aligned}$$

(ii) Now

$$
\begin{aligned}
&\|(f-({}_xS_{n,k}\oplus {}_yS_{m,l})f)^{(2,1)}\|_\infty\\
&\le \Gamma_{1,2}(x)\cdot\left[\Gamma_{1,1}(y)\cdot\omega_{1,1}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)+\Gamma_{2,1}(y)\cdot\omega_{1,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)\right]\\
&\qquad+\Gamma_{2,2}(x)\cdot\left[\Gamma_{1,1}(y)\cdot\omega_{2,1}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)+\Gamma_{2,1}(y)\cdot\omega_{2,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)\right]\\
&=\frac{3}{2}\left(1+\sqrt{\frac{l}{12}}\right)^2\cdot\omega_{1,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)\\
&\qquad+\frac{3}{2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2\cdot\frac{3}{2}\left(1+\sqrt{\frac{l}{12}}\right)^2\cdot\omega_{2,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)\\
&=\mathcal{O}\left(\omega_{1,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)+\omega_{2,2}\left(f^{(2,1)};\frac{1}{n},\frac{1}{m}\right)\right)\\
&=\mathcal{O}\left(\frac{1}{n}\cdot\frac{1}{m^2}\right)\quad\text{for } f\in C^{3,3}[0,1]^2.
\end{aligned}
$$

(iii) In this case

$$
\begin{aligned}
&\|(f-({}_xS_{n,k}\oplus {}_yS_{m,l})f)^{(2,2)}\|_\infty\\
&\le \Gamma_{1,2}(x)\cdot\left[\Gamma_{1,2}(y)\cdot\omega_{1,1}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)+\Gamma_{2,2}(y)\cdot\omega_{1,2}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)\right]\\
&\qquad+\Gamma_{2,2}(x)\cdot\left[\Gamma_{1,2}(y)\cdot\omega_{2,1}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)+\Gamma_{2,2}(y)\cdot\omega_{2,2}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)\right]\\
&=\omega_{1,1}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)+\frac{3}{2}\left(1+\sqrt{\frac{l-1}{12}}\right)^2\cdot\omega_{1,2}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)\\
&\qquad+\frac{3}{2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2\cdot\omega_{2,1}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)\\
&\qquad+\frac{3}{2}\left(1+\sqrt{\frac{k-1}{12}}\right)^2\cdot\frac{3}{2}\left(1+\sqrt{\frac{l-1}{12}}\right)^2\cdot\omega_{2,2}\left(f^{(2,2)};\frac{1}{n},\frac{1}{m}\right)\\
&=\mathcal{O}\left(\frac{1}{n}\cdot\frac{1}{m}+\frac{1}{n}\cdot\frac{1}{m}+\frac{1}{n}\cdot\frac{1}{m}+\frac{1}{n}\cdot\frac{1}{m}\right)=\mathcal{O}\left(\frac{1}{n}\cdot\frac{1}{m}\right)\quad\text{for } f\in C^{3,3}[0,1]^2.\square
\end{aligned}
$$

Remark 36 *The fact that the partials of orders* $(2,0)$ *and* $(2,1)$ *are approximated with the same order* $\mathcal{O}\left(\frac{1}{n}\cdot\frac{1}{m^2}\right)$ *is due to the fact that on the small interval* $\left[\frac{l-1}{n},1-\frac{l-1}{n}\right]$ *for the univariate operator* $S_{m,l}$ *both a function* $f\in C^3[0,1]$ *and its derivative are approximated with the same order. When it comes to the second derivative a power of one is lost in the univariate case.*

6.2 Tensor products

For the tensor product case the situation is the same as for Boolean sums: the reader should compare our results with those of Munteanu and Schumaker in order to confirm that the consequent use of ω_2 provides more insight. Further papers on tensor products were provided by Coman and Frenţiu [15] and Felicia Stancu [77], for example.

Also in the tensor product case we will use an inheritance principle. It is stated in a form which exactly suits our purposes and makes the same assumptions concerning the univariate building blocks as in the Boolean sum case.

Theorem 37 *Let $I' \subseteq I, J' \subseteq J$ be non–trivial compact intervals of the real axis $\mathbb{R}$. For $p, q \in I\!N_0$ let $L : C^p(I) \to C^p(I')$ and $M : C^q(J) \to C^q(J')$ be discretely defined operators as given above and such that for fixed $r, s \in I\!N_0$*

$$\left|(g - Lg)^{(p)}(x)\right| \;\leq\; \sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x) \cdot \omega_\rho(g^{(p)}; \Lambda_{\rho,p,L}(x)), \;\; x \in I', g \in C^p(I),$$

and

$$\left|(h - Mh)^{(q)}(y)\right| \;\leq\; \sum_{\sigma=0}^{s} \Gamma_{\sigma,q,M}(y) \cdot \omega_\sigma(h^{(q)}; \Lambda_{\sigma,q,M}(y)), \;\; y \in J', h \in C^q(J).$$

Here, Γ and Λ are bounded functions.

(i) Then for $(x, y) \in I' \times J'$ and $f \in C^{p,q}(I \times J)$ the following hold:

$$\begin{aligned}\left|\left[f - ({}_xL \circ {}_yM)f\right]^{(p,q)}(x, y)\right| \;&\leq\; \sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x) \cdot \omega_\rho(f^{(p,q)}; \Lambda_{\rho,p,L}(x), 0)\\ &+\; \|D^p \circ L\|^* \cdot \sum_{\sigma=0}^{s} \Gamma_{\sigma,q,M}(y) \cdot \omega_\sigma(f^{(p,q)}; 0, \Lambda_{\sigma,q,M}(y)).\end{aligned}$$

Here

$$\|D^p \circ L\|^* := \inf\{c : \|(D^p \circ L)g\|_{\infty,I'} \leq c \cdot \|g^{(p)}\|_{\infty,I}, \; \forall \; g \in C^p(I)\}.$$

(ii) A symmetric upper bound is given by

$$\sum_{\sigma=0}^{s} \Gamma_{\sigma,q,M}(y) \cdot \omega_\sigma(f^{(p,q)}; 0, \Lambda_{\sigma,q,M}(y)) + \|D^q \circ M\|^* \cdot \sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x) \cdot \omega_\rho(f^{(p,q)}; \Lambda_{\rho,p,L}(x), 0).$$

Proof:

Recall first that

$$D^{(0,q)} \circ {}_xL = {}_xL \circ D^{(0,q)} \text{ on } C^{(p,q)}(I \times J).$$

Then

$$
\begin{aligned}
&\left|[f-({}_xL\circ{}_yM)f]^{(p,q)}(x,y)\right| \\
&= \left|D^{(p,0)}\circ D^{(0,q)}\circ(Id-{}_xL\circ{}_yM)(f;x,y)\right| \\
&= \left|[D^{(p,0)}\circ D^{(0,q)}\circ(Id-{}_xL)+D^{(p,0)}\circ D^{(0,q)}\circ{}_xL\circ(Id-{}_yM)](f;x,y)\right| \\
&= \left|[D^{(p,0)}\circ(Id-{}_xL)\circ D^{(0,q)}+D^{(p,0)}\circ{}_xL\circ D^{(0,q)}\circ(Id-{}_yM)](f;x,y)\right| \\
&\le \left|D^{(p,0)}\circ(Id-{}_xL)\circ D^{(0,q)}(f;x,y)\right|+\left|D^{(p,0)}\circ{}_xL\circ D^{(0,q)}\circ(Id-{}_yM)(f;x,y)\right| \\
&=: E_1(x,y)+E_2(x,y).
\end{aligned}
$$

Now, for $x\in I'$,

$$
\begin{aligned}
E_1(x,y) &= \left|D^{(p,0)}\circ(Id-L)((f^{(0,q)})_y;x)\right| \le \sum_{\rho=0}^{r}\Gamma_{\rho,p,L}(x)\cdot\omega_\rho((f^{(p,q)})_y;\Lambda_{\rho,p,L}(x)) \\
&\le \sum_{\rho=0}^{r}\Gamma_{\rho,p,L}(x)\cdot\omega_\rho(f^{(p,q)};\Lambda_{\rho,p,L}(x),0).
\end{aligned}
$$

Furthermore, with $F:=D^{(0,q)}\circ(Id-{}_yM)f$, we have

$$
E_2(x,y)=\left|\left(D^{(p,0)}\circ{}_xL\right)(F;x)\right|=|(D^p\circ L)(F_y;x)|\le\|(D^p\circ L)(F_y)\|_{\infty,I'}\,.
$$

Here again $F_y\in C^p(I)$ for all $y\in J'$. By our assumption on L we have for any $g\in C^p(I)$ that

$$
\|(D^p\circ L)g\|_{\infty,I'}\le\left(1+\sum_{\rho=0}^{r}2^\rho\cdot\|\Gamma_{\rho,p,L}\|_{\infty,I'}\right)\cdot\|g^{(p)}\|_{\infty,I}.
$$

Hence

$$
\|D^p\circ L\|^*:=\inf\{c:\|D^p\circ L)g\|_{\infty,I'}\le c\cdot\|g^{(p)}\|_{\infty,I}\quad\forall\, g\in C^p(I)\}<\infty.
$$

In our present situation we have

$$
\begin{aligned}
\|F_y^{(p)}\|_{\infty,I} &= \left\|\frac{d^p}{dx^p}\left[D^{(0,q)}\circ(Id-{}_yM)f\right]_y(x)\right\|_{\infty,I}=\left\|D^{(p,0)}\circ D^{(0,q)}\circ(Id-{}_yM)f(\cdot,y)\right\|_{\infty,I} \\
&= \left\|D^{(0,q)}\circ(Id-{}_yM)f^{(p,0)}(\cdot,y)\right\|_{\infty,I}=\left\|\frac{d^q}{dy^q}\circ(Id-{}_yM)\left(f^{(p,0)}\right)_x(y)\right\|_{\infty,I} \\
&\le \left\|\sum_{\sigma=0}^{s}\Gamma_{\sigma,q,M}(y)\cdot\omega_\sigma\left(\frac{d^q}{dy^q}\left(f^{(p,0)}\right)_x;\ \Lambda_{\sigma,q,M}(y)\right)\right\|_{\infty,I} \\
&\le \sum_{\sigma=0}^{s}\Gamma_{\sigma,q,M}(y)\cdot\sup_{x\in I}\omega_\sigma\left(\frac{d^q}{dy^q}\left(f^{(p,0)}\right)_x;\Lambda_{\sigma,q,M}(y)\right) \\
&= \sum_{\sigma=0}^{s}\Gamma_{\sigma,q,M}(y)\cdot\omega_\sigma\left(f^{(p,q)};0,\Lambda_{\sigma,q,M}(y)\right).
\end{aligned}
$$

Hence

$$
\begin{aligned}
E_1(x,y)+E_2(x,y) \;&\leq\; \sum_{\rho=0}^{r} \Gamma_{\rho,p,L}(x)\cdot\omega_\rho\left(f^{(p,q)};\Lambda_{\rho,p,L}(x),0\right)\\
&+\; \|D^p\circ L\|^*\cdot\sum_{\sigma=0}^{s}\Gamma_{\sigma,q,M}(y)\cdot\omega_\sigma\left(f^{(p,q)};0,\Lambda_{\sigma,q,M}(y)\right).
\end{aligned}
$$

The second upper bound can be obtained in an analogous fashion. □

For the tensor product of two Schoenberg spline operators we first state

Theorem 38 *For $n,m\geq 1$ and $k,l\geq 1$ we have*

$$
\begin{aligned}
&\|f-({}_xS_{\Delta_{n,k}}\circ{}_yS_{\Delta_{m,k}})f\|_{\infty,I\times J}\\
&\leq\left(1+\frac{k+1}{24}\right)\cdot\omega_2(f;\|\Delta_n\|,0)+\left(1+\frac{l+1}{24}\right)\cdot\omega_2(f;0,\|\Delta_m\|)\\
&\leq\left(2+\frac{k+l+2}{24}\right)\cdot\omega_2(f;\|\Delta_n\|,\|\Delta_m\|).
\end{aligned}
$$

Proof:

This is the case $p=q=0,\ r=s=2$. With $\Gamma_{0,0}(x)=\Gamma_{1,0}(x)=0,\ \Gamma_{2,0}(x)=1+\frac{k+1}{24},\ \Lambda_{2,0}(x)=\|\Delta_n\|$ and analogous choices with respect to the variable y we arrive at the above upper bound, also observing that $\|D^0\circ S_{\Delta_{n,k}}\|^*=1$. □

For the partial derivatives up to order $(1,1)$ we arrive at

Theorem 39 *For $n,m\geq 1$ and $k,l\geq 2$ we have the following inequalities for any $f\in C^{2,2}[0,1]^2$.*

(i) $\|f-({}_xS_{\Delta_{n,k}}\circ{}_yS_{\Delta_{m,l}})f\|_\infty=\mathcal{O}(\|\Delta_n\|^2+\|\Delta_m\|^2)$;

(ii) $\|(f-({}_xS_{\Delta_{n,k}}\circ{}_yS_{\Delta_{m,l}})^{(1,0)}\|_\infty=\mathcal{O}(\|\Delta_n\|+\|\Delta_m\|^2)$;

(iii) $\|(f-({}_xS_{\Delta_{n,k}}\circ{}_yS_{\Delta_{m,l}})^{(0,1)}\|_\infty=\mathcal{O}(\|\Delta_n\|^2+\|\Delta_m\|)$;

(iv) $\|(f-({}_xS_{\Delta_{n,k}}\circ{}_yS_{\Delta_{m,l}})^{(1,1)}\|_\infty=\mathcal{O}(\|\Delta_n\|+\|\Delta_m\|)$.

In all four cases $\mathcal{O}$ depends in k and l, and the sup *norms are those over* $[0,1]^2$.

Proof:

The Γ's and the Λ's are again the same as in the Boolean sum case (see the proof of Theorem 33).

(i) This is an immediate consequence from Theorem 38.

(ii) With $r = s = 2,\ p = 1,\ q = 0$, in the general theorem we have

$$\begin{aligned}
&\|(f - ({}_xS_{\Delta_{n,k}} \circ {}_yS_{\Delta_{m,l}})f)^{(1,0)}\|_{\infty, I\times J}\\
&\leq 1 \cdot \omega_1\left(f^{(1,0)}; \|\Delta_n\|, 0\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k}{12}}\right)^2 \cdot \omega_2\left(f^{(1,0)}; \|\Delta_n\|, 0\right)\\
&\qquad + \|D^1 \circ S_{\Delta_{n,k}}\|^* \cdot \left(1 + \frac{l+1}{24}\right) \cdot \omega_2\left(f^{(1,0)}; 0, \|\Delta_m\|\right).
\end{aligned}$$

From the representation of $(D^1 \circ S_{\Delta_{n,k}})g$ for $g \in C^1[0,1]$ it follows that $\|D^1 \circ S_{\Delta_{n,k}}\|^* \leq 1$.

Thus we obtain an upper bound of order

$$\begin{aligned}
&\mathcal{O}\left(\omega_1(f^{(1,0)}; \|\Delta_n\|, 0) \ + \ \omega_2(f^{(1,0)}; \|\Delta_n\|, 0) \ + \ \omega_2(f^{(1,0)}; 0, \|\Delta_m\|)\right)\\
&= \mathcal{O}\left(\|\Delta_n\| + \|\Delta_m\|^2\right) \ \text{ for } f \in C^{2,2}[0,1]^2.
\end{aligned}$$

(iii) The proof for the partial of order $(0,1)$ is 'symmetric' to that for $(1,0)$.

(iv) Again with $r = s = 2,\ p = q = 1$, we have

$$\begin{aligned}
&\left\|(f - ({}_xS_{\Delta_{n,k}} \circ {}_yS_{\Delta_{m,l}})f)^{(1,1)}\right\|_{\infty, I\times J}\\
&\leq 1 \cdot \omega_1\left(f^{(1,1)}; \|\Delta_n\|, 0\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k}{12}}\right)^2 \cdot \omega_2\left(f^{(1,1)}; \|\Delta_n\|, 0\right)\\
&\qquad + 1 \cdot \left(\omega_1\left(f^{(1,1)}; 0, \|\Delta_m\|\right) + \frac{3}{2}\left(1 + \sqrt{\frac{l}{12}}\right)^2 \cdot \omega_2\left(f^{(1,1)}; 0, \|\Delta_m\|\right)\right)\\
&= \mathcal{O}\left(\|\Delta_n\| + \|\Delta_m\|\right) \ \text{ for } f \in C^{2,2}[0,1]^2. \ \square
\end{aligned}$$

For the remaining partials up to order $(2,2)$ again we consider only the case of equidistant knots and the smaller intervals $\left[\frac{k-1}{n}, 1 - \frac{k-1}{n}\right] \times \left[\frac{l-1}{m}, 1 - \frac{l-1}{m}\right]$. We now have

Theorem 40 *For $3 \leq k \leq \frac{n}{2} + 1,\ 3 \leq l \leq \frac{m}{2} + 1$ the following are true for $f \in C^{3,3}[0,1]^2$:*

(i) $\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,0)}\right\|_\infty = \mathcal{O}\left(\frac{1}{n} + \frac{1}{m^2}\right)$;

(ii) $\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,1)}\right\|_\infty = \mathcal{O}\left(\frac{1}{n} + \frac{1}{m^2}\right)$;

(iii) $\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,2)}\right\|_\infty = \mathcal{O}\left(\frac{1}{n} + \frac{1}{m}\right)$.

Analogous statements hold for the partials of orders $(0,2)$ and $(1,2)$; $\mathcal{O}$ depends on k and l in all cases and the sup *norms are those over the smaller subinterval given above.*

Proof:

The Γ's and the Λ's are again the same as in the Boolean sum case (see the proof of Theorem 35).

(i) From the general statement we obtain with $p = 2, q = 0, r = s = 2$,

$$\begin{aligned}
&\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,0)}\right\|_\infty \\
&\leq \Gamma_{1,2}(x) \cdot \omega_1\left(f^{(2,0)}; \Lambda_{1,2}(x), 0\right) + \Gamma_{2,2}(x) \cdot \omega_2\left(f^{(2,0)}; \Lambda_{2,2}(x), 0\right) \\
&\qquad + \left\|D^2 \circ S_{n,k}\right\|^* \cdot \Gamma_{2,0}(y) \cdot \omega_2\left(f^{(2,0)}; 0, \Lambda_{2,0}(y)\right) \\
&\leq \omega_1\left(f^{(2,0)}; \frac{1}{n}, 0\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k-1}{12}}\right)^2 \cdot \omega_2\left(f^{(2,0)}; \frac{1}{n}, 0\right) \\
&\qquad + \frac{3}{2} \cdot \frac{k-1}{k} \cdot \left(1 + \frac{l+1}{24}\right) \cdot \omega_2\left(f^{(2,0)}; 0, \frac{1}{m}\right) \\
&= \mathcal{O}\left(\frac{1}{n} + \frac{1}{m^2}\right) \text{ for } f \in C^{3,2}[0,1]^2.
\end{aligned}$$

Note that $\|D^2 \circ S_{n,k}\|^* \leq \frac{3}{2} \cdot \frac{k-1}{k}$ was shown in the section on global smoothness preservation.

(ii) For $p = 2,\ q = 1,\ r = s = 2$ we have

$$\begin{aligned}
&\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,1)}\right\|_\infty \\
&\leq \Gamma_{1,2}(x) \cdot \omega_1\left(f^{(2,1)}; \Lambda_{1,2}(x), 0\right) + \Gamma_{2,2}(x) \cdot \omega_2\left(f^{(2,1)}; \Lambda_{2,2}(x), 0\right) \\
&\qquad + \left\|D^2 \circ S_{n,k}\right\|^* \cdot \left\{\Gamma_{1,1}(y) \cdot \omega_1\left(f^{(2,1)}; 0, \Lambda_{1,1}(y)\right) + \Gamma_{2,1}(y) \cdot \omega_2\left(f^{(2,1)}; 0, \Lambda_{2,1}(y)\right)\right\} \\
&\leq \omega_1\left(f^{(2,1)}; \frac{1}{n}, 0\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k-1}{12}}\right)^2 \cdot \omega_2\left(f^{(2,1)}; \frac{1}{n}, 0\right) \\
&\qquad + \frac{3}{2} \cdot \frac{k-1}{k} \cdot \frac{3}{2}\left(1 + \sqrt{\frac{l}{12}}\right)^2 \cdot \omega_2\left(f^{(2,1)}; 0, \frac{1}{m}\right) \\
&= \begin{cases} \mathcal{O}\left(\frac{1}{n} + \frac{1}{m}\right) & \text{for } f \in C^{3,2}, \\ \mathcal{O}\left(\frac{1}{n} + \frac{1}{m^2}\right) & \text{for } f \in C^{3,3}. \end{cases}
\end{aligned}$$

(iii) Now $p = q = 2,\ r = s = 2$. Thus

$$\begin{aligned}
&\left\|(f - ({}_xS_{n,k} \circ {}_yS_{m,l})f)^{(2,2)}\right\|_\infty \\
&\leq \omega_1\left(f^{(2,2)}; \frac{1}{n}, 0\right) + \frac{3}{2}\left(1 + \sqrt{\frac{k-1}{12}}\right)^2 \cdot \omega_2\left(f^{(2,2)}, \frac{1}{n}, 0\right)
\end{aligned}$$

$$+\frac{3}{2}\cdot\frac{k-1}{k}\cdot\left[\omega_1\left(f^{(2,2)};0,\frac{1}{m}\right)+\frac{3}{2}\left(1+\sqrt{\frac{l-1}{12}}\right)^2\cdot\omega_2\left(f^{(2,2)};0,\frac{1}{m}\right)\right]$$

$$=\mathcal{O}\left(\omega_1\left(f^{(2,2)},\frac{1}{n},0\right)+\omega_2\left(f^{(2,2)};\frac{1}{n},0\right)+\omega_1\left(f^{(2,2)};0,\frac{1}{m}\right)+\omega_2\left(f^{(2,2)};0,\frac{1}{m}\right)\right)$$

$$=\mathcal{O}\left(\frac{1}{n}+\frac{1}{m}\right)\text{ for } f\in C^{3,3}[0,1]^2. \ \square$$

7. Concluding remarks and open problems

1. For the case of equidistant knots we were able to show
$$(S_{n,k}(e_1-x)^2)(x)\le 1\cdot\frac{\min\left\{2x(1-x),\frac{k}{n}\right\}}{n+k-1}.$$
For $n=1,\ k\ge 1$ and $n\ge 2,\ k=1$ the constant 1 can be replaced by $\frac{1}{2}$. It should be clearified if 1 is globally optimal.

   It would likewise be desirable to have an analogous inequality for general knot sequences.

   Instructive exact representations (and thus lower bounds) for $(S_{n,k}(e_1-x)^2)(x)$ are only known in a few exceptional cases. It would thus be of interest to find such representations or lower bounds for more general combinations of n and k.

2. Strong converse inequalities also seem to be known only in very special cases. For piecewise linear interpolation at equidistant knots see Ditzian and Ivanov [23], for Bernstein and related operators consult the papers by Zhou, Totik, Knoop, Ivanov and Ditzian [51, 79, 23]. It seems as if there is no strong converse inequality even for the case of quadratic Schoenberg splines, rather a popular tool (cf. their use in packages such as MacDraw, for example).

3. Likewise it would be desirable to prove at least inverse and saturation results including such for derivatives.

4. We have only given estimates for the approximation of derivatives up to order 2. We are not aware of corresponding estimates for derivatives of order $l\ge 3$.

5. The preservation of global smoothness is well understood for Bernstein operators and their derivatives (see [17]). For the Schoenberg operator this appears to be much more difficult since derivatives of order $l\ge 3$ would have to be also represented appropriately in order to come up with inequalities such as
$$\begin{aligned}\omega_2(D^1S_{n,k}f;\delta) &\le \ldots\omega_2(f';\ldots), \text{ or}\\ \omega_2(D^2S_{n,k}f;\delta) &\le \ldots\omega_2(f'';\ldots).\end{aligned}$$

Acknowledgments

The present article was significantly improved and finalized while all four authors were RiP–fellows at the Mathematical Research Institute in Oberwolfach. They gratefully acknowledge the support of the Volkswagen–Stiftung for this opportunity. The fourth author's contribution was also made possible by a DAAD–fellowship for visiting the University of Duisburg in the fall of 2001.

References

[1] A. Akopjan, B. Bojanov: *The Theory of Spline Functions* (Bulgarian). Sofia: Nauka i Izkustv 1990.

[2] G.A. Anastassiou, C. Cottin, H.H. Gonska: *Global smoothness of approximating functions*, Analysis **11** (1991), 43–57.

[3] G.A. Anastassiou, S.G. Gal: *Approximation theory. Moduli of continuity and global smoothness preservation.* Boston, MA: Birkhäuser 2000.

[4] R.E. Barnhill, R.F. Riesenfeld (eds.): *Computer Aided Geometric Design.* Proc. Int. Conf. on Computer Aided Design. New York: Acad. Press 1974.

[5] R.K. Beatson: *Monotone and convex approximation by splines: error estimates and a curve–fitting algorithm*, SIAM J. Num. Anal. **19** (1982), 1278–1285.

[6] L. Beutel, H.H. Gonska: *Quantitative inheritance properties for simultaneous approximation by tensor product operators.* Submitted. *Temporary reference*: Duisburg: Schriftenreihe des Instituts für Mathematik der Gerhard–Mercator–Universität **SM–DU–504** (2001).

[7] B.D. Bojanov, H.A. Hakopian, A.A. Sahakian: *Spline Functions and Multivariate Interpolations.* Dordrecht: Kluwer 1993.

[8] P. Blaga, Gh. Coman: *On some bivariate spline operators*, Anal. Numér. Théor. Approx. **8** (1979), 143–153.

[9] J.–D. Cao: *Kantorovich type operators and integral Schoenberg splines* (Chinese), Kexue Tonbao **27** (1982), no. 7, 385–388.

[10] J.–D. Cao: *On some results of integral Schoenberg splines* (Chinese), Ziran Zazhi–Nature Journal (1983), no. 7, 558.

[11] J.–D. Cao: *Operators of Kantorovich type and integral Schoenberg splines* (Russian), Acta Math. Hungar. **42** (1983), no. 3–4, 189–203.

[12] J.–D. Cao: *Operators of Kantorovich type and integral Schoenberg splines* (Russian), Kexue Tonbao (English Ed.) **28** (1983), no. 1, 5–9.

[13] E. Cohen, R.F. Riesenfeld, G. Elber: *Geometric Modeling with Splines. An Introduction.* Natick, MA: A K Peters 2001.

[14] Gh. Coman: *On the approximation of multivariate functions.* MRC Technical Summary Report 1254, University of Wisconsin, Madison, 1974.

[15] Gh. Coman, M. Frenţiu: *Bivariate spline approximation*, Studia Univ. Babeş–Bolyai, Ser. Math. Mech. **19** (1) (1974), 59–64.

[16] C. Cottin: *Quantitative Aussagen zur Blending–Typ–Approximation.* Ph.D. Thesis, University of Duisburg 1988.

[17] C. Cottin, H.H. Gonska: *Simultaneous approximation and global smoothness preservation*, Rend. Circ. Mat. Palermo (2) Suppl. **33** (1993), 259–279.

[18] H.B. Curry, I.J. Schoenberg: *On Pólya frequency functions. IV: The fundamental spline functions and their limits*, J. Anal. Math. **17** (1966), 71–107.

[19] C. de Boor: *A Practical Guide to Splines.* New York et al.: Springer 1978.

[20] C. de Boor, R.A. DeVore: *A geometric proof of total positivity for spline interpolation*, Math. Comp. **172** (1985), 497–504.

[21] R.A. DeVore: *The Approximation of Continuous Functions by Positive Linear Operators. Lecture Notes in Mathematics.* Berlin et al.: Springer 1972.

[22] Z. Ditzian: *Direct estimates for Bernstein polynomials*, J. Approx. Theory **79** (1994), 165–166.

[23] Z. Ditzian, K.G. Ivanov: *Strong converse inequalities*, J. d'Analyse Math. **61** (1993), 61–111.

[24] Z. Ditzian, V. Totik: *Moduli of Smoothness.* New York: Springer 1987.

[25] H. Esser: *On pointwise convergence estimates for positive linear operators on $C[a,b]$*, Indag. Math. **38** (1976), 189–194.

[26] G. Farin: *Curves and Surfaces for Computer–Aided Geometric Design. A Practical Guide.* San Diego et al.: Academic Press 1997.

[27] M. Felten: *Direct and inverse estimates for Bernstein polynomials*, Constr. Approx. **14** (1998), 459–468.

[28] M. Felten: *Local and global approximation theorems for positive linear operators*, J. Approx. Theory **94** (1998), 396–419.

[29] Y.–Y. Feng, G.–Z. Chang: *A simple proof of the iterated limit for variation diminishing operator of spline functions* (Chinese), J. Systems Sci. Math. Sci. **5** (1984), no. 3, 165–172.

[30] B. Germain Bonne, P. Sablonnière: *Comparaison des formes de courbes parametrées et de leurs approximants–spline.* Publ. n. **76** du Laboratoire de Calcul de l'Université des Sciences et Techniques de Lille 1976.

[31] H.H. Gonska: *Quantitative Aussagen zur Approximation durch positive lineare Operatoren.* Ph.D. Thesis, University of Duisburg 1979.

[32] H.H. Gonska: *Quantitative Korovkin–type theorems on simultaneous approximation*, Math. Z. **186** (1984), 419–433.

[33] H.H. Gonska: *On approximation by linear operators: improved estimates.* Anal. Numér. Théor. Approx. **14** (1985), 7–32.

[34] H.H. Gonska: *Quantitative Approximation in* $C(X)$. Habilitationsschrift, Universität Duisburg 1985. Duisburg: Schriftenreihe des Fachbereichs Mathematik der Universität Duisburg **SM–DU–94** (1986).

[35] H.H. Gonska: *Simultaneous approximation by algebraic blending functions.* In: "Proc. Alfred Haar Memorial Conf.", Budapest 1985 (ed. by J. Szabados, K. Tandori). Colloq. Soc. János Bolyai **49**, 363–382. Amsterdam: North–Holland 1987.

[36] H.H. Gonska: *Degree of simultaneous approximation of bivariate functions by Gordon operators*, J. Approx. Theory **62** (1990), 170–191.

[37] H.H. Gonska, R. Păltănea: *General estimates for the Ditzian–Totik modulus*, manuscript 2001.

[38] H.H. Gonska, G. Tachev: *The second Ditzian–Totik modulus revisited: refined estimates for positive linear operators.* To appear in Anal. Numér. Théor. Approx. **30** (2001).

[39] T.N.T. Goodman: *Shape preserving representations.* In: "Mathematical Methods in Computer Aided Geometric Design" (ed. by T. Lyche and L.L. Schumaker), 333–351. Boston: Acad. Press 1989.

[40] T.N.T. Goodman: *Asymptotic formulas for multivariate Bernstein–Schoenberg operators*, Constr. Approx. **11** (1995), 439–453.

[41] T.N.T. Goodman, S.L. Lee: *Spline approximation operators of Bernstein–Schoenberg type in one and two variables*, J. Approx. Theory **33** (1981), 248–263.

[42] T.N.T. Goodman, S.L. Lee, A. Sharma: *Asymptotic formula for the Bernstein–Schoenberg operator*, Approx. Theory Appl. **4** (1988), 67–86.

[43] T.N.T. Goodman, A. Sharma: *A property of Bernstein–Schoenberg spline operators*, Proc. Edinburgh Math. Soc. **28** (1985), 333–340.

[44] T.N.T. Goodman, A. Sharma: *A modified Bernstein–Schoenberg operator.* In: "Constructive Theory of Functions" (Varna, 1987), 166–173. Bulgar. Acad. Sci., Sofia 1988.

[45] T.N.T. Goodman, A. Sharma: *A Bernstein type operator on the simplex*, Math. Balkanica **5** (1991), 129–145.

[46] W.J. Gordon, R.F. Riesenfeld: *B–spline curves and surfaces.* In: "Computer Aided Geometric Design" (Proc. Int. Conf. on Computer Aided Design; ed. by R.E. Barnhill, R.F. Riesenfeld), 95–126. New York: Acad. Press 1974.

[47] J. Hoschek, D. Lasser: *Fundamentals of Computer Aided Geometric Design.* Wellesley, MA: A K Peters 1993.

[48] D. Kacsó: *On the degree of simultaneous approximation.* In: "RoGer 2000 – Braşov" (Proc. 4th Romanian–German Seminar on Approximation Theory and its Applications, Braşov 2000; ed. by H. Gonska, D. Kacsó and L. Beutel), 62–69. Duisburg: Schriftenreihe des Fachbereichs Mathematik der Gerhard–Mercator–Universität **SM–DU–485** (2000).

[49] D. Kacsó: *Simultaneous approximation by almost convex operators.* To appear in: "Rend. Circ. Mat. Palermo Suppl." (Proc. 4th International Conference on Functional Analysis and Approximation Theory, Maratea 2000).

[50] H.–B. Knoop, P. Pottinger: *Ein Satz vom Korovkin–Typ für C^k–Räume*, Math. Z. **148** (1976), 23–32.

[51] H.–B. Knoop, X.–l. Zhou: *The lower estimate for linear positive operators (II)*, Resultate Math. **25** (1994), 315–330.

[52] J.M. Lane, R.F. Riesenfeld: *A geometric proof of the variation diminishing property of B–spline approximation*, J. Approx. Theory **37** (1983), 1–4.

[53] D. Leviatan: *On the representation of the remainder in the variation–diminishing spline approximation*, J. Approx. Theory **7** (1973), 63–70.

[54] A. Lupaş: *Contribuţii la teoria aproximării prin operatori liniari.* Ph.D. Thesis, Babeş–Bolyai University, Cluj–Napoca 1975.

[55] M.J. Marsden: *An identity for spline functions with applications to variation–diminishing spline approximation*, J. Approx. Theory **3** (1970), 7–49.

[56] M.J. Marsden: *On uniform spline approximation*, J. Approx. Theory **6** (1972), 249–253.

[57] M.J. Marsden: *A Voronovskaya theorem for variation–diminishing spline approximation*, Canad. J. Math. **38** (1986), 1081–1093.

[58] M.J. Marsden, I.J. Schoenberg: *On variation diminishing spline approximation methods*, Mathematica **31** (1966), 61–82.

[59] M.J. Marsden, S.D. Riemenschneider: *Asymptotic formulae for variation–diminishing splines.* In: "Second Edmonton Conference on Approximation Theory" (Proc. 1982 Sem. on Approximation Theory, Edmonton, Alberta; ed. by Z. Ditzian et al.), 255–261. Providence, R.I.: Amer. Math. Soc. 1983.

[60] W.W. Meyer, D.H. Thomas: *Variations on a theme by Schoenberg*, J. Approx. Theory **18** (1976), 39–49.

[61] W.W. Meyer, D.H. Thomas: *Erratum: Variations on a theme by Schoenberg*, J. Approx. Theory **25** (1979), 91.

[62] M.J. Munteanu, L.L. Schumaker: *Some multidimensional spline approximation methods*, J. Approx. Theory **10** (1974), 23–40.

[63] M.W. Müller: *Degree of L_p approximation by integral Schoenberg splines*, J. Approx. Theory **21** (1977), 385–393.

[64] M.W. Müller: *On the degree of L_p–approximation by integral Schoenberg splines.* In: "Fourier Analysis and Approximation Theory" (Proc. Conf. Budapest 1976; ed. by G. Alexits, P. Turán), 565–570. Amsterdam: North Holland 1978.

[65] M.W. Müller: *Über die Methode der integrierten Schoenberg–Splines.* In: "Constructive Function Theory" (Proc. Conf. Varna 1981; ed. by Bl. Sendov et al.), 442–446. Sofia: Publishing House of the Bulgarian Academy of Sciences 1983.

[66] D.J. Newman: *The Zygmund condition for polygonal approximation*, Proc. Amer. Math. Soc. **45** (1974), 303–304.

[67] J. Nitsche: *Sätze vom Jackson–Bernstein–Typ für die Approximation mit Spline–Funktionen*, Math. Z. **109** (1969), 97–106.

[68] R. Păltănea: *Optimal estimates with moduli of continuity*, Result. Math. **32** (1997), 318–331.

[69] R.F. Riesenfeld: *Applications of B–spline Approximation to Geometric Problems of Computer Aided Design.* Ph.D. Thesis, Syracuse University 1973.

[70] P. Sablonnière: *Positive spline operators and orthogonal splines*, J. Approx. Theory **52** (1988), 28–42.

[71] K. Scherer: *Über die beste Approximation von L_p–Funktionen durch Splines.* In: "Constructive Theory of Functions" (Proc. Int. Conf. Varna 1970, ed. by B. Penkov, D. Vacov), 277–286. Sofia: Izdat. Bolgar. Akad. Nauk 1972.

[72] I.J. Schoenberg: *On spline functions, with a supplement by T.N.E. Greville.* In: "Inequalities" (Proc. Symposium Wright–Patterson Air Force Base, August 1965, ed. by O. Shisha) 255–291. New York: Acad. Press 1967.

[73] L.L. Schumaker: *Spline Functions: Basic Theory.* New York: Wiley 1981.

[74] M.Sh. Shabozov: *Exact bounds for simultaneous approximation of functions of two variables and their derivatives by bilinear splines*, Mathematical Notes **59** (1996), 104–111.

[75] O. Shisha: *A characterization of functions having Zygmund's property*, J. Approx. Theory **9** (1973), 395–397.

[76] D.D. Stancu: *A generalization of the Schoenberg approximation spline operator*, Studia Univ. Babeş–Bolyai Math. **26** (1981), no. 2, 37–42.

[77] F. Stancu: *Approximation of functions of several variables by means of a class of spline operators.* In: "Itinerant Seminar on Functional Equations, Approximation and Convexity" (Cluj–Napoca 1983), 153–158. Preprint 83–2, Univ. Babeş–Bolyai, Cluj–Napoca 1983.

[78] G. Tachev: *Piecewise linear interpolation with nonequidistant nodes*, Num. Funct. Anal. Appl. **21** (2000), 945–953.

[79] V. Totik: *Approximation by Bernstein polynomials*, Amer. J. Math. **116** (1994), 995–1018.

[80] H.–J. Wenz: *On the limits of (linear combinations of) iterates of linear operators*, J. Approx. Theory **89** (1997), 219–237.

[81] V.V. Žuk: *Functions of the Lip 1 class and S.N. Bernstein's polynomials* (Russian), Vestnik Leningrad. Univ. Mat. Mekh. Astronom., (1) 1989, 25–30, 122–123.

Monografías de la Academia de Ciencias de Zaragoza. **20:** 59–67, (2002).

On saturation in conservative approximation*

D. Cárdenas-Morales† and P. Garrancho‡

† Departamento de Matemáticas. Universidad de Jaén. 23071 Jaén. Spain.

email: `cardenas@ujaen.es`

‡ Departamento de Matemáticas. I.E.S. Martín Halaja. Jaén. Spain.

Abstract

In this paper we state a pointwise saturation result for sequences of linear operators that preserve the sign of the k-th derivative of the functions. We apply it to some well known sequences of operators.

1. Notation and introduction

Let $A \subset \mathbb{R}$, $i \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. As usual, we denote by $C^i(A)$ the space of all real-valued i-times continuously differentiable functions defined on A and by D^i the i-th differential operator. $C^i_B(A)$ denotes the subspace formed by the functions of $C^i(A)$ which are bounded on A, and we write e_i for the polynomial $e_i(t) = t^i$. A function $f \in C^i(A)$ is said to be i-convex if $D^i f \geq 0$ on A and a linear operator is said to be i-convex if it maps i-convex functions onto i-convex functions.

Now let I be a closed real interval, let $k \in \mathbb{N}_0$ and let $L_n : C^k(I) \longrightarrow C^k(I)$ be a sequence of linear operators satisfying the following asymptotic condition:

A) there exist a sequence λ_n of real positive numbers, and a function $p \in C^k(I)$ strictly positive on $\mathrm{Int}(I)$ such that for all $g \in C^k_B(I)$, $k+2$-times differentiable in some neighborhood of a point $x \in \mathrm{Int}(I)$,

$$\lim_{n \to +\infty} \lambda_n \left(D^k L_n g(x) - D^k g(x) \right) = D^k \left(p D^2 g \right)(x). \tag{1}$$

---

[*] This work is partially supported by Junta de Andalucía, Grupo de Investigación FQM 0178 and by Ministerio de Ciencia y Tecnología, Proyecto de Investigación BFM 2000-0911.

Some recent papers have proved that this formula is satisfied for many known operators (see [1], [4], [7]). It informs that the speed of convergence of $D^kL_ng(x)-D^kg(x)$ to 0 cannot overcome in general the one of λ_n^{-1}. Those functions f that satisfy $D^kL_nf(x)-D^kf(x)=o(\lambda_n^{-1})$ for $x\in(a,b)$ with $a,b\in Int(I)$ form the trivial class for the local saturation problem of D^kL_n, while the functions that satisfy $D^kL_nf(x)-D^kf(x)=O(\lambda_n^{-1})$ for $x\in(a,b)$ form the so-called saturation class. Recently, in [5] the authors have found these classes assuming also the following shape preserving property:

B) for all $n\in\mathbb{N}, L_n$ is k-convex.

They have extended the results obtained by Mühlbach [11] for the case $k=0$, taking into account the outstanding work by Lorent and Schumaker [9] in 1972. Simultaneously, Berens [3] dealt with this matter from a more general point of view but also just for $k=0$. Here the basic tools were convexity arguments through the use of the theory of extended complete Tchebycheff systems (ECT-Systems) and a generalization of the parabola technique introduced by Bajsanski and Bojanić [2].

In the present paper we assume A) and B) and prove a pointwise saturation result for D^kL_n that extends this last one in the sense of considering $k>0$. We will also apply it in the last section to the well-known operators of Bernstein and Szász-Mirakjan. Firstly, we prove the general formulation of the parabola technique we shall use here.

Lemma 1 *Let $0=Ly:=D^2y+a_1(t)D^1y+a_2(t)y=0$ be a second-order linear differential equation with $a_1,a_2\in C(I)$ and assume that it has a unique solution taking any two given real values at any two given points within $Int(I)$. Let $g\in C(I)$ and $t_1,t_2\in Int(I)$. If $f\in C(I)$ verifies that $f(t_1)=f(t_2)=0$ and $f(t_0)>0$ for some $t_0\in(t_1,t_2)$, then there exist a real constant $\alpha<0$ and a solution of the previous differential equation, $\tilde{y}$, such that for all $t\in[t_1,t_2]$, $\alpha g(t)+\tilde{y}(t)>f(t)$ and for some $s\in(a,b)$, $\alpha g(s)+\tilde{y}(s)=f(s)$.*

Proof Let L_g be the unique solution of $Ly=0$ satisfying $L_g(t_i)=g(t_i),\ i=1,2$, let y_0 be a solution of $Ly=0$ satisfying $y_0(t)>0\ \forall t\in[t_1,t_2]$ (whose existence is garanteed taking into account that $t_1,t_2\in Int(I)$), and let $\epsilon>0$ be sufficiently small so that $f(t_0)-\epsilon\left(L_g(t_0)-g(t_0)\right)>0$. Then the function

$$\frac{f-\epsilon\left(L_g-g\right)}{y_0}$$

is continuous in $[t_1,t_2]$, it vanishes at the end points of this interval and it is strictly positive at the point t_0, so it reaches a maximum value, say M, at a point $s\in(t_1,t_2)$. Consequently, for all $t\in[t_1,t_2]$

$$\epsilon\left(L_g(t)-g(t)\right)+My_0(t)\geq f(t)$$

and

$$\epsilon\left(L_g(s)-g(s)\right)+My_0(s)=f(s).$$

Now the proof is over taking $\alpha=-\epsilon$ and $\tilde{y}=\epsilon L_g+My_0$. $\square$

2. The results

Along this section we assume that the operators L_n defined in Section 1. satisfy A) and B). In [5] we proved the following result

Lemma 2 *a) For $k\in\mathbb{N}$ the ordinary linear differential equation*

$$D^k(pD^2y)\equiv 0, \tag{2}$$

has a fundamental system of solutions of the form $\{e_0,\dots,e_{k-1},y_0,y_1\}$, and using the change of variable $z=D^ky$ it can be reduced to

$$D^2z+\frac{kD^1p}{p}D^1z+\frac{k(k-1)D^2p}{2p}z\equiv 0 \tag{3}$$

(p is necessarily a polynomial of degree less than or equal to 2).

b) If $f\in C_B^k(I)$ is a solution of (2) on some neighborhood of a point $x\in Int(I)$, then

$$D^kL_nf(x)-D^kf(x)=o(\lambda_n^{-1}).$$

c) Let $f,g\in C_B^k(I)$. If $D^kf\le D^kg$ on some neighborhood of a point $x\in Int(I)$, then

$$D^kL_nf(x)\le D^kL_ng(x)+o(\lambda_n^{-1}).$$

In the sequel, if not specified in other sense, solutions of equation (2) and (3) are understood on $\mathrm{Int}(I)$.

Take $a,b\in Int(I)$ with $a<b$ and assume that (3) has a fundamental system of solutions, say $\{z_0,z_1\}$, which form an ECT-System on (a,b). We write it as in [6], from the functions w_0 and w_1:

$$z_0(t)=w_0(t),\quad z_1(t)=w_0(t)\int_a^t w_1(s)ds. \tag{4}$$

The next lemma shows the relation between convexity respect to this ECT-System and approximation by D^kL_n.

Lemma 3 *Let $f\in C^k(I)$. If*

$$\limsup_{n\to\infty}\lambda_n\left(D^kL_nf(t)-D^kf(t)\right)\ge 0,\ t\in(a,b),$$

then D^kf is convex on (a,b) with respect to z_0,z_1.

Proof Assuming the contrary, there exist $a < t_1 < t_0 < t_2 < b$ such that $D^k f(t_0) > z(t_0)$, being $z = z(t)$ the unique solution of (3) satisfying $z(t_i) = D^k f(t_i)$, $i = 1, 2$. Let us apply Lemma 1 to (3) and $D^k f - z$, with $g = D^k w$, being $w \in C^k(I)$ such that $D^2 w(t) = e_k(t)/p(t)$ for all $t \in [t_1, t_2]$. Then there exist a solution of (3), say $\tilde{z}$, and a constant $\alpha < 0$ verifying that for all $t \in [t_1, t_2]$

$$D^k f(t) - z(t) \leq \alpha D^k w(t) + \tilde{z}(t)$$

and for some $s \in (t_1, t_2)$

$$D^k f(s) - z(s) = \alpha D^k w(s) + \tilde{z}(s).$$

Now if we take $y, \tilde{y} \in C^k_B(I)$, solutions of (2) on (t_1, t_2) such that $D^k y(t) = z(t)$ and $D^k \tilde{y}(t) = \tilde{z}(t)$, then using c), Lemma 2,

$$\lambda_n \left(D^k L_n f(s) - D^k f(s) \right) \leq \alpha \lambda_n \left(D^k L_n w(s) - D^k w(s) \right)$$

$$+ \lambda_n \left(D^k L_n \tilde{y}(s) - D^k \tilde{y}(s) \right) + \lambda_n \left(D^k L_n y(s) - D^k y(s) \right) + o(1).$$

Using b), Lemma 2 for y and $\tilde{y}$ and the asymptotic condition A) for w we obtain

$$\lambda_n \left(D^k L_n f(s) - D^k f(s) \right) \leq \alpha k! + o(1)$$

what is a contradiction that ends the proof. □

Now we prove the main result. We shall obtain some information about functions f that verify $D^k L_n f(x) - D^k f(x) = o(\lambda_n^{-1})$ and $D^k L_n f(x) - D^k f(x) = O(\lambda_n^{-1})$ for $x \in (a, b)$, though we shall consider a more general framework. For this purpose we define $\varphi_x(t) := (t - x)^{k+2}/(k+2)!$ and $\mu_n(x) := D^k L_n \varphi_x(x)$. Notice that from A) $\mu_n(x) = O(\lambda_n^{-1})$, specifically

$$\lambda_n \mu_n(x) = p(x) + o(1). \tag{5}$$

Firstly we prove the following technical lemma which shall be used in the proof of Theorem 1. In both of them we use the functions $w_2 := 1/w_0 w_1$ and $W_2(t) := \int_a^t w_2(s) ds$.

Lemma 4 *Let $h \in C[a, b]$ and $H \in C^k(I)$ be such that for all $t \in (a, b)$, $D^k H(t) = w_0(t) \int_a^t h(s) w_1(s) ds$. Then, for $x \in (a, b)$,*

$$\limsup_{n \to \infty} \frac{D^k L_n H(x) - D^k H(x)}{\mu_n(x)} \leq \limsup_{t \to x} \frac{h(t) - h(x)}{W_2(t) - W_2(x)}$$

and

$$\liminf_{t \to x} \frac{h(t) - h(x)}{W_2(t) - W_2(x)} \leq \liminf_{n \to \infty} \frac{D^k L_n H(x) - D^k H(x)}{\mu_n(x)}.$$

Proof We shall only prove the first inequality; the other one works similarly. Let $x \in (a, b)$. We assume that there exists a real number m such that

$$\limsup_{t \to x} \frac{h(t) - h(x)}{W_2(t) - W_2(x)} < m$$

because if this is not the case there is nothing to prove. Then for some $\delta > 0$ whenever $|t - x| < \delta$,

$$\frac{h(t) - h(x)}{W_2(t) - W_2(x)} < m.$$

So for a sufficiently small δ we have

$$\frac{h(t) - h(x)}{(t - x)w_2(x)} < m.$$

Multiplying by $w_1(t)$ and integrating we have

$$\int_x^t (h(s) - h(x))\, w_1(s)ds < m w_2(x) \int_x^t (s - x) w_1(s)ds,$$

which, taking into account that

$$\frac{D^k H(t)}{w_0(t)} - \frac{D^k H(x)}{w_0(x)} = \int_x^t h(s) w_1(s)ds,$$

provides

$$\frac{D^k H(t)}{w_0(t)} - \frac{D^k H(x)}{w_0(x)} - h(x) \int_x^t w_1(s)ds < m w_2(x) \int_x^t (s - x) w_1(s)ds.$$

Multiplying now by $w_0(t)$ and considering $W_1(t) := \int_a^t w_1(s)ds$ we obtain

$$D^k H(t) - \frac{D^k H(x) w_0(t)}{w_0(x)} - h(x) \left(z_1(t) - W_1(x) w_0(t)\right)$$

$$< m w_2(x) w_0(t) \int_x^t (s - x) w_1(s)ds.$$

Equivalently, taking $y_0, y_1, Y \in C^k(I)$ such that their k-th derivatives coincide respectively with z_0, z_1 and $w_2(x) w_0(t) \int_x^t (s - x) w_1(s)ds$ in the neighborhood of the point x we are dealing with,

$$D^k H(t) - \frac{D^k H(x) D^k y_0(t)}{w_0(x)} - h(x) \left(D^k y_1(t) - W_1(x) D^k y_0(t)\right) < m D^k Y(t).$$

Applying c), Lemma 2

$$D^k L_n H(x) - \frac{D^k H(x) D^k L_n y_0(x)}{w_0(x)} - h(x) \left(D^k L_n y_1(x) - W_1(x) D^k L_n y_0(x)\right)$$

$$\leq m D^k L_n Y(x) + o(\lambda_n^{-1}).$$

Introducing the zero terms $-D^kH(x)+\frac{D^kH(x)}{z_0(x)}D^ky_0(x)$ and $-D^ky_1(x)+D^ky_0(x)W_1(x)$ (notice that $z_1(x)=z_0(x)W_1(x)$), and regrouping,

$$D^kL_nH(x)-D^kH(x)-\frac{D^kH(x)}{z_0(x)}\left(D^kL_ny_0(x)-D^ky_0(x)\right)$$

$$-h(x)\left(D^kL_ny_1(x)-D^ky_1(x)-W_1(x)\left(D^kL_ny_0(x)-D^ky_0(x)\right)\right)$$

$$\leq mD^kL_nY(x)+o(\lambda_n^{-1}).$$

Applying b), Lemma 2 to the functions y_0 and y_1, and hypothesis A) to Y,

$$\lambda_n\left(D^kL_nH(x)-D^kH(x)\right)\leq mD^k\left(pD^2Y\right)(x)+o(1)$$

$$=mp(x)w_2(x)w_0(x)w_1(x)+o(1)=mp(x)+o(1),$$

where for the last equalities we have done some calculations taking into account the definitions of Y, w_2 and that p is a polynomial of degree less than or equal to 2. Finally, using (5) and taking $\limsup_{n\to\infty}$ we obtained

$$\limsup_{n\to\infty}\frac{D^kL_nH(x)-D^kH(x)}{\mu_n(x)}\leq m,$$

and the proof is over. □

Theorem 1 *Let $f\in C^k(I)$ and suppose that ψ is a finitely valued function in $L_1[a,b]$ for which*

$$\liminf_{n\to\infty}\frac{D^kL_nf(x)-D^kf(x)}{\mu_n(x)}\leq\psi(x)\leq\limsup_{n\to\infty}\frac{D^kL_nf(x)-D^kf(x)}{\mu_n(x)}.$$

Then there exist two constants α_0 and α_1 such that for all $t\in(a,b)$,

$$D^kf(t)=\alpha_0z_0(t)+\alpha_1z_1(t)+w_0(t)\int_a^t w_1(s)\int_a^s\psi(v)w_2(v)dvds.$$

Proof Let $G\in C^k(I)$ such that for all $t\in(a,b)$

$$D^kG(t)=D^kf(t)-w_0(t)\int_a^t w_1(s)\int_a^s\psi(v)w_2(v)dvds.$$

We shall prove that D^kG is convex and concave in (a,b) with respect to z_0 and z_1.

For $q\in\mathbb{N}$ let m_q and M_q be respectively the minor and major functions of ψ with respect to w_2, such that

$$\left|m_q(t)-\int_a^t\psi(s)w_2(s)ds\right|<\frac{1}{q},\ t\in(a,b),$$

$$\left|M_q(t)-\int_a^t\psi(s)w_2(s)ds\right|<\frac{1}{q},\ t\in(a,b),$$

whose existence is well known from the theory of Lebesgue integration (see for instance [12]). In particular it follows that

$$\limsup_{t\to x}\frac{m_q(t)-m_q(x)}{W_2(t)-W_2(x)}\le\psi(x)\le\liminf_{t\to x}\frac{M_q(t)-M_q(x)}{W_2(t)-W_2(x)}.$$

From Lemma 4 and the hypothesis, if we consider $\tilde{m}_q \in C^k(I)$ such that for all $t\in(a,b)$ $D^k\tilde{m}_q(t)=w_0(t)\int_a^t m_q(s)w_1(s)ds$, we have that

$$\limsup_{n\to\infty}\frac{D^kL_n\tilde{m}_q(x)-D^k\tilde{m}_q(x)}{\mu_n(x)}\le\limsup_{t\to x}\frac{m_q(t)-m_q(x)}{W_2(t)-W_2(x)}$$

$$\le\psi(x)\le\limsup_{n\to\infty}\frac{D^kL_nf(x)-D^kf(x)}{\mu_n(x)},$$

so

$$\limsup_{n\to\infty}\frac{D^kL_n\left(f-\tilde{m}_q\right)(x)-D^k\left(f-\tilde{m}_q\right)(x)}{\mu_n(x)}\ge 0.$$

From (5) and Lemma 3, we deduce that for all $q\in\mathbb{N}$ $D^k\left(f-\tilde{m}_q\right)$ is convex in (a,b) with respect to z_0 and z_1. Letting q tend to infinity we conclude that this also holds for D^kG. Analogously from M_q we obtain that D^kG is concave in (a,b) with respect to z_0 and z_1. □

Remark This theorem recovers the converse result of b), Lemma 2 that was stated in [5]. Indeed, if $D^kL_nf(x)-D^kf(x)=o(\lambda_n^{-1})$, then $D^kL_nf(x)-D^kf(x)=o(\mu_n(x))$ and the theorem applies with $\psi\equiv 0$.

3. Applications

In this section we apply the previous result to the Bernstein and Szász-Mirakjan operators defined as follows respectively on $C[0,1]$ and $C[0,\infty)$:

$$B_nf(t)=\sum_{p=0}^{n}f\left(\frac{p}{n}\right)\binom{n}{p}t^p(1-t)^{n-p},$$

$$S_nf(t)=e^{-nt}\sum_{p=0}^{\infty}f\left(\frac{p}{n}\right)\frac{n^pt^p}{p!}.$$

It is very well-known (see [8], [10]) that they are convex of any order , i.e. B) holds true for any value of $k\in\mathbb{N}_0$. The validity of A) for B_n with $k=0$, $\lambda_n=2n$, $p(t)=t(1-t)$, and for S_n with $k=0$, $\lambda_n=2n$, $p(t)=t$ follows from classical results of Voronovskaya [14] and Szász [13]. Specifically, under the aforementioned conditions,

$$\lim_{n\to+\infty}2n\left(B_ng(x)-g(x)\right)=x(1-x)D^2g(x), \tag{6}$$

and

$$\lim_{n\to+\infty} 2n\left(S_n g(x) - g(x)\right) = xD^2 g(x). \tag{7}$$

Roughly speaking, one can apply the differential operator D^k for any $k \in \mathbb{N}$ to both sides of the identities, what yields that A) holds true for B_n, S_n and all $k \in \mathbb{N}$ taking for λ_n and p the corresponding values above (see [4], [7], [1]).

Hence we can apply our result to these operators. The following table contains for each operator and for $k > 0$ the values of λ_n and $p(t)$, and a choice for $w_0(t)$ and $w_1(t)$. We do not apply our result to the case $k = 0$ because this can be done from [3].

	B_n	S_n
λ_n	$2n$	$2n$
$p(t)$	$t(1-t)$	t
$w_0(t)$	$1/t^{k-1}$	1
$w_1(t)$	$t^{k-2}/(1-t)^k$	$1/t^k$
$w_2(t)$	$t(1-t)^k$	t^k

From Theorem 1, the following corollaries are easily obtained.

Corollary 1 *Let $k \in \mathbb{N}$, $0 < a < b < 1$, $f \in C^k[0,1]$, $\mu_n(x) = D^k B_n \varphi_x(x)$ and suppose that ψ is a finitely valued function in $L_1[a,b]$ such that*

$$\liminf_{n\to\infty} \frac{D^k B_n f(x) - D^k f(x)}{\mu_n(x)} \le \psi(x) \le \limsup_{n\to\infty} \frac{D^k B_n f(x) - D^k f(x)}{\mu_n(x)}.$$

Then there exist two constants α_0 and α_1 such that for all $t \in (a,b)$, one has

$$D^1 f(t) = \alpha_0 + \alpha_1 \log \frac{t}{1-t} + \int_a^t \frac{1}{s(1-s)} \int_a^s \psi(v) v(1-v)\, dv ds,$$

for $k = 1$ and

$$D^k f(t) = \frac{\alpha_0}{t^{k-1}} + \frac{\alpha_1}{(1-t)^{k-1}} + \frac{1}{t^{k-1}} \int_a^t \frac{s^{k-2}}{(1-s)^k} \int_a^s \psi(v) v(1-v)^k\, dv ds,$$

for $k > 1$.

Corollary 2 *Let $k \in \mathbb{N}$, $0 < a < b$, $f \in C_B^k[0,\infty)$, $\mu_n(x) = D^k S_n \varphi_x(x)$ and suppose that ψ is a finitely valued function in $L_1[a,b]$ such that*

$$\liminf_{n\to\infty} \frac{D^k S_n f(x) - D^k f(x)}{\mu_n(x)} \le \psi(x) \le \limsup_{n\to\infty} \frac{D^k S_n f(x) - D^k f(x)}{\mu_n(x)}.$$

Then there exist two constants α_0 and α_1 such that for all $t \in (a,b)$, one has

$$D^1 f(t) = \alpha_0 + \alpha_1 \log t + \int_a^t \frac{1}{s} \int_a^s \psi(v) v\, dv ds,$$

for $k = 1$ and

$$D^k f(t) = \alpha_0 + \frac{\alpha_1}{t^{k-1}} + \int_a^t \frac{1}{s^k} \int_a^s \psi(v) v^k\, dv ds,$$

for $k > 1$.

References

[1] U. Abel, M. Ivan, Asymptotic expansion of the Jakimovski-Leviatan operators and their derivatives, Proc. of the F. Alexits Conference, Budapest (1999) (to appear).

[2] B. Bajsanski, R. Bojanić, A note on approximation by Bernstein polynomials Bull. Amer. Math. Soc., 70 (1964), 675-677.

[3] H. Berens, Pointwise saturation of positive operators, J. Approx. Theory, 6 (1972), 135-146.

[4] D. Cárdenas-Morales, P. Garrancho, F. J. Muñoz-Delgado, A result on asymptotic formulae for linear k-convex operators, Int. J. Differ. Equ. Appl., 2, no.3, (2001), 335-347.

[5] D. Cárdenas-Morales, P. Garrancho, Local saturation of conservative operators, Acta Math. Hung. (to appear).

[6] S. J. Karlin, W. J. Studden, Tchebycheff Systems, Interscience, New York (1966).

[7] A. J. López-Moreno, Expresiones y estimaciones de operadores lineales conservativos, Doctoral Thesis, University of Jaén, Spain (2001).

[8] G. G. Lorentz, Bernstein Polynomials, Chelsea Publishing Company, New York (1986).

[9] G. G. Lorentz, L. L. Schumaker, Saturation of Positive Operators, J. Approx. Theory, 5 (1972), 413-424.

[10] A. Lupas, Some properties of the linear positive operators (I), Mathematica, Cluj, 9 (1967), 77-83.

[11] G. Mühlbach, Operatoren vom Bernsteinschen Typ, J. Approx. Theory, 3 (1970), 274-292.

[12] I. Pesin, Classical and Modern Integration Theories, Academic Press, New York (1951).

[13] O. Szász, Generalization of S. Bernstein's polynomials to the infinite interval, J. Res. Nat. Bur. Standards, 45 (1950), 239-245: Collected Mathematical Works, Cincinnati (1955), 1401-1407.

[14] E. Voronovskaya, Détermination de la forme asymptotique d'approximation des fonctions par les polynômes de S. Bernstein, Dokl. Akad. Nauk. USSR, A (1932), 79-85.

Monografías de la Academia de Ciencias de Zaragoza. **20:** 69–76, (2002).

An iterative scheme to approach the asymptotic behaviour of a Kolmogorov system *

M. Gámez and M^a. I. López

Department of Statistics and Applied Mathematics.

University of Almería. Spain.

Abstract

In this paper we consider a biological system consisting of several preys and several predators. We study the existence of a global attractor for such a system and obtain an approximation to the model's solution by means a numerical method.

1 Introduction.

In the last few years many papers have been devoted to the dynamics of applied populations to Biology. From a mathematical point of view it's of a great interest to determine qualitative properties on these differential systems (see [1], [2], [3]) which give information about the behaviour of the solutions, due to the impossibility, in most cases to solve these systems explicitly. For this reason, the numerical methods are reaching an important role in this subject, in order to determine some properties of the solutions of these systems.

We consider a biological Kolmogorov system consisting of several preys and several predators. The case of the usual predator-prey system has been extensively studied by many authors. For instance, see [1] for optimal results.

We show some results about the logistic equation adapted to the notation necessary for the Kolmogorov system studied. Using an iterative scheme, we find the existence of a global attractor for the positive solutions of Kolmogorov system studied, which determine an approximation of the solution of this system.

Finally, we present concrete examples determining these approximations with the help of MATHEMATICA and we compare the results with the numerical resolution of them using the Populus software. Therefore, we verify how these systems can be used to model a process of biological fight and we get in this way another tool which help us to know the development of some biological species.

---

*This paper was sponsored by the UAL-CAJAMAR 2001 CR-UAL-011

2 The Logistic Equation.

In this section we introduce some notations and we state some interesting properties, which will be basic in our study, of the periodic logistic equation,

$$x' = xF(t,x) \quad , \; x \geq 0. \tag{1}$$

Given $T > 0$, we denote by $\mathcal{C}_T$ the set of all continuous function $F : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ such that:

a) $F(t,x)$ is T-periodic in t and locally Lipschitz continuous in x.
b) $F(t,x)$ is decreasing in x.
c) $F(\tau,x)$ is strictly decreasing in x for some $\tau = \tau(F) \in \mathbb{R}$.
e) There exists $R = R(F) > 0$ satisfying $\int_0^T F(t,R)dt < 0$.

In "Iterative schemes for some population models"(Nonlinear Word, 3, (1996), 695-708), the following results were proved by A. Tineo.

Theorem 2.1 *If $F \in \mathcal{C}_T$ then equation (1) has a T-periodic solution U^F, which is globally asymptotically stable. That is, if u is a solution of (1) and $u(0) > 0$, then u is defined on $[0,\infty)$ and*

$$u(t) - U^F(t) \to 0 \quad as \quad t \to +\infty.$$

Moreover, $U^F > 0$ if $\int_0^T F(t,0)dt > 0$, and $U^F \equiv 0$ if $\int_0^T F(T,0)dt \leq 0$.

We say that U^F is the "*global attractor*" of (1).

Corollary 2.2 *Let $F,\ G \in \mathcal{C}_T$ and suppose $F \leq G$. Then $U^F \leq U^G$.*

Theorem 2.3 *Let $\{F_n\}$ be a sequence in $\mathcal{C}_T$ converging to $F \in \mathcal{C}_T$ uniformly on compact sets. Then $U^{F_n}(t) \to U^F(t)$, uniformly on $\mathbb{R}$.*

3 Kolmogorov System.

In this section we study a predator-prey model for a biological community consisting of $n-$prey and $m-$predators developed under a Kolmogorov system. In a more precise way we suppose the following system,

$$\begin{aligned} x_i' &= x_i f_i(t, x_1, \dots, x_n, y_1, \dots, y_m) \quad 1 \leq i \leq n \\ y_j' &= y_j g_j(t, x_1, \dots, x_n, y_1, \dots, y_m) \quad 1 \leq j \leq m. \end{aligned} \tag{2}$$

where $f_i, g_j : \mathbb{R} \times \mathbb{R}_+^n \times \mathbb{R}_+^m \to \mathbb{R}$ are continuous functions, which are T-periodic in t and locally Lipschitz continuous in (x,y).

We shall assume that:

P_1) $f_i(t,x,y)$ is decreasing in $(x,y) \in \mathbb{R}^{n+m}_+$ and $g_j(t,x,y)$ is increasing in $x \in \mathbb{R}^n_+$ and decreasing in $y \in \mathbb{R}^m_+$.

P_2) There exist $\tau_i;\ \theta_j \in \mathbb{R}$ such that $f_i(\tau_i,x,y);\ g_j(\theta_j,x,y)$ are strictly decreasing in x_i, y_j respectively $(i=1,\dots,n;\ j=1,\dots,m)$.

P_3) There exists $R>0$ satisfying,

$$\int_0^T f_i(t,Re_i,0)dt < 0 \qquad 1 \le i \le n$$

$$\int_0^T g_j(t,U^1(t),R\nu_j)dt < 0 \qquad 1 \le j \le m.$$

Here $\{e_1,\dots,e_n\}$, $\{\nu_1,\dots,\nu_m\}$ denote the canonical vector basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and $U^1 = (U^1_1,\dots,U^1_n) : \mathbb{R} \to \mathbb{R}^n_+$, where $U^1_i;\ \ 1 \le i \le n$; is the global attractor of the equation,

$$z' = zf_i(t,ze_i,0) \quad 1 \le i \le n, \tag{3}$$

See Theorem 2.1.

4 An Iterative Scheme.

Associated to system (2), we have two sequences of nonnegative T-periodic functions $\{U^N = (U^N_1,\dots,U^N_n)\}$ and $\{V^N = (V^N_1,\dots,V^N_m)\}$, $N \in \mathbb{N}$, defined inductively as follows: $U^0 = V^0 \equiv 0$, and $U^{N+1}_i;\ \ 1 \le i \le n$; is the global attractor of the logistic equation,

$$z' = zf_i(t,U^N_1(t),\dots,U^N_{i-1}(t),z,U^N_{i+1}(t),\dots,U^N_n(t),V^N(t)), \tag{4}$$

and $V^N_j \ \ 1 \le j \le m$ the global attractor of the equation

$$z' = zg_j(t,U^{N+1}(t),V^N_1(t),\dots,V^N_{j-1}(t),z,V^N_{j+1}(t),\dots,V^N_m(t)). \tag{5}$$

Remark. The above scheme is obtained, using some ideas in Lopez-Gomez, Ortega and Tineo in "The Periodic Predator-Prey Lotka-Volterra Model"(Avances in differential Equ. vol 1, 3, 1996, 403-423, section 3). In fact, the scheme in that paper is obtained from (4)-(5) when $m=n=1$

These sequences are well defined and we easily get:

$$\begin{aligned} 0 &\le U^2 \le U^4 \le \ldots \le U^{2N} \le U^{2N-1} \le \ldots \le U^3 \le U^1 \\ 0 &\le V^2 \le V^4 \le \ldots \le V^{2N} \le V^{2N-1} \le \ldots \le V^3 \le V^1. \end{aligned} \tag{6}$$

By (6), $\{U^{2N-1}\}$, $\{U^{2N}\}$, $\{V^{2N-1}\}$, $\{V^{2N}\}$; $\quad N \in \mathbb{N}$ are monotone and uniformly bounded sequences. So, we have well defined functions:

$$\begin{array}{lcl} \bar{U}(t) = \lim_{N\to\infty} U^{2N-1}(t) & ; & \underline{U}(t) = \lim_{N\to\infty} U^{2N}(t); \\ \bar{V}(t) = \lim_{N\to\infty} V^{2N-1}(t) & ; & \underline{V}(t) = \lim_{N\to\infty} V^{2N}(t). \end{array} \tag{7}$$

Analogously, if $(u(t), v(t))$ is a solution of (2) such that $u(0) > 0$; $v(0) > 0$.

Inductively we can construct two sequences $\{u^N = (u_1^N, \dots, u_n^N)\}$ and $\{v^N = (v_1^N, \dots, v_m^N)\}$, defined on $[0, \infty)$ as follows: $u^0 = v^0 \equiv 0$,

$$\begin{aligned} &(u_i^N)' = u_i^N f_i(t, u_1^{N-1}(t), \dots, u_{i-1}^{N-1}(t), u_i^N, u_{i+1}^{N-1}(t), \dots, u_n^{N-1}(t), v^{N-1}(t)) \\ &(v_j^N)' = v_j^N g_j(t, u^N(t), v_1^{N-1}(t), \dots, v_{j-1}^{N-1}(t), v_i^N, v_{j+1}^{N-1}(t), \dots, v_m^{N-1}(t)) \\ &u_i^N(0) = u_i(0); \quad v_j^N(0) = v_j(0); (i = 1, \dots, n; j = 1, \dots, m; \qquad N \in \mathbb{N}) \end{aligned} \tag{8}$$

It is not difficult to show that using the theorem 1.3.7 [2],

$$\begin{aligned} 0 \ &\le \ u^2 \le u^4 \le \ldots \le u^{2N} \le u \le u^{2N-1} \le \ldots \le u^3 \le u^1 \\ 0 \ &\le \ v^2 \le v^4 \le \ldots \le v^{2N} \le v \le v^{2N-1} \le \ldots \le v^3 \le v^1. \end{aligned}$$

On the other hand, using induction and Theorem 2.3 it is easy to show the following result.

Corollary 4.1 *For all $N \in \mathbb{N}$, we have*

$$u^N(t) - U^N(t) \to 0 \quad \text{as} \quad t \to +\infty,$$

$$v^N(t) - V^N(t) \to 0 \quad \text{as} \quad t \to +\infty,$$

where $u^N; v^N; U^N; V^N$ are defined in (8),(4) and (5).

Theorem 4.2 *Let $(u(t), v(t))$ be a positive solution of (2).Then, (u, v) is defined on a terminal interval of $\mathbb{R}$ and,*

$$\limsup_{t\to\infty}[u_i(t) - \bar{U}_i(t)] \le 0 \le \liminf_{t\to\infty}[u_i(t) - \underline{U}_i(t)], \quad 1 \le i \le n; \quad t \ge t_0,$$

$$\limsup_{t\to\infty}[v_j(t) - \bar{V}_j(t)] \le 0 \le \liminf_{t\to\infty}[v_j(t) - \underline{V}_j(t)], \quad 1 \le j \le m; \quad t \ge t_0.$$

That is, $[\underline{U}, \overline{U}] \times [\underline{V}, \overline{V}]$ is an approximation of the solution of the system (2).

Example 1

The following example shows the case of an autonomous system with $n = 2$ and $m = 1$ in (2):

$$x_i' = x_i \left[a_i - \sum_{j=1}^{n} b_{ij} x_j - d_i y \right] \qquad y' = y \left[-\alpha + \sum_{i=1}^{n} \beta_i x_i - \gamma y \right] \tag{9}$$

with,

$$B \equiv (b_{ij}) = \begin{pmatrix} 1.4 & -0.5 \\ -0.5 & 1.7 \end{pmatrix}$$

$\alpha = 1$, $\gamma = 14$, $\vec{\beta} = (1.2, 1.6)$, $\vec{d} = (2, 8)$ and $\vec{a} = (0.7, 0.8)$.

Then, with the help of the MATHEMATICA software, the algorithm done in the annexe I obtains the chain of global attractors or approximate solutions expressed in the next list,

	ODD				
Iteration k	1	3	...	47	49
x_1^k	0.746	0.745	...	0.6181	0.6180
x_2^k	0.690	0.687	...	0.479	0.478
y^k	0.0763	0.0758	...	0.0397	0.0396

	EVEN				
Iteration k	2	4	...	48	50
x_1^k	0.481	0.482	...	0.6088	0.6088
x_2^k	0.253	0.256	...	0.463	0.463
y^k	0.0005	0.001	...	0.0369	0.0369

In the Figure 1 we present the numerical resolution of the system done by the "Populus" software. Note that the result agrees with the approximation that we have obtained.

Example 2.

Also we can pose the autonomous case supposing that the predator breed, for that we use the model analogy to (9), changing the second equation:

$$x_i' = x_i \left[a_i - \sum_{j=1}^{n} b_{ij} x_j - d_i y \right], \qquad y' = y \left[\alpha + \sum_{i=1}^{n} \beta_i x_i - \gamma y \right] \tag{10}$$

Like in the Example 1 we are going to see a concrete case for the model (10), using the next coefficients,

$$B \equiv \begin{pmatrix} 1.3 & -0.2 \\ -0.1 & 1 \end{pmatrix},$$

and that, $\alpha = 0.2$, $\gamma = 8.5$, $\vec{\beta} = (0.2, 0.7)$, $\vec{d} = (13, 5)$, and $\vec{a} = (1.3, 0.5)$. Again by the algorithm of the Annexe I we obtain,

	ODD				
Iteration k	1	3	...	49	50
x_1^k	1.094	0.832	...	0.475445	0.475444
x_2^k	0.609	0.463	...	0.264891	0.26489
y^k	0.009	0.081	...	0.0565309	0.0565308

	EVEN				
Iteration k	2	4	...	48	50
x_1^k	0.006	0.205	...	0.475443	0.475444
x_2^k	0.003	0.114	...	0.26489	0.26489
y^k	0.0239	0.0378	...	0.0565307	0.0565308

Again the graphic expression of the numerical resolution, using the same software, end up as the next form.

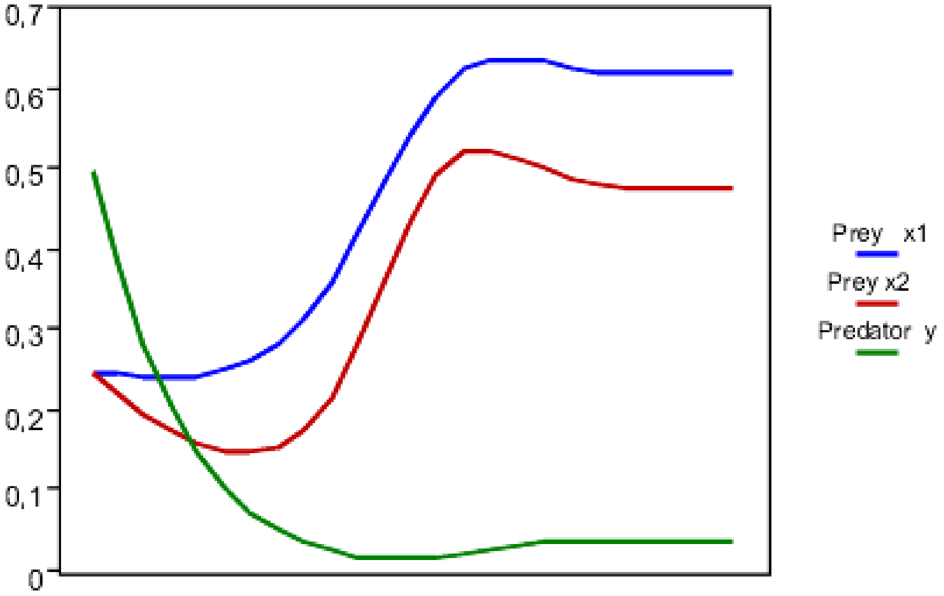

Figure 1: Numerical resolution to example 1

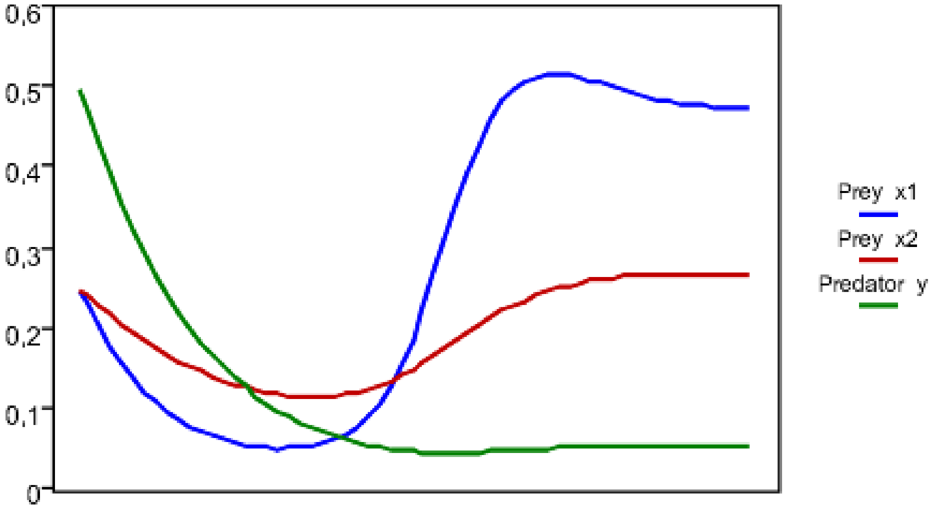

Figure 2: Numerical resolution to example 2

ANNEX I

Using Mathematica, we have developed the next program which has been used for the iterative schemes of the examples 1 and 2 posed in the beginning of this chapter.

- **Main Procedure.**

```
Clear ["Global"];
Presa[mb_,a_,d_,aalfa_,bbeta_,ggamma_,n_,error_,flag_]:=
Module[{sx,sy=0,y,iter=0,listax={},listay={},er=109̂},
   While [iter<n &&er>error, iter++; sx=LinearSolve[mb,a-sy*d];
   If[EvaluacionSinAlfa, sy=Solve[ggamma*y==-aalfa+({bbeta}.
   Transpose[{sx}])[[1,1]],y][[1,1,2]], sy=Solve[ggamma*y==aalfa+({bbeta}.
   Transpose[{sx}])[[1,1]],y][[1,1,2]] ];
   If[iter>1,
      er=Max[Abs[{sx-Last[listax], sy-Last[Listay]}]] ];
   AppendTo[listax,sx];AppendTo[listay,sy]; ];
  Return[If[flag==1,{listax,listay,er},{sx,sy,er}]]]
```

- **Receipts of Data.**

```
file=Input["Archivo de datos:"];
datos=ReadList[file,Number,NullRecord->True,RecordList->True];
n=Lenght[datos[[1]]]; mb=Table[datos[[i]],{i,1,n}];a=datos[[n+1]];
d=datos[[n+2]]; bbeta=datos[[n+3]];niter=datos[[n+4,1]];
error=datos[[n+5,1]];flag=datos[[n+6]];
If[datos[[n+7]]=={},EvaluacionSinAlfa=True, aalfa=datos[[n+7]];flag=0;
  (*Fin entrada*)

  If[EvaluacionSinAlfa,
  (*Case in which the variable "alpha" haven't taken of the card index of the data*)
  ss=Solve[({bbeta}.Inverse[mb].Transpose[{a}])[[1,1]]-x==0,x];
  mensaje1="Alfa(<"<>ToString[ss[[1,1,2]]]<>"):"; aalfa=Input[mensaje1];
  var2=Max[{Max[Table[(({bbeta}.Inverse[mb].Transpose[{a}])[[1,1]]-aalfa*
     (Inverse[mb].Transpose[{d}])[[i,1]]/(Inverse[mb].Transpose[{a}])[[i,1]],
     {i,1,Length[a]}]], ({bbeta}.Inverse[mb].Transpose[{d}])[[1,1]]}],

  (*Case in which the variable "alpha" can take only positive values*)
  var2=Max[Table[(({bbeta}.Inverse[mb].Transpose[{a}])[[1,1]]+aalfa*
  (Inverse[mb].Transpose[{d}])[[i,1]]/(Inverse[mb].Transpose[{a}])[[i,1]],
  {i,1,Length[a]}]]];
  mensaje2="Gamma(>"<>ToString[var2]<>"):";
  ggamma=Input[mensaje2];
```

- **Numerical Resolution. Presentation of Results.**

```
(*Resolution by Iterative Methods of Differential System*)
sol=Presa[mb,a,d,aalfa,bbeta,ggamma,niter,error,flag];
(*Presentation of Results*)
If[flag==1,
   (*Visualization of the sequence of the iterations obtained*)
   Print["Lista x:",MatrixForm[sol[[1]]]]; Print[""]; Ip={};li={};
   For[i=1,i<=Length[sol[[2]]],i++, If[Mod[i,2]==0,
   AppendTo[Ip,sol[[2,i]]],AppendTo[li,sol[[2,i]]]]]; Print[li,lp],
   (*Final solution adjusted to the level of error stablished.*)
   Print["x=",sol[[1]]]; Print["y=",sol[[2]]]; Print["Errores=",sol[[3]]];]
```

References

[1] Burton T.A. and Huston V., *Permanence for Non-Autonomous Predator-Prey Systems.* Diff. and Int. Equations, Vol.4, N 6 (1991), 1269-1280.

[2] Gámez, M., *Modelo depredador-presa. Aplicaciones al control biológico.* Tesis doctoral, Universidad de Almeria, Spain (1999) 134 pp.

[3] Hallam, T.G. and Zhien, Ma., *Persistence in population models with demographic fluctuations.* Jour. Math. Biol., 24, (1986), 327-339.

[4] López-Gómez J.; Ortega R. and Tineo A., *The Periodic Predator-Prey Lotka Model.* Avances in Differential Equations.Vol.1,N 3 (1996), 403-423.

[5] Pinghua Y.and Rui X., *Global Attractivity of the Periodic Lotka-Volterra System.* Jour. Math. Anal. and Appl., 233 (1999), 221-232.

[6] Tineo A., *Global On the Asymptotic Behavior of Some Population Models.* Jour. Math. Anal. and Appl.,Vol 167, N 2 (1992), 516-529.

[7] Tineo A., *Iterative Schemes for Some Population Models.* Nonli. World. 3 (1996), 695-708.

[8] Zanolin F., *Pernanence and Positive Periodic Solution for Kolmogorov Competing Species Systems.* Results in Mathematics, 21 (1992), 224-250.

Monografías de la Academia de Ciencias de Zaragoza. **20:** 77–85, (2002).

A tauberian theorem for a class of function spaces *

Miguel A. Jiménez

Benemérita Universidad Autónoma de Puebla. Puebla, México.

`mjimenez@fcfm.buap.mx`

Abstract

Let $(E, \|.\|_E)$ and $(F, \|.\|_F)$, $F \subset E$, be Banach spaces. Assume that $\|.\|_F := \|.\|_E + \theta(.)$, where θ is a seminorm. It is proved that sequences in F that converge in $\|.\|_E$ and whose elements satisfy certain equicontinuous behavior, also converge in $\|.\|_F$ to the same limit points. Quantitative estimates of the degree of convergence are obtained. Examples of applications to different function spaces are presented.

Mathematics Subject Classification: 41A65
Keywords and phrases: tauberian theorem, Hölder approximation, Lipschitz function, total variation, absolutely continuous function, equilipschitzian set, Hölder space, Besov space, Bernstein polynomial.

1 Introduction

Let X be either the real interval [0,1] or the multiplicative group $T = \{z \in \mathbb{C} : |z| = 1\}$. Let $Lip_\infty^\alpha X$ $(Lip_\infty^\alpha$ for short$)$, $0 < \alpha < 1$, be the Hölder space of continuous real (or complex) functions $f \in C(X)$, which satisfy the Hölder (also called Lipschitz) condition

$$\theta_\infty^\alpha(f) := \sup_{\delta>0} \theta_\infty^\alpha(f,\delta) < \infty, \tag{1}$$

where

$$\theta_\infty^\alpha(f,\delta) := \sup\{|f(x) - f(y)| / d(x,y)^\alpha : 0 < d(x,y) \le \delta\}. \tag{2}$$

Here $d(x,y) := |x-y|$ if $X = [0,1]$ or equal to the length of the shortest arc which joins x and y if $X = T$. In the last case, if functions on T are identified with 2π-periodic functions on $\mathbb{R}$, d should be the semidistance between elements of $\mathbb{R}$, given by

$$d(x + 2j\pi, y + 2k\pi) := \min\{|x-y|, 2\pi - |x-y| : x, y \in [0, 2\pi[;\ j, k \in \mathbb{Z}\}. \tag{3}$$

---

*This paper has been partially supported by CONACyT Project 32181-E, Mexico and University of Jaén, Spain.

Setting

$$\|f\|_{\alpha,\infty} := \|f\|_{\infty} + \theta^{\alpha}_{\infty}(f) \tag{4}$$

or another equivalent norm, the linear space Lip^{α}_{∞} becomes a Banach space. Further, denote by $lip^{\alpha}_{\infty}X$ $(lip^{\alpha}_{\infty}$ for short$)$, $0 < \alpha < 1$, the Banach subspace of those functions $f \in Lip^{\alpha}_{\infty}$, for which

$$\theta^{\alpha}_{\infty}(f,\delta) \longrightarrow 0 \text{ as } \delta \longrightarrow 0. \tag{5}$$

Basic results on Hölder spaces can be found in [4] and [5]. A recent survey of approximation in these spaces is given in [3].

From (4), a sequence that converges in Lip^{α}_{∞} also converges in the sup-norm $\|.\|_{\infty}$, with the same limit. The converse is false, of course. However, there is a certain tauberian condition (*) which lets us to prove the following assertion:

$$(f_n) \subset lip^{\alpha}_{p}\ ,\quad \|f_n - f\|_p \longrightarrow 0 \text{ and } (*) \Longrightarrow \|f_n - f\|_{\alpha,p} \longrightarrow 0. \tag{6}$$

In fact, in 1985, Leindler, Meir and Totik proved a first result of type (6) for X being the group T and (f_n) defined by a convolution process $K_n * f$, $f \in lip^{\alpha}_{\infty}$ (see [8]). They also estimated the degree of convergence. Later, Bustamante-Jiménez [2] introduced the following tauberian condition: A sequence $(f_n) \subset lip^{\alpha}_{\infty}X$, $0 < \alpha < 1$, is called *equilipschitzian* if (5) holds uniformly in n, i.e. if

$$\sup\{\ \theta^{\alpha}_{\infty}(f_n,\delta) : n \in \mathbb{N}\} \longrightarrow 0 \text{ as } \delta \longrightarrow 0. \tag{7}$$

The main theorem in [2] states that any equilipschitzian sequence (f_n) in lip^{α}_{p} converges in this space whenever it converges in the sup-norm, i.e. (6). Since sequences defined by convolution processes $(K_n * f)$, $f \in lip^{\alpha}_{p}(T)$ and (K_n) bounded in $L^1(T)$, are equilipschitzian, we get another view of the qualitative part of paper [8].

When $1 \leq p < \infty$, one defines Lip^{α}_{p} and lip^{α}_{p} in L_p, through standard procedures. Leindler, Meir and Totik announced the possibility of extending their results to $lip^{\alpha}_{p}(T)$. Further, in [7], Jiménez-Martínez extended most of results in [2] to these spaces.

With these antecedents at hand, one should expect a more general theorem that covers and unifies these particular results. In fact, in the next section, using a concept similar to (7), we establish and prove such a theorem. Estimates of the degree of convergence will also be obtained. The last section is devoted to applications in different function spaces.

2 Definitions and results

In order to follow the ideas of this section, let us keep our mind on the examples given by lip_p^α.

Set $\mathbb{R}_+ := \{t \in \mathbb{R}\colon\, t \geqslant 0\}$, $\mathbb{R}_+^* := \{t \in \mathbb{R}\colon\, t > 0\}$ and denote by I the real open interval $]0,b[$ (or semi-open $]0,b]$) where $I = \mathbb{R}_+^*$ is possible. Let E be a real or complex linear space and

$$\theta : E \times I \longrightarrow \mathbb{R}_+ \cup \{\infty\}, \tag{8}$$

a family $\theta\,(.,\delta)$, $\delta \in I$, of quasi-seminorms on E, i.e. the subadditivity of usual seminorms is substituted by the most general assertion that there exists a constant $C \geq 1$ (that here we assume is independent of δ), such that for every pair of elements $f, g \in E$, one has $\theta\,(f+g,\delta) \leq C\,(\theta\,(f,\delta) + \theta\,(g,\delta))$. Without loss of generality it is also assumed that for every fixed $f \in E$, $\theta\,(f,.)$ is an increasing function (in the large sense) of δ. Set

$$\theta\,(f) := \sup\{\theta\,(f,\delta) : \delta \in I\}. \tag{9}$$

Consider

$$\begin{aligned} \mathbb{F} \ &: = \{f \in E : \theta\,(f) < \infty\ \} && (10)\\ F \ &: = \{f \in \mathbb{F} : \theta\,(f,\delta) \longrightarrow 0 \ \text{ as } \delta \longrightarrow 0\} && (11) \end{aligned}$$

Then, $\mathbb{F}$ and F are linear subspaces of E, that are quasi-seminormed by (9) and that, eventually, could coincide .

We remark that F is a closed subspace of $(\mathbb{F},\theta)$. In fact, let $(f_n) \subset F$ be a sequence that converges to $f \in \mathbb{F}$. Fix $\varepsilon > 0$. First, take n such that $\theta\,(f_n - f) \leq \varepsilon$ and then $\delta_0 > 0$ such that $\theta\,(f_n,\delta) \leq \varepsilon$, for every $\delta \leq \delta_0$ Thus $\theta\,(f,\delta) \leq C\,(\theta\,(f_n - f) + \theta\,(f_n,\delta)) \leq 2C\varepsilon$.

Definition 1 *A set $G \subset F$ is called 0-equicontinuous if*

$$\theta\,(G,\delta) := \sup\{\ \theta\,(g,\delta) : g \in G\} \longrightarrow 0 \ \text{as}\ \delta \longrightarrow 0. \tag{12}$$

A sequence (f_n) is called 0-equicontinuous if the set $\{f_n : n \in \mathbb{N}\}$ is. In that case we simplify the notation by writing

$$\theta\,((f_n),\delta) : =\theta\,(\{f_n : n \in \mathbb{N}\},\delta).$$

Of course, equilipschitzian sets in our introductory section not only are examples of 0-equicontinuous sets but also the starting point of the present definition.

Proposition 2 *Let (f_n) be a convergent sequence in the quasi-seminormed space (F, θ). Then such a sequence is 0-equicontinuous.*

Proof: Suppose $\theta(f_n - f) \longrightarrow 0$ for some $f \in F$. Fix $\varepsilon > 0$ and choose N such that $\theta(f_n - f) \leq \varepsilon$ whenever $n > N$. Also choose $\delta_0 \in I$, such that $\theta(f, \delta_0) \leq \varepsilon$. Then, for any $0< \tau \leq \delta_0$ and $n > N$,

$$\theta(f_n, \tau) \leq C \ \left(\theta(f_n - f, \tau) + \theta(f, \tau)\right) \leq C \ \left(\theta(f_n - f) + \theta(f, \delta_0)\right) \ \leq 2C \ \varepsilon.$$

For $i = 1, 2, ..., N$, choose δ_i such that $\theta(f_i, \delta_i) \leq \varepsilon$. Set $\delta := \min\{\delta_i : 0 \leq i \leq N\}$. Thus $\sup\{ \ \theta(f_n, \delta) : n \in \mathbb{N}\} \leq 2C \ \varepsilon$. ■

In the remainder of this section we assume E to be a topological vector space whose topology is defined by a distance d_E, which is complete and translation invariant. We define another distance or quasi-distance in $\mathbb{F}$ by setting

$$d_{\mathbb{F}}(f, g) := d_E(f, g) + \theta(f - g). \tag{13}$$

Write $d_{\Xi}(f)$ instead of $d_{\Xi}(f, 0)$, where Ξ could be either E or $\mathbb{F}$. Then, $d_{\Xi}(f - g) = d_{\Xi}(f, g)$.

From (13), a sequence that converges in $(\mathbb{F}, d_{\mathbb{F}})$ also converges in (E, d_E) and to the same limit. The converse assertion is false in general. However, as we have already pointed out, we shall prove a certain converse result. In order to establish it we need a link between d_E and θ.

Definition 3 *The family of quasi-seminorms $\theta(., \delta)$, $\delta \in I$, defined above, is said to be admissible with respect to the distance d_E if the following conditions are satisfied:*

i) $(F, d_{\mathbb{F}})$ is complete

ii) There exists a constant $K > 0$ and a function $\Psi : I \times \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that for each $\delta \in I$,

$$\lim_{t \to 0} \Psi(\delta, t) = \Psi(\delta, 0) := 0$$

and for every $f \in F$,

$$\theta(f) \leq K \ \theta(f, \delta) + \Psi(\delta, d_E(f)). \tag{14}$$

With respect to condition i), since F is a closed subspace of $(\mathbb{F}, \theta)$, it follows from (13) that F is also a closed subspace of $(\mathbb{F}, d_{\mathbb{F}})$. Then, if $(\mathbb{F}, d_{\mathbb{F}})$ is complete, so is $(F, d_{\mathbb{F}})$.

Theorem 4 (*tauberian*) *Suppose that* $(F, d_{\mathbb{F}})$ *has been defined by a family of admissible quasi-seminorms* $\theta(., \delta)$, $\delta \in I$, *on* (E, d_E). *Let* $(f_n) \subset F$ *be a convergent sequence in* (E, d_E), *to an element* f. *If* (f_n) *is 0-equicontinuous , then* $f \in F$ *and* (f_n) *converges to* f *in* $(F, d_{\mathbb{F}})$. *Moreover, if for each* $\delta \in I$, $\Psi(\delta, .)$ *is continuous in* $\mathbb{R}_+$, *then*

$$\theta(f_n - f) \leq 2C\ K\ \ \theta((f_n), \delta) + \Psi(\delta, d_E(f_n - f)). \tag{15}$$

Proof: Assume we have already proved that $(\theta(f_n))$ is a real Cauchy sequence. Since the hypothesis of the theorem include that (f_n) is a Cauchy sequence in E, it would follow from (13) that (f_n) is a Cauchy sequence in $(F, d_{\mathbb{F}})$. But F is a complete metric space, then there exits $g \in F$ such that $d_{\mathbb{F}}(f_n - g) \longrightarrow 0$ as $n \to \infty$. Also by (13), $d_E(f_n - g) \leq d_{\mathbb{F}}(f_n - g)$, then $d_E(f_n - g) \to 0$. But $d_E(f_n - f) \longrightarrow 0$ as $n \to \infty$. That forces $f = g$. In order to prove that $(\theta(f_n))$ is a Cauchy sequence, fix $\varepsilon > 0$. For every $\delta \in I$, we use (14) to obtain,

$$\theta(f_n - f_m) \leq K\ \theta(f_n - f_m, \delta) + \Psi(\delta, d_E(f_n - f_m)). \tag{16}$$

Take δ such that $\theta((f_n), \delta) \leq \varepsilon$. Further, take N such that for every $n > N$ and $m > N$, $\Psi(\delta, d_E(f_n - f_m)) \leq \varepsilon$. By substituting into (16),

$$\theta(f_n - f_m)\ \ \leq (2CK + 1)\,\epsilon.$$

The qualitative part of the theorem has been proved. In particular $\theta(f_n - f_m) \longrightarrow \theta(f_n - f)$ as $m \longrightarrow \infty$. Then, using (16) and the continuity of $\Psi(\delta, .)$ we deduce (15). ■

Equivalent distances to (13) are given by

$$d_{\mathbb{F}}(f) := \left(d_E(f)^p + \theta(f)^p\right)^{1/p}, \quad 1 < p < \infty, \tag{17}$$

$$d_{\mathbb{F}}(f) := \max\{d_E(f), \theta(f)\ \}, \qquad\qquad p = \infty. \tag{18}$$

In those cases, using (15), we remark that

$$d_{\mathbb{F}}(f_n - f) \leq\ \left(d_E(f_n - f)^p +\ \left[2CK\ \theta((f_n), \delta) + \Psi(\delta, d_E(f_n - f))\right]^p\right)^{1/p}, \tag{19}$$

if $1 \leq p < \infty$; or

$$d_{\mathbb{F}}(f_n - f)\ \leq \max\{\ d_E(f_n - f), 2C\ K\ \theta((f_n), \delta) + \Psi(\delta, d_E(f_n - f))\}, \tag{20}$$

if $p = \infty$.

Also we remark that formula (15) is a general one. Therefore its accuracy could be improved in particular problems. In the same way, optimal values for δ depend on the problem on hand.

Theorem 5 *Suppose that* $(F, d_{\mathbb{F}})$ *has been defined from* (E, d_E) *by a family of admissible quasi-seminorms* $\theta(., \delta)$, $\delta \in I$. *Then a set* $A \subset F$ *is compact with respect to the topology induced by* $d_{\mathbb{F}}$ *if and only if* A *is compact in* (E, d_E) *and 0-equicontinuous.*

Proof. Let $(f_n) \subset A$. If A is a compact set of (E, d_E), there exists a subsequence (f_{n_k}) that converges to an element $f \in A$ with respect to d_E. If A is a 0-equicontinuous set, then (f_{n_k}) converges to f with respect to $d_{\mathbb{F}}$. Reciprocally, if A is a compact set of $(F, d_{\mathbb{F}})$, there exists a subsequence (f_{n_k}) that converges to an element $f \in A$ with respect to $d_{\mathbb{F}}$ Then (f_{n_k}) converges to the same limit with respect to d_E. ■

3 Examples and Applications

In this section we show that well known function spaces are included in the class of spaces defined above . Of course, it is impossible to examine here the great variety of important function spaces not even to examine only a few of them in their general setting (see Triebel [9], for instance). Thus the particular examples below are conceived just to conform an illustrative sample of applications.

Example 6 *Set* $E := C(X)$. *Taking* $\theta(f, \delta) := \theta_\infty^\alpha(f, \delta)$, *defined in* (2), *we obtain* $\mathbb{F} = Lip_\infty^\alpha$ *and* $F = lip_\infty^\alpha$. *Set* $K := 1$. *Thus , with* $\Psi(\delta, t) := 2t/\delta^\alpha$, *the family of seminorms is admissible. An application of* (15) *leads to*

$$\|f_n - f\|_{\alpha,\infty} \leq (1+2/\delta^\alpha) \|f_n - f\|_\infty + 2\theta((f_n), \delta) \tag{21}$$

The qualitative part of this application is the main theorem in Bustamante-Jiménez [2]. In particular, the sequence of Bernstein polynomials $(B_n f)$, $f \in lip_\infty^\alpha([0,1])$ is *0*-equicontinuous . In fact, Bustamante-Jiménez proved that $(B_n f)$ converges to f in $lip_p^\alpha[0,1]$, i.e. in the norm (4) which implies convergence in the seminorm (1). Then Proposition 2 asserts that $(B_n f)$ is *0*-equicontinuous. On the other hand, theorem 5 characterizes the compact sets in lip_∞^α in the same way that it was done in [2].

Example 7 *In the last example, take* $X := T$ *and change* (2) *by*

$$\begin{aligned} \theta(f, \delta) &: = \sup\{\zeta(f, t) : 0 < t \leq \delta\}, \\ \zeta(f, t) &: = \sup\{|f(x+t) - f(x)| / \varphi(t) : x \in T\}, \end{aligned}$$

where $\varphi : \mathbb{R}_+^* \to \mathbb{R}_+^*$, *is an increasing function. For* $f \in F$, *define the sequence* $f_n := K_n * f$, *where* $K_n \in L^1(T)$ *and* $M := \sup\{\|K_n\|_1 : n \in \mathbb{N}\} < \infty$. *Then* (f_n) *is 0-equicontinuous with* $\theta((f_n), \delta) \leq M\ \theta(f, \delta)$. *Assume that* $f_n \longrightarrow f$ *in uniform norm. Set* $K := 1$ *and* $\Psi(\delta, t) := 2t/\varphi(\delta)$. *In this situation* (21) *is transformed into*

$$\|f_n - f\|_{\mathbb{F}} \leq (1+2/\varphi(\delta))\ \|f_n - f\|_\infty + 2\ M\ \theta(f, \delta).$$

This is the estimate given by Leindler, Meir and Totik, from which several implications to Fourier series follow ([8])

Example 8 *We can use modulus of smoothness of higher order* r. *For instance, set* $X := T$ *and* $E := L^p(T)$, *1*$\leq p < \infty$. *Define*

$$\begin{aligned} \theta(f,\delta) &: = \sup\{\zeta(f,t) : 0 < t \leq \delta\}, \\ \zeta(f,t) &: = \left\{\left(\frac{1}{2\pi}\int_0^{2\pi} |\Delta_t^r f(x)|^p \, d(x)\right)^{1/p} /t^\alpha\right\}. \end{aligned}$$

Here $\Delta_t f := \Delta_t^1 f := f(.+t) - f$, $\Delta_t^r f := \Delta_t\left(\Delta_t^{r-1} f\right)$. *Set* $K := 1$. *Then, with the function* $\Psi(\delta,t) := 2^r t/\delta^\alpha$, *the family of semi-norms is admissible in definition 3.*

Example 9 *Set* $E := L_p(T)$, *1*$\leq p < \infty$. *In* [6], *the author has defined homogeneous Hölder spaces* B_p^α, $\alpha > 0$, *which are equivalent in norm to certain Besov spaces. A function* $f \in L_p(T)$ *is in* B_p^α, *if* $F_\alpha(x,y) := (f(x) - f(y))/d(x,y)^\alpha \in L_p(T^2)$. *A crucial point here is that* d *is given by* (3) *and then* F_α *has period 2*π *in each variable. Set*

$$\theta(f) := \left(\frac{1}{4\pi^2}\int_0^{2\pi}\left(\int_0^{2\pi} |F_\alpha(x,y)|^p \, dx\right) dy\right)^{1/p}.$$

Then B_p^α *becomes a homogeneous Banach (Hilbert if* $p = 2$*) space under the norm*

$$\|f\|_{\alpha,p} := \left(\|f\|_{L^p(T)}^p + \|F_\alpha\|_{L^p(T^2)}^p\right)^{1/p}.$$

Taking

$$\theta(f,\delta) := \left(\frac{1}{2\pi^2}\int_0^{\delta}\left(\int_0^{2\pi} |\Delta_t f(x)/t^\alpha|^p \, dx\right) dt\right)^{1/p},$$

we can show that $\theta(f) = \theta(f,\pi)$.

Thus $\mathbb{F} = F = B_p^\alpha$. *Set* $K := 1$. *Therefore, with the function*

$$\Psi(\delta,t) := \left(\frac{2}{\pi}\int_\delta^\pi \frac{dx}{x^{\alpha p}}\right)^{1/p} t,$$

the family of seminorms is admissible.

With the following two examples, we show the connection of section 2 with the theory of Measure and Integration and also the convenience of considering the general scope in which the tauberian theorem above has been established.

Example 10 *Let* E *be the complex linear space of all bounded complex functions* f *on* $\mathbb{R}$ *that are continuous to the right and such that* $f(x) \longrightarrow 0$ *as* $x \to -\infty$. *For all* $\delta > 0$, *set*

$$(22) \quad \theta(f,\delta) := \sup\left\{\begin{array}{c} \sum_{1\leq i\leq m} |f(y_i) - f(x_i)| : x_1 < y_1 \leq x_2 < \ldots < y_m; \\ m = 1,2,..; \quad \sum_{1\leq i\leq m} y_i - x_i \leq \delta \end{array}\right\}.$$

Thus $\theta(f)$ stands for the total variation of f in $\mathbb{R}$; $(\mathbb{F},\theta)$ is defined to be the Banach space of functions of bounded variation and F is its closed subspace of absolutely continuous functions.

We remark that for a given function f, it could happen that $\theta(f,\delta)\longrightarrow 0$ as $\delta\to 0$, but $\theta(f)=\infty$. For instance, $f(x):=\sin(x)/x$. However such a function is not in F by [10] and [11].

On the other hand, since (22) is equal to

$$\sup\left\{\int_A |f\,'(x)|\,d(x):\ \text{meas}(A)=\delta\right\},\ f\in F,$$

this example is connected with the next one, for which the theoretical background can be found in chapter 4 of [1]. However, to avoid technical difficulties that are not any objective at present, we restrict ourself to a set of finite measure.

Example 11 *Let E be the complex linear space of all measurable complex functions f on $[0,1]$. We identify functions that are equal Lebesgue almost everywhere and consider any complete and translation invariant distance d_E which characterize the convergence in measure. .*

For any $f\in E$, $0<p<\infty$ and $0<\delta\leq 1$, define

$$\theta(f,\delta):=\sup\left\{\left(\int_A |f|^p\,d(x)\right)^{1/p}:\ meas(A)=\delta\right\}.$$

Then $\mathbb{F}=F=L^p[0,1]$. A sequence (f_n) is 0-equicontinuous if and only if it is equi-integrable and it is known that convergence of (f_n) in $L^p[0,1]$, occurs if and only if (f_n) is a Cauchy sequence in measure and equi-integrable. In this example, the function Ψ depends on the particular distance d_E at hands. In fact, for a given function $f\in F$ and $0<\delta\leq 1$, fix a measurable set A, with meas$(A)=\delta$, such that for any pair $x\in A$ and $y\in A^c$, $f(y)\leq f(x)$. Using typical procedures, we obtain

$$\theta(f)=\left[\int_0^1 |f|^p\,d(x)\right]^{1/p}\leq K\left[\theta(f,\delta)+\left[\int_{A^c}|f|^p\,d(x)\right]^{1/p}\right],$$

with $K:=C=1$ if $1\leq p<\infty$ or $K:=C=2^{1/p}$ if $0<p<1$. Then, in terms of the sequence (f_n) and its limit in measure f,

$$\theta(f_n-f)\leq 2K^2\ \theta((f_n),\delta)+K\ \beta_n\ (1-\delta)^{1/p},$$

where the sequence β_n, that converges to 0 when $n\to\infty$, can be expressed in terms of $d_E(f_n-f)$.

References

[1] Bourbaki, N. , *Éléments de Mathématique, Livre VI Intégration*, Chapitres 1, 2, 3 et 4. Hermann, Paris (1975).

[2] Bustamante, J. ; Jiménez, M. A., *Chebyshev and Hölder approximation*, Aportaciones Matemáticas, Serie Comunicaciones, **27** (2000), 23-31.

[3] Bustamante, J. ; Jiménez, M. A., *Trends in Hölder approximation*, Approximation, Optimization and Mathematical Economics, Physica-Verlag (2001), 81-95.

[4] Butzer P. L. ; Berens, H., *Semi-Groups of Operators and Approximation*, Springer-Verlag (1967).

[5] DeVore, R. A. ; Lorentz, G. G., *Constructive Approximation*, Grundlehren der mathematischen Wissenschaften **303**, Springer-Verlag, (1993).

[6] Jiménez, M. A., *A new approach to Lipschitz spaces of periodic integrable functions*, Aportaciones Matemáticas, Serie Comunicaciones, **25** (1999), 153-157.

[7] Jiménez, M. A.; Martínez, G., *Equilipschitzian sets of Hölder integrable functions,* Aportaciones Matemáticas, Serie Comunicaciones 29 (2001), 55-60.

[8] Leindler, L., Meir A., and Totik, V., *On approximation of continuous functions in Lipschitz norms*, Acta Math. Hung., **45** (3-4) (1985), 441-443.

[9] Triebel, H. *Theory of Function spaces,* Birkhaüser Verlag, Basel, Vol. **1** (1983), Vol. **2** (1992).

Monografías de la Academia de Ciencias de Zaragoza. **20:** 87–96, (2002).

Some Properties of Linear Positive Operators Defined in Terms of Finite Differences *

A.J. López-Moreno, J. Martínez-Moreno,
F.J. Muñoz-Delgado and J.M. Quesada

Departamento de Matemáticas. Universidad de Jaén. Spain

Abstract

In this paper, we study some properties of Mastroianni operators [4] and generalized Baskakov operators [3].

1 Introduction and Notation

In [4] Mastroianni introduced and studied a generalization of the classical Bernstein operators consisting in replacing the functions $(1-x)^{n-k}$ by more general ones satisfying suitable relations. His work was motivated by the development of a general expression that cover other Bernstein-type operators.

In this paper we study some properties of these Mastroianni operators. We obtain some recursive properties of the derivatives of the operators, that allow us to give a characterization of the Szász operators. Also, we consider the linear combination of iterates $I-(I-L_n)^m$ of Mastroianni operators of fixed degree n for increasing order of iteration m and prove that these Boolean sums have a good behaviour for polynomials.

In the same manner, we consider a generalization of the Baskakov operators [3] which are related to certain functions. We study the convergence properties of the sequence of these operators and give an asymptotic expansion for them.

We will use Stirling numbers, S_j^i y σ_j^i, of first and second kind defined, respectively, by: $x^{\underline{j}} = \sum_{i=0}^{j} S_j^i x^i$ and $x^j = \sum_{i=0}^{j} \sigma_j^i x^{\underline{i}}$, with $j \in \mathbb{N}_0$. Here $x^{\underline{j}} = x(x-1)\dots(x-j+1)$ if $j>0$ and $x^{\underline{0}} = 1$.

---

*This work is partially supported by Junta de Andalucía, Research Groups FQM178 and FQM268 and by Ministerio de Ciencia y Tecnología, Project BFM2000-0911.

Throughout this paper we also use the following notation:

$$E_2 = \{f \in C[0, +\infty) : \ \frac{f(x)}{1+x^2} \text{ is convergent as } x \to +\infty\},$$

$\mathbb{P}_r$ denotes the space of real polynomials of degree at most r and t^r is the monomial $t^r(x) = x^r$. For convenience, we define $\sum_{i=1}^{0} = \sum_{i=0}^{-1} = 1$.

Here we simply recall the Mastroianni operators. First we start with a sequence $\{\phi_n\}_{n\in\mathbb{N}}$ of real functions on $I = [0, \infty)$ which are infinitely differentiable and strictly monotone satisfying the following additional conditions:

A1) $\phi_n(0) = 1$, for every $n \in \mathbb{N} = \{1, 2, ...\}$.

A2) $(-1)^i \phi_n^{(i)}(x) \geq 0$, for every $n \in \mathbb{N}$, $x \in I$ and $i \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

A3) For every $(n, i) \in \mathbb{N}\times\mathbb{N}_0$ there exists a positive integer $p(n, i) \in \mathbb{N}$ and a real function $\alpha_{n,i} : I \to \mathbb{R}$ such that

$$\phi_n^{(i+k)}(x) = (-1)^i \phi_{p(n,i)}^{(k)}(x)\alpha_{n,i}(x),$$

for every $k \in \mathbb{N}_0$ and $x \in I$ and

$$\lim_{n\to\infty} \frac{n}{p(n, i)} = \lim_{n\to\infty} \frac{\alpha_{n,i}(x)}{n^i} = 1.$$

For short, we will denote $\alpha_{n,i} = \alpha_{n,i}(0)$.

By **A3)**, $(-1)^{i+k}\phi_n^{(i+k)}(x) = (-1)^k \phi_{p(n,i)}^{(k)}(x)\alpha_{n,i}(x)$ and by **A2)** we conclude that necessarily $\alpha_{n,i}(x) \geq 0$.

Also from **A3)**, we deduce that $(-1)^i \phi_n^{(i)}(x) = \phi_{p(n,i)}(x)\alpha_{n,i}(x)$, and, in particular, **A1)** implies that $(-1)^i \phi_n^{(i)}(0) = \alpha_{n,i}$.

To the above sequence Mastroianni [4] associates a sequence of positive linear operators $\{L_n : E_2 \to C^\infty(I)\}_{n\in\mathbb{N}}$ defined by

$$L_n f(x) = \sum_{k=0}^{\infty} (-1)^k f\left(\frac{k}{n}\right) x^k \frac{\phi_n^{(k)}(x)}{k!},$$

for every $f \in E_2$ and $x \in [0, \infty)$. L_n can be represented in the form

$$L_n f(x) = \sum_{i=0}^{\infty} (-1)^i \frac{\phi_n^{(i)}(0)}{i!} \ \Delta_{\frac{1}{n}}^{i} f(0) \ x^i. \tag{1}$$

2 Derivatives.

It is straightforward to check the following relation of the derivatives.

Proposition 1. For every $k \in \mathbb{N}_0$ and $x \in [0, \infty)$

$$L_n^{(k)} f(x) = \alpha_{n,k} \sum_{i=0}^{\infty} (-1)^i \frac{\phi_{p(n,k)}^{(i)}(0)}{i!} \; \Delta_{\frac{1}{n}}^{i+k} f(0) \; x^i. \tag{2}$$

In the particular case, $p(n,k) = n$, then

$$L_n^{(k)} f(x) = \alpha_{n,k} L_n(\Delta_{\frac{1}{n}}^{k} f)(x). \tag{3}$$

Proof.

$$\begin{aligned} L_n^{(k)} f(x) = \sum_{i=k}^{\infty} (-1)^i \frac{\phi_n^{(i)}(0)}{i!} \; \Delta_{\frac{1}{n}}^{i} f(0) \; \frac{i!}{(i-k)!} \; x^{i-k} \\ = \sum_{i=0}^{\infty} (-1)^{i+k} \frac{\phi_n^{(i+k)}(0)}{(i+k)!} \; \Delta_{\frac{1}{n}}^{i+k} f(0) \; \frac{(i+k)!}{i!} \; x^i. \end{aligned}$$

From property **A3)** we get (2). Then (3) follows from (1) and (2). □

We will use this identity to compute the moments $L_n(t^s)$. Formula (3) is similar to $(B_n f)^{(k)}(x) = n^{\underline{k}} \; B_{n-k}(\Delta_{\frac{1}{n}}^{k} f)(x)$, which is valid for Bernstein polynomials. And it is also similar to the identity

$$S_n^{(k)} f(x) = n^k S_n(\Delta_{\frac{1}{n}}^{k} f)(x),$$

valid for Szász operators. In fact, the last one is a particular case of (2) and (3).

L_n operators have classical shape preserving properties. Recall the notion of higher order convexity. Given $k \in \mathbb{N}_0$, a function f is said to be convex of order k (k-convex) if for all $h > 0$ one has that $\Delta_h^k f \geq 0$. A function $f \in C^k[0, \infty)$ is convex of order k if and only if $f^{(k)} \geq 0$.

We can use (1) to deduce some properties of L_n, such as preservation of n-convexity. In this way, the following result is immediate:

Proposition 2. For any $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$ if f is convex of order k, then $L_n f$ is convex of order k.

Of course, Mastroianni's operator L_n has the degree-preserving property

$$L_n \mathbb{P}_r \subset \mathbb{P}_r, \qquad (0 \leq r \leq n).$$

3 Moments and asymptotic expansion of L_n.

Now, we give explicit expressions for the moments $L_n(t^i)$ and central moments $L_n\left((t-x)^i\right)(x)$, for $i \in \mathbb{N}$ and $x \in [0, \infty)$.

Lemma 3. *For any* $n \in \mathbb{N}$, $r \in \mathbb{N}_0$, *one has*

$$L_n(t^r) = n^{-r} \sum_{i=1}^{r} \alpha_{n,i}\ \sigma_r^i\ t^i. \tag{4}$$

Proof. By (1) we have

$$L_n(t^r) = \sum_{i=0}^{r} (-1)^i \frac{\phi_n^{(i)}(0)}{i!}\ \Delta_{\frac{1}{n}}^i (t^r)(0)\ t^i.$$

On the other hand, it is known that $\Delta_{\frac{1}{n}}^i t^r(0) = i!\sigma_r^i n^{-r}$, (see [1, Section 24.1.4.II.C]). Now, replacing this expression in the above formula we get

$$L_n(t^r) \ = \ \sum_{i=0}^{r} (-1)^i \frac{\phi_n^{(i)}(0)}{i!}\ i!\sigma_r^i n^{-r}\ t^i = \sum_{i=0}^{r} (-1)^i \phi_n^{(i)}(0)\ \sigma_r^i n^{-r}\ t^i.$$

Finally, properties **A3)** and **A1)** are used. ☐

We would like to remark that in the above Lemma in the special case $r = 0$, we have $L_n(1) = 1$.

Lemma 4. *For any* $n \in \mathbb{N}$, $p \in \mathbb{N}_0$, $x \in [0, \infty)$, *one has*

$$L_n\left((t-x)^p\right)(x) = n^{-p} \sum_{j=0}^{p} (-1)^j (j - nx)^p G(j, n, x),$$

where $G(j,n,x) = \sum_{i=j}^{p} \frac{\phi_n^{(i)}(0)}{i!} \binom{i}{j}\ x^i = \sum_{i=j}^{p} (-1)^i \frac{\alpha_{n,i}}{i!} \binom{i}{j}\ x^i$.

Proof. First by (1)

$$L_n\left((t-x)^p\right)(x) = \sum_{i=0}^{\infty} (-1)^i \frac{\phi_n^{(i)}(0)}{i!} \Delta_{\frac{1}{n}}^i ((t-x)^p)(0)\ x^i.$$

From the definition of Δ

$$\begin{aligned} \Delta_{\frac{1}{n}}^i ((t-x)^p)(0) &= \sum_{j=0}^{i} \binom{i}{j} (-1)^{i+j} ((t-x)^p) \left(\frac{j}{n}\right) \\ &= \sum_{j=0}^{i} \binom{i}{j} (-1)^{i+j} \left(\frac{j}{n} - x\right)^p = n^{-p} \sum_{j=0}^{i} \binom{i}{j} (-1)^{i+j} (j - nx)^p. \end{aligned}$$

It is evident that $\Delta_{\frac{1}{n}}^i ((t-x)^p)(0) = 0$ for $i \geq p$. Replacing it in the above expression we get

$$\begin{aligned} L_n\left((t-x)^p\right)(x) &= \sum_{i=0}^{p} (-1)^i \frac{\phi_n^{(i)}(0)}{i!} n^{-p} \sum_{j=0}^{i} \binom{i}{j} (-1)^{i+j} (j - nx)^p\ x^i \\ &= n^{-p} \sum_{j=0}^{p} (-1)^j (j - nx)^p \sum_{i=j}^{p} \frac{\phi_n^{(i)}(0)}{i!} \binom{i}{j}\ x^i. \end{aligned}$$

☐

Now, we give the asymptotic expansion of the sequence of L_n operators.

Theorem 5. *Let $f \in E_2$ r times differentiable at $x \in [0,\infty)$, then*

$$L_n f(x) = \sum_{j=0}^{r} (-1)^j G(j,n,x) \sum_{p=j}^{r} \frac{f^{(p)}(x)}{p!} (\frac{j}{n} - x)^p + o(L_n((t-x)^r)),$$

where $G(j,n,x)$ is given as in Lemma 4.

Proof. The proof of Sikkema Theorem in [6] is valid to check that

$$L_n f(x) = \sum_{p=0}^{r} \frac{f^{(p)}(x)}{p!} L_n\left((t-x)^p\right)(x) + o(L_n((t-x)^r)), \tag{5}$$

for every $x \in [0,\infty)$. We finish the proof using Lemma 4. □

4 Some Limiting Properties.

Mastroianni [4] proves that the sequence of iterations $\{I-(I-L_n)^m\}_{n\in\mathbb{N}}$ converges as m tends to infinity under certain assumptions. More precisely, he showed that

$$\lim_{m\to\infty} \left(I-(I-L_n)^m\right)(f)(x) = f(0) + \sum_{i=1}^{\infty} \frac{(nx)^{\underline{i}}}{n^i} \Delta_{\frac{1}{n}}^{i-1} f(0) \tag{6}$$

holds for all $f \in C[0,b]$ and $x \in [0,b]$ if and only if $\left|\frac{\phi_n^{(i)}(0)}{n^2}\right| < 2$ for every $i \geq 2$. In the general case, such a convergence does not always hold but we are going to see that at least we can obtain a good behavior for polynomials. For this purpose we employ a modification of the technique used by Sevy [5] for Bernstein operators.

Theorem 6. *Given $n, r \in \mathbb{N}$, let us suppose that $0 < \frac{\alpha_{n,i}}{n^i} < 2$ for every $i \in \{0,\ldots,r\}$. Then, for any $p \in \mathbb{P}_r$ we have*

$$\lim_{m\to\infty} \left(I-(I-L_n)^m\right)(p) = p$$

uniformly on compact sets.

Proof. From Lemma 3 it is straightforward that L_n is a linear map, $L_n : \mathbb{P}_r \to \mathbb{P}_r$, whose eigenvalues are $\lambda_i^{(n)} = \frac{\alpha_{n,i}}{n^i}$, $i = 0\ldots,r$. It is also clear that for every eigenvalue, $\lambda_i^{(n)}$, we can find an eigenvector $p_i^{(n)}$ which is a polynomial with exact degree i and also that $p_0^{(n)} = 1$.

Take the operator $V = I - L_n$. It is not difficult to check that for all $m, N \in \mathbb{N}$,

$$V^{m+N} - V^m = -(I-L_n)^m \left(\sum_{i=0}^{N-1} (I-L_n)^i\right) L_n. \tag{7}$$

If $\lambda_i^{(n)} = 1$ then $V^m(p_i^{(n)}) = 0$ for every $m \in \mathbb{N}$. If $\lambda_i^{(n)} \neq 1$ from (7) we have

$$V^{m+N}(p_i^{(n)}) - V^m(p_i^{(n)}) = -(1-\lambda_i^{(n)})^m \lambda_i^{(n)} \frac{1-(\lambda_i^{(n)})^N}{1-\lambda_i^{(n)}} p_i^{(n)}.$$

Since $0 < \lambda_i^{(n)} < 2$, one has $|1-\lambda_i^{(n)}| < 1$. In both cases, as $p_i^{(n)}$ is continuous, taking into account the preceding identity we can conclude that the sequence $\{V_m(p_i^{(n)})\}_{n\in\mathbb{N}}$ satisfies the Cauchy condition on compact subsets from which we deduce that such a sequence converges towards $g \in C[0,\infty)$. Furthermore, it is immediate that $V^m(p_i^{(n)}) \in \mathbb{P}_r$ for any $m \in \mathbb{N}$ so that $g \in \mathbb{P}_r$ because $C[0,a]$ is closed and the convergence is uniform in $[0,a]$.

We know that the linear operator $L_n : \mathbb{P}_r \to \mathbb{P}_r$ is always continuous because $\mathbb{P}_r$ is a finite dimensional space. Then, $V(g) = \lim_{m\to\infty} V(V^m(p_i^{(n)})) = \lim_{m\to\infty} V^{m+1}(p_i^{(n)}) = g$ from which $g = V(g) = g - L_n(g)$ and then $L_n(g) = 0$, that is possible only when $g(\frac{i}{n}) = 0$ for all $i \in \{0,\dots,r\}$ (see Remark 8). Hence, g is a polynomial of degree at most r that vanishes at $r+1$ points which implies that $g = 0$. Therefore

$$\lim_{m\to\infty} \left(I - (I-L_n)^m\right)(p_i^{(n)}) = \lim_{m\to\infty} \left(p_i^{(i)} - V^m(p_i^{(n)})\right) = p_i^{(n)}.$$

If $p \in \mathbb{P}_r$ then we can write $p = \sum_{i=0}^r A_i p_i^{(n)}$. Since $I-(I-L_n)^m$ is a linear and continuous operators the results obtained for $p_i^{(n)}$ also hold for p. ☐

Corollary 7. *Given $r \in \mathbb{N}$ and $p \in \mathbb{P}_r$ there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,*

$$\lim_{m\to\infty} \left(I - (I-L_n)^m\right)(p) = p.$$

Proof. From the definition of the Mastroianni operators, for every $i \in \mathbb{N}$,

$$\lim_{n\to\infty} \lambda_i^{(n)} = \lim_{n\to\infty} \frac{\alpha_{n,i}}{n^i} = 1.$$

Therefore, we can find n_0 large enough such that $|1-\lambda_i^{(n)}| < 1$ for $i \in \{0,\dots,r\}$ and $n \geq n_0$. ☐

Remark 8. From the hypotheses of Theorem 6 we know that $0 < \alpha_{n,i}$ so that $\phi_n^{(i)}(0) \neq 0$, $i \in \{0,\dots,r\}$. By means of (1), if $L_n(g) = 0$ we easily deduce that

$$(-1)^i \frac{\phi_n^{(i)}(0)}{i!} \Delta_{\frac{1}{n}}^i g(0) = 0, \ \forall i \in \mathbb{N}_0$$

and hence $\Delta_{\frac{1}{n}}^i g(0) = 0$ for $i \in \{0,\dots,r\}$ (because the corresponding $\phi_n^{(i)}(0)$ does not vanish) from which it is straightforward that $g\left(\frac{i}{n}\right) = 0$ for $i \in \{0,\dots,r\}$.

So, we can conclude that under the conditions of Theorem 6, given $g \in C[0,\infty)$ such that $L_n(g) = 0$ we have that $g\left(\frac{i}{n}\right) = 0$, $i \in \{0,\dots,r\}$.

5 Generalized Baskakov Operators.

For every $f \in E_2$ and $x \in [0, \infty)$, Baskakov operators are defined by

$$M_n f(x) = \sum_{k=0}^{\infty} (-1)^k f\left(\frac{k}{n}\right) x^k \frac{\phi_n^{(k)}(x)}{k!},$$

(cf. Martini [3]) and they are generated by a sequence of analytic functions on $[0, \infty)$, $\phi_n : \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}$, satisfying **A1)**, **A2)** and

$$\phi_n^{(k)}(x) = -(n + l_1)\phi_{n+l_2}^{(k-1)}(x)$$

for all $k \in \mathbb{N}$ and $x \in [0, \infty)$, where $l_1, l_2 \in \mathbb{N}_0$ are independent of n, k and x.

By induction **A3)** is also verified for $\alpha_{n,i} = \alpha_{n,i}(x) = (n+l_1)^{(i,-l_2)}$ and $p(n,i) = n+il_2$, where $x^{(0,l)} = 1$ and $x^{(i,-l)} = x(x+l)...(x+(i-1)l)$, for $i > 0$.

Furthermore, from suitable choices of the sequences $\{\phi_n\}_{n\in\mathbb{N}_0}$ we obtain some well known operators. For example, choosing $l_2 = 0$, we have the Schurer-Szász-Mirakjan operators [2] $S_{l_1,n}$ with

$$S_{l_1,n} f(x) = e^{-(n+l_1)x} \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) \frac{(n + l_1)^k}{k!} x^k.$$

From (2) and (3) we get

Proposition 9. For every $k \in \mathbb{N}_0$ and $x \in [0, \infty)$

$$M_n^{(k)} f(x) = (n + l_1)^{(k,-l_2)} \sum_{i=0}^{\infty} (-1)^i \frac{\phi_{n+kl_2}^{(i)}(0)}{i!} \, \Delta_{\frac{1}{n}}^{i+k} f(0) \, x^i. \tag{8}$$

In particular,

$$S_{l_1,n}^{(k)} f(x) = (n + l_1)^k S_{l_1,n}(\Delta_{\frac{1}{n}}^k f)(x). \tag{9}$$

Lemma 3 allows us to state the moments of the operators.

Proposition 10. For any $n \in \mathbb{N}$, $r \in \mathbb{N}_0$, one has

$$M_n(t^r) = \sum_{\beta=0}^{r} n^{-\beta} \sum_{s=0}^{min\{r-1,\beta\}} A(r - s, r - \beta) \, \sigma_r^{r-s} \, t^{r-s}, \tag{10}$$

where $A(i, \alpha) = \sum_{j=\alpha}^{i} \binom{j}{\alpha} l_2^{i-j} S_i^j \, (l_1 + l_2(i - 1))^{j-\alpha}$.

Proof. Observe that

$$x^{(i,-l)} = l^i \left(\frac{x}{l} + i - 1\right)^{\underline{i}} = l^i \sum_{j=0}^{i} S_i^j \left(\frac{x}{l} + i - 1\right)^j$$

and in particular

$$(n+l_1)^{(i,-l_2)} = l_2^i \sum_{j=0}^{i} S_i^j \left(\frac{n+l_1}{l_2}+i-1\right)^j = \sum_{j=0}^{i} l_2^{i-j} S_i^j \left(n+l_1+l_2(i-1)\right)^j$$

$$= \sum_{j=0}^{i} l_2^{i-j} S_i^j \sum_{\alpha=0}^{j} \binom{j}{\alpha} n^\alpha \left(l_1+l_2(i-1)\right)^{j-\alpha}$$

$$= \sum_{\alpha=0}^{i} n^\alpha \sum_{j=\alpha}^{i} \binom{j}{\alpha} l_2^{i-j} S_i^j \left(l_1+l_2(i-1)\right)^{j-\alpha}.$$

Lemma 3 yields

$$M_n(t^r) = n^{-r} \sum_{i=1}^{r} (n+l_1)^{(i,-l_2)}\ \sigma_r^i\ t^i; \tag{11}$$

i.e.,

$$M_n(t^r) \quad = \quad n^{-r} \sum_{i=1}^{r} \sum_{\alpha=0}^{i} n^\alpha A(i,\alpha) \quad \sigma_r^i \quad t^i \quad = \quad \sum_{\alpha=0}^{r} n^{\alpha-r} \sum_{i=max\{1,\alpha\}}^{r} A(i,\alpha) \quad \sigma_r^i \quad t^i.$$

Replacing $\alpha = r-\beta$ and $i = r-s$, the proof is concluded. □

A characterization of $S_{l_1,n}$ operators is obtained using Proposition 10. (In fact we use (11)). We show the operator M_n satisfies (9) if and only if M_n is the Schurer-Szász-Mirakjan operator.

Corollary 11.

$$M_n^{(k)} f = (n+l_1)^k M_n(\Delta_{\frac{1}{n}}^k f), \ \textit{for every}\ k \in \mathbb{N}_0 \Leftrightarrow l_2 = 0.$$

Proof. If we suppose that M_n satisfies (9), then in particular (9) holds for $f = t^k$. Now, (11) implies that $(n+l_1)^{(k,-l_2)} = (n+l_1)^k$ and so $l_2 = 0$. □

In order to give the asymptotic expansion for the M_n operators, we study their central moments.

Proposition 12. Given $p \in \mathbb{N}_0$, $n \in \mathbb{N}$ and $x \in [0,\infty)$ the following identity holds:

$$M_n\left((t-x)^p\right)(x) = \sum_{\beta=0}^{p} n^{-\beta} \sum_{r=\beta}^{p} (-1)^{p-r} \binom{p}{r} \sum_{s=0}^{min\{r-1,\beta\}} A(r-s, r-\beta)\ \sigma_r^{r-s}\ x^{p-s}, \tag{12}$$

where $A(\ ,\ )$ is given in Proposition 10.

Proof. From Proposition 10 we obtain

$$\begin{aligned}
M_n\left((t-x)^p\right)(x) &= \sum_{r=0}^{p} \binom{p}{r} (-x)^{p-r} M_n\left(t^r\right)(x) \\
&= \sum_{r=0}^{p} \binom{p}{r} (-1)^{p-r} \sum_{\beta=0}^{r} n^{-\beta} \sum_{s=0}^{min\{r-1,\beta\}} A(r-s, r-\beta)\ \sigma_r^{r-s}\ x^{p-s} \\
&= \sum_{\beta=0}^{p} n^{-\beta} \sum_{r=\beta}^{p} (-1)^{p-r} \binom{p}{r} \sum_{s=0}^{min\{r-1,\beta\}} A(r-s, r-\beta)\ \sigma_r^{r-s}\ x^{p-s}.
\end{aligned}$$

□

Remark 13. Using the same arguments as Sikkema [6] for Szász operators, we can deduce that, in fact, in (12) β runs from $[\frac{p+1}{2}]$, the greatest integer less than or equal to $\frac{p+1}{2}$, to p.

The main result of this section is:

Theorem 14. *Let $f \in E_2$ r times differentiable at $x \in [0, \infty)$. Then*

$$M_n f(x) = f(x) + \sum_{\beta=1}^{r} n^{-\beta} a(\beta, r, f, x) + o(n^{-r}), \tag{13}$$

where

$$a(\beta, r, f, x) = \sum_{p=\beta}^{r} \frac{f^{(p)}(x)}{p!} \sum_{r=\beta}^{p} (-1)^{p-r} \binom{p}{r} \sum_{s=0}^{min\{r-1,\beta\}} A(r-s, r-\beta)\ \sigma_r^{r-s}\ x^{p-s}.$$

Proof. By (5), the proof follows from Proposition 12. □

For the convenience of the reader we list the initial summands of the expansion (13), for $r = 3$;

$$\begin{aligned}
M_n f(x) = f(x) + \frac{l_1 x}{n} f'(x) &+ \frac{nx(1+l_2x) + x(l_1 + l_1(l_1+l_2)x)}{2n^2} f''(x) \\
&+ \left[\frac{nx\left(1 + 3(l_1+l_2)x + l_2(3l_1+2l_2)x^2\right)}{6n^3} \right. \\
&+ \left. \frac{x(l_1 + 3l_1(l_1+l_2)x + l_1(l_1+l_2)(l_1+2l_2)x^2)}{6n^3} \right] f'''(x) + o(n^{-3}),
\end{aligned}$$

as $n \to \infty$.

References

[1] Abramowitz, M. and A. Stegun, "Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables", Dover Publications, Inc. New York, (1972).

[2] Altomare, F. and M. Campiti, “Korovkin-type Approximation Theory and its Applications”, De Gruyter Studies in Mathematics, 17, Walter de Gruyter, Berlin-New York, (1994).

[3] Martini, R., *On the Approximation of Functions together with their derivatives by Certain Linear Positive Operators,* Indag. Math. 31, (1969), 473–481.

[4] Mastroianni, G., *Su un operatore lineare e positivo*, Rend. Accad. Sci. Fis. Mat. Napoli (4) 46 (1979), 161–176 (1980).

[5] Sevy, J.C., *Lagrange and least-squares polynomials as limits of linear combinations of iterates of Bernstein and Durrmeyer polynomials*, J. Approx. Theory 80 (5), (1995), 267–271.

[6] Sikkema, P. C., *On some linear positive operators,* Indag. Math., 32, (1970), 327–337.

Monografías de la Academia de Ciencias de Zaragoza. **20:** 97–110, (2002).

Ideal bases for graded polynomial rings and applications to interpolation

Thomas Sauer

Lehrstuhl für Numerische Mathematik, Justus–Liebig–Universität Gießen

Heinrich–Buff–Ring 44, D–35392 Gießen, Germany

Abstract

Based on a generalized algorithm for the division with remainder of polynomials in several variables, a method for the construction of standard bases for polynomial ideals with respect to arbitrary grading structures is derived. In the case of ideals with finite codimension, which can be viewed upon as a polynomial interpolation problem, an explicit representation for the interpolation space of reduced polynomials can be given.

1 Introduction

We consider polynomial rings in several variables, equipped with a graded structure induced by an arbitrary grading monoid. The goal is a construction method for ideal bases which leads, depending on the underlying grading structure, to the H-bases introduced by Macaulay [9] in 1916 as well as to the Gröbner bases which have been developed by Buchberger [4] in 1965.

For that purpose, let

$$\Pi = \mathbb{K}\,[x_1, \dots, x_d]$$

denote the ring of polynomials over the infinite field $\mathbb{K}$. Let Γ be a monoid (i.e., a semigroup with neutral element) and let "$<$" be a total order on Γ. The monoid Γ is called a *grading monoid* for Π if there is a direct sum decomposition

$$\Pi = \bigoplus_{\gamma \in \Gamma} \Pi_\gamma$$

such that each Π_γ is an abelian group with respect to addition and that

$$\Pi_\gamma \Pi_{\gamma'} \subset \Pi_{\gamma+\gamma'}.$$

Note that this implies that $\mathbb{K} \subset \Pi_0$. Indeed, suppose that there exist $0 \neq K \in \mathbb{K}$ and $0 < \gamma \in \Gamma$ such that $K \in \Pi_\gamma$. Then $1/K \in \Pi_{\gamma'}$ for some $\gamma' \in \Gamma$ and, consequently, we have $1 = K \cdot 1/K \in \Pi_\kappa$, where $\kappa = \gamma + \gamma' \geq \gamma > 0$. But this also yields that $1 \in \Pi_{\kappa+\kappa}$ and since $\kappa + \kappa > \kappa$ this contradicts the fact that the homogeneous spaces form a direct sum decomposition.

The canonical examples for grading monoids are $\mathbb{N}_0$ and grading by total degree, i.e.,

$$\Pi_k = \left\{ \sum_{|\alpha|=n} c_\alpha x^\alpha : c_\alpha \in \mathbb{K} \right\}$$

and $\mathbb{N}_0^d$ where

$$\Pi_\alpha = \left\{ c x^\alpha : c \in \mathbb{K} \right\}.$$

One can decompose any polynomial $p \in \Pi$ into its *homogeneous terms* p_γ, $\gamma \in \Gamma$, writing it as

$$p = \sum_{\gamma \in \Gamma} p_\gamma,$$

where only finitely many terms of the above sum are not zero. A well–ordering "$<$" on Γ naturally determines the notion of the *degree* $\delta : \Pi \to \Gamma$ which is defined for a polynomial $p \in \Pi$ as

$$\delta(p) = \max_< \left\{ \gamma : p_\gamma \neq 0 \right\}.$$

The *leading term* $\Lambda_\Gamma(p)$ of a polynomial p is its maximal homogeneous component; in other words,

$$\Lambda_\Gamma(p) = p_{\delta(p)}.$$

For any set of polynomials $\mathcal{P} \subset \Pi$, the ideal $\langle \mathcal{P} \rangle$ generated by $\mathcal{P}$ is

$$\langle \mathcal{P} \rangle = \left\{ \sum_{g \in \mathcal{P}} q_g g : q_g \in \Pi \right\}.$$

If $\mathcal{I} \subset \Pi$ is any polynomial ideal, then Hilbert's Basissatz tells us that there always exists a finite basis $\mathcal{B}_\mathcal{I} \subset \Pi$ such that $I = \langle \mathcal{B}_\mathcal{I} \rangle$. However, often one is interested in ideal bases which have additional desirable properties and which are also computationally effective. The probably best–known ones are the H–bases and the Gröbner bases which are, in this notation, described by the requirement that any polynomial $p \in \langle \mathcal{G} \rangle$ can be written as

$$p = \sum_{g \in \mathcal{G}} q_g g, \qquad \delta(p) \geq \delta\left(q_g g\right), \, g \in \mathcal{G}, \tag{1}$$

where Γ is either $\mathbb{N}_0$ together with ordering by total degree (in the case of H–bases) or $\mathbb{N}_0^d$ together with a term order (in the case of Gröbner bases). It is exactly the type of basis, characterized by (1), which we want to construct in this paper for arbitrary grading monoids, but without having to refine the grading to a term order. This type of basis was introduced and investigated under the name "*standard basis*" by Robbiano [11] for even more abstract and axiomatically defined graded structures on commutative rings. For our purpose here which deals with the ring of polynomials, we make the following formal definition.

Definition 1 *A (finite) set* $\mathcal{G} \subset \Pi$ *is called a* Γ–basis *(for the ideal* $\langle \mathcal{G} \rangle$*), if any* $p \in \langle \mathcal{G} \rangle$ *can be written as*

$$p = \sum_{g \in \mathcal{G}} q_g g, \qquad \delta(p) \geq \delta\left(q_g g\right), \, g \in \mathcal{G}. \tag{2}$$

Since any ideal has a finite basis and since a Γ–basis is, in particular, a basis for the ideal generated by its member polynomials, we can always assume that a Γ–basis is a finite set. Moreover, we will simply say that "$\mathcal{G}$ is a Γ–basis" instead of "$\mathcal{G}$ is a Γ–basis for $\langle \mathcal{G} \rangle$".

We also recall that an equivalent definition for a Γ–basis would be the requirement that

$$\{\Lambda_\Gamma(p) : p \in \langle \mathcal{G} \rangle\} = \langle \Lambda_\Gamma(g) : g \in \mathcal{G} \rangle. \tag{3}$$

2 A reduction algorithm

In order to formulate the reduction algorithm, we first have to introduce some more notation. Throughout this section, let $\mathcal{P}$ denote a finite set of polynomials and write $\#\mathcal{P}$ for its cardinality. When writing $(\mathcal{P})$ we want to view $\mathcal{P}$ as an *ordered* set in the sense that there exists an increasing chain of subsets

$$\emptyset = \mathcal{P}_0 \subset \mathcal{P}_1 \subset \mathcal{P}_2 \subset \cdots \subset \mathcal{P}_{\#\mathcal{P}} = \mathcal{P}, \qquad \#\mathcal{P}_j = j, \, j = 1, \ldots, \#\mathcal{P},$$

where the order is arbitrary but fixed.

Also, let $(\cdot, \cdot) : \Pi \times \Pi \to \mathbb{K}$ be the scalar product (i.e., the positive definite bilinear form) given by

$$(p, q) = \left(p(D)q\right)(0) = \sum_{\alpha \in \mathbb{N}_0^d} \frac{p_\alpha q_\alpha}{\alpha!}, \tag{4}$$

provided that

$$p(x) = \sum_{\alpha \in \mathbb{N}_0^d} p_\alpha \frac{x^\alpha}{\alpha!} \qquad \text{and} \qquad q(x) = \sum_{\alpha \in \mathbb{N}_0^d} q_\alpha \frac{x^\alpha}{\alpha!}.$$

If $\mathbb{K} = \mathbb{C}$, one has to add complex conjugation and consider the respective sesquilinear form instead; however, we will not dwell on this explicitly. Making use of this scalar product, the Taylor expansion of a polynomial p at the origin becomes

$$p(x) = \sum_{\alpha \in \mathbb{N}_0^d} \left(x^\alpha, p\right) \frac{x^\alpha}{\alpha!}. \tag{5}$$

Though at various points in this paper, in particular in this section, an arbitrary scalar product may be admissible, we want to restrict ourselves to the above "standard" scalar product defined in (4).

For $\gamma \in \Gamma$ we define the homogeneous subspace

$$V_\gamma(\mathcal{P}) = \left\{ \sum_{p \in \mathcal{P}} q_p \Lambda_\Gamma(p) : q_p \in \Pi_{\gamma - \delta(p)},\, p \in \mathcal{P} \right\} \subset \Pi_\gamma$$

with the convention that $\Pi_{\gamma-\delta(p)} = \{0\}$ if $\gamma - \delta(p) \notin \Gamma$. Using the above order on $\mathcal{P}$ and the Hilbert space structure which the scalar product $(\cdot,\cdot)$ defines on Π, we obtain an orthogonal decomposition of $V_\gamma(\mathcal{P})$ as

$$V_\gamma(\mathcal{P}) = \bigoplus_{j=1}^{\#\mathcal{P}} W_\gamma\left(\mathcal{P}_j\right),$$

where

$$W_\gamma\left(\mathcal{P}_j\right) = V_\gamma\left(\mathcal{P}_j\right) \ominus V_\gamma\left(\mathcal{P}_{j-1}\right), \qquad j = 1, \ldots, \#\mathcal{P},$$

i.e.,

$$(W_\gamma(\mathcal{P}_j), V_\gamma(\mathcal{P}_{j-1})) = 0.$$

The goal of the reduction algorithm is to decompose a given $f \in \Pi$ into

$$f = \sum_{p \in \mathcal{P}} q_p p + r, \qquad \delta(p) \geq \delta(q_p p),\, p \in \mathcal{P},$$

where the remainder $r \in \Pi$ should be in a "normalized" or "reduced" form. In the well–known context of Gröbner bases ($\Gamma = \mathbb{N}_0^d$) this means that none of the (monomial) leading terms of $\mathcal{P}$ divides any (monomial) term of r. However, when working with the grading by total degree, for example, the above requirement has to be weakened. It will turn out that orthogonality of any homogeneous term of r yields the "proper" generalization in the sense that the respective reduction process leads to an "algorithmic" characterization of Γ–bases.

Algorithm 2 **Given:** $f \in \Pi$ *and* $\mathcal{P} \subset \Pi$.

While $f \neq 0$*:*

1. *Set* $\gamma = \delta(f)$.

2. *For* $j = 1, \ldots, \#\mathcal{P}$*:*

    - (Orthogonal projections) *Find*

$$q_j^\gamma \in W_\gamma\left(\mathcal{P}_j\right), \qquad q_j^\gamma = \sum_{p \in \mathcal{P}_j} q_{j,p}^\gamma \Lambda_\Gamma(p), \tag{6}$$

such that

$$\left(\Lambda_\Gamma(f) - \sum_{k=1}^{j} q_k^\gamma, W_\gamma(\mathcal{P}_j) \right) = 0. \tag{7}$$

3. Set

$$r_\gamma := \Lambda_\Gamma(f) - \sum_{j=1}^{\#\mathcal{P}} q_j^\gamma. \tag{8}$$

4. Set

$$f := f - \sum_{p\in\mathcal{P}} \left(\sum_{\{j : p \in \mathcal{P}_j\}} q_{j,p}^\gamma \right) p - r_\gamma. \tag{9}$$

Result: *Decomposition*

$$f = \sum_{p\in\mathcal{P}} \left(\sum_{\gamma\in\Gamma} \sum_{\{j : p \in \mathcal{P}_j\}} q_{j,p}^\gamma \right) p + r, \tag{10}$$

where

$$(r_\gamma, V_\gamma(\mathcal{P})) = 0, \qquad \gamma \in \Gamma. \tag{11}$$

Motivated by equation (11) we call a polynomial $q \in \Pi$ *reduced* or *in normal form* with respect to a finite set $\mathcal{P} \subset \Pi$ if all the homogeneous terms of q are perpendicular to the respective $V_\gamma(\mathcal{P})$, i.e., if

$$(q_\gamma, V_\gamma(\mathcal{P})) = 0, \qquad \gamma \in \Gamma, \qquad q = \sum_{\gamma\in\Gamma} q_\gamma.$$

Note, however, that this notion of being reduced depends on the underlying scalar product. Nevertheless, the question whether a polynomial p is reduced with respect to $\mathcal{P}$ does *not* depend on an ordering of $\mathcal{P}$.

Proposition 3 *Algorithm (2) finishes after finitely many steps and the remainder polynomial r satisfies (11).*

Proof: We first remark that the algorithm is well–determined since the orthogonal projection $q_j^\gamma \in W_\gamma(\mathcal{P}_j)$ is unique. For the termination of the algorithm, we only have to notice that

$$r_\gamma = \Lambda_\Gamma(f) - \sum_{p\in\mathcal{P}} \sum_{\{j : p \in \mathcal{P}_j\}} q_{j,p} \Lambda_\Gamma(p),$$

hence the terms of degree γ on the right hand side of (9) are

$$\Lambda(f) - \sum_{p\in\mathcal{P}} \sum_{\{j : p \in \mathcal{P}_j\}} q_{j,p} \Lambda_\Gamma(p) - r_\gamma = 0,$$

which shows that the degree of f is strictly reduced in each step. Therefore, the algorithm terminates after a finite number of steps.

For the second claim, it is easily observed by induction that for $j = 1, \ldots, \#\mathcal{P}$ we have

$$\left(\Lambda_\Gamma(f) - \sum_{k=1}^{j} q_k, V_\gamma(\mathcal{P}_j) \right) = 0.$$

Indeed, for $j = 1$ this is exactly the requirement in the construction of q_1^γ, while for $j > 1$ the induction hypothesis and

$$q_j \in W_\gamma\left(\mathcal{P}_j\right) \perp V_\gamma\left(\mathcal{P}_{j-1}\right)$$

yield

$$\left(\Lambda_\Gamma(f) - \sum_{k=1}^{j} q_k, V_\gamma(\mathcal{P}_{j-1})\right) = 0.$$

Equation (7) and

$$V_\gamma(\mathcal{P}_j) = V_\gamma(\mathcal{P}_{j-1}) \oplus W_\gamma\left(\mathcal{P}_j\right)$$

finally advance the induction hypothesis. □

3 Reduction and Γ–bases

The first result shows that reduction with respect to a Γ–basis has a more "deterministic" outcome than reduction by a general finite set of polynomials.

Theorem 4 *Let $\mathcal{G}$ be a Γ–basis and suppose that $p \in \Pi$ can be written as*

$$p = \sum_{g \in \mathcal{G}} q_g g + r,$$

where r is reduced with respect to $\mathcal{G}$. Then

$$r = p \xrightarrow[(\mathcal{G})]{}$$

Remark 5 *In particular, the above theorem says that for Γ–bases the remainder of the reduction algorithm does* not *depend on the ordering we impose on $\mathcal{G}$. In this case we will simply write $\xrightarrow[\mathcal{G}]{}$.*

Proof of Theorem (4): Let

$$p = \sum_{g \in \mathcal{G}} \tilde{q}_g g + \tilde{r}, \qquad \tilde{q}_g = \sum_{\gamma \in \Gamma} q_{\gamma,g}, \quad \tilde{r} = p \xrightarrow[(\mathcal{G})]{},$$

be the decomposition obtained by Algorithm (2). Then

$$r - \tilde{r} = \sum_{g \in \mathcal{G}} \left(\tilde{q}_g - q_g\right) g \in \langle \mathcal{G} \rangle.$$

Now suppose that $q := r - \tilde{r} \neq 0$. Since each homogeneous term of r and $\tilde{r}$ of any degree γ is orthogonal to the respective $V_\gamma(\mathcal{G})$, the same holds true for q_γ, $\gamma \in \Gamma$, and, in particular,

$$\Big(\Lambda_\Gamma(q), V_{\delta(q)}(\mathcal{G})\Big) = 0. \tag{12}$$

On the other hand, since $q \in \langle \mathcal{G} \rangle$ and since $\mathcal{G}$ is a Γ–basis, we also conclude from (3) that

$$\Lambda_\Gamma(q) \in \langle \Lambda_\Gamma(\mathcal{G}) \rangle \cap \Pi_{\delta(q)} = V_{\delta(q)}(\mathcal{G}). \tag{13}$$

But (12) and (13) are contradictory if $q \neq 0$, hence we must have $q = 0$ or $r = \tilde{r} = p_{\underset{(\mathcal{G})}{\longrightarrow}}$. □

Next, let us recall that, for a finite set $\mathcal{P} \subset \Pi$ of polynomials, a *syzygy* for $\mathcal{P}$ is a vector of polynomials $\boldsymbol{q} \in \Pi^{\mathcal{P}}$ such that

$$\boldsymbol{q} \cdot \mathcal{P} = \sum_{p \in \mathcal{P}} q_p p = 0.$$

The set of all syzygies for $\mathcal{P}$ forms a module, denoted by $S(\mathcal{P})$. It is well–known (cf. [8]) that this module is finitely generated, i.e., there exists a finite basis $\mathcal{S} \subset S(\mathcal{P})$ such that any syzygy $\boldsymbol{q} \in S(\mathcal{P})$ can be written as

$$\boldsymbol{q} = \sum_{\boldsymbol{s} \in S(\mathcal{P})} q_{\boldsymbol{s}} \boldsymbol{s}, \qquad q_{\boldsymbol{s}} \in \Pi,\ \boldsymbol{s} \in S(\mathcal{P}).$$

We also remark that such a basis can be constructed explicitly making use of a reduced Gröbner basis of $\langle \mathcal{P} \rangle$. This has been pointed out by Buchberger in Method 6.17 of his survey paper [5].

Now, there is a Γ–bases analogue of the classical characterization of Gröbner bases via the reduction of the syzygies of leading terms. This result reads as follows.

Theorem 6 *Let $\mathcal{G} \subset \Pi$ be a finite set of polynomials and let $\mathcal{S}$ be a basis of $S(\Lambda_\Gamma(\mathcal{G}))$. Then $\mathcal{G}$ is a Γ–basis if and only if*

$$\boldsymbol{s} \cdot \mathcal{G}_{\underset{(\mathcal{G})}{\longrightarrow}} = 0, \qquad \boldsymbol{s} \in S(\Lambda_\Gamma(\mathcal{G})). \tag{14}$$

Proof: Since $p_{\underset{(\mathcal{G})}{\longrightarrow}}$ is unique for a Γ–basis $\mathcal{G}$ by Theorem (4) and since $\boldsymbol{s} \cdot \mathcal{G} \in \langle \mathcal{G} \rangle$, the direction “⇒” is clear.

The proof of “⇐” follows the argumentation in [10]. Pick any $p \in \langle \mathcal{G} \rangle$ which can be written as

$$p = \sum_{g \in \mathcal{G}} p_g g. \tag{15}$$

We have to show that the polynomials p_g, $g \in \mathcal{G}$, in (15) can be chosen such that $\delta(p) \geq \delta(p_g g)$, $g \in \mathcal{G}$. Assume that this is not the case in equation (15) and set

$$\gamma = \max_{<} \{\delta(p_g g) : g \in \mathcal{G}\},$$

then there is a nonempty subset $\mathcal{G}' \subset \mathcal{G}$ such that

$$\mathcal{G}' = \{g \in \mathcal{G} : \delta(p_g g) = \gamma\} \qquad \text{and} \qquad \delta(p_g g) > \delta(p),\ g \in \mathcal{G}'.$$

Consequently, the leading terms of these polynomials, which belong to Π_γ, have to cancel, i.e.

$$\sum_{g\in\mathcal{G}'} \Lambda_\Gamma(p_g g) = \sum_{g\in\mathcal{G}'} \Lambda_\Gamma(p_g)\, \Lambda_\Gamma(g) = 0$$

and therefore

$$\boldsymbol{q} = (q_g : g \in \mathcal{G})\,, \qquad q_g = \begin{cases} \Lambda_\Gamma(p_g) & g \in \mathcal{G}', \\ 0 & g \in \mathcal{G} \setminus \mathcal{G}', \end{cases}$$

belongs to $S\left(\Lambda_\Gamma(\mathcal{G})\right)$. By assumption,

$$\boldsymbol{q} \cdot \mathcal{G} \underset{(\mathcal{G})}{\rightarrow} = \sum_{\boldsymbol{s}\in\mathcal{S}} q_{\boldsymbol{s}} \left(\boldsymbol{s}\cdot\mathcal{G}\right) \underset{(\mathcal{G})}{\rightarrow} = 0,$$

hence, there exist polynomials $\tilde{p}_g \in \Pi$, $g \in \mathcal{G}$, such that

$$\sum_{g\in\mathcal{G}'} \Lambda_\Gamma(p_g)\, g = \boldsymbol{q}\cdot\mathcal{G} = \sum_{g\in\mathcal{G}} \tilde{p}_g g,$$

where $\delta(p_g g) < \gamma$, since $\delta(\boldsymbol{q}\cdot\mathcal{G}) < \gamma$. This yields that

$$p = \sum_{g\in\mathcal{G}'} \left(p_g - \Lambda_\Gamma(p_g) + \tilde{p}_g\right) g + \sum_{g\in\mathcal{G}\setminus\mathcal{G}'} \left(p_g + \tilde{p}_g\right) g = \sum_{g\in\mathcal{G}} \widehat{p}_g\, g,$$

which is again a representation of the form (15), but now with the property that $\delta(\widehat{p}_g\, g) < \gamma$, $g \in \mathcal{G}$. Repeating this process, we arrive, after finitely many steps, at a "minimal" representation of the form (15), where $\delta(p_g g) \le \delta(p)$, $g \in \mathcal{G}$, which shows that $\mathcal{G}$ is a Γ–basis. ☐

This allows us to finally formulate a crude version of Buchbergers algorithm for the computation of a Γ–basis for the ideal $\langle\mathcal{P}\rangle$, where $\mathcal{P}$ is any finite set of polynomials.

Algorithm 7 **Given:** *Finite set* $\mathcal{P} \subset \Pi$.

```
1. Set $\mathcal{G} = \mathcal{P}$ and $\mathcal{G}' = \emptyset$.
2. While $\mathcal{G}' \neq \mathcal{G}$:
   (a) Set $\mathcal{G}' = \mathcal{G}$.
   (b) Compute a basis $\mathcal{S}$ of $S\left(\Lambda_\Gamma(\mathcal{G})\right)$.
   (c) For $\boldsymbol{s} \in \mathcal{S}$:
       i. Compute $h = \boldsymbol{s}\cdot\mathcal{G}' \underset{(\mathcal{G}')}{\rightarrow}$.
       ii. If $h \neq 0$ then set $\mathcal{G} = \mathcal{G} \cup \{h\}$.
```

Result: Γ–*basis* $\mathcal{G}$.

Theorem 8 *Algorithm (7) generates a* Γ*–basis after finitely many steps.*

Proof: The argument is identical with the one for termination of Buchberger's algorithm for Gröbner bases, cf. [6]. Let $\mathcal{G}_k$, $k \in \mathbb{N}_0$, denote the set $\mathcal{G}$ after the kth step of the algorithm. Then $\mathcal{G}_k \subset \langle \mathcal{P} \rangle$, $k \in \mathbb{N}_0$, and, by construction, as long as $\mathcal{G}_k \subset \mathcal{G}_{k+1}$ is a strict inclusion, then the inclusion of homogeneous ideals $\langle \Lambda_\Gamma (\mathcal{G}_k) \rangle \subset \langle \Lambda_\Gamma (\mathcal{G}_k) \rangle$ is also a strict one. However, after finitely many steps the sequence of homogeneous ideal $\langle \Lambda_\Gamma (\mathcal{G}_k) \rangle$ must stabilise which means that there exists $k_0 \in \mathbb{N}$ such that $\mathcal{G}_k = \mathcal{G}_{k+1}$ for all $k \geq k_0$. But then Γ_{k_0} is a Γ–basis since all syzygies reduce to zero. $\square$

4 Least interpolation

In this section we will connect the technique of Γ–bases to multivariate polynomial interpolation of (Γ–) minimal degree and of Γ–least interpolation which extends and generalizes the approach from [12, 13]. To clarify these notions, we have to introduce some more terminology.

Let $\Theta \subset \Pi'$ be a finite set of linearly independent linear functionals defined on Π. Following a terminology of G. Birkhoff [1], we say that Θ *admits an ideal interpolation scheme* if

$$\ker \Theta = \{p : \Theta(p) = 0\} \subset \Pi$$

is an ideal in Π.

It is well–known that $\theta \in \Pi'$ can be identified with a formal power series $f_\theta \in \mathbb{K}[\![x_1, \dots, x_d]\!]$ by the assignment

$$\theta(p) = (p, f_\theta)\,;$$

clearly, the scalar product $(\cdot, \cdot)$ can be extended to $\Pi \times \mathbb{K}[\![x_1, \dots, x_d]\!]$ since the sum of coefficients still runs over a finite index set only. For example, to the point evaluation functional $\theta = \delta_x$ the power series of $f_\theta(y) = e^{x \cdot y}$ is associated. The following characterization of ideal interpolation schemes has been given by de Boor and Ron [3].

Theorem 9 *A finite set Θ of linear functionals admits an ideal interpolation scheme if and only if the subspace*

$$f_\Theta = span_{\mathbb{K}} \{f_\theta : \theta \in \Theta\} \subset \mathbb{K}[\![x_1, \dots, x_d]\!]$$

is closed under formal differentiation.

Assume that Θ admits an ideal interpolation scheme. We say that a linear subspace $\mathcal{P} \subset \Pi$ is an *interpolation space* with respect to Θ if for any $y \in \mathbb{K}^\Theta$ there exists a *unique* polynomial $p \in \mathcal{P}$ such that

$$\Theta(p) = y.$$

If $\mathcal{P}$ is an interpolation space with respect to Θ, then we denote by

$$L(\mathcal{P};\cdot):\mathbb{K}^\Theta\to\mathcal{P}$$

the *interpolation operator* (which is a clearly linear operator). Moreover, $L(\mathcal{P};\Theta(\cdot)):\Pi\to\mathcal{P}$ is a linear projection from $\Pi\to\mathcal{P}$. We say that $\mathcal{P}$ is a Γ*–minimal degree interpolation space* with respect to Θ if $\mathcal{P}$ is an interpolation space and the projection $L(\mathcal{P};\Theta(\cdot))$ is *degree reducing*, i.e.,

$$\delta\left(L\left(\mathcal{P};\Theta(p)\right)\right)\le\delta(p),\qquad p\in\Pi,$$

which is a desirable behaviour of polynomial projections.

The following result tells us that the normal forms (or, reduced polynomials) with respect to a Γ–basis $\mathcal{G}$ for ker Θ are always a canonical minimal degree interpolation space with respect to Θ.

Theorem 10 *Suppose that $\Theta\subset\Pi'$ admits an ideal interpolation scheme and let $\mathcal{G}$ be a Γ–basis of* ker Θ*. Then $\mathcal{P}_\Theta=\Pi_{\underset{\mathcal{G}}{\rightarrow}}$ is a Γ–minimal degree interpolation space with respect to Θ and*

$$L\left(\mathcal{P}_\Theta;\Theta(p)\right)=p_{\underset{\mathcal{G}}{\rightarrow}}.$$

Proof: Since the functionals in Θ are linearly independent, there exist *dual polynomials* $p_\Theta\in\Pi^\Theta$ such that

$$\theta\left(p_{\theta'}\right)=\delta_{\theta,\theta'},\qquad\theta,\theta'\in\Theta.$$

Since $\Theta\left(p-p_{\underset{\mathcal{G}}{\rightarrow}}\right)=0$ for any $p\in\Pi$, the polynomials $p_{\Theta\underset{\mathcal{G}}{\rightarrow}}$ are also dual to Θ and therefore also linearly independent. Consequently, for any data $y\in\mathbb{K}^\Theta$, the polynomial

$$p_y=\sum_{\theta\in\Theta}y_\theta p_{\theta\underset{\mathcal{G}}{\rightarrow}}\in\Pi_{\underset{\mathcal{G}}{\rightarrow}}$$

satisfies $\Theta(p_y)=y$. In addition, all polynomials $p\in\Pi$ which have the property that $\Theta(p)=y$, differ by an element of $\langle\mathcal{G}\rangle$ and therefore the p_y above is the *unique* (because $\mathcal{G}$ is a Γ–basis) reduced interpolant which proves that $\Pi_{\underset{\mathcal{G}}{\rightarrow}}$ is an interpolation space where the interpolation operator is given by reduction. Since the reduction process is also degree–reducing, we finally find that $\Pi_{\underset{\mathcal{G}}{\rightarrow}}$ is a Γ–minimal degree interpolation space. $\square$

For a power series $f\in\mathbb{K}[\![x_1,\dots,x_d]\!]$ we denote its Γ*–least term* by

$$\lambda_\Gamma(f)=\min_\gamma\left\{f_\gamma:f_\gamma\ne0\right\},$$

in other words, its homogeneous term of Γ–minimal degree. The vector space of the least terms of a subspace F of $\mathbb{K}[\![x_1,\dots,x_d]\!]$ will be denoted by $\lambda_\Gamma(F)$. The *least space* related to Θ, $\lambda_\Gamma\left(f_\Theta\right)$, has been discovered as very useful in interpolation by de Boor and Ron [2] for the case of grading by total degree.

Finally, we assume that the grading induced by Γ is a *monomial grading* which requires that for $\gamma \in \Gamma$

$$\Pi_\gamma = \operatorname{span}_{\mathbb{K}} \{x^\alpha : \alpha \in I_\gamma\}, \qquad I_\gamma \subset \mathbb{N}_0^d,\ \gamma \in \Gamma.$$

In particular, each monomial x^α, $\alpha \in \mathbb{N}_0^d$, belongs to some homogeneous space Π_γ; we denote the respective index γ by $\gamma(\alpha)$.
Then we have the following result.

Theorem 11 *Suppose that* $\Theta \subset \Pi'$ *admits an ideal interpolation scheme and let* $\mathcal{G}$ *be a* Γ*–basis of* $\ker \Theta$*. Then*

$$\Pi_{\overrightarrow{\mathcal{G}}} = \lambda_\Gamma (f_\Theta).$$

Corollary 12 *If* $\Theta \subset \Pi'$ *admits an ideal interpolation scheme and* Γ *is a monomial grading, then* $\lambda_\Gamma (f_\Theta)$ *is a* Γ*–minimal degree interpolation space.*

For the prove of the theorem we begin with collecting some auxiliary results. First we remark that it follows directly from (4) that for $f, g, p \in \Pi$ one has

$$(f, p(D)g) = (fp, g). \tag{16}$$

Next, we give a simple observation on monomial gradings.

Lemma 13 *Suppose that* Γ *induces a monomial grading. If* $p, q \in \Pi$ *satisfy* $\delta(p) < \delta(q)$*, then* $\Lambda_\Gamma(q)(D)p = 0$*.*

Proof: Pick any $\beta \in \mathbb{N}_0^d$ such that $\gamma(\beta) > \delta(p)$. By (5) and (16) we obtain that

$$D^\beta p = \sum_{\alpha \in \mathbb{N}_0^d} \left(x^\alpha, D^\beta p\right) \frac{x^\alpha}{\alpha!} = \sum_{\alpha \in \mathbb{N}_0^d} \left(x^{\alpha+\beta}, p\right) \frac{x^\alpha}{\alpha!}.$$

On the other hand, equation (5) yields that $\left(x^{\alpha+\beta}, p\right)$ is the coefficient of p with respect to the monomial

$$\frac{x^{\alpha+\beta}}{(\alpha+\beta)!} \in \Pi_{\gamma(\alpha)+\gamma(\beta)}.$$

Since $\gamma(\alpha) + \gamma(\beta) \geq \gamma(\beta) > \delta(p)$ and since the grading yields a direct sum decomposition, this coefficient has to be zero. Since

$$\Lambda(q) \in \operatorname{span}_{\mathbb{K}} \left\{x^\alpha : \alpha \in I_{\delta(q)}\right\},$$

the result follows. □

To prove the theorem, we make use of the following additional characterization of reduced polynomials with respect to a Γ–basis $\mathcal{G}$.

Proposition 14 *Let $\mathcal{G}$ be a Γ–basis for $\langle \mathcal{G} \rangle$ and let Γ induce a monomial grading. Then a polynomial $p \in \Pi$ is reduced with respect to $\mathcal{G}$ if and only if*

$$p \in \bigcap_{g \in \mathcal{G}} \ker \Lambda_\Gamma(g)(D). \tag{17}$$

Proof: We first remark that

$$\bigcap_{g \in \mathcal{G}} \ker \Lambda_\Gamma(g)(D) = \bigcap_{g \in \langle \mathcal{G} \rangle} \ker \Lambda_\Gamma(g)(D). \tag{18}$$

Indeed, the inclusion "$\supset$" is trivial since $\mathcal{G} \subset \langle \mathcal{G} \rangle$. For "$\subset$" we pick any $q \in \langle \mathcal{G} \rangle$, and write it as $q = \sum_g q_g g$. Since $\mathcal{G}$ is a Γ–basis for $\langle \mathcal{G} \rangle$, we know that $\delta(q_g g) \le \delta(q)$ and defining the subset $\mathcal{G}' \subset \mathcal{G}$ as

$$\mathcal{G}' = \{ g \in \mathcal{G} : \delta(q_g g) = \delta(q) \}$$

we have that

$$\Lambda_\Gamma(q) = \Lambda_\Gamma \left(\sum_{g \in \mathcal{G}'} q_g g \right) = \sum_{g \in \mathcal{G}'} \Lambda_\Gamma(q_g) \Lambda_\Gamma(g),$$

hence,

$$\Lambda_\Gamma(q)(D)p = \sum_{g \in \mathcal{G}'} \Lambda_\Gamma(q_g)(D) \underbrace{(\Lambda_\Gamma(g)(D)p)}_{=0} = 0.$$

Now, pick any *homogeneous* polynomial $p \in \Pi_\gamma$ for some $\gamma \in \Gamma$ and $q \in \langle \mathcal{G} \rangle$. If $\delta(p) < \delta(q)$, then $\Lambda_\Gamma(q)(D)p = 0$ by Lemma (13). If, on the other hand, $\delta(p) \ge \delta(q)$, then

$$\Lambda_\Gamma(q)(D)p = \sum_{\alpha) \in \mathbb{N}_0^d} (x^\alpha, \Lambda_\Gamma(q)(D)p) \frac{x^\alpha}{\alpha!} = \sum_{\{\alpha : \delta(q) + \gamma(\alpha) = \gamma\}} (x^\alpha \Lambda_\Gamma(q), p) \frac{x^\alpha}{\alpha!}.$$

Hence, $\Lambda_\Gamma(q)(D)p = 0$ holds if and only if

$$(x^\alpha \Lambda_\Gamma(q), p), \qquad \gamma(\alpha) + \delta(q) = \gamma.$$

However,

$$V_\gamma(\mathcal{G}) = \bigoplus_{g \in \mathcal{G}} \operatorname{span}_{\mathbb{K}} \{ x^\alpha g : \gamma(\alpha) + \delta(g) = \gamma \}$$

and therefore $\Lambda_\Gamma(q)(D)p = 0$, $q \in \langle \mathcal{G} \rangle$ is equivalent to

$$(p, V_\gamma(\mathcal{G})) = 0.$$

This immediately yields the statement of the proposition. □

Hence, the proof of Proposition (14) also yields the following description of the joint kernels of homogeneous differential operators, cf. [7].

Corollary 15 *Suppose that Θ admits an ideal interpolation scheme and let $\mathcal{G}$ be a Γ–basis of* $\ker \Theta$. *Then*

$$\Pi_{\overrightarrow{\mathcal{G}}} = \bigcap_{g \in \mathcal{G}} \ker \Lambda_\Gamma(g)(D)$$

Another immediate consequence is the following "algorithmic" description of the joint kernels of homogeneous differential operators with constant coefficients.

Corollary 16 *Let $\mathcal{P} \subset \Pi$ be a finite set of homogeneous polynomials and let $\mathcal{G}$ be a Γ–basis for $\langle \mathcal{P} \rangle$. Then*

$$\bigcap_{p \in \mathcal{P}} \ker p(D) = \Pi_{\overrightarrow{\mathcal{G}}}.$$

Proof of Theorem (11): By Corollary (15) it suffices to prove that

$$\bigcap_{g \in \langle \mathcal{G} \rangle} \ker \Lambda_\Gamma(g)(D) = \lambda_\Gamma \left(f_\Theta \right). \tag{19}$$

To prove the inclusion "$\supset$", we assume that there exists some $f \in \mathbb{K}[\![x_1, \dots, x_d]\!]$, $f = \sum_\theta c_\theta f_\theta$, and $q \in \langle \mathcal{G} \rangle = \ker Theta$ such that

$$\Lambda_\Gamma(q)(D)\lambda_\Gamma(f) \neq 0.$$

Hence, $\delta(q) = \delta\left(\lambda_\Gamma(f)\right)$ and therefore, by Lemma (13) we have that

$$\Lambda_\Gamma(q)(D)\lambda_\Gamma(f) = \left(q(D)f\right)(0) = (f, q) = \sum_{\theta \in \Theta} c_\theta \left(f_\theta, q\right) = \sum_{\theta \in \Theta} c_\theta \theta(q) = 0$$

since $q \in \ker \Theta$, which is a contradiction.

Conversely, since the functionals in Θ were assumed to be linearly independent and therefore

$$\lambda_\Gamma \left(\sum_{\theta \in \Theta} c_\theta f_\theta \right) = 0 \quad \Leftrightarrow \quad \sum_{\theta \in \Theta} c_\theta f_\theta = 0 \quad \Leftrightarrow \quad c_\theta = 0,\ \theta \in \Theta,$$

we conclude that

$$\dim \lambda_\Gamma(f_\Theta) = \dim f_\Theta = \dim \operatorname{span}_{\mathbb{K}} \Theta = \#\Theta = \dim \mathcal{P}_\Theta.$$

Hence, $\lambda_\Gamma\left(f_\Theta\right)$ is a linear subspace of $\mathcal{P}_\Theta$ which has the same dimension. Consequently, the spaces are the same. □

References

[1] G. Birkhoff. The algebra of multivariate interpolation. In C.V. Coffman and G.J. Fix, editors, *Constructive Approaches to Mathematical Models*, pages 345–363. Academic Press Inc., 1979.

[2] C. de Boor and A. Ron. On multivariate polynomial interpolation. *Constr. Approx.*, **6** (1990), 287–302.

[3] C. de Boor and A. Ron. The least solution for the polynomial interpolation problem. *Math. Z.*, **210** (1992), 347–378.

[4] B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal.* PhD thesis, Innsbruck, 1965.

[5] B. Buchberger. Gröbner bases: An algorithmic method in polynomial ideal theory. In N. K. Bose, editor, *Multidimensional Systems Theory*, pages 184–232. D. Reidel Publishing Company, 1985.

[6] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms.* Undergraduate Texts in Mathematics. Springer–Verlag, 2. edition, 1996.

[7] W. Dahmen and C. A. Micchelli. Local dimension of piecewise polynomial spaces, syzygies, and the solutions of systems of partial differential operators. *Math. Nachr.*, **148** (1990), 117–136.

[8] W. Gröbner. *Algebraische Geometrie II.* Number 737 in B.I–Hochschultaschenbücher. Bibliographisches Institut Mannheim, 1970.

[9] F. S. Macaulay. *The Algebraic Theory of Modular Systems.* Number 19 in Cambridge Tracts in Math. and Math. Physics. Cambridge Univ. Press, 1916.

[10] H. M. Möller. On the construction of Gröbner bases using syzygies. *J. Symbolic Comput.*, **6** (1988), 345–359.

[11] L. Robbiano. On the theory of graded structures. *J. Symbolic Computation*, **2** (1986), 139–170.

[12] T. Sauer. Polynomial interpolation of minimal degree and Gröbner bases. In B. Buchberger and F. Winkler, editors, *Groebner Bases and Applications (Proc. of the Conf. 33 Years of Groebner Bases)*, volume 251 of *London Math. Soc. Lecture Notes*, pages 483–494. Cambridge University Press, 1998.

[13] T. Sauer. Gröbner bases, H–bases and interpolation. *Trans. Amer. Math. Soc.*, **353** (2001), 2293–2308.

Monografías de la Academia de Ciencias de Zaragoza. **20:** 111–135, (2002).

Testing Methods for 3D Scattered Data Interpolation

Mira Bozzini and Milvia Rossini

Department of Mathematics and Applications. Univ. of Milano Bicocca
via Bicocca degli Arcimboldi 8 I-20126 Milano Italy
email: `bozzini@matapp.unimib.it, rossini@matapp.unimib.it`

Abstract

This paper is concerned with the evaluation of methods for 3D scattered data interpolation. In particular, we discuss the computational performances in 3D of methods which give superior results for the two-dimensional case. The testing process was carried out by considering the accuracy, the graphical behaviour of the interpolant and the timing. In addition we have taken into account the computational efficiency and the sensitivity respect to the sample. Moreover, in order to evaluate the graphical behaviour, we present an evolutive visualization of the interpolant.

1 Introduction

The problem of constructing a smooth function $g(x, y, z)$, $g : Q \subset \mathbb{R}^3 \to \mathbb{R}$ which takes on certain prescribed values

$$g(x_i, y_i, z_i) = f_i, \quad i = 1, \ldots, N, \quad (x_i, y_i, z_i) \in Q, \tag{1}$$

arises in many applied fields. We mention some examples that can be useful for a computational analysis:

- behaviour of the temperature in a furnace,
- concentrations of a mineral in the soil,
- behaviour of precipitations in a geographical area,
- electroencephalogram (EEG).

In the literature there are methods which can be extended to any dimension d (for instance the radial basis methods) and the theoretic properties of some of these have been studied (see [17]), but little or nothing is known on the computational results for $d = 3$.

For the case $d = 2$, Franke has evaluated (see his well–known paper [8]) the numerical performances of a wide class of methods. For $d = 3$, there is the paper [13] which gives a first answer to the problem for data sets with a moderate dimension.

Our aim is to study the computational behaviour of methods to interpolate $3D$ scattered data. Taking into account [8], we have analyzed those methods which have shown, at least in two dimensions, the better performances.

In the testing process we have considered, as usual, the accuracy and, in addition, all those aspects that in the $3D$ case can give more problems than in two dimensions. Namely we have considered the stability and the computational costs. By the definition of two indices, we provide the computational efficiency of the different methods. In addition we think it was important to analyze the conditioning of the sample respect to the approximating function. Finally we provide a new visualization method, useful for the comparison of the graphical behaviour of the interpolants. For sake of brevity, we will report only a few examples. The full experimentation can be found in [5].

2 The testing process

2.1 The methods considered

In [8] it is shown that the methods giving the superior performances (see table 1 of [8]), belong to the classes of inverse distance weighted methods (the modified quadratic Shepard Method), triangle based methods (Nielson minimum norm network), radial basis methods (thin-plate splines, multiquadrics).

The triangle based methods need a triangulation of the convex hull of the point set and, to achieve a C^1 function, they need also a scheme for estimating some derivatives at the data points. The quality of the corresponding interpolating function (as measured in terms of visual appearance, smoothness, and accuracy) depends critically on the "accuracy" of the derivative estimates. Moreover, it is known that triangulations have to satisfy some optimal conditions that, at the moment, do not exist for three dimensions.

For these reasons, we have considered only inverse distance weighted methods (the modified quadratic Shepard method (MQSM)), and the standard radial basis functions

$$\phi(r) = r^3, \tag{2}$$

$$\phi(r) = r^2 logr, \tag{3}$$

$$\phi(r) = \sqrt{r^2 + c^2} \tag{4}$$

$$\phi(r) = (1-r)_+^4(1+4r). \tag{5}$$

The reader can see the Appendix A for a short description of these methods.

2.2 Test functions

The methods have been tested on functions that exhibit a variety of behaviours. Namely, we have considered the $3D$ extension of the functions used in [8]

$$\begin{aligned}
\mathbf{F2}(x,y,z) &= \left(\tanh(9z-9x-9y)+1\right)/9 \\
\mathbf{F3}(x,y,z) &= \cos(6z)\left(1.25+\cos(5.4y)\right)/\left(6+6\left(3x-1\right)^2\right) \\
\mathbf{F4}(x,y,z) &= \exp\left\{-\frac{81}{16}\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]\right\}/3 \\
\mathbf{F5}(x,y,z) &= \exp\left\{-\frac{81}{4}\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]\right\}/3 \\
\mathbf{F6}(x,y,z) &= \sqrt[2]{64-81\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]}/9-0.5.
\end{aligned}$$

In addition we have considered some functions with particular behaviours. A “relative” of the two-dimensional Franke’s function

$$\begin{aligned}
\mathbf{F1}(x,y,z) = {} & .75\exp\left[-\frac{(9x-2)^2+(9y-2)^2+(9z-2)^2}{4}\right] \\
& +0.75\exp\left[-\frac{(9x+1)^2}{49}-\frac{(9y+1)^2}{10}-\frac{(9z+1)^2}{10}\right] \\
& +0.5\exp\left[-\frac{(9x-7)^2+(9y-3)^2+(9z-5)^2}{4}\right] \\
& -0.2\exp\left[-(9x-4)^2-(9y-7)^2-(9z-5)\right],
\end{aligned}$$

an extension of the sigmoidal function

$$\mathbf{F7}(x,y,z) = 1/\sqrt{1+2exp(-3(\sqrt{x^2+y^2+z^2}-6.7))},$$

and the peak function

$$\begin{aligned}
\mathbf{F8}(x,y,z) &= 50exp(-200((x-0.3)^2+(y-0.3)^2)) \\
&+ exp(-50((x-0.5)^2+(y-0.5)^2)).
\end{aligned}$$

These functions have been considered in $Q=[0,1]^3$.

2.3 Data configurations

The test functions have been sampled at N scattered points of Q. Let

$$S_N = \{P_i(x_i, y_i, z_i) \in Q = [0,1]^3,\ i = 1, \ldots, N\}$$

be the point set.

The word "*scattered*" may have different meanings.

In [8], scattered means that the points of S_N are not assumed to satisfy any particular condition as spacing or density. (In a simulation process, the point are obtained from the generation of pseudo random numbers.)

In practice, many people use points that satisfy some condition. We can have

- *equidistributed scattered points*. For instance, they can be obtained by a pseudo random number generation so that one point falls in each subcube of side $1/\sqrt[3]{N}$.
- *Perturbed grid points*. If G_i are the points of a grid G, the points of S_N are

  $$P_i = G_i + e_i,$$

  where e_i are random variables with mean $E(e_i) = 0$ and variance (which measures the distortion from the grid points) $E(e_i^2) = \sigma^2$.

We call all these point sets *volumetric data*.

In many applications, the data points are scattered but with some structure. For instance, when we study mineral concentrations in the subsoil, we pick up the data by drillings. This mean that the points P_i are scattered along some straight lines of Q. In other situations, the points of S_N can be scattered on certain number of planes or, more generally, on some surfaces (for instance when we study the monthly or seasonal precipitations).

These kind of data (say *structured scattered data*) have been discussed in [6].

2.4 Comparison

It is now important to choose the characteristics on which the methods are to be evaluated and compared.

In our opinion, two fundamental aspects are the accuracy and the graphical behavior. In addition, we have also considered the computation times and the condition of the interpolation matrices.

Accuracy. We measure the accuracy with the root mean square error e_2 and the maximum error e_∞:

$$e_2 = \sqrt{\frac{\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{l=1}^{n_z}(g(x_i,y_j,z_l) - f(x_i,y_j,z_l))^2}{n_x n_y n_z}}, \tag{6}$$

$$e_\infty = \max_{(x_i,y_j,z_l)\in G} |g(x_i,y_j,z_l) - f(x_i,y_j,z_l)|, \tag{7}$$

where $g(x, y, z)$ is the interpolant function, $f(x, y, z)$ is the test function, and (x_i, y_j, z_l) are the points of a grid G of Q with size $n_x \times n_y \times n_z$.

These indices provide a global information on the resulting approximation.

Graphical visualization. The interpolant graphic gives an immediate indication on the approximation goodness. In fact, different methods could give the same accuracy but some may reproduce the function behaviour more faithfully than the others.

Since trivariate function representations describe hypersurfaces in $\mathbb{R}^4$, it is obvious that we cannot render them directly. It is possible to visualize their behaviour by displaying one or more projected surfaces in $\mathbb{R}^3$. For example, we can work with isoparametric surfaces which correspond to constant values of one variable x, y or z. Alternatively, we can create contour plots of hypersurfaces which correspond to constant function values. Surfaces of constant parameter values are considered in [12], [7], [1]. [15] and [14] suggest using a combination of three isoparametric surfaces $g(x_i, y, z)$, $g(x, y_j, z)$ and $g(x, y, z_k)$, associated with the value $g(x_i, y_j, z_k)$. This idea involves an axonometric view of the domain with the planes $x = x_i$, $y = y_j$ and $z = z_k$, along with another copy of the domain with the graphs of the three surfaces located on the faces of the cubical domain. The paper [14] contains another example of visualization named *slice viewer.* A survey of other techniques can be found in [11].

Although these techniques are very useful, they still provide a static representation which do not allow to have a dynamic vision of the problem we are studying.

Usually, in a real problem, the variation of one independent variable represents the evolution of some phenomenon. For instance, to study mineral concentrations in the subsoil means to study how the concentration changes in relation to depth.

Therefore, we believe that the visualization should be connected to the problem features and point out the evolution of the phenomenon with respect to the variable describing its variation.

For this reason, we have considered an evolutive representation. We pick up the variable describing the evolution (for instance z), we take some values of it $(z_1, \dots, z_m)$, and we consider the evolution of the surfaces $g(x, y, z_k)$, $k = 1, \dots, m$. (See figs. 1–10 in the Appendix).

3 Volumetric data

In this section we deal with the interpolant behavior when the sample dimension N increases.

First we will compare the method performances both with respect to the accuracy (§3.1) and to the graphical visualization (§3.2).

Then we will consider the condition numbers of the interpolation matrices and the computational efficiency in relation to the sample size N. Finally, in §3.4 we will discuss the sensitivity of the methods with respect to the point locations (for point sets S_N with the same size).

Here we have considered the case of scattered equidistributed data, because, for this case, the theoretical results for radial basis functions surely hold. Examples for the other point sets (scattered and perturbed grid points) can be found in [5].

3.1 Accuracy

The values of e_2 and e_∞ for N=343, 1000, 3375 are shown in tables 1, 2 and 3. These indices have been computed using the points of a grid of size $n_x = n_y = n_z = 21$.

The tables show that the methods give equivalent results. We can also say that, in general, the radial basis function (2) provides the better accuracy.

Moreover, the error decreases, as N increases, according to the theory [18].

Remark. For the multiquadric, we know that the optimal choice of c is still an open problem. There are some suggestions for the two–dimensional case, including methods based only on the points distribution. [20] suggests a value that takes account of how the points are dispersed in both the x and y directions

$$c = \sqrt{1/10 \max\{\max_{i,j}|x_i - x_j|, \max_{i,j}|y_i - y_j|\}}. \tag{8}$$

[10] uses

$$c = 0.851d, \tag{9}$$

where d is the average distance of the points to their near neighbors. Franke replaced d by $D/\sqrt{N}$, where D is the diameter of the minimal circle enclosing all data points and suggests to use

$$c = 1.25D/\sqrt{N}. \tag{10}$$

These techniques can be trivially extended to the $3D$ case and they provide more or less the same results (both for the accuracy and the graphical behaviour [5]). We have not considered techniques that take account of the function values f_i (see, for instance, [16]) because we believe they are too expensive for our case.

In tables 1, 2, 3, the c parameter has been computed using (9).

3.2 Graphical visualization

Let us now compare the graphical performances. For sake of brevity, we will show only the graphs related to the test functions $F1$ and $F2$ (which exhibit particular behaviors),

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	4.6e-3	8.5e-3	4.2e-3	5.1e-4	6.5e-4	5.4e-3	2.1e-2	2.6
	e_∞	4.2e-2	7.8e-2	4.5e-2	5.6e-3	8.3e-3	1.3e-1	2.2e-1	42e
$r^2\log(r)$	e_2	5.9e-3	9.1e-3	6.9e-3	1.1e-3	1.3e-3	8.7e-3	2.2e-2	2.6
	e_∞	9.3e-2	7.5e-2	7.1e-2	1.4e-2	1.7e-2	1.8e-1	2.1e-1	42
$\sqrt{r^2+c^2}$	e_2	6.6e-3	9.2e-3	8.2e-3	8.3e-4	1.2e-3	1.2e-2	2.3e-2	2.7
	e_∞	1.5e-1	8.2e-2	8.2e-2	1.0e-2	1.5e-2	2.3e-1	2.2e-1	42
$(1-r)_+^4(1+4r)$	e_2	5.3e-3	8.5e-3	3.9e-3	7.3e-4	8.6e-4	1.2e-2	2.1e-2	2.8
	e_∞	1.2e-1	7.3e-2	4.6e-2	8.7e-3	8.4e-3	2.3e-2	2.2e-1	43
MQSM	e_2	9.4e-3	9.7e-3	5.9e-3	2.2e-3	2.1e-3	4.5e-3	2.3e-2	3.5
	e_∞	1.3e-1	9.1e-2	6.1e-2	1.4e-2	1.9e-02	1.2e-1	1.9e-1	46

Table 1: Errors for $\mathbf{N} = \mathbf{343}$.

and to $F5$ (for which we get the better accuracy). In fig. 1 their evolutions respect to the z–variable are shown.

We begin to consider the interpolants achieved with (2), which provides the better accuracy.

Figures 2, 3, 4 point out that, generally, samples of size $N = 343$ do not assure a good phenomenon reproduction. Only for $F5$, we get adeguate results.
In this case, for $N = 1000$, we have a remarkable improvement at the boundary (see fig. 2).
This does not happen for $F1$ and $F2$. For $F1$, the interpolant does not reproduce the correct behaviour near the boundary $z = 1$. The approximation in this zone becomes better when we consider $N = 3375$. But on the boundary $z = 1$ the function behavior has not yet been reproduced (fig. 3).

For $F2$ (fig. 4), we may observe an improvement from $N = 343$ to $N = 1000$, even if we have many oscillations that are eliminated, only in part, considering $N = 3375$.

Let us now consider the other methods. As already said, the interpolants achieved from the radial basis (3)–(5), provide, in mean, the same accuracy of (2), but their graphical behaviors can be worse. See, for instance, fig. 5 where the $F1$ interpolants are shown ($N = 1000$).

Finally, let us consider the modified quadratic Shepard method. The experimentation has shown that, near the boundary, it can perform worse than the other methods. But for samples with $N > 1000$, it provides approximations comparable to those given by (2) and with lower computational costs (see §3.3). In fig. 6 we show the graphs of the $F1$ MQSM interpolants ($N = 343, 1000, 3375$).

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	1.5e-3	4.7e-3	1.3e-3	1.5e-4	1.1e-4	3.3e-3	7.8e-3	1.4
	e_∞	4.0e-2	7.3e-2	1.4e-2	1.8e-3	1.3e-3	9.6e-2	1.1e-1	34
$r^2\log(r)$	e_2	2.2e-3	5.4e-3	2.2e-3	4.4e-4	2.4e-4	5.3e-3	1.0e-2	1.5
	e_∞	5.3e-2	8.0e-2	2.0e-2	7.2e-3	3.5e-3	1.4e-1	1.3e-1	35
$\sqrt{r^2+c^2}$	e_2	2.9e-3	5.6e-3	3.0e-3	4.0e-4	2.7e-4	7.9e-3	1.1e-2	1.6
	e_∞	9.1e-2	8.3e-2	3.4e-2	6.7e-3	4.3e-3	1.8e-1	1.4e-1	35
$(1-r)_+^4(1+4r)$	e_2	1.8e-3	4.7e-3	9.7e-4	2.9e-4	2.5e-4	6.7e-3	8.4e-3	1.4
	e_∞	5.8e-2	7.0e-2	1.2e-2	4.5e-3	2.6e-3	1.7e-1	1.2e-1	33
MQSM	e_2	2.7e-3	4.8e-3	2.1e-3	4.6e-4	7.9e-4	2.7e-3	1.0e-2	1.1
	e_∞	3.9e-2	9.0e-2	2.7e-2	4.3e-3	8.7e-3	8.0e-2	1.0e-1	30

Table 2: Errors for $\mathbf{N} = \mathbf{1000}$.

We can evaluate the graphical performances also considering the absolute error graphs. We will see that this can be useful, but at the same, time misleading.
In fig. 7, the errors of the $F1$ interpolants are shown ($N = 343$, 1000, 3375, radial basis (2)).
In each graph, we have used the scale given by the maximum error obtained for $N = 343$. We can notice that the higher errors are at the extrema, while near the boundary $z = 1$, the approximation seems to be accurate. But we have remarked that, just in this zone, the interpolant does not reproduce the correct $F1$ behavior (see fig. 1 and 3).

3.3 Conditioning, computation times and efficiency

In this section we shall see how the condition number $K_2(A)$ of the interpolation matrix, the computation times and the efficiency change with regard to the sample dimension N.

- **Condition Number.** For the global methods (2)–(5), $K_2(A)$ increases as shown in table 4.

For the local method MQSM, we do not have this problem because we solve a linear system of small dimension ($n = 9$).

- **Computation times**. For the methods (2)–(5), the increase of sample size leads to a remarkable increase of computation times. This is due both to the solution of a linear system with full matrix of order N, and to interpolant evaluation. But efficient methods exist for the evaluation (see, for instance, [2]). Therefore, in table 5[1], we show only the system solution times (radial bases (2)). For MQSM, the times include also the interpolant

---

[1] The computations have been performed on a PC with a AMD 800 Mhz processor and 256 Mb Ram.

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	3.7e-4	1.3e-3	5.2e-4	3.7e-5	2.2e-5	1.9e-3	2.4e-3	4.1e-1
	e_∞	8.0e-3	2.7e-2	9.5e-3	4.5e-4	2.6e-4	6.6e-2	3.9e-2	13
$r^2 \log(r)$	e_2	8.3e-4	1.6e-3	1.1e-3	1.6e-4	4.6e-5	3.3e-3	3.7e-3	4.9e-1
	e_∞	3.5e-2	3.4e-2	1.7e-2	3.5e-3	6.6e-4	9.8e-2	6.0e-2	15
$\sqrt{r^2+c^2}$	e_2	1.8e-3	1.8e-3	1.7e-3	2.0e-4	7.6e-5	4.8e-3	4.6e-3	4.8e-1
	e_∞	8.0e-2	3.7e-2	2.6e-2	4.3e-3	1.1e-3	1.4e-1	7.8e-2	16
$(1-r)_+^4(1+4r)$	e_2	7.4e-4	1.2e-3	3.9e-4	9.5e-5	7.3e-5	3.3e-3	2.6e-3	4.1e-1
	e_∞	4.0e-2	2.7e-2	7.0e-3	2.1e-3	7.3e-4	1.1e-1	4.4e-2	13
MQSM	e_2	8.7e-4	1.6e-3	6.5e-4	1.4e-4	2.4e-4	1.5e-3	3.4e-3	5.6e-1
	e_∞	1.5e-2	2.4e-2	1.2e-2	1.3e-3	2.5e-3	4.9e-2	5.3e-1	15

Table 3: Errors for $\mathbf{N} = \mathbf{3375}$.

$\phi(r)$	$N = 343$	$N = 1000$	$N = 3375$
r^3	3.9e+06	2.8e+07	9.5e+08
$r^2 \log(r)$	5.4e+04	3.4e+05	5.3e+06
$\sqrt{r^2+c^2}$	2.3e+05	9.0e+05	1.2e+07
$(1-r)_+^4(1+4r)$	7.7e+04	5.2e+05	1.7e+07

Table 4: Condition numbers for the radial basis interpolation matrices.

evaluation.

Le us notice that to solve the system with $N = 3375$, we need a computation time which is approximately equal to $1000 t_{343}$. MQSM is more advantageous: in fact $t_{3375} \approx 5 t_{343}$.

N	r^3	MQSM
343	0.55	1.04
1000	14.94	1.96
3375	547.28	5.5

Table 5: Computation times.

- **Computational efficiency.** The computational efficiency can be evaluated by two different indices.

1. Computational efficiency, defined as the inverse of the product between the compu-

tation time t_N and the relative error $e_\infty^r(N)$

$$Ef_N = \frac{1}{t_N e_\infty^r(N)}, \tag{11}$$

2. Ratio quality (maximum error $e_\infty(N)$) to cost (computation time t_N)

$$R_N = \frac{e_\infty(N)}{t_N}. \tag{12}$$

The values of (11) e (12) for the the radial basis (2) and MQSM are shown in table 6 (test function $F1$).

When we use a radial basis function method, the increase of N leads to high computation times not rewarded by an appreciable reduction of the error. This causes a considerable loss of efficiency.

According to definition (11), to use a sample of size $N = 1000$ is ten times more efficient than to use a sample with $N = 3375$. If we consider definition (12), using $N = 1000$ is approximately one hundred and eighty times more efficient than using $N = 3375$.

MQSM is more efficient. In fact we have $Ef_{3375} \approx Ef_{1000}$, while, according to (12), to use $N = 1000$ is five times more efficient than to use $N = 3375$.

Increasing the sample dimension is not always convenient. In fact, on one hand, the interpolant may not reproduce anyway the correct behavior in some zones and on the other we may have a considerable loss of efficiency.

Method	Efficiency	$N = 343$	$N = 1000$	$N = 3375$
r^3	Ef_N	25.97	0.99	0.14
	R_N	7.6e-2	2.6e-3	1.5e-5
MQSM	Ef_N	4.4	7.8	7.3
	R_n	1.2e-1	1.2e-2	2.7e-3

Table 6: Computational efficiency.

3.4 Sensitivity respect to the point set

We conclude our analysis by considering the method sensitivity with regard to the point sets S_N. That is we want to see how the point locations influence the interpolant problem solution.

Here, we report the results obtained for three different equidistributed point sets S_N^1 , S_N^2 e S_N^3, with $N = 343$, 1000, 3375 and show what we get for $F1$ and the radial basis (2). In [5] we have considered also scattered points and distorted grid points.

The error does not change very much: in fact it is a function of N. The graphs, indeed, point out that the point locations condition the local behavior of the interpolants, especially for $N = 343$ and $N = 1000$ (figs. 8, 9). For $N = 3375$ the graphical behavior is always the same (fig. 10). That is we have stability respect to the data set.

N	e	S_N^1	S_N^2	S_N^3
343	e_2	4.6e-3	4.0e-3	3.9e-3
	e_∞	7.0e-2	5.5e-2	6.1e-2
1000	e_2	9.5e-4	1.4e-3	1.4e-3
	e_∞	1.2e-2	3.3e-2	3.8e-2
3375	e_2	3.7e-4	2.5e-4	3.1e-4
	e_∞	8.0e-3	5.3e-3	5.1e-3

Table 7: Sensitivity respect to point sets. Errors for the radial basis function (2).

This analysis has also pointed out that there can be data configurations of moderate size which may provide an interpolant function with a good graphic behavior. See, for instance, fig. 8. The interpolant obtained from S_{343}^2 has a graph comparable to that of S_{1000}.

4 Conclusions

We have tested the computational performances of some methods for interpolating $3D$ scattered data and we have measured the performances, considering the accuracy (e_2, e_∞), the graphic behaviour, the computation times, and the computational efficiency.

In our opinion, the evaluation of the interpolant graphs is a fundamental aspect, even if quite subjective. Ratings by different persons will give somewhat different results.

From our study it is came out that

- In general, the global methods(2)–(5) provide better results than MQSM.

Among the considered radial basis functions, (2) gives the better results in terms of accuracy and graphic behavior (even if it has the worst condition number).

- The disadvantage of global methods is that an increase of N leads to high condition numbers and high computation times.
- Increasing the sample dimension is not always convenient. In fact, on one hand, the interpolant may not reproduce anyway the correct behavior in some zones and, on the other, we may have a considerable loss of efficiency.

- MQSM usually reproduces the qualitative features of test functions quite well. Near the boundary, it may perform worse than the methods (2)–(5). The interpolant is only C^1. But for samples of size $N > 1000$, it gives a graphic quality comparable to that provided by (2), with the advantage of being more and more efficient.
- For samples with moderate dimension, the interpolation problem solution is strongly influenced by the point locations. There can be data configurations of moderate size which may provide an interpolant with a good behavior.

This remark suggests a possible way to get a satisfactory solution with low costs. When we have large samples of size N, we can extract from it a subsample of dimension $\bar{N} << N$ which allows a correct reproduction of the unknown function.

On this subject, there are some techniques for one and two dimensions. (see [3], [4]). For three dimensions, the problem is still open.

References

[1] C.L Bajaj. Rational hypersurface display. *Adv. Comput. Math*, 68:117–127, 1990.

[2] R.K. Beatson and G.N. Newsam. Fast evaluation of radial basis functions: moment based methods. *SIAM J. Comput.*, 19:1428–1449, 1998.

[3] L. Bozzini, M. Lenarduzzi and R. Schaback. Adaptive interpolation by scaled multiquadrics. To appear in Adv. Comp. Math, 2001.

[4] M. Bozzini and L. Lenarduzzi. Selection of data and interpolation by multiquadrics translate with different elliptical shapes. IAMI 22, 2000.

[5] M. Bozzini and M. Rossini. A comparison of some methods for interpolation of trivariate data. Technical Report, 2001.

[6] M. Bozzini and M. Rossini. Approximating 3D "structured" scattered data. Preprint, 2002.

[7] T.A. Foley. Interpolation and approximation of 3-d and 4-d scattered data. *Comput. Math. Appl.*, 13:711–740, 1987.

[8] R. Franke. Scattered data interpolation. Test of some methods. *Mathematics of Computation*, 48:181–199, 1982.

[9] R. Franke and G. M. Nielson. Smooth interpolation of large sets of scattered data. *Internat. J. Numerical Method in Engineering*, 15:1691–1704, 1980.

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

[11] J. Hoschek and D. Lasser. *Computer aided geometric design.* A K Peters, Wellesley, Massachusetts, 1993.

[12] D. Lasser. Bernstein–Bézier representation of volumes. *Comput. Aided Geom. Design*, 2:145–150, 1985.

[13] G. M. Nielson. Scattered data modeling. *IEEE Computer Graphics & Applications*, 1:60–70, 1993.

[14] G.M. Nielson and al. Visualizing and modeling scattered multivariate data. *IEEE Computer Graphics & Applications*, 11:47–55, (1991).

[15] G.M. Nielson and B. Hamann. Techniques for the visualization of volumetric data. In A. Kaufman, editor, *Visualization'90*, pages 45–50. IEEE Computer Society Press, 1990.

[16] S. Rippa. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Adv. Comput. Math*, 11(vol. 2–3):193–210, (1999).

[17] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264, 1995.

[18] R. Schaback. Improved error bounds for scattered data interpolation by radial basis functions. *Mathematics of Computation*, 68:201–216, 1999.

[19] R. Schaback and H. Wendland. Characterization and construction of radial basis functions. In N. Dyn, D. Leviatan, and D Levin, editors, *Eilat proceedings.* Cambridge University Press, 2000.

[20] S. E. Stead. Estimation of gradients from scattered data. *Rocky Mountain J. Math.*, 14:265–279, (1984).

A Description of the methods

A.1 Radial basis function interpolation

Let $\Omega \subset \mathbb{R}^d$ be a compact set, and let us denote the space of d-variate polynomials of order not exceeding m by $I\!P_m^d$. We consider multivariate interpolation by conditionally positive definite radial functions

$$\phi \, : \, I\!R_{\geq 0} \rightarrow I\!R$$

of order $m \geq 0$. This means that for all possible choices of sets

$$X = \{X_1, \ldots, X_N\} \subset \Omega$$

of N distinct points the quadratic form induced by the $N \times N$ matrix

$$A = (\phi(\|X_j - X_k\|_2))_{1 \leq j,k \leq N} \tag{13}$$

is positive definite on the subspace

$$V := \left\{ \alpha \in I\!R^N \ : \ \sum_{j=1}^{N} \alpha_j p(X_j) = 0 \text{ for all } p \in I\!P_m^d \right\}.$$

Note that $m = 0$ implies $V = I\!R^d$ because of $I\!P_m^d = \{0\}$, and then the matrix A in (13) is positive definite. The most prominent examples of conditional positive definite radial basis functions of order m on $I\!R^d$ are

$$\begin{array}{lcll}
\phi(r) & = & (-1)^{\lceil \beta/2 \rceil} r^{\beta}, \beta > 0, \beta \notin 2I\!N_0 & m \geq \lceil \beta/2 \rceil \\
\phi(r) & = & (-1)^{k+1} r^{2k} \log(r), \quad k \in I\!N & m \geq k+1 \\
\phi(r) & = & (c^2 + r^2)^{\beta/2}, \quad \beta < 0 & m \geq 0 \\
\phi(r) & = & (-1)^{\lceil \beta/2 \rceil} (c^2 + r^2)^{\beta/2}, \quad \beta > 0, \beta \notin 2I\!N_0 & m \geq \lceil \beta/2 \rceil \\
\phi(r) & = & e^{-\alpha r^2}, \quad \alpha > 0 & m \geq 0 \\
\phi(r) & = & (1-r)_+^4 (1+4r) & m \geq 0, d \leq 3.
\end{array}$$

See e.g. [19] for a comprehensive derivation of the properties of these functions. Interpolation of real values $f_1, \ldots, f_N$ on a set $X = \{X_1, \ldots, X_N\}$ of N distinct scattered points of Ω by such a function $\phi(\cdot)$ is done by solving the $(N+q) \times (N+q)$ system

$$\begin{array}{rcl}
A\alpha & + & P\beta = f \\
P^T\alpha & + & 0 = 0
\end{array}$$

where $Q = \dim I\!P_m^d$ and

$$P = (p_i(X_j))_{1 \leq j \leq N, 1 \leq i \leq q}$$

for a basis $p_1, \ldots, p_q$ of $I\!P_m^d$. In fact, if the additional assumption

$$\text{rank } (P) = Q \leq N$$

holds, then the system is uniquely solvable. The resulting interpolant has the form

$$s(x) = \sum_{j=1}^{N} \alpha_j \phi(\|X_j - x\|_2) + \sum_{i=1}^{q} \beta_i p_i(x)$$

with the additional condition $\alpha \in V$.

A.2 The modified quadratic Shepard Method

The ideas of the bivariate modified quadratic Shepard method extend directly to the trivariate case (see for instance [9] and [13]). This method has the general form

$$s(X) = \frac{\sum_{k=1}^{N} \frac{Q_k(X)}{\rho_k^2(X)}}{\sum_{k=1}^{N} \frac{1}{\rho_k^2(X)}}, \tag{14}$$

where

$$\frac{1}{\rho_k(X)} = \frac{(R_w - ||X - X_k||_2)_+}{R_w||X - X_k||_2} \tag{15}$$

for some constant R_w. $Q_k(X)$ are the quadratic polynomials, obtained by a weighted least squares fit and constrained to take on the value f_i at the trivariate points X_i. The weights in the least squares process are the same form as the weight functions (15), but with a different constant R_q.

We select two values N_q and N_w and define

$$R_w = \frac{D}{2}\sqrt{\frac{N_w}{N}}, \quad R_q = \frac{D}{2}\sqrt{\frac{N_q}{N}},$$

where $D = \max_{X_i, X_j \in X} ||X_i - X_j||_2$.

We consider the weight functions

$$\begin{aligned} \frac{1}{\rho_k(X)} &= \frac{(R_w - ||X - X_k||_2)_+}{R_w||X - X_k||_2} \\ \frac{1}{\nu_i(X)} &= \frac{(R_q - ||X - X_k||_2)_+}{R_q||X - X_k||_2}. \end{aligned}$$

We define the local basis

$$\begin{aligned} Q_k(x, y, z) &= f_k + a_{k2}(x - x_k) + a_{k3}(y - y_k) + a_{k4}(z - z_k) \\ &+ a_{k5}(x - x_k)^2 + a_{k6}(y - y_k)^2 + a_{k7}(z - z_k)^2 \\ &+ a_{k8}(x - x_k)(y - y_k) + a_{k9}(x - x_k)(z - z_k) \\ &+ a_{k10}(y - y_k)(z - z_k) \end{aligned}$$

solving the following least squares problem

$$\begin{aligned} \min_{a_{kj}\, j=2,\dots,10} \sum_{i=1}^{N} \frac{1}{\nu_i^2(x_k, y_k, z_k)} \{f_k &+ a_{k2}(x_i - x_k) + a_{k3}(y_i - y_k) \\ &+ a_{k4}(z_i - z_k) + a_{k5}(x_i - x_k)^2 \\ &+ a_{k6}(y_i - y_k)^2 + a_{k7}(z_i - z_k)^2 \\ &+ a_{k8}(x_i - x_k)(y_i - y_k) \\ &+ a_{k9}(x_i - x_k)(z_i - z_k) \\ &+ a_{k10}(y_i - y_k)(z_i - z_k) - f_i\}^2. \end{aligned}$$

The interpolant is locally determined, the influence of any point not extending further than a distance $R_w + R_q$ from each data point.

Assuming that the data are equidistributed, constant values for R_q and R_w (that is constant values for N_q and N_w)are appropriate. A good choice (suggested in the literature, see [13]) is $N_q = 54$, $N_w = 27$. If the data density is not reasonably uniform, we might want to let the radii R_q and R_w depend on i.

B Figures

For the graphs, we have chosen the evolutive representation described in §2. In the examples, the evolution variable is z. We have evaluated the interpolant $g(x, y, z)$ in the points of a grid of size $21 \times 21 \times 21$. In the pictures, we show the evolutions of some surfaces $g(x, y, z_k)$. Namely, we have considered $z_k = (k-1)/20$ with $k \in K$, $K = \{1, 3, 6, 9, 13, 16, 19, 21\}$ for $F1$, $F5$ and with $k \in K \setminus \{11\}$ for $F2$. In all the figures, except fig. 7 (in which the errors are shown), the surfaces are represented using a gray scale.

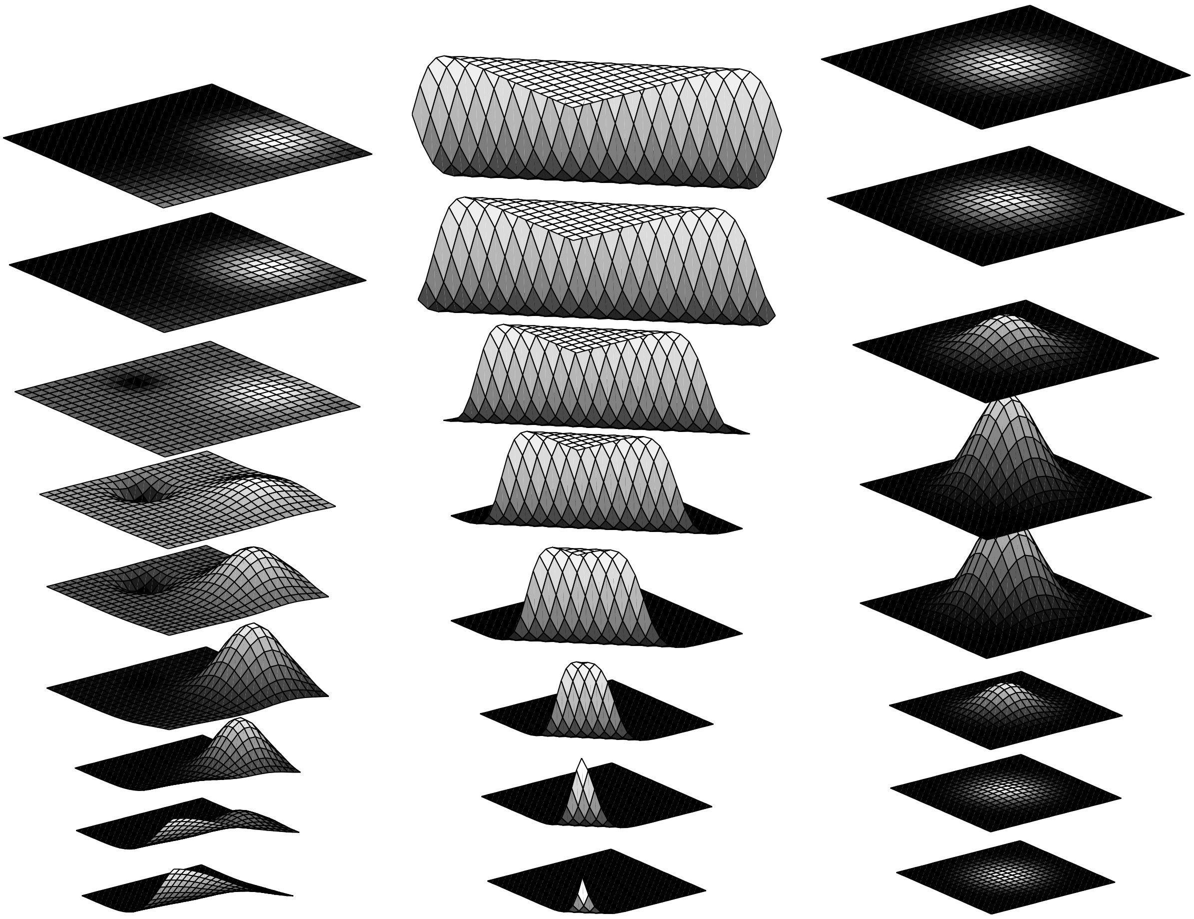

Figure 1: Exact functions. Left: $F1$, center: $F2$, right: $F5$

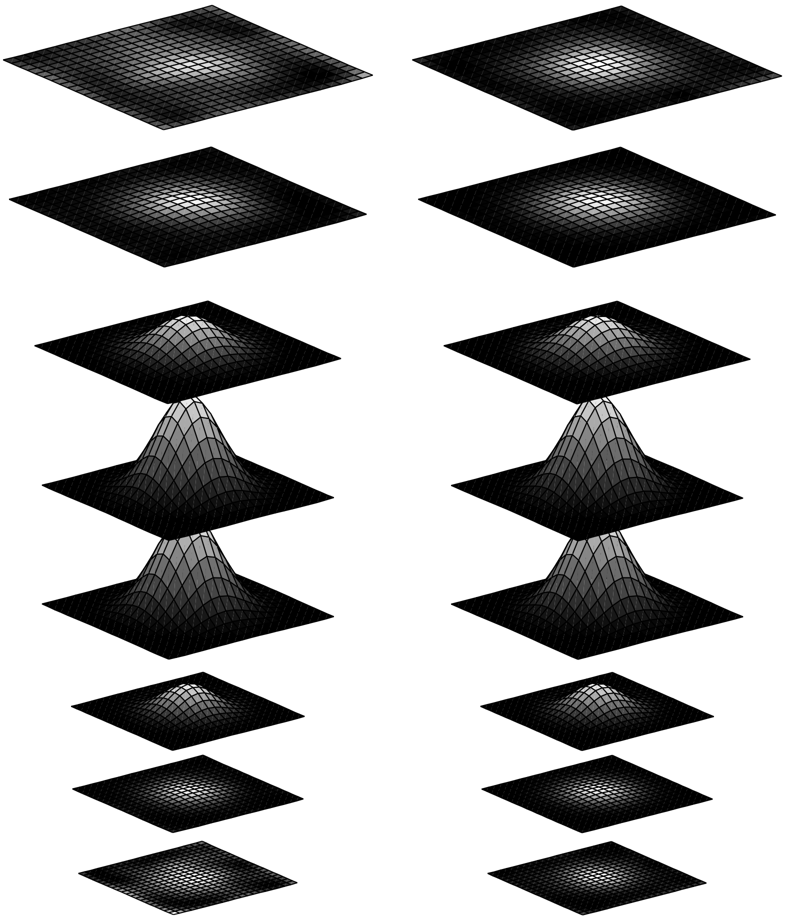

Figure 2: Test function $F5$. Interpolation with r^3. Left: $N = 343$, right: $N = 1000$

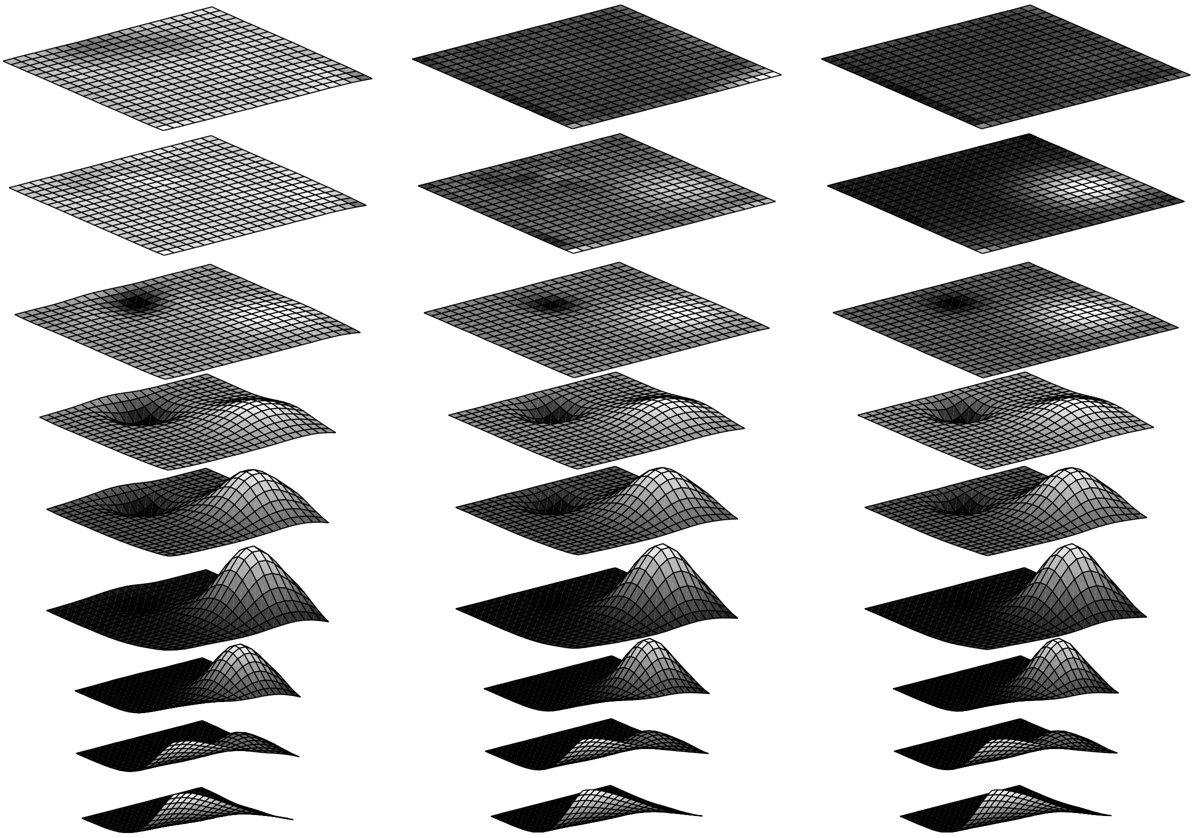

Figure 3: Test function $F1$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

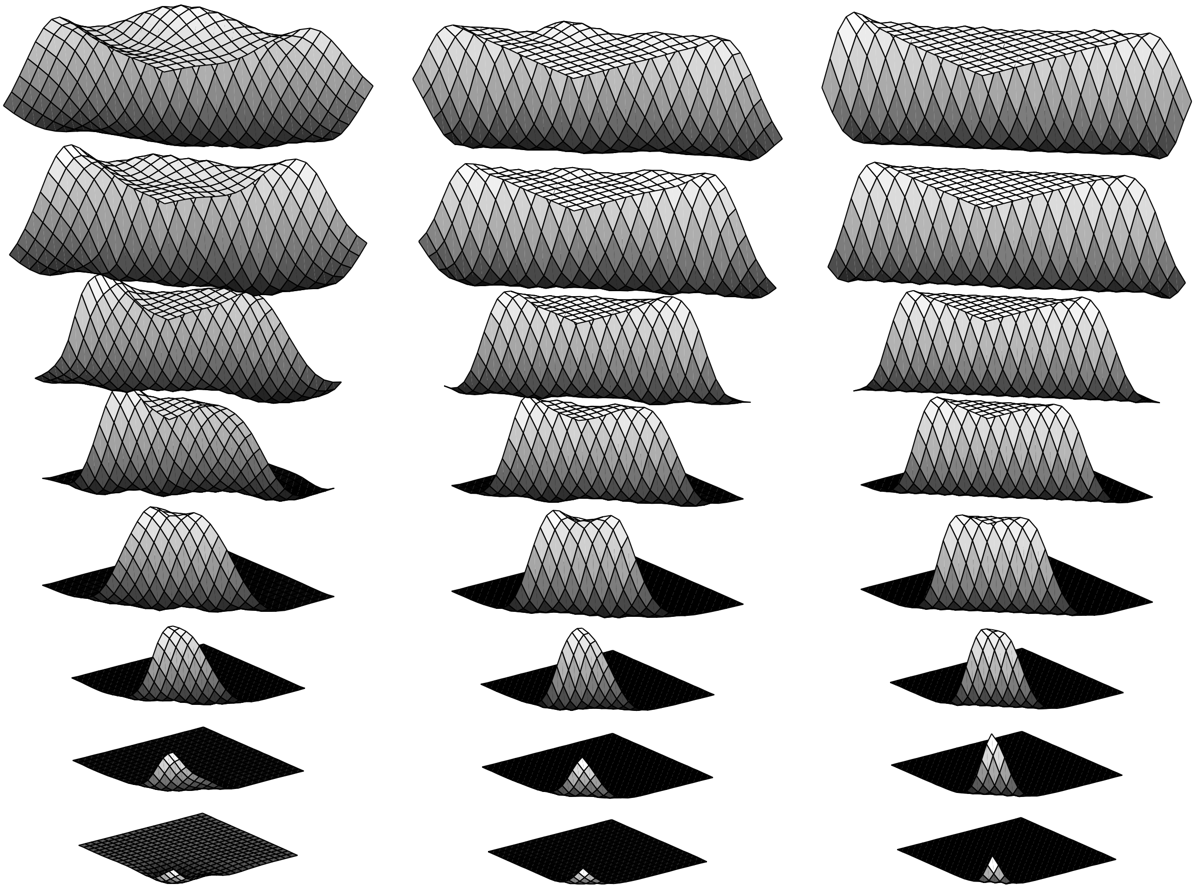

Figure 4: Test function $F2$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

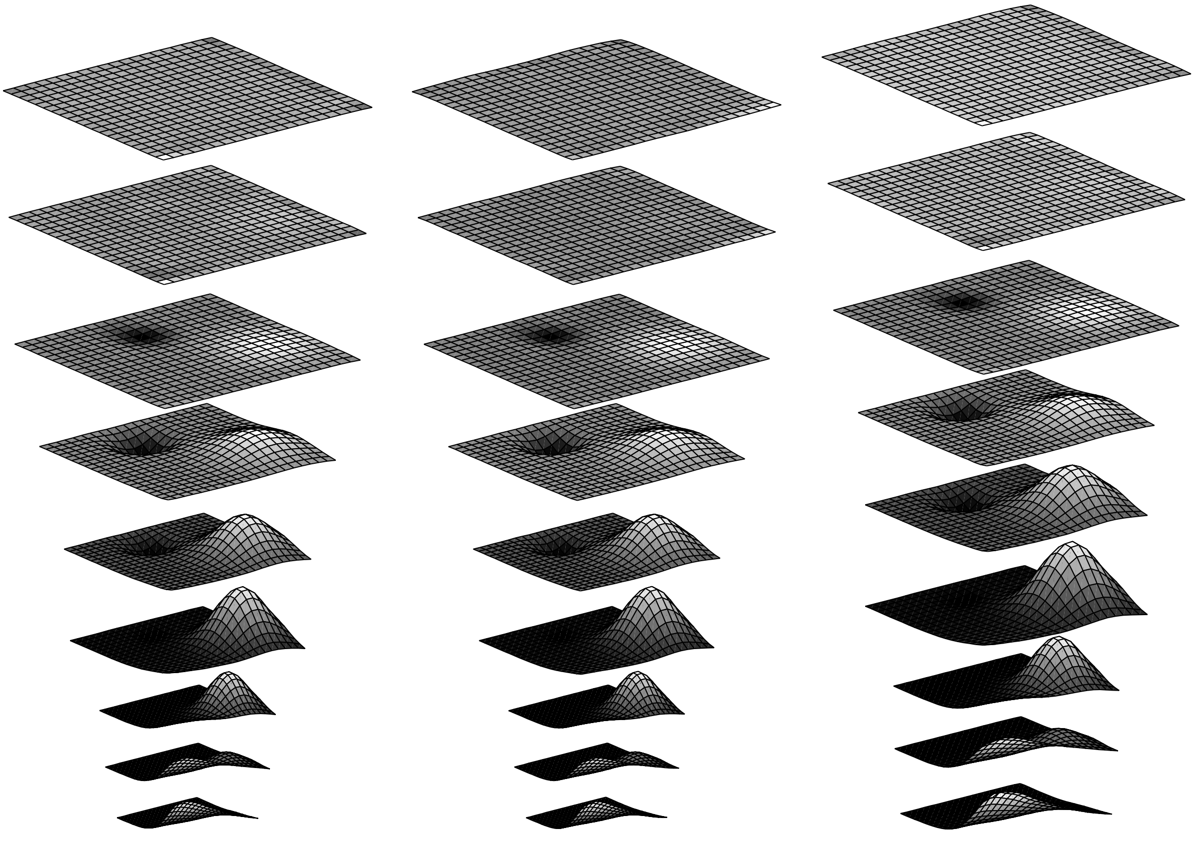

Figure 5: Test function $F2$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

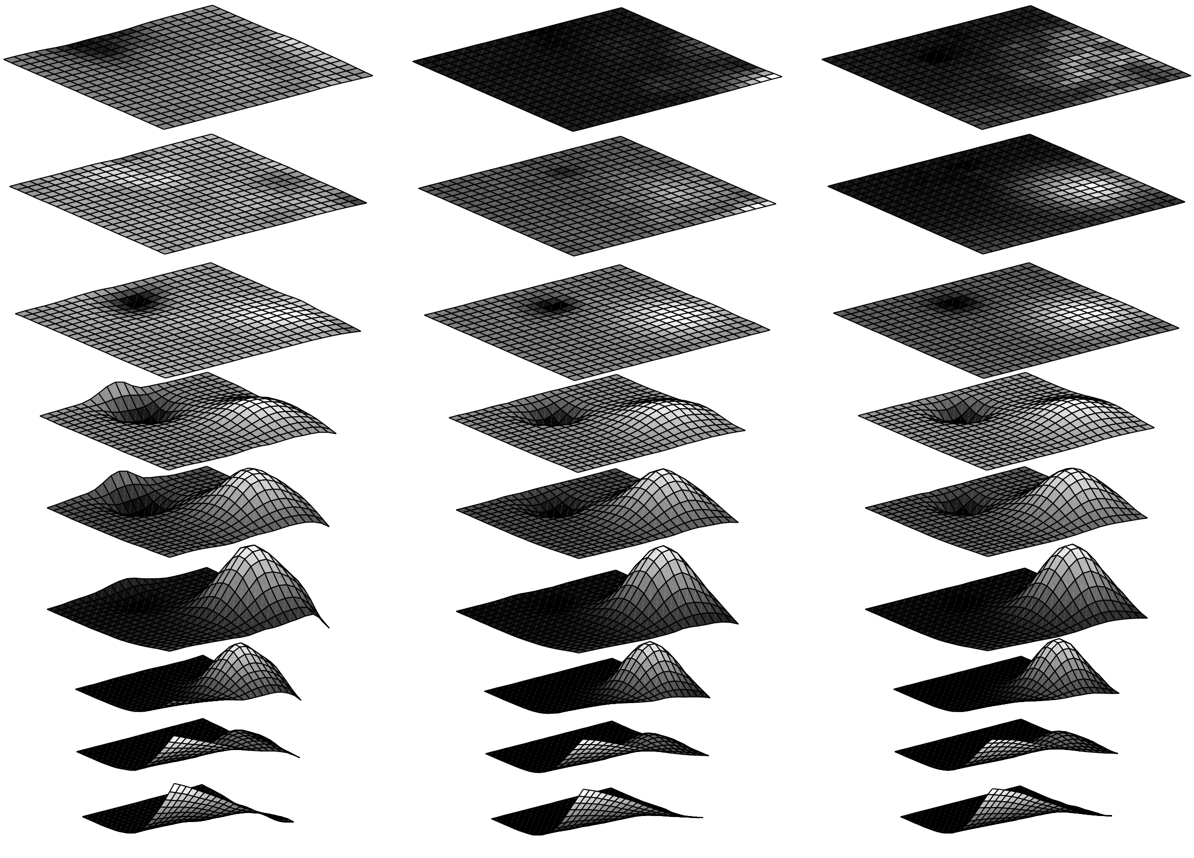

Figure 6: Test function $F1$. Interpolation with MQSM. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

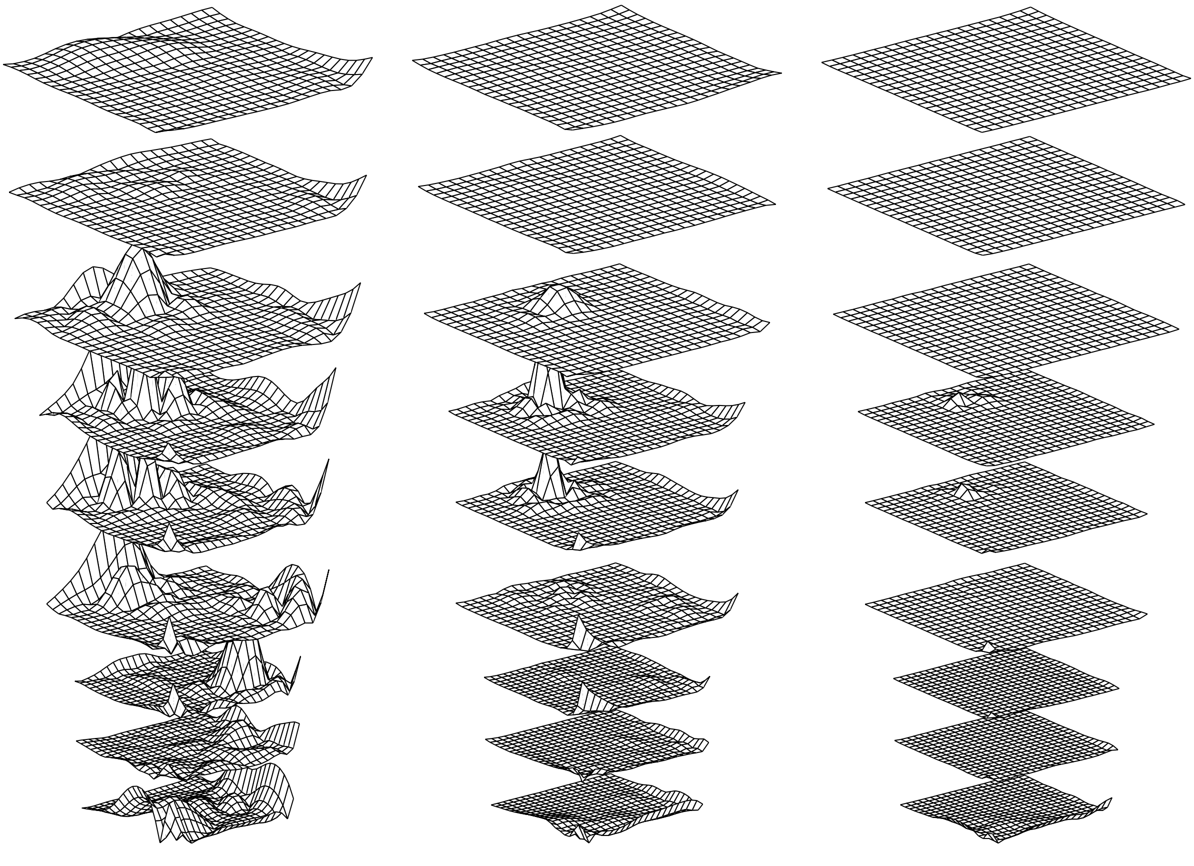

Figure 7: Test function $F1$. Absolute errors for the interpolants obtained with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

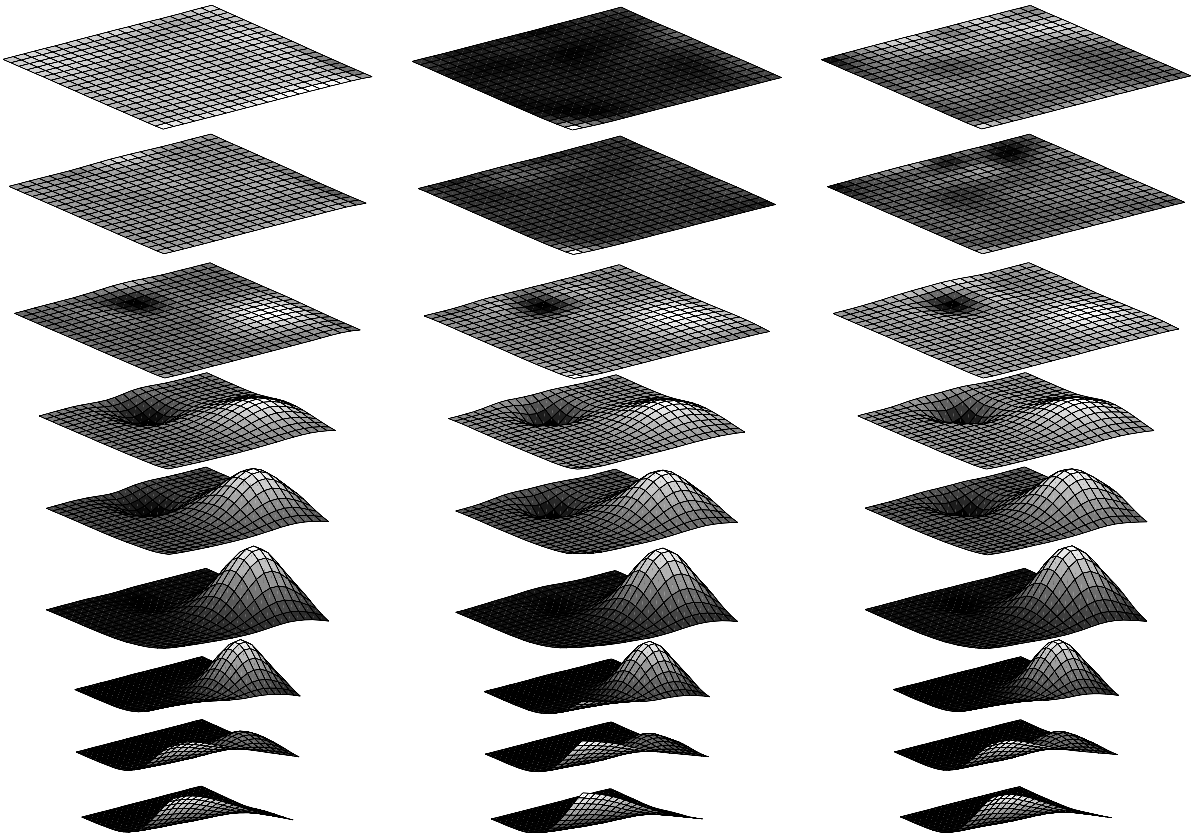

Figure 8: Test function $F1$. Interpolants on different sets. Left: S^1_{343}, center: S^2_{343}, right: S^3_{343}

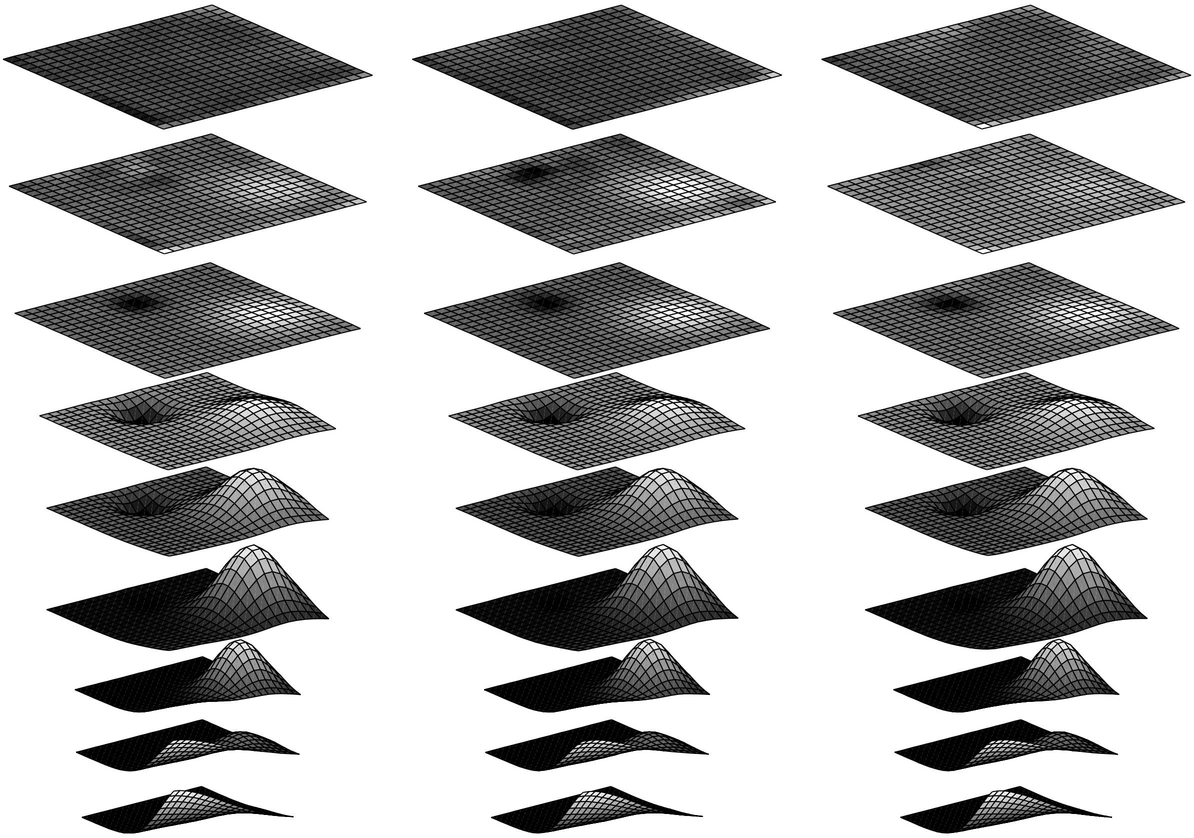

Figure 9: Test function $F1$. Interpolants on different data sets. Left: S^1_{1000}, center: S^2_{1000}, right: S^3_{1000}

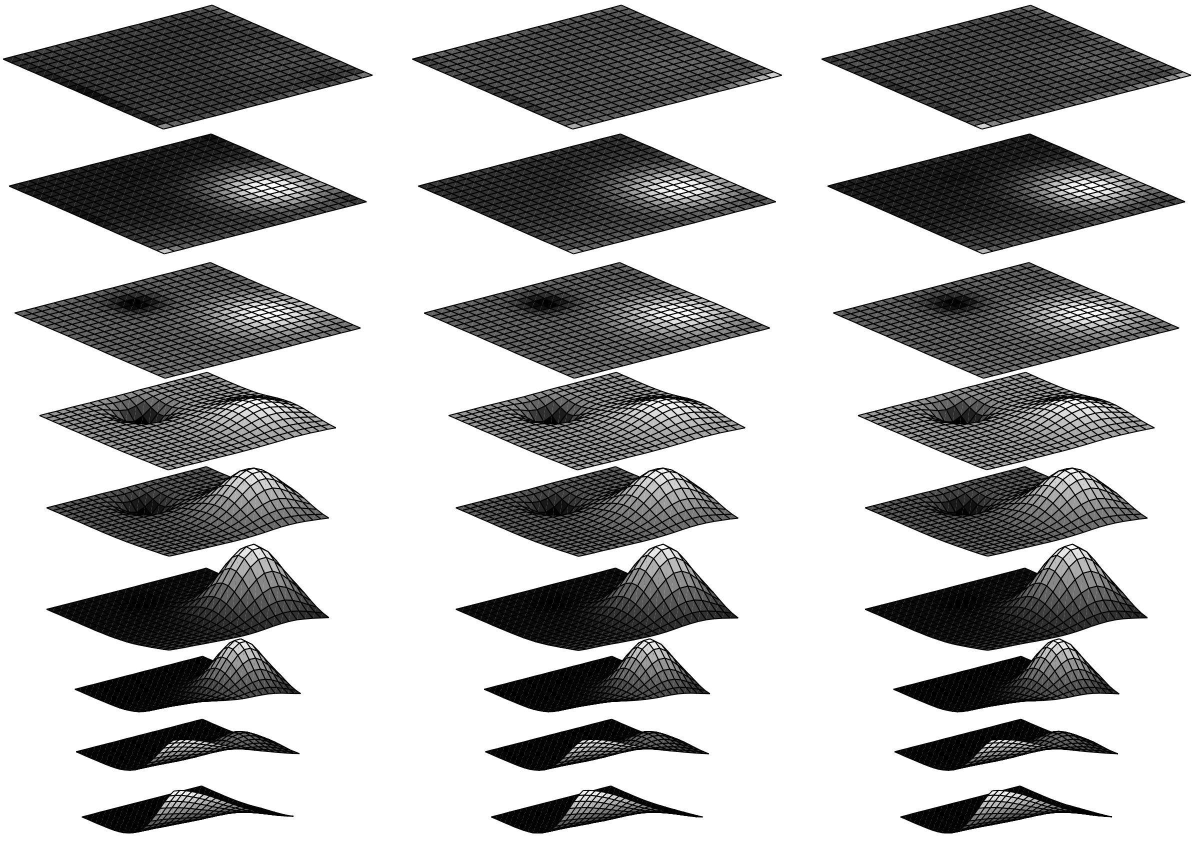

Figure 10: Test function $F1$. Interpolants on different data sets. Left: S^1_{3375}, center: S^2_{3375}, right: S^3_{3375}