Anul LI **2005**

S T U D I A
UNIVERSITATIS BABEŞ-BOLYAI
MATHEMATICA
3

Redacţia: 400084 Cluj-Napoca, str. M. Kogălniceanu nr. 1

Telefon: 405300

SUMAR – CONTENTS – SOMMAIRE

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

TWO THEOREMS ON KENMOTSU HYPERSURFACES IN A W_3-MANIFOLD

ABU-SALEEM AHMAD MAHMOUD AL-SALEH   MIHAIL BANARU

Abstract. A criterion of the minimality of a Kenmotsu hypersurface in a special Hermitian manifold is established. It is also proved that a Kenmotsu hypersurface in a special Hermitian manifold is minimal if and only if its type number is even.

1. **Introduction**

The theory of almost contact metric structures occupies one of the leading places in modern differential-geometrical researches. It is due to a number of its applications in mathematical physics (for example, in classical mechanics [1] and in theory of geometrical quantization [7]). Furthermore, we mark out the richness of the internal contents of the theory of almost contact metric structures as well as the close connection of this theory with other sections of geometry.

We recall that an almost contact metric structure on an odd-dimensional manifold N is defined by the system of tensor fields $\{\Phi, \xi, \eta, g\}$ on this manifold, where ξ is a vector, η is a covector, Φ is a tensor of the type $(1,1)$ and $g = \langle \cdot, \cdot \rangle$ is the Riemannian metric. Moreover, the following conditions are fulfilled:

$$\eta(\xi) = 1, \quad \Phi(\xi) = 0, \quad \eta \circ \Phi = 0, \quad \Phi^2 = -id + \xi \otimes \eta,$$

$$\langle \Phi X, \Phi Y \rangle = \langle X, Y \rangle - \eta(X)\eta(Y), \quad X, Y \in \aleph(N),$$

Received by the editors: 09.09.2005.

2000 *Mathematics Subject Classification.* 53C40.

Key words and phrases. Almost Hermitian manifold, almost contact metric manifold, Kenmotsu manifold, hypersurface.

where $\aleph(N)$ is the module of smooth vector fields on N. As an example of an almost contact metric structure we can consider the cosymplectic structure, that is characterized by the following condition:

$$\nabla\eta = 0, \quad \nabla\Phi = 0,$$

where ∇ is the Levi-Civita connection of the metric. It has been proved that the manifold, admitting the cosymplectic structure, is locally equivalent to a product $M \times R$, where M is a Kählerian manifold [10].

The almost contact metric structures are closely connected to the almost Hermitian structures. For instance, if $(N, \{\Phi, \xi, \eta, g\})$ is an almost contact metric manifold, then an almost Hermitian structure is induced on $N \times R$ [5]. If this almost Hermitian structure is integrable, then the input almost contact metric structure is called normal. As it is known, a normal contact metric structure is called Sasakian [5]. On the other hand, we can characterize the Sasakian struture by the following condition:

$$\nabla_X(\Phi)Y = \langle X, Y\rangle\xi - \eta(Y)X, \quad X, Y \in \aleph(N). \tag{1}$$

For example, Sasakian structures are induced on totally umbilical hypersurfaces in a Kählerian manifold [5]. As it is well known, the Sasakian structures have many remarkable properties and play a fundamental role in contact geometry.

In 1972 Katsuei Kenmotsu has introduced a new class of almost contact metric structures [8], defined by the condition

$$\nabla_X(\Phi)Y = \langle \Phi X, Y\rangle\xi - \eta(Y)\Phi X, \quad X, Y \in \aleph(N). \tag{2}$$

The Kenmotsu manifolds are normal and integrable, but they are not contact, consequently, they can not be Sasakian. In spite of the fact that the conditions (1) and (2) are similar, the properties of Kenmotsu manifolds are to some extent antipodal to the Sasakian manifolds properties [9]. Note that the new investigation [9] in this field contains a detailed description of Kenmotsu manifolds as well as a collection of examples of such manifolds.

In the present paper, Kenmotsu hypersurfaces in W_3-manifolds are considered. This note is a continuation of research of the authors (for example, the second author studied six-dimensional W_3-manifolds before [3], [4]). We remark that the class of W_3-manifolds is one of the most important classes of almost Hermitian manifolds [6]. However, it has been studied not so detailed as other so-called "small" classes of almost Hermitian manifolds. Some dozens of significant works are devoted to the nearly-Kählerian, almost Kählerian and locally conformal Kählerian manifolds, but much less of articles are written about W_3-manifolds.

2. **Preliminaries**

We consider an almost Hermitian manifold M^{2n}, i.e. a $2n$-dimensional manifold with a Riemannian metric $g = \langle \cdot, \cdot \rangle$ and an almost complex structure J. Moreover, the following condition must hold:

$$\langle JX, JY \rangle = \langle X, Y \rangle, \quad X, Y \in \aleph(M^{2n}),$$

where $\aleph(M^{2n})$ is the module of smooth vector fields on M^{2n}. All considered manifolds, tensor fields and similar objects are assumed to be of the class C^∞. We recall that the fundamental (or Kählerian) form of an almost Hermitian manifold is determined by

$$F(X, Y) = \langle X, JY \rangle, \quad X, Y \in \aleph(M^{2n}).$$

Let $\big(M^{2n}, \{J, g = \langle \cdot, \cdot \rangle\}\big)$ be an arbitrary almost Hermitian manifold. We fix a point $p \in M^{2n}$. As $T_p(M^{2n})$ we denote the tangent space at the point p, $\{J_p, g_p = \langle \cdot, \cdot \rangle\}$ is the almost Hermitian structure at the point p induced by the structure $\{J, g = \langle \cdot, \cdot \rangle\}$. The frames adapted to the structure (or the A-frames) look as follows:

$$(p, \varepsilon_1, \ldots, \varepsilon_n, \varepsilon_{\widehat{1}}, \ldots, \varepsilon_{\widehat{n}}),$$

where ε_a are the eigenvectors corresponded to the eigenvalue $i = \sqrt{-1}$, and $\varepsilon_{\widehat{a}}$ are the eigenvectors corresponded to the eigenvalue $-i$ [2]. Here the index a ranges from 1 to n, and we state $\widehat{a} = a + n$.

The matrix of the operator of the almost complex structure written in an A-frame looks as follows:

$$(J_j^k) = \left(\begin{array}{c|c} iI_n & 0 \\ \hline 0 & -iI_n \end{array}\right),$$

where I_n is the identity matrix; $k, j = 1, \ldots, 2n$. By direct computing it is easy to obtain that the matrices of the metric g and of the fundamental form F in an A-frame look as follows, respectively:

$$(g_{kj}) = \left(\begin{array}{c|c} 0 & I_n \\ \hline I_n & 0 \end{array}\right), \quad (F_{kj}) = \left(\begin{array}{c|c} 0 & iI_n \\ \hline -iI_n & 0 \end{array}\right).$$

An almost Hermitian manifold is called special Hermitian, if

$$\delta F = 0, \quad \nabla_X(F)(Y, Z) - \nabla_{JX}(F)(JY, Z) = 0, \quad X, Y, Z \in \aleph(M^{2n}),$$

where δ is the codifferentiation operator. The first group of the Cartan structural equations of a special Hermitian manifold written in an A-frame looks as follows:

$$d\omega^a = \omega_b^a \wedge \omega^b + B^{ab}{}_c \omega^c \wedge \omega_b,$$

$$d\omega_a = -\omega_a^b \wedge \omega_b + B_{ab}{}^c \omega_c \wedge \omega^b,$$

and moreover,

$$B^{ab}{}_b = 0, \quad B_{ab}{}^b = 0, \tag{3}$$

where $\{B^{ab}{}_c\}$ and $\{B_{ab}{}^c\}$ are components of the Kirichenko tensors of M^{2n} [2], $a, b, c = 1, \ldots, n$.

3. **The main results**

Theorem 3.1. Let N be a Kenmotsu hypersurface in a special Hermitian manifold M^{2n}, and let σ be the second fundamental form of the immersion of N into M^{2n}. Then N is a minimal submanifold of M^{2n} if and only if $\sigma(\xi, \xi) = 0$.

Proof. Let us use the Cartan structural equations of an almost contact metric structure on a hypersurface in a Hermitian manifold [4]:

$$d\omega^\alpha = \omega^\alpha_\beta \wedge \omega^\beta + B^{\alpha\beta}{}_\gamma \omega^\gamma \wedge \omega_\beta + (\sqrt{2}B^{\alpha n}{}_\beta + i\sigma^\alpha_\beta)\omega^\beta \wedge \omega +$$

$$+(-\frac{1}{\sqrt{2}}B^{\alpha\beta}{}_n + i\sigma^{\alpha\beta})\omega_\beta \wedge \omega,$$

$$d\omega_\alpha = -\omega^\beta_\alpha \wedge \omega_\beta + B_{\alpha\beta}{}^\gamma \omega_\gamma \wedge \omega^\beta + (\sqrt{2}B_{\alpha n}{}^\beta - i\sigma^\beta_\alpha)\omega_\beta \wedge \omega +$$

$$+(-\frac{1}{\sqrt{2}}B_{\alpha\beta}{}^n - i\sigma_{\alpha\beta})\omega^\beta \wedge \omega,$$

$$d\omega = (\sqrt{2}B^{n\alpha}{}_\beta - \sqrt{2}B_{n\beta}{}^\alpha - 2i\sigma^\alpha_\beta)\omega^\beta \wedge \omega_\alpha + (B_{n\beta}{}^n + i\sigma_{n\beta})\omega \wedge \omega^\beta +$$

$$+(B^{n\beta}{}_n - i\sigma^\beta_n)\omega \wedge \omega_\beta.$$

Here and further the indices α, β, γ range from 1 to $n-1$.

Taking into account that the Cartan structural equations of a Kenmotsu structure look as follows [9]:

$$d\omega^\alpha = \omega^\alpha_\beta \wedge \omega^\beta + \omega \wedge \omega^\alpha,$$

$$d\omega_\alpha = -\omega^\beta_\alpha \wedge \omega_\beta + \omega \wedge \omega_\alpha,$$

$$d\omega = 0,$$

we get the conditions, whose simultaneous fulfillment is a criterion for the structure on N to be Kenmotsu:

$$1)\ B^{\alpha\beta}{}_\gamma = 0; \quad 2)\ \sqrt{2}B^{\alpha n}{}_\beta + i\sigma^\alpha_\beta = -\delta^\alpha_\beta; \quad 3)\ -\frac{1}{\sqrt{2}}B^{\alpha\beta}{}_n + i\sigma^{\alpha\beta} = 0;$$

$$4)\ \sqrt{2}B^{n\alpha}{}_\beta - \sqrt{2}B_{n\beta}{}^\alpha - 2i\sigma^\alpha_\beta = 0; \quad 5)\ B^{n\beta}{}_n - i\sigma^\beta_n = 0 \qquad (4)$$

and the formulae obtained by the complex conjugation (no need to write them explicitly). From $(4)_3$ we have:

$$\sigma^{\alpha\beta} = -\frac{i}{\sqrt{2}}B^{\alpha\beta}{}_n.$$

Since

$$0 = \sigma^{[\alpha\beta]} = -\frac{i}{\sqrt{2}}B^{[\alpha\beta]}{}_n = -\frac{i}{2\sqrt{2}}(B^{\alpha\beta}{}_n - B^{\beta\alpha}{}_n) = -\frac{i}{\sqrt{2}}B^{\alpha\beta}{}_n,$$

we get ${B^{\alpha\beta}}_n = 0$, and that is why

$$\sigma^{\alpha\beta} = 0.$$

Similarly, from $(4)_5$ we obtain

$$\sigma^{\beta}_{n} = 0.$$

Therefore we can rewrite the conditions (4) as follows:

$$1)\ {B^{\alpha\beta}}_\gamma = 0;\quad 2)\ \sigma^{\alpha\beta} = 0;\quad 3)\ \sigma^{\beta}_{n} = 0;\quad 4)\ \sigma^{\alpha}_{\beta} = i\sqrt{2}{B^{\alpha n}}_\beta + i\delta^{\alpha}_{\beta} \tag{5}$$

and the formulae obtained by the complex conjugation.

Now, let us use a criterion of the minimality of an arbitrary hypersurface [11]:

$$g^{ps}\sigma_{ps} = 0, \quad p, s = 1, 2, \dots, 2n-1.$$

Knowing how the matrix of the contravariant metric tensor on N looks [3]:

$$(g^{ps}) = \left(\begin{array}{c|c|c} 0 & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & I_{n-1} \\ \hline 0 \dots 0 & 1 & 0 \dots 0 \\ \hline I_{n-1} & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & 0 \end{array}\right),$$

we obtain:

$$g^{ps}\sigma_{ps} = g^{\alpha\beta}\sigma_{\alpha\beta} + g^{\widehat{\alpha}\widehat{\beta}}\sigma_{\widehat{\alpha}\widehat{\beta}} + g^{\widehat{\alpha}\beta}\sigma_{\widehat{\alpha}\beta} + g^{\alpha\widehat{\beta}}\sigma_{\alpha\widehat{\beta}} + g^{nn}\sigma_{nn} =$$

$$= g^{\widehat{\alpha}\beta}\sigma_{\widehat{\alpha}\beta} + g^{\alpha\widehat{\beta}}\sigma_{\alpha\widehat{\beta}} + g^{nn}\sigma_{nn}.$$

By force of (3) and (5) we have

$$g^{ps}\sigma_{ps} = i\sqrt{2}{B^{\alpha n}}_\alpha + i(n-1) - i\sqrt{2}{B_{\alpha n}}^\alpha - i(n-1) + \sigma_{nn} = \sigma_{nn}.$$

That is why $g^{ps}\sigma_{ps} = 0 \Leftrightarrow \sigma_{nn} = 0$. The last equality means that

$$\sigma(\xi, \xi) = 0. \tag{6}$$

So, a Kenmotsu hypersurface in a W_3-manifold is minimal precisely when (6) holds, Q.E.D.

Since the class of special Hermitian manifolds contains all Kählerian manifolds [6], by force of THEOREM A we come to the following result.

Corollary 3.1. A Kenmotsu hypersurface in a Kählerian manifold is minimal if and only if

$$\sigma(\xi, \xi) = 0.$$

Now, let N be a totally umbilical Kenmotsu hypersurface in a W_3-manifold M^{2n}. Then $\sigma = \lambda g$, λ — *const*, therefore the matrix of the second fundamental form looks as follows:

$$(\sigma_{ps}) = \left(\begin{array}{c|c|c} 0 & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & \lambda I_{n-1} \\ \hline 0 \ \dots \ 0 & \lambda & 0 \ \dots \ 0 \\ \hline \lambda I_{n-1} & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & 0 \end{array}\right),$$

As it has been proved, the hypersurface will be minimal if and only if $\lambda = 0$. Evidently, the matrix (σ_{ps}) vanishes in this case, therefore we conclude that N will be a totally geodesic hypersurface in M^{2n}. That is why we have such an additional result.

Corollary 3.2. A totally umbilical Kenmotsu hypersurface in a W_3-manifold is minimal if and only if it is totally geodesic.

As it is well-known (see, for example, [12] or [13]), when we give a Riemannian manifold and its submanifold, the rank of the determined second fundamental form is called the type number. Now, we can state the second main result of this work:

Theorem 3.2. A Kenmotsu hypersurface in a W_3-manifold is minimal if and only if its type number is even.

Proof. Considering the matrix of the second fundamental form of a Kenmotsu hypersurface in a special Hermitian manifold, it is easy to see that this hypersurface is minimal precisely when the following condition holds:

$$(\sigma_{ps}) = \left(\begin{array}{c|c|c} 0 & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & \sigma_{\alpha\hat{\beta}} \\ \hline 0 \dots 0 & 0 & 0 \dots 0 \\ \hline \sigma_{\hat{\alpha}\beta} & \begin{matrix} 0 \\ \dots \\ 0 \end{matrix} & 0 \end{array}\right),$$

Taking into account that [14] $\sigma_{\hat{\alpha}\beta} = \overline{\sigma_{\alpha\hat{\beta}}}$, we have:

$$rank(\sigma_{ps}) = 2rank(\sigma_{\alpha\hat{\beta}}).$$

On the other hand, if the Kenmotsu hypersurface is not minimal, then

$$rank(\sigma_{ps}) = 2rank(\sigma_{\alpha\hat{\beta}}) + 1.$$

Thus, a Kenmotsu hypersurface in a W_3-manifold is minimal precisely when its type number is even, Q.E.D.

References

[1] Arnol'd V.I., *Mathematical methods of classical mechanics,* Springer-Verlag, Berlin–Heidelberg–New-York, 1989.

[2] Arsen'eva O.E., Kirichenko V.F., *Self-dual geometry of generalized Hermitian surfaces,* Mat. Sbornik, **189**(1998), N1, 21–44 (in Russian).

[3] Banaru M., *On six-dimensional Hermitian submanifolds of Cayley algebra satisfying the G-cosymplectic hypersurfaces axiom,* Annuaire de l'Universite de Sofia "St. Kl. Ohridski", **94**(2000), 91–96.

[4] Banaru M.,*Two theorems on cosymplectic hypersurfaces of six-dimensional Hermitian submanifolds of Cayley algebra,* J. Harbin Inst. Techn., **8**(2001), N1, 38–40.

[5] Blair D.E.,*Contact manifolds in Riemannian geometry,* Lect. Notes Math. **509**(1976),1–145.

[6] Gray A., Hervella L.M., *The sixteen classes of almost Hermitian manifolds and their linear in variants,* Ann. Mat., Pura Appl., **123**(1980), N4, 35–58.

[7] Hurt N., *Geometric quantization in action, Reidel Publishing Company*, Dordrecht–Boston–London, 1983.

[8] Kenmotsu K., *A class of almost contact Riemannian manifolds,* Tôhoku Math.J., **24**(1972), 93–103.

[9] Kirichenko V.F., *On geometry of Kenmotsu manifolds,* DAN Math., 380(2001), N5, 585–587 (in Russian).

[10] Kiritchenko V.F., *Sur la geometrie des varietes approximativement cosymplectiques,* C.R. Acad. Sci. Paris, Ser.1, 295(1982), N12, 673–676.

[11] Kobayashi S., Nomizu K., *Foundations of differential geometry, Interscience*, New-York–London–Sydney, 1969.

[12] Kurihara H., *The type number on real hypersurfaces in a quaternionic space form,* Tsukuba Journal Math., 24(2000), 127–132.

[13] Kurihara H., Takagi R.,*A note on the type number of real hypersurfaces in $P_n(C)$,* Tsukuba Journal Math., **22**(1998), 793-802.

[14] Stepanova L.V.,*Contact geometry of hypersurfaces of quasi-Kählerian manifolds,* MSPU, Moscow, 1995 (in Russian).

Departament of Mathematics
Al al-Bayt University
Mafrag, JORDAN
Departament of Computer Technologes
Smolensk University of Humanities,
Gertsen str., 2
Smolensk 214014, RUSSIA
E-mail address: dr_ahmad57@yahoo.ru, banaru@keytown.com

PROPERTIES OF SOME NEW SEMINORMED SEQUENCE SPACES DEFINED BY A MODULUS FUNCTION

YAVUZ ALTIN AYŞEGÜL GÖKHAN HIFSI ALTINOK

Abstract. In this paper we introduce the sequence spaces $\hat{c}_0(p,f,q,s)$, $\hat{c}(p,f,q,s)$ and $\hat{m}(p,f,q,s)$ using a modulus function f and defined over a seminormed space (X,q) seminormed by q. We study some properties of these sequence spaces and obtain some inclusion relations.

1. Introduction

Let m, c and c_0 be the Banach spaces of bounded, convergent and null sequences $x = (x_k)$ with the usual norm $\|x\| = \sup_{k\geq 0} |x_k|$. Let D be the shift operator on s, that is, $Dx = (x_k)_{k=1}^{\infty}$, $D^2x = (x_k)_{k=2}^{\infty}$ and so on. It may be recalled that a Banach limit (see Banach [1]) L is a nonnegative linear functional on m such that L is invariant under shift operator (that is, $L(Dx) = L(x)$ for $x \in m$) and $L(e) = 1$, where $e = (1, 1, ...)$. A sequence $x \in m$ is almost convergent (see Lorentz [8]) if all Banach limits of x coincide. Let $\hat{c}$ denote the space of almost convergent sequences. It is proved by Lorentz [8] that

$$\hat{c} = \left\{ x : \lim_{m\to\infty} t_{m,n}(x) \text{ exists uniformly in } n \right\}$$

where

$$t_{m,n}(x) = \frac{1}{m+1} \sum_{i=0}^{m} D^i x_n, \quad \left(D^0 = 1\right).$$

Several authors including Duran [5] , King [7] and Nanda ([12] ,[13]) have studied almost convergent sequences.

Received by the editors: 11.09.2005.
2000 *Mathematics Subject Classification.* 40A05.
Key words and phrases. Modulus function, seminorm, almost convergence, sequence spaces.

The notion of a modulus function was introduced by Nakano [11] in 1953. We recall that a modulus f is a function from $[0,\infty)$ to $[0,\infty)$ such that *(i)* $f(x)=0$ if and only if $x=0$, *(ii)* $f(x+y)\leq f(x)+f(y)$, for all $x\geq 0$, $y\geq 0$, *(iii)* f is increasing, *(iv)* f is continuous from the right at 0.

Since $|f(x)-f(y)|\leq f(|x-y|)$, it follows from condition (iv) that f is continuous on $[0,\infty)$. Furthermore, we have $f(nx)\leq nf(x)$ for all $n\in\mathbb{N}$, from condition (ii), and so

$$f(x)=f\left(nx\frac{1}{n}\right)\leq nf\left(\frac{x}{n}\right),$$

hence

$$\frac{1}{n}f(x)\leq f\left(\frac{x}{n}\right) \text{ for all } n\in\mathbb{N}.$$

A modulus may be bounded or unbounded. For example, $f(x)=x^p$, $(0<p\leq 1)$ is unbounded and $f(x)=\frac{x}{1+x}$ is bounded. Maddox [10] and Ruckle [14] used a modulus function to construct some sequence spaces.

After then some sequence spaces, defined by a modulus function, were introduced and studied by Bhardwaj [2], Bilgin [3], Connor [4], Esi [6], and many others.

Definition 1.1. *Let q_1, q_2 be seminorms on a vector space X. Then q_1 is said to be stronger than q_2 if whenever (x_n) is a sequence such that $q_1(x_n)\to 0$, then also $q_2(x_n)\to 0$. If each is stronger than the other q_1 and q_2 are said to be equivalent (one may refer to Wilansky* [15]*).*

Lemma 1.1. *Let q_1 and q_2 be seminorms on a linear space X. Then q_1 is stronger than q_2 if and only if there exists a constant M such that $q_2(x)\leq Mq_1(x)$ for all $x\in X$ (see for instance Wilansky* [15]*).*

Let $p=(p_m)$ be a sequence of strictly positive real numbers and X be a seminormed space over the field $\mathbb{C}$ of complex numbers with the seminorm q. We define the sequence spaces as follows:

$$\begin{aligned}
\hat{c}_0(p,f,q,s) &= \left\{x \in X : \lim_{m\to\infty} m^{-s}\left[\left(f\left(q\left(t_{m,n}(x)\right)\right)\right)\right]^{p_m} = 0 \text{ uniformly in } n\right\},\\
\hat{c}(p,f,q,s) &= \left\{x \in X : \lim_{m\to\infty} m^{-s}\left[\left(f\left(q\left(t_{m,n}(x-\ell e)\right)\right)\right)\right]^{p_m} = 0 \text{ for some } \ell,\right.\\
&\qquad \left.\text{uniformly in } n\right\},\\
\hat{m}(p,f,q,s) &= \left\{x \in X : \sup_{m,n} m^{-s}\left[\left(f\left(q\left(t_{m,n}(x)\right)\right)\right)\right]^{p_m} < \infty\right\}.
\end{aligned}$$

where f is a modulus function.

The following inequalities will be used throughout the paper. Let $p = (p_m)$ be a bounded sequence of strictly positive real numbers with $0 < p_m \le \sup p_m = H$, $C = \max\left(1, 2^{H-1}\right)$, then

$$|a_m + b_m|^{p_m} \le C\left\{|a_m|^{p_m} + |b_m|^{p_m}\right\}, \tag{1.1}$$

where $a_m,\, b_m \in \mathbb{C}$.

2. Main results

Theorem 2.1. *Let $p = (p_m)$ be a bounded sequence, then $\hat{c}_0(p,f,q,s)$, $\hat{c}(p,f,q,s)$, $\hat{m}(p,f,q,s)$ are linear spaces.*

Proof. We give the proof for $\hat{c}_0(p,f,q,s)$ only. The others can be treated similarly. Let $x, y \in \hat{c}_0(p,f,q,s)$. For $\lambda, \mu \in \mathbb{C}$, there exist positive integers M_λ and N_λ such that $|\lambda| \le M_\lambda$ and $|\mu| \le N_\mu$. Since f is subadditive and q is a seminorm $m^{-s}\left[f\left(q\left(t_{m,n}(\lambda x + \mu y)\right)\right)\right]^{p_m} \le C\left(M_\lambda\right)^H m^{-s}\left[f\left(q\left(t_{m,n}(x)\right)\right)\right]^{p_m} + C\left(N_\mu\right)^H m^{-s}\left[f\left(q\left(t_{m,n}(y)\right)\right)\right]^{p_m} \to 0$,uniformly in n. This proves that $\hat{c}_0(p,f,q,s)$ is a linear space.

Theorem 2.2. *The space $\hat{c}_0(p,f,q,s)$ is a paranormed space, paranormed by*

$$g(x) = \sup_{m,n} m^{-s}\left(\left[f\left(q\left(t_{m,n}(x)\right)\right)\right]^{p_m}\right)^{\frac{1}{M}},$$

where $M = \max(1, \sup p_m)$. The spaces $\hat{c}(p,f,q,s)$, $\hat{m}(p,f,q,s)$ are paranormed by g, if $\inf p_m > 0$.

Proof. Omitted.

Theorem 2.3. *Let f be modulus function, then*

(i) $\hat{c}_0(p,f,q,s) \subseteq \hat{m}(p,f,q,s)$,

(ii) $\hat{c}(p,f,q,s) \subseteq \hat{m}(p,f,q,s)$.

Proof. We prove the second inclusion, since the first inclusion is obvious. Let $x \in \hat{c}(p,f,q,s)$, by definition of a modulus function (the inequality *(ii)*), we have

$$m^{-s}\left[f\left(q\left(t_{m,n}(x)\right)\right)\right]^{p_m} \leq Cm^{-s}\left[f\left(q\left(t_{m,n}(x-\ell)\right)\right)\right]^{p_m} + Cm^{-s}\left[f\left(q(\ell)\right)\right]^{p_m}.$$

Then there exists an integer K_ℓ such that $q(\ell) \leq K_\ell$. Hence, we have

$$m^{-s}\left[f\left(q\left(t_{m,n}(x)\right)\right)\right]^{p_m} \leq Cm^{-s}\left[f\left(q\left(t_{m,n}(x-\ell)\right)\right)\right]^{p_m} + Cm^{-s}\max(1,\left[(K_\ell)f(1)\right]^H), \tag{1}$$

so $x \in \hat{m}(p,f,q,s)$.

Theorem 2.4. *Let f, f_1, f_2 be modulus functions q, q_1, q_2 seminorms and $s, s_1, s_2 \geq 0$. Then*

(i) If $s > 1$ then $Z(f_1,q,s) \subseteq Z(f \circ f_1,q,s)$,

(ii) $Z(p,f_1,q,s) \cap Z(p,f_2,q,s) \subseteq Z(p,f_1+f_2,q,s)$,

(iii) $Z(p,f,q_1,s) \cap Z(p,f,q_2,s) \subseteq Z(p,f,q_1+q_2,s)$,

(iv) If q_1 is stronger than q_2 then $Z(p,f,q_1,s) \subseteq Z(p,f,q_2,s)$,

(v) If $s_1 \leq s_2$ then $Z(p,f,q,s_1) \subseteq Z(p,f,q,s_2)$,

(vi) If $q_1 \cong$(equivalent to) q_2, then $Z(p,f,q_1,s) = Z(p,f,q_2,s)$,

where $Z = \hat{m}$, $\hat{c}$ and $\hat{c}_0$.

Proof. (i) We prove this part for $Z = \hat{c}$ and the rest of the cases will follow similarly. Let $x \in \hat{c}(p,f,q,s)$, so that

$$S_m = m^{-s}\left[f_1\left(q\left(t_{m,n}(x-\ell)\right)\right)\right] \to 0.$$

Let $\varepsilon > 0$ and choose δ with $0 < \delta < 1$ such that $f(c) < \varepsilon$ for $0 \leq t \leq \delta$. Now we write

$$\begin{aligned} I_1 &= \{m \in \mathbb{N} : f_1 (q (t_{m,n} (x - \ell))) \leq \delta\} \\ I_2 &= \{m \in \mathbb{N} : f_1 (q (t_{m,n} (x - \ell))) > \delta\}. \end{aligned}$$

For $f_1 (q (t_{m,n} (x - \ell))) > \delta$,

$$f_1 (q (t_{m,n} (x - \ell))) < f_1 (q (t_{m,n} (x - \ell))) \delta^{-1} < 1 + \left[\left| f_1 (q (t_{m,n} (x - \ell))) \delta^{-1} \right| \right]$$

where $m \in I_2$ and $[|u|]$ denotes the integer part of u. By the definition of f we have for $f_1 (q (t_{m,n} (x - \ell))) > \delta$,

$$\begin{aligned} f (f_1 (q (t_{m,n} (x - \ell)))) &\leq \left(1 + \left[\left| f_1 (q (t_{m,n} (x - \ell))) \delta^{-1} \right| \right] \right) f (1) \\ &\leq 2f (1) f_1 (q (t_{m,n} (x - \ell))) \delta^{-1}. \end{aligned} \tag{2.1}$$

For $f_1 (q (t_{m,n} (x - \ell))) \leq \delta$,

$$f (f_1 (q (t_{m,n} (x - \ell)))) < \varepsilon \tag{2.2}$$

where $m \in I_1$. By (2.1) and (2.2) we have

$m^{-s} [f (f_1 (q (t_{m,n} (x - \ell))))] \leq m^{-s}\varepsilon + \left[2f (1) \delta^{-1}\right] S_m \rightarrow 0.$ as $m \rightarrow \infty$,uniformly n.

Hence $\hat{c} (p, f_1, q, s) \subseteq \hat{c} (p, f \circ f_1, q, s)$.

(ii) The proof follows from the following inequality

$$m^{-s} [(f_1 + f_2) (q (t_{m,n} (x)))]^{p_m} \leq Cm^{-s} [f_1 (q (t_{m,n} (x)))]^{p_m} + Cm^{-s} [f_2 (q (t_{m,n} (x)))]^{p_m}.$$

(iii), (iv) (v)and (vi) follow easily.

Corollary 2.1. *Let f be a modulus function, then we have*

(i) If $s > 1$, $Z (p, q, s) \subseteq Z (p, f, q, s)$,

(ii) $Z (p, f, q) \subseteq Z (p, f, q, s)$,

(iii) $Z (p, q) \subseteq Z (p, q, s)$,

(iv) $Z (f, q) \subseteq Z (f, q, s)$

where $Z = \hat{m}$, $\hat{c}$ *and* $\hat{c}_0$.

The proof is straightforward.

Theorem 2.5. *For any two sequences* $p = (p_k)$ *and* $r = (r_k)$ *of positive real numbers and for any two seminorms* q_1 *and* q_2 *on* X *we have* $Z(p, f, q_1, s) \cap Z(r, f, q_2, s) \neq \emptyset$.

Proof. The proof follows from the fact that the zero element $\bar{\theta}$ belongs to each of the classes of sequences involved in the intersection.

Theorem 2.6. *For any two sequences* $p = (p_m)$ *and* $r = (r_m)$, *we have* $\hat{c}_0(r, f, q, s) \subseteq \hat{c}_0(p, f, q, s)$ *if and only if* $\liminf \frac{p_m}{r_m} > 0$.

Proof. If we take $y_m = f(q(t_{m,n}(x)))$ for all $m \in \mathbb{N}$, then using the same technique of lemma 1 of Maddox [9], it is easy to prove the theorem.

Theorem 2.7. *For any two sequences* $p = (p_m)$ *and* $r = (r_m)$, *we have* $\hat{c}_0(r, f, q, s) = \hat{c}_0(p, f, q, s)$ *if and only if* $\liminf \frac{p_m}{r_m} > 0$ *and* $\liminf \frac{r_m}{p_m} > 0$.

Theorem 2.8. *Let* $0 < p_m \leq r_m \leq 1$. *Then* $\hat{m}(r, f, q, s)$ *is closed subspace of* $\hat{m}(p, f, q, s)$.

Proof. Let $x \in \hat{m}(r, f, q, s)$. Then there exists a constant $B > 1$ such that

$$k^{-s}\left[f\left(t_{m,n}(x)\right)\right]^{r_m/M} \leq B \quad \text{for all } m, n$$

and so

$$k^{-s}\left[f\left(t_{m,n}(x)\right)\right]^{p_m/M} \leq B \quad \text{for all } m, n.$$

Thus $x \in \hat{m}(p, f, q, s)$. To show that $\hat{m}(r, f, q, s)$ is closed, suppose that $x^i \in \hat{m}(r, f, q, s)$ and $x^i \to x \in \hat{m}(p, f, q, s)$. Then for every $0 < \varepsilon < 1$, there exists N such that for all m, n

$$k^{-s}\left[f\left(t_{m,n}\left(x^i - x\right)\right)\right]^{p_m/M} \leq B \quad \text{for all } i > N.$$

Now

$$k^{-s}\left[f\left(t_{m,n}\left(x^i - x\right)\right)\right]^{r_m/M} < k^{-s}\left[f\left(t_{m,n}\left(x^i - x\right)\right)\right]^{p_m/M} < \varepsilon \quad \text{for all } i > N.$$

Therefore $x \in \hat{m}(r, f, q, s)$. This completes the proof.

References

[1] Banach, S. : *Théorie des operations linéaires*, Warszawa, 1932.

[2] Bhardwaj, V.K. : *A generalization of a sequence space of Ruckle*, Bull. Calcutta Math. Soc. **95** (5) (2003), 411-420.

[3] Bilgin, T : *The sequence space* $l(p, f, q, s)$ *on seminormed spaces.* Bull. Calcutta Math. Soc. **86** (4) (1994), 295-304.

[4] Connor, J.S. : *On strong matrix summability with respect to a modulus and statistical convergence*. Canad. Math. Bul. **32** (2) (1989), 194-198.

[5] Duran, J.P. : *Infinite matrices and almost convergence*, Math. Zeit., **128** (1972), 75-83.

[6] Esi, A. : *A new sequence space defined by a modulus function.* J. Anal. **8** (2000), 31-37.

[7] King, J.P. : *Almost summable sequences,* Proc. Amer. Math. Soc., **16** (1966), 1219-1225.

[8] Lorentz, G.G. : *A contribution the theory of divergent series*, Acta Math. **80** (1948), 167-190.

[9] Maddox, I.J. : *Spaces of strongly summable sequences,* Ouart. J.Math. Oxford **2** (18) (1967), 345-355.

[10] Maddox, I.J. : *Sequence spaces defined by a modulus,* Math. Proc. Camb. Phil. Soc. **100** (1986), 161-166.

[11] Nakano, H. : *Concave modulars*, J. Math. Soc. Japan. **5** (1953), 29-49.

[12] Nanda, S. : *Infinite matrices and almost convergence*, Journal of Indian Math.Soc. **40** (1976), 173-184.

[13] Nanda, S. : *Matrix transformations and almost boundedness,* Glas. Mat., III. Ser. **14** (34) (1979), 99-107.

[14] Ruckle, W.H. : *FK spaces in which the sequence of coordinate vectors is bounded,* Canad. J. Math. **25** (1973), 973-978.

[15] Wilansky, A. : *Functional Analysis*, Blaisdell Publishing Company, New York, 1964.

Department of Mathematics, Firat University, 23119, Elaziğ-TURKEY

E-mail address: yaltin23@yahoo.com, agokhan1@firat.edu.tr, hifsialtinok@yahoo.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

EINSTEIN EQUATIONS IN THE GEOMETRY OF SECOND ORDER

GH. ATANASIU N. VOICU

Abstract. In [7], R. Miron and Gh. Atanasiu wrote the Einstein equations of a metric structure G on the tangent bundle of order two, T^2M (previously named "2-osculator bundle" and denoted by Osc^2M), endowed with a nonlinear connection N and a linear connection D such that the 2-tangent structure J be absolutely parallel to D.

In the present paper, the authors determine the Einstein equations by making use of the concept of N-linear connection defined by Gh. Atanasiu, [1], this is, a linear connection which is not neccesarily compatible with J, but only preserves the distributions generated by the nonlinear connection N.

1. The Tangent Bundle T^2M

Let M be a real n- dimensional manifold of class $\mathcal{C}^\infty$, $\left(T^2M, \pi^2, M\right)$ its second order tangent bundle and let $\widetilde{T^2M}$ be the space T^2M without its null section. For a point $u \in T^2M$, let $(x^a, y^{(1)a}, y^{(2)a})$ be its coordinates in a local chart.

Let N be a nonlinear connection, [3, 8-13], and denote its coefficients by $\left(\underset{1}{N}{}^a_b, \underset{2}{N}{}^a_b\right)$, a, $b = 1, ..., n$.Then, N determines the direct decomposition

$$T_uT^2M = N_0(u) \oplus N_1(u) \oplus V_2(u), \forall u \in T^2M. \tag{1}$$

Received by the editors: 15.09.2005.

2000 *Mathematics Subject Classification.* 53C60, 58 B20, 70 G45.

Key words and phrases. 2-tangent bundle, nonlinear connection, N-linear connection, Riemannian metric, Ricci tensor, Einstein equations.

The adapted basis to (1) is $(\delta_a, \delta_{1a}, \delta_{2a})$ and its dual basis is $(dx^a, \delta y^{(1)a}, \delta y^{(2)a})$, where

$$\begin{cases} \delta_a = \dfrac{\delta}{\delta x^a} = \dfrac{\partial}{\partial x^a} - \underset{1}{N}{}^c_a \dfrac{\partial}{\partial y^{(1)c}} - \underset{2}{N}{}^c_a \dfrac{\partial}{\partial y^{(2)c}} \\ \delta_{1a} = \dfrac{\delta}{\delta y^{(1)a}} = \dfrac{\partial}{\partial y^{(1)a}} - \underset{1}{N}{}^c_a \dfrac{\partial}{\partial y^{(2)c}} \\ \delta_{2a} = \dfrac{\partial}{\partial y^{(2)a}}, \end{cases} \tag{2}$$

respectively,

$$\begin{cases} \delta y^{(1)a} = dy^{(1)a} + \underset{1}{M}{}^c_a dx^c \\ \delta y^{(2)a} = dy^{(2)a} + \underset{1}{M}{}^c_a dy^{(1)c} + \underset{2}{M}{}^c_a dx^c, \end{cases} \tag{3}$$

where $\underset{1}{M}{}^c_a, \underset{2}{M}{}^c_a$ are the dual coefficients of the nonlinear connection N.

Then, a vector field $X \in \mathcal{X}\left(T^2M\right)$ is represented in the local adapted basis as

$$X = X^{(0)a}\delta_a + X^{(1)a}\delta_{1a} + X^{(2)a}\delta_{2a}, \tag{4}$$

with the three right terms (called *d-vector fields*) belonging to the distributions N, N_1 and V_2 respectively.

A 1-form $\omega \in \mathcal{X}^*\left(T^2M\right)$ will be decomposed as

$$\omega = \omega^{(0)}_a dx^a + \omega^{(1)}_a \delta y^{(1)a} + \omega^{(2)}_a \delta y^{(2)a}.$$

Similarly, a tensor field $T \in \mathcal{T}^r_s\left(T^2M\right)$ can be split with respect to (1) into components ,which will be called *d-tensor fields.*

The $\mathcal{F}\left(T^2M\right)$-linear mapping $J : \mathcal{X}\left(T^2M\right) \to \mathcal{X}\left(T^2M\right)$given by

$$J\left(\delta_a\right) = \delta_{1a}, J\left(\delta_{1a}\right) = \delta_{2a}, J\left(\delta_{2a}\right) = 0 \tag{5}$$

is called the *2-tangent structure on* T^2M,[8-13].

2. N-linear connections. d-tensors of curvature

An *N-linear connection D*, [1], is a linear connection on T^2M, which preserves by parallelism the distributions N, N_1 and V_2. Let us notice that an N-linear connection, in the sense of the definition above, is not necessarily compatible to the

2-tangent structure J (an N-linear connection which is also compatible to J is called, [1], a *JN-linear connection*).

An N-linear connection is locally given by its coefficients

$$D\Gamma\left(N\right)=\left(\underset{(00)}{L}{}^{a}{}_{bc},\underset{(10)}{L}{}^{a}{}_{bc},\underset{(20)}{L}{}^{a}{}_{bc},\underset{(01)}{C}{}^{a}{}_{bc},\underset{(11)}{C}{}^{a}{}_{bc},\underset{(21)}{C}{}^{a}{}_{bc},\underset{(02)}{C}{}^{a}{}_{bc},\underset{(12)}{C}{}^{a}{}_{bc},\underset{(22)}{C}{}^{a}{}_{bc}\right), \tag{6}$$

where

$$\left\{\begin{array}{l} D_{\delta_c}\delta_b=\underset{(00)}{L}{}^{a}{}_{bc}\delta_a,\ D_{\delta_c}\delta_{1b}=\underset{(10)}{L}{}^{a}{}_{bc}\delta_{1a},\ D_{\delta_c}\delta_{2b}=\underset{(20)}{L}{}^{a}{}_{bc}\delta_{2a} \\ D_{\delta_{1c}}\delta_b=\underset{(01)}{C}{}^{a}{}_{bc}\delta_a,\ D_{\delta_{1c}}\delta_{1b}=\underset{(11)}{C}{}^{a}{}_{bc}\delta_{1a},\ D_{\delta_{1c}}\delta_{2b}=\underset{(21)}{C}{}^{a}{}_{bc}\delta_{2a} \\ D_{\delta_{2c}}\delta_b=\underset{(02)}{C}{}^{a}{}_{bc}\delta_a,\ D_{\delta_{2c}}\delta_{1b}=\underset{(12)}{C}{}^{a}{}_{bc}\delta_{1a},\ D_{\delta_{2c}}\delta_{2b}=\underset{(22)}{C}{}^{a}{}_{bc}\delta_{2a} \end{array}\right. . \tag{7}$$

In the particular case when D is J-compatible, we have

$$\begin{aligned} \underset{(00)}{L}{}^{a}{}_{bc} &= \underset{(10)}{L}{}^{a}{}_{bc}=\underset{(20)}{L}{}^{a}{}_{bc}=:L^{a}{}_{bc}, \\ \underset{(01)}{C}{}^{a}{}_{bc} &= \underset{(11)}{C}{}^{a}{}_{bc}=\underset{(21)}{C}{}^{a}{}_{bc}=\underset{(1)}{C}{}^{a}{}_{bc}, \\ \underset{(02)}{C}{}^{a}{}_{bc} &= \underset{(12)}{C}{}^{a}{}_{bc}=\underset{(22)}{C}{}^{a}{}_{bc}=\underset{(2)}{C}{}^{a}{}_{bc}. \end{aligned}$$

For an N-linear connection, let

$$\left\{\begin{array}{r} \underset{0}{D}{}^{H}_{X}Y=D_{X^H}Y^H,\ \underset{0}{D}{}^{V_1}_{X}Y=D_{X^{V_1}}Y^H,\ \underset{0}{D}{}^{V_2}_{X}Y=D_{X^{V_2}}Y^H \\ \underset{\beta}{D}{}^{H}_{X}Y=D_{X^H}Y^{V_\beta},\ \underset{\beta}{D}{}^{V_1}_{X}Y=D_{X^{V_1}}Y^{V_\beta},\ \underset{\beta}{D}{}^{V_2}_{X}Y=D_{X^{V_2}}Y^{V_\beta}, \\ \beta=1,2. \end{array}\right.$$

$\underset{\alpha}{D}{}^{H}$, $\underset{\alpha}{D}{}^{V_1}$, $\underset{\alpha}{D}{}^{V_2}$ are called respectively, h_α-, $v_{1\alpha}$- and $v_{2\alpha}$-*covariant derivatives*, $\alpha=0,1,2$. In local coordinates, for a d-tensor field

$$T=T^{a_1\ldots a_r}_{b_1\ldots b_s}\left(x,y^{(1)},y^{(2)}\right)\delta_{a_1}\otimes\ldots\otimes\delta_{2a_r}\otimes dx^{b_1}\otimes\ldots\otimes\delta y^{(2)b_s}.$$

we have

$$\underset{\alpha}{D}{}^{H}_{X}T=X^{(0)m}\,T^{a_1\ldots a_r}_{b_1\ldots b_s|\alpha m}\delta_{a_1}\otimes\ldots\otimes\delta_{2a_r}\otimes dx^{b_1}\otimes\ldots\otimes\delta y^{(2)b_s},$$

where

$$T^{a_1...a_r}_{b_1...b_s|\alpha m} = \delta_m T^{a_1...a_r}_{b_1...b_s} + \underset{(\alpha 0)}{L}{}^{a_1}_{hm} T^{ha_2...a_r}_{b_1...b_s} + ... \underset{(\alpha 0)}{L}{}^{a_r}_{hm} T^{a_1...a_{r-1}h}_{b_1...b_s} -$$
$$- \underset{(\alpha 0)}{L}{}^{h}_{b_1 m} T^{a_1...a_r}_{hb_2...b_s} - ... - \underset{(\alpha 0)}{L}{}^{h}_{b_s m} T^{a_1...a_r}_{b_1...b_{s-1}h}.$$

and

$$\underset{\alpha}{D}{}^{V_\beta}_{X} T = X^{(1)m}\, T^{a_1...a_r}_{b_1...b_s} \overset{(\beta)}{|}_{\alpha m}\ \delta_{a_1} \otimes ... \otimes \delta_{2a_r} \otimes dx^{b_1} \otimes ... \otimes \delta y^{(2)b_s},$$

where

$$T^{a_1...a_r}_{b_1...b_s} \overset{(\beta)}{|}_{\alpha m} = \delta_{\beta m} T^{a_1...a_r}_{b_1...b_s} + \underset{(\alpha\beta)}{C}{}^{a_1}_{hm} T^{ha_2...a_r}_{b_1...b_s} + ... \underset{(\alpha\beta)}{C}{}^{a_r}_{hm} T^{a_1...a_{r-1}h}_{b_1...b_s} -$$
$$- \underset{(\alpha\beta)}{C}{}^{h}_{b_1 m} T^{a_1...a_r}_{hb_2...b_s} - ... - \underset{(\alpha\beta)}{C}{}^{h}_{b_s m} T^{a_1...a_r}_{b_1...b_{s-1}h}.$$

The curvature of the N-linear connection D,

$$R(X,Y)Z = D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z$$

is completely determined by its components (which are d-tensors) $R(\delta_{\gamma l}, \delta_{\beta k})\,\delta_{\alpha j}$. Namely, the 2-forms of curvature of an N- linear connection are, [1],

$$\underset{(\alpha)}{\Omega}{}^a_b = \frac{1}{2} \underset{(0\alpha)}{R}{}_b{}^a{}_{cd} dx^c \wedge dx^d + \underset{(1\alpha)}{P}{}_b{}^a{}_{cd} dx^c \wedge \delta y^{(1)d} + \underset{(2\alpha)}{P}{}_b{}^a{}_{cd} dx^c \wedge \delta y^{(2)d} + \tag{8}$$
$$\frac{1}{2} \underset{(1\alpha)}{S}{}_b{}^a{}_{cd} \delta y^{(1)c} \wedge \delta y^{(1)d} + \underset{(2\alpha)}{Q}{}_b{}^a{}_{cd} dy^{(1)c} \wedge \delta y^{(2)d} + \frac{1}{2} \underset{(2\alpha)}{S}{}_b{}^a{}_{cd} \delta y^{(2)c} \wedge \delta y^{(2)d},$$

$\alpha = 0, 1, 2$, where the coefficients $\underset{(0\alpha)}{R}{}_b{}^a{}_{cd}$, $\underset{(\beta\alpha)}{P}{}_b{}^a{}_{cd}$, $\underset{(\beta\alpha)}{Q}{}_b{}^a{}_{cd}$, $\underset{(\beta\alpha)}{S}{}_b{}^a{}_{cd}$ are d-tensors, named the *d-tensors of curvature* of the N-linear connection D. For a JN- linear connection, there holds

$$\underset{(0)}{\Omega}{}^a_b = \underset{(1)}{\Omega}{}^a_b = \underset{(2)}{\Omega}{}^a_b,$$

this is,

$$\begin{aligned} \underset{(00)}{R}{}_b{}^a{}_{cd} &= \underset{(01)}{R}{}_b{}^a{}_{cd} = \underset{(02)}{R}{}_b{}^a{}_{cd} = \underset{(0)}{R}{}_b{}^a{}_{cd}; \\ \underset{(\beta 0)}{P}{}_b{}^a{}_{cd} &= \underset{(\beta 1)}{P}{}_b{}^a{}_{cd} = \underset{(\beta 2)}{P}{}_b{}^a{}_{cd} = \underset{(\beta)}{P}{}_b{}^a{}_{cd} \end{aligned} \tag{9}$$

$$\underset{(20)}{Q}{}_{b\,cd}^{\;a} = \underset{(21)}{Q}{}_{b\,cd}^{\;a} = \underset{(22)}{Q}{}_{b\,cd}^{\;a} = Q_{b\,cd}^{\;a}$$

$$\underset{(\beta 0)}{S}{}_{b\,cd}^{\;a} = \underset{(\beta 1)}{S}{}_{b\,cd}^{\;a} = \underset{(\beta 2)}{S}{}_{b\,cd}^{\;a} = \underset{(\beta)}{S}{}_{b\,cd}^{\;a}, \beta = 1, 2.$$

The detailed expressions of the d-tensors of curvature can be found in [1].

3. **Metric structures on** T^2M

A *Riemannian metric* on T^2M is a tensor field G of type $(0,2)$, which is nondegenerate in each $u \in T^2M$ and is positively defined on T^2M.

In this paper, we shall consider metrics in the form

$$G = \underset{(0)}{g}{}_{ab} dx^a \otimes dx^b + \underset{(1)}{g}{}_{ab} \delta y^{(1)a} \otimes \delta y^{(1)b} + \underset{(2)}{g}{}_{ab} \delta y^{(2)a} \otimes \delta y^{(2)b}, \tag{10}$$

where $\underset{(\alpha)}{g}{}_{ab} = \underset{(\alpha)}{g}{}_{ab}(x, y^{(1)}, y^{(2)})$; this is, such that the distributions N, N_1 and V_2 generated by the nonlinear connection N be orthogonal with respect to G.

An N-linear connection D is called a *metrical* N-linear connection if $D_X G = 0$, $\forall X \in \mathcal{X}(T^2M)$.

This means

$$\underset{(\alpha)}{g}{}_{ab|\alpha c} = \underset{(\alpha)}{g}{}_{ab} \overset{\beta}{|}_{\alpha c} = 0,\ \alpha = 0,1,2,\ \beta = 1,2.$$

The existence of metrical $N-$linear connections is proved in [2].

4. **The Ricci tensor** $Ric\,(D)$

Let us notice that, if D is not J- compatible, we could expect that the components of the Ricci tensor look in a more complicated way that the ones in the Miron-Atanasiu theory, [7].

Indeed, if we consider the Ricci tensor $Ric\,(D)$,[14], as the trace of the linear operator

$$V \mapsto R\,(V, X)\,Y,\ \forall V = V^{(0)a}\delta_a + V^{(1)a}\delta_{1a} + V^{(2)a}\delta_{2a} \in \mathcal{X}\left(T^2M\right), \tag{11}$$

then we have:

$$Ric\,(D)\,(X,Y) = trace(V \mapsto R\left(V^{H},X\right)Y + R\left(V^{V_1},X\right)Y + +R\left(V^{V_1},X\right)Y). \tag{12}$$

By a straightforward calculus, we obtain:

Theorem 4.1. *The Ricci tensor $Ric\,(d)$ has the following components:*

$$\begin{aligned}
Ric\,(D)\left(\frac{\delta}{\delta x^b},\frac{\delta}{\delta x^a}\right) &= \underset{(00)}{R}{}_{a\ bc}^{\ c} =: R_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(1)b}},\frac{\delta}{\delta x^a}\right) &= -\underset{(10)}{P}{}_{a\ cb}^{\ c} =: -\underset{(10)}{\overset{2}{P}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(2)b}},\frac{\delta}{\delta x^a}\right) &= -\underset{(20)}{P}{}_{a\ cb}^{\ c} =: -\underset{(20)}{\overset{2}{P}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta x^b},\frac{\delta}{\delta y^{(1)a}}\right) &= \underset{(11)}{P}{}_{a\ bc}^{\ c} =: \underset{(11)}{\overset{1}{P}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(1)b}},\frac{\delta}{\delta y^{(1)a}}\right) &= \underset{(11)}{S}{}_{a\ bc}^{\ c} =: \underset{(1)}{S}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(2)b}},\frac{\delta}{\delta y^{(1)a}}\right) &= -\underset{(21)}{Q}{}_{a\ cb}^{\ c} =: -\underset{(21)}{\overset{2}{Q}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta x^b},\frac{\delta}{\delta y^{(2)a}}\right) &= \underset{(22)}{P}{}_{a\ bc}^{\ c} =: \underset{(22)}{\overset{1}{P}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(1)b}},\frac{\delta}{\delta y^{(2)a}}\right) &= \underset{(22)}{Q}{}_{a\ bc}^{\ c} =: \underset{(22)}{\overset{1}{Q}}{}_{ab};\\
Ric\,(D)\left(\frac{\delta}{\delta y^{(2)b}},\frac{\delta}{\delta y^{(2)a}}\right) &= \underset{(22)}{S}{}_{a\ bc}^{\ c} =: \underset{(2)}{S}{}_{ab}.
\end{aligned}$$

The Ricci scalar $Sc\,(D)$ is, thus,

$$Sc\,(D) = g^{ab}R_{ab} + \underset{(1)}{g}{}^{ab}\underset{(1)}{S}{}_{ab} + \underset{(2)}{g}{}^{ab}\underset{(2)}{S}{}_{ab}, \tag{13}$$

where g^{ab}, $\underset{(1)}{g}{}^{ab}$, $\underset{(2)}{g}{}^{ab}$ are the coefficients of the inverse matrix of G.

In the particular case of a JN-linear connection, taking into account $(8')$, with the notations in [7], we have

$$\begin{aligned} \overset{1}{\underset{(\beta\beta)}{P}}{}_{ab} &= \overset{1}{\underset{(\beta)}{P}}{}_{ab},\ \overset{2}{\underset{(\beta 0)}{P}}{}_{ab} = \underset{(\beta)}{P}{}_{ab},\ \overset{1}{\underset{(22)}{Q}}{}_{ab} = \underset{(21)}{P}{}^{1}_{ab}\,(= Q_{a\ bc}^{\ c}), \\ \overset{2}{\underset{(21)}{Q}}{}_{ab} &= \underset{(21)}{P}{}^{2}_{ab}\,(= Q^{c}_{a\ cb}). \end{aligned} \tag{14}$$

5. Einstein equations

The Einstein equations associated to the metrical N-linear connection D are

$$Ric\,(D) - \frac{1}{2} Sc\,(D)\,G = \kappa \mathcal{T}, \tag{15}$$

where κ is a constant and $\mathcal{T}$ is the energy-momentum tensor, given by its components

$$\underset{(\alpha\beta)}{\mathcal{T}}{}_{ab} = \mathcal{T}(\delta_{\beta b}, \delta_{\alpha a})$$

Expressing the above relation in the adapted frame (2), we obtain

Theorem 5.1. *The Einstein equations associated to the metrical N- linear connection D are*

$$\begin{aligned} & R_{ab} - \tfrac{1}{2} Sc\,(D)\, g_{ab} = \kappa\, \underset{(00)}{\mathcal{T}}{}_{ab}; \\ & \overset{1}{\underset{(\beta\beta)}{P}}{}_{ab} = \kappa\, \underset{(\beta 0)}{\mathcal{T}}{}_{ab},\ \beta = 1,2; \\ & \overset{2}{\underset{(\beta 0)}{P}}{}_{ab} = -\kappa\, \underset{(0\beta)}{\mathcal{T}}{}_{ab},\ \beta = 1,2; \\ & \underset{(\beta)}{S}{}_{ab} - \tfrac{1}{2} Sc\,(D)\ \underset{(\beta)}{g}{}_{ab} = \kappa\, \underset{(\beta\beta)}{\mathcal{T}}{}_{ab},\ \alpha = 1,2; \\ & \overset{1}{\underset{(22)}{Q}}{}_{ab} = \kappa\, \underset{(21)}{\mathcal{T}}{}_{ab}; \\ & \overset{2}{\underset{(21)}{Q}}{}_{ab} = -\kappa\, \underset{(12)}{\mathcal{T}}{}_{ab}. \end{aligned}$$

In the case when D is a JN-linear connection, one obtains the result in [7].

In order to avoid confusions when raising and lowering indices, because of the fact that the components g^{ab}, $\underset{(1)}{g}{}^{ab}$, $\underset{(2)}{g}{}^{ab}$ are different, we will denote in the following by $i, j, ...$ the indices corresponding to the horizontal distribution, by $a, b, ...$ those corresponding to N_1, and by $p, q, ...$ those corresponding to V_2. Thus, if we impose the condition that the divergence of the energy- momentum tensor vanish, in the adapted frame we will obtain

Theorem 5.2. *The law of conservation on T^2M endowed with the metrical N-linear connection D is given by*

$$\left(R^i_j - \frac{1}{2}Sc\,(D)\,\delta^i_j\right)_{|i} + \underset{(11)}{\overset{1}{P}}{}^a_j \overset{(1)}{|}_a - \underset{(10)}{\overset{2}{P}}{}^a_j \overset{(1)}{|}_a + \underset{(22)}{\overset{1}{P}}{}^p_j \overset{(2)}{|}_p - \underset{(20)}{\overset{2}{P}}{}^p_j \overset{(2)}{|}_p = 0;$$

$$\underset{(11)}{\overset{1}{P}}{}^i_{b|i} - \underset{(10)}{\overset{2}{P}}{}^i_{b|i} + \left(\underset{(1)}{S}{}^a_b - \frac{1}{2}Sc\,(D)\,\delta^a_b\right) \overset{(1)}{|}_a + \underset{(22)}{\overset{1}{Q}}{}^p_b \overset{(2)}{|}_p - \underset{(21)}{\overset{2}{Q}}{}^p_b \overset{(2)}{|}_p = 0;$$

$$\underset{(22)}{\overset{1}{P}}{}^i_{p\,|i} - \underset{(20)}{\overset{2}{P}}{}^i_{p|i} + \underset{(22)}{\overset{1}{Q}}{}^a_p \overset{(2)}{|}_a - \underset{(21)}{\overset{2}{Q}}{}^a_p \overset{(2)}{|}_a + \left(\underset{(2)}{S}{}^a_b - \frac{1}{2}Sc\,(D)\,\delta^p_q\right) \overset{(2)}{|}_p = 0.$$

In the same way, one can deduce the Maxwell equations associated to the metrical N-linear connection D.

References

[1] Atanasiu, Gh.: *New aspects in the Differential Geometry of second order,* Sem. de Mecanică, Univ de Vest, Timişoara, 2001.

[2] Atanasiu, Gh., *The homogeneous prolongation to the second order tangent bundle of a Riemannian metric* (to appear);

[3] Miron, R., *The Geometry of Higher Order Lagrange Spaces.* Applications to Mechanics and Physics, Kluwer Acad. Publ. FTPM no. 82, 1997;

[4] Miron, R., *The Geometry of Higher Order Finsler Spaces*, Hadronic Press, Inc. USA, (1998).

[5] Miron, R., *The homogeneous lift of a Riemannian metric*, An. St. Univ "Al. I. Cuza", Iaşi;

[6] Miron, R. and Anastasiei, M., *The Geometry of Lagrange Spaces.* Theory and Applications, Kluwer Acad. Publ. no. 59, 1994;

[7] Miron, R. and Atanasiu, Gh., *Geometrical theory of gravitational and electromagnetic fields in higher order Lagrange spaces*, Tsukuba J. of Math., vol 20, 1996;

[8] Miron, R., and Atanasiu, Gh., *Compendium on the higher-order Lagrange spaces: The geometry of k-osculator bundles. Prolongation of the Riemannian, Finslerian and Lagrangian structures. Lagrange spaces,* Tensor N.S. 53, 1993.

[9] Miron, R., and Atanasiu, Gh.,*Compendium sur les espaces Lagrange d'ordre superieur: La geometrie du fibre k-osculateur. Le prolongement des structures Riemanniennes, Finsleriennes et Lagrangiennes. Les espaces $L^{(k)n}$,* Univ. Timişoara, Seminarul de Mecanică, no. 40, 1994, 1-27.

[10] Miron, R., and Atanasiu, Gh.,*Lagrange Geometry of Second Order*, Math. Comput Modelling, **20**, no. 4, 1994, 41-56.

[11] Miron, R., and Atanasiu, Gh., *Differential Geometry of the k-Osculator Bundle*, Rev. Roumaine Math. Pures et Appl., **41**, 3/4, 1996, 205-236.

[12] Miron, R., and Atanasiu, Gh., *Higher-order Lagrange Spaces,* Rev. Roumaine Math. Pures et Appl., **41**, 3/4, 1996, 251-262.

[13] Miron, R., and Atanasiu, Gh., *Prolongations of the Riemannian, Finslerian and Lagrangian Structures*, Rev. Roumaine Math. Pures et Appl., **41**, 3/4, 1996, 237-249.

[14] Munteanu, Gh. and Bălan, V., *Lecţii de Teoria relativităţii,* Ed. Bren, Bucureşti, 2000.

"Transilvania" University,
Brasov, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

ITERATES OF SOME MULTIVARIATE APPROXIMATION PROCESSES, VIA CONTRACTION PRINCIPLE

CLAUDIA BACOŢIU

Abstract. In this paper we study a general class of linear positive operators, using the theory of weakly Picard operators. The convergence of the iterates of the defined operators will be proven.

1. **Introduction**

In [2] and [1] Agratini and Rus applied the theory of weakly Picard operators to prove the convergence of iterates of a certain class of linear positive operators. In some particular cases, these operators are well known approximation operators, such as Bernstein or Stancu operators. In the above mentioned papers, the authors have considered the univariate, respectivelly the bivariate cases. In the present paper we give a generalization of these results to a class of linear positive operators defined on $C([0,1]^p)$, $p \in \mathbb{N}$.

2. **Weakly Picard operators**

Let $(X, \rightarrow)$ be an L-space and $A : X \rightarrow X$ an operator. In this paper we will use the following notations:

$$F_A := \{x \in X : A(x) = x\};$$

$$I(A) := \{Y \in P(X) : A(Y) \subset Y\};$$

$$A^0 := 1_X \,,\; A^{n+1} := A \circ A^n \;\forall n \in \mathbb{N}.$$

Received by the editors: 11.09.2005.

2000 *Mathematics Subject Classification.* 41A36, 47H10.

Key words and phrases. Linear positive operators, contraction principle, weakly Picard operators.

Definition 2.1. *(Rus [7]) The operator A is said to be:*

(i) weakly Picard operator (WPO) if $\forall x_0 \in X$ $A^n(x_0) \to x_0^*$, and the limit x_0^* is a fixed point of A, which may depend on x_0;

(ii) Picard operator (PO) if $F_A = \{x^*\}$ and $\forall x_0 \in X$ $A^n(x_0) \to x^*$.

If A is an WPO, we consider the operator A^∞ defined by

$$A^\infty : X \to X, \quad A^\infty(x) := \lim_{n\to\infty} A^n(x).$$

We have the next characterization theorem of WPOs:

Theorem 2.1. *(Rus [7]) The operator A is WPO if and only if there exists a partition of X, $X = \bigcup\limits_{\lambda\in\Lambda} X_\lambda$ such that:*

(i) $X_\lambda \in I(A)$, $\forall \lambda \in \Lambda$;

(ii) $A|_{X_\lambda} : X_\lambda \to X_\lambda$ is PO, $\forall \lambda \in \Lambda$.

3. Main results

Let $p \geq 1$ be a fixed integer and

$$D := [0,1] \times [0,1] \times ... \times [0,1] = [0,1]^p.$$

$C(D) = \{f : D \to \mathbb{R} \ : \ f - \text{continuous}\}$.

We introduce the next notations: $\alpha^{\langle 0\rangle} := (0,0,...,0) = 0_{\mathbb{R}^p}$ is the null vector. For all $k \in \overline{1,p}$ and for all $1 \leq i_1 < i_2 < ... < i_k \leq p$, denote by $\alpha^{\langle k\rangle}_{i_1,i_2,...,i_k}$ the vector from $\mathbb{R}^p$ defined as follows: on positions $i_1, i_2, ..., i_k$ the value 1 appears and on all other positions the value 0 is displayed.

$$M_k := \{(i_1,i_2,...,i_k) \ : \ 1 \leq i_1 < i_2 < ... < i_k \leq p\} \subset \mathbb{N}^k \quad \forall k \in \overline{1,p}$$

$$\nu_D := \left\{\alpha^{\langle 0\rangle}\right\} \bigcup \left\{\alpha^{\langle k\rangle}_{i_1,i_2,...,i_k} \ : \ k \in \overline{1,p} \text{ and } (i_1,i_2,...,i_k) \in M_k\right\}.$$

Denote by e_α, $\alpha \in \nu_D$ the test functions

$$e_\alpha : D \to \mathbb{R}_+; \quad e_\alpha(x_1,x_2,...,x_p) := \prod_{k=1}^p x_k^{\alpha_k} \quad \forall (x_1,x_2,...,x_p) \in D,$$

with the convention that, if in a component, α_k is null, then $x_k^{\alpha_k}$ will be replaced by 1.

We notice that

$$Card(M_k) = \begin{pmatrix} p \\ k \end{pmatrix}, \ \forall k = \overline{1,p} \quad \text{and} \quad Card(\nu_D) = \sum_{k=0}^{p} \begin{pmatrix} p \\ k \end{pmatrix} = 2^p := N.$$

Remark 3.1. *Any* $\alpha \in \nu_D$ *is* $\alpha^{\langle 0 \rangle}$ *or there exist* $k \in \overline{1,p}$ *and* $(i_1, i_2, ..., i_k) \in M_k$ *such that* $\alpha = \alpha^{\langle k \rangle}_{i_1,i_2,...,i_k}$.

Remark 3.2. *Because* $Card(\nu_D) = Card\{1, 2, ..., N\}$*, it follows that there exists a bijective function*

$$\omega : \nu_D \to \{1, 2, ..., N\}.$$

More precisely:

- for $k = 0$ *there exists a unique* $j \in \overline{1,N}$ *such that* $\omega(\alpha^{\langle 0 \rangle}) = j$ *and*

- for any $k \in \overline{1,p}$ *and for any* $(i_1, i_2, ..., i_k) \in M_k$ *there exists a unique* $j \in \overline{1,N}$ *such that* $\omega(\alpha^{\langle k \rangle}_{i_1,i_2,...,i_k}) = j$.

For all $(m_1, m_2, ..., m_p) \in \mathbb{N}^p$ consider the next p-dimensional net

$$\Delta^k_{m_k} := (0 = x_{k,m_k,0} < x_{k,m_k,1} < ... < x_{k,m_k,m_k} = 1) \quad \forall k = \overline{1,p}.$$

We also consider the next systems of real positive functions

$$0 \le \psi_{k,m_k,i} \in C[0,1], \quad \forall i = \overline{0,m_k} \quad \forall k = \overline{1,p}.$$

Let the next assumptions be satisfied:

$$\sum_{i=0}^{m_k} \psi_{k,m_k,i}(x) = 1, \quad \forall x \in [0,1], \quad \forall k = \overline{1,p}; \tag{1}$$

$$\sum_{i=0}^{m_k} x_{k,m_k,i}\psi_{k,m_k,i}(x) = x, \quad \forall x \in [0,1], \quad \forall k = \overline{1,p}; \tag{2}$$

$$\psi_{k,m_k,0}(0) = \psi_{k,m_k,m_k}(1) = 1, \quad \forall k = \overline{1,p}. \tag{3}$$

We also introduce the next notation:

$$K := \{0, 1, ..., m_1\} \times \{0, 1, ..., m_2\} \times ... \times \{0, 1, ..., m_p\}.$$

Clearly,

$$\partial K = \{(0,0,...,0),(m_1,0,...,0),...,(0,0,...,m_p),...,(m_1,m_2,...,m_p)\} \subset \mathbb{R}^p.$$

Notice that $Card\partial K = N$ and

$$(x_{1,m_1,i_1},...,x_{p,m_p,i_p}) \in \nu_D, \quad \forall (i_1,...,i_p) \in \partial K. \tag{4}$$

Let $u_{m_1,...,m_p} : D \to \mathbb{R}$ be the function given by

$$u_{m_1,...,m_p}(x_1,...,x_p) := \sum_{(i_1,...,i_p)\in \partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p) \tag{5}$$

and

$$\sigma_{m_1,...,m_p} := \inf \left\{ u_{m_1,...,m_p}(x_1,...,x_p) \,:\, (x_1,...,x_p) \in D \right\}. \tag{6}$$

We define now the operators:

$$L_{m_1,m_2,...,m_p} : C(D) \to C(D)$$

by

$$(L_{m_1,m_2,...,m_p} f)(x_1,x_2,...,x_p) :=$$
$$= \sum_{i_1=0}^{m_1} ... \sum_{i_k=0}^{m_k} ... \sum_{i_p=0}^{m_p} \psi_{1,m_1,i_1}(x_1)...\psi_{k,m_k,i_k}(x_k)...\psi_{p,m_p,i_p}(x_p)\cdot \tag{7}$$
$$\cdot f(x_{1,m_1,i_1},...,x_{k,m_k,i_k},...,x_{p,m_p,i_p})$$

for all $f \in C(D)$, $\forall (x_1,x_2,...,x_p) \in D$.

Proposition 3.1. *The operators $L_{m_1,m_2,...,m_p}$ have the next properties:*
(i) $L_{m_1,m_2,...,m_p}(e_\alpha) = e_\alpha$, for all $\alpha \in \nu_D$;
(ii) $(L_{m_1,m_2,...,m_p} f)(\alpha) = f(\alpha)$, for all $f \in C(D)$, $\forall \alpha \in \nu_D$;
(iii) $L_{m_1,m_2,...,m_p}$ are linear and positive.

Proof: The first statement follows from (1) and (2). The second follows from (1) and (3). The last statement is obvious.□

For all $\Lambda = (\lambda_1,\lambda_2,...,\lambda_N) \in \mathbb{R}^N$, consider the sets

$$X_\Lambda := \left\{ f \in C(D) \,:\, f(\alpha) := \lambda_{\omega(\alpha)},\ \forall \alpha \in \nu_D \right\} \tag{8}$$

Lemma 3.1. *(i) For all* $\Lambda \in \mathbb{R}^N$*, the sets* X_Λ *are closed in* $C(D)$*;*
(ii) $X_\Lambda \in I\left(L_{m_1,m_2,...,m_p}\right)$*;*
(iii) $C(D) = \bigcup_{\Lambda\in\mathbb{R}^N} X_\Lambda$ *is a partition of the space* $C(D)$*.*

The main result is given by the next theorem.

Theorem 3.1. *If* $\sigma_{m_1,...,m_p}$ *given by (6) is non-zero, then the operators* $L_{m_1,m_2,...,m_p}$ *defined by (7) are WPOs and for all* $(m_1, m_2, ..., m_p) \in \mathbb{N}^p$*, we have:*

$$L^{\infty}_{m_1,m_2,...,m_p}(f) = \varphi^*_f, \quad \forall f \in C(D).$$

The function φ^*_f *is defined by*

$$\varphi^*_f(x_1, x_2, ..., x_p) = C^0_0 + \sum_{i_1\in M_1} C^1_{i_1} x_{i_1} + \sum_{(i_1,i_2)\in M_2} C^2_{i_1,i_2} x_{i_1} x_{i_2} + ...+$$

$$+ \sum_{(i_1,i_2,...,i_k)\in M_k} C^k_{i_1,i_2,...,i_k} x_{i_1} x_{i_2} ... x_{i_k} + ... + C^p_{1,2,...,p} x_1 x_2 ... x_p \quad \forall f \in C(D)\ \forall (x_1, x_2, ..., x_p) \in D$$

where C^0_0 *and* $C^k_{i_1,i_2,...,i_k}$*,* $\forall k \in \overline{1,p}$*,* $\forall (i_1, i_2, ..., i_k) \in M_k$ *are real numbers which depend of* f*, given by*

$$C^0_0 := f(\alpha^{\langle 0\rangle});$$

$$C^k_{i_1,i_2,...,i_k} := (-1)^k f(\alpha^{\langle 0\rangle}) + (-1)^{k-1} \sum_{s_1=1}^{k} f(\alpha^{\langle 1\rangle}_{i_{s_1}}) + (-1)^{k-2} \sum_{1\le s_1<s_2\le k} f(\alpha^{\langle 2\rangle}_{i_{s_1},i_{s_2}}) +$$

$$+ ... + (-1)^{k-l} \sum_{1\le s_1<s_2<...<s_l\le k} f(\alpha^{\langle l\rangle}_{i_{s_1},i_{s_2},...,i_{s_l}}) + ... + (-1)^0 f(\alpha^{\langle k\rangle}_{i_1,i_2,...,i_k});$$

for all $k \in \overline{1,p}$*,* $\forall (i_1, i_2, ..., i_k) \in M_k$*.*

Proof: By virtue of Lemma 3.1, the sets X_Λ are closed, $X_\Lambda \in I\left(L_{m_1,m_2,...,m_p}\right)$, and $C(D) = \bigcup_{\Lambda\in\mathbb{R}^N} X_\Lambda$ is a partition of the space $C(D)$.
Denote by $\|\cdot\|_{C(D)}$ the Cebysev norm in $C(D)$, i.e.

$$\|v\|_{C(D)} := \sup_{(x_1,...,x_p)\in D} |v(x_1, ..., x_p)|, \quad \forall v \in C(D).$$

For all $\Lambda \in C(D)$ and for all $f, g \in X_\Lambda$ we have:

$$|(L_{m_1,m_2,...,m_p} f)(x_1, x_2, ..., x_p) - (L_{m_1,m_2,...,m_p} g)(x_1, x_2, ..., x_p)| =$$

$$= \Big| \sum_{i_1=0}^{m_1} ... \sum_{i_k=0}^{m_k} ... \sum_{i_p=0}^{m_p} \psi_{1,m_1,i_1}(x_1) ... \psi_{k,m_k,i_k}(x_k) ... \psi_{p,m_p,i_p}(x_p) \cdot$$

$$\cdot(f-g)(x_{1,m_1,i_1},...,x_{k,m_k,i_k},...,x_{p,m_p,i_p})\Big| =$$

$$= \Big| \sum_{(i_1,...,i_p)\in K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\cdot(f-g)(x_{1,m_1,i_1},...,x_{p,m_p,i_p})\Big| \leq$$

$$\leq \Big| \sum_{(i_1,...,i_p)\in K-\partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\cdot(f-g)(x_{1,m_1,i_1},...,x_{p,m_p,i_p})\Big| +$$

$$+\Big| \sum_{(i_1,...,i_p)\in \partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\cdot(f-g)(x_{1,m_1,i_1},...,x_{p,m_p,i_p})\Big| \overset{(4)}{=}$$

$$= \Big| \sum_{(i_1,...,i_p)\in K-\partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\cdot(f-g)(x_{1,m_1,i_1},...,x_{p,m_p,i_p})\Big| \leq$$

$$\leq \left[\sum_{(i_1,...,i_p)\in K-\partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\right]\cdot \|f-g\|_{C(D)} =$$

$$= \left[\sum_{(i_1,...,i_p)\in K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p) - \sum_{(i_1,...,i_p)\in \partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\right]\cdot$$

$$\cdot\|f-g\|_{C(D)} \overset{(1)}{=} \left[1 - \sum_{(i_1,...,i_p)\in \partial K} \psi_{1,m_1,i_1}(x_1)...\psi_{p,m_p,i_p}(x_p)\right]\cdot \|f-g\|_{C(D)} \overset{(5)}{=}$$

$$= \left[1-u_{m_1,...,m_p}(x_1,...,x_p)\right]\cdot\|f-g\|_{C(D)} \overset{(6)}{\leq} (1-\sigma_{m_1,...,m_p})\cdot\|f-g\|_{C(D)}.$$

Because $\sigma_{m_1,...,m_p}$ in non-zero, the restrictions $L_{m_1,m_2,...,m_p}|_{X_\Lambda}$ are contractions with the same constant $1-\sigma_{m_1,...,m_p}\in[0,1[$. Consequently, they are POs.

It can be proven that for all $\Lambda\in\mathbb{R}^N$, $\varphi_f^*\in X_\Lambda$ $\forall f\in X_\Lambda$. For any $\Lambda\in\mathbb{R}^N$, the restriction $L_{m_1,m_2,...,m_p}|_{X_\Lambda}$ has a unique fixed point which is φ_f^* (it follows from Proposition 3.1).

From Theorem 2.1 it follows that $L_{m_1,m_2,...,m_p}: C(D)\to C(D)$ are WPOs. Besides, for all $f\in C(D)$, the limit operator is φ_f^*. □

Remark 3.3. *In the case* $p=2$ *we have* $D=[0,1]\times[0,1]$, $N=4$,

$$\alpha^{\langle 0\rangle}=(0,0),\quad \alpha_1^{\langle 1\rangle}=(1,0),\quad \alpha_2^{\langle 1\rangle}=(0,1),\quad \alpha_{1,2}^{\langle 2\rangle}=(1,1)$$

and

$$\nu_D=\{(0,0),(1,0),(0,1),(1,1)\}$$

There exists a bijective function $\omega : \nu_D \to \{1, 2, 3, 4\}$.

For all $\Lambda := (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4$ *consider the sets:*

$$X_\Lambda := \{f \in C(D) \,:\, f(0,0) = \lambda_{\omega(0,0)}, f(1,0) = \lambda_{\omega(1,0)}, f(0,1) = \lambda_{\omega(0,1)}, f(1,1) = \lambda_{\omega(1,1)}\}$$

For all $m_1 := m \in \mathbb{N}$, $m_2 := n \in \mathbb{N}$, *the operators* $L_{m,n}$ *are WPOs and*

$$L_{m,n}^{\infty}(f) = \varphi_f^*, \quad \forall f \in C(D)$$

where

$$\varphi_f^*(x,y) = \underbrace{f(\alpha^{\langle 0\rangle})}_{C_0^0} + \left(\underbrace{[f(\alpha_1^{\langle 1\rangle}) - f(\alpha^{\langle 0\rangle})]}_{C_1^1} x + \underbrace{[f(\alpha_2^{\langle 1\rangle}) - f(\alpha^{\langle 0\rangle})]}_{C_2^1} y \right) +$$

$$+ \underbrace{\left(f(\alpha^{\langle 0\rangle}) - [f(\alpha_1^{\langle 1\rangle}) + f(\alpha_2^{\langle 1\rangle})] + f(\alpha_{1,2}^{\langle 2\rangle}) \right)}_{C_{1,2}^2} xy$$

So, we reobtain [1; Remark 1 - Theorem 9] in the particular case $a_1 = a_2 = 0$ *and* $b_1 = b_2 = 1$.

4. Applications

4.1. Bernstein operators of $(m_1, ..., m_p)$ order.

For all $(m_1, ..., m_p) \in \mathbb{N}^p$ consider the next system of points:

$$\Delta_{m_k}^k := \left(0 = \frac{0}{m_k} < \frac{1}{m_k} < ... < \frac{m_k}{m_k} = 1 \right) \quad \forall k = \overline{1, p}.$$

Let the functions $\psi_{k,m_k,i}$ be the fundamental polynomials of Bernstein

$$\psi_{k,m_k,i}(x) := b_{m_k,i}(x) = \begin{pmatrix} m_k \\ i \end{pmatrix} x^i (1-x)^{m_k - i} \quad \forall x \in [0,1]$$

for all $i = \overline{0, m_k}$, $k = \overline{1, p}$.

Then the polynomials $L_{m_1,m_2,...,m_p} : C(D) \to C(D)$ from (7) are the Bernstein polynomials of $(m_1, ..., m_p)$ order, given by

$$(L_{m_1,m_2,...,m_p} f)(x_1, x_2, ..., x_p) :=$$

$$= \sum_{i_1=0}^{m_1} ... \sum_{i_p=0}^{m_p} b_{m_1,i_1}(x_1) ... b_{m_p,i_p}(x_p) \cdot f\left(\frac{i_1}{m_1}, ..., \frac{i_p}{m_p} \right).$$

The next theorem states the convergence of the iterates of the generalized Bernstein operator.

Theorem 4.1. *The Bernstein operators* $L_{m_1,m_2,...,m_p}$ *are WPOs and*

$$L^{\infty}_{m_1,m_2,...,m_p}(f) = \varphi^*_f, \quad \forall f \in C(D)$$

with φ^*_f *as in Theorem 3.1.*

Remark 4.1. *In the particular case* $p = 2$, $m_1 := m$, $m_2 := n$ *we reobtain the estimation*

$$\lambda_{m,n} = \frac{1}{2^{m+n-2}}$$

(see [1; §4.1.])

4.2. **Stancu modified operators of** $(m_1, ..., m_p)$ **order.** For all $(m_1, ..., m_p) \in \mathbb{N}^p$ consider the systems of points: $\Delta^k_{m_k}, k = \overline{1,p}$ as in the previous application.
The functions $\psi_{k,m_k,i}$ are the fundamental polynomials of Stancu:

$$\psi_{k,m_k,i}(x) := w_{m_k,i,\alpha_k}(x) = \frac{\begin{pmatrix} m_k \\ i \end{pmatrix} x^{[i,-\alpha_k]}(1-x)^{[m_k-i,-\alpha_k]}}{1^{[m_k,-\alpha_k]}} \quad \forall x \in [0,1]$$

for all $i = \overline{0, m_k}$, $k = \overline{1,p}$. α_k are real positive numbers.
Then $L_{m_1,m_2,...,m_p}$ from (7) are the Stancu modified polynomials of $(m_1, ..., m_p)$ order, given by

$$(L_{m_1,m_2,...,m_p} f)(x_1, x_2, ..., x_p) := (S^{\langle \alpha_1,\alpha_2,...,\alpha_p\rangle}_{m_1,m_2,...,m_p} f)(x_1, x_2, ..., x_p) =$$

$$= \sum_{i_1=0}^{m_1} ... \sum_{i_p=0}^{m_p} w_{m_1,i_1,\alpha_1}(x_1)...w_{m_p,i_p,\alpha_p}(x_p) \cdot f\left(\frac{i_1}{m_1}, ..., \frac{i_p}{m_p}\right)$$

The next theorem states the convergence of the iterates of the generalized Stancu operators:

Theorem 4.2. *The Stancu operators* $S^{\langle \alpha_1,\alpha_2,...,\alpha_p\rangle}_{m_1,m_2,...,m_p}$ *are WPOs and*

$$\left(S^{\langle \alpha_1,\alpha_2,...,\alpha_p\rangle}_{m_1,m_2,...,m_p}\right)^{\infty}(f) = \varphi^*_f, \quad \forall f \in C(D)$$

where φ^* *is as in Theorem 3.1.*

Remark 4.2. *In the particular case $p = 2$, $m_1 := m$, $m_2 := n$ we obtain:*

$$\lambda_{m,n} \geq \frac{1}{2^{m+n-2} \cdot 1^{[m,-\alpha_1]} \cdot 1^{[n,-\alpha_2]}}$$

which is the estimation given in [1; §4.2.].

References

[1] O. Agratini and I. A. Rus, *Iterates of Some Bivariate Approximation Process Via Contraction Principle*, Nonlinear Analysis Forum, **8**, 2(2003), 159–168.

[2] O. Agratini and I. A. Rus, *Iterates of a class of discrete linear operators via contraction principle*, Comment. Math. Univ. Carolinae, **44**, 3(2003), 555–563.

[3] O. Agratini, Stancu modified operators revisited, Rev. Anal. Numér. Théor. Approx., **31**, 1(2002), 9–16.

[4] I. A. Rus, *Iterates of Stancu operators, via contraction principle*, Studia Univ. Babeş Bolyai (Mathematica), **47**, 4(2000), 101–104.

[5] I. A. Rus,*Iterates of Bernstein operators, via contraction principle*, J. Math. Anal. Appl., **292** (2004), 259–261.

[6] I. A. Rus,*Some application of weakly Picard operators*, Studia Univ. Babeş Bolyai (Mathematica), **48**, 1(2003), 101–107.

[7] I. A. Rus, *Weakly Picard mappings, Comment. Math. Univ. Carolinae*, **34**, 4(1993), 769–773.

[8] I. A. Rus, *Picard operators and applications*, Scientiae Mathematicae Japonicae, **58**, 1(2003), 191–219.

"Samuel Brassai" High School,
Cluj-Napoca, Romania
E-mail address: `Claudia.Bacotiu@clujnapoca.ro`

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

BIERMANN INTERPOLATION WITH HERMITE INFORMATION

MARIUS BIROU

Abstract. If $P_1, P_2, ..., P_r$ and $Q_1, Q_2, ..., Q_r$ are Lagrange univariate projectors which form the chains i.e.

$$P_1 \le P_2 \le \cdots \le P_r, \quad Q_1 \le Q_2 \le \cdots \le Q_r$$

then the Biermann operator is defined by

$$B_r = P_1' Q_r'' \oplus P_2' Q_{r-1}'' \oplus \cdots \oplus P_r' Q_1''$$

where $P_1', \ldots, P_r', Q_1'', \ldots, Q_r''$ are the parametric extension (see [5])

In this paper, we construct projectors of Hermite type which form the chains. The representation of Biermann interpolation projector of Hermite type and the corresponding remainder term are given. Using Hermite information we increase the approximation order. We give some examples.

1. Preliminaries

LetX, Y be the liniar spaces on $\mathbb{R}$ or $\mathbb{C}$.

The liniar operator P defined on space X is called projector if and only if $P^2 = P$.

The operator $P^c = I - P$, where I is identity operator, is called the remainder projector of P.

Received by the editors: 22.09.2005.

2000 *Mathematics Subject Classification.* 41A63, 41A05, 41A80.

Key words and phrases. Biermann interpolation, Hermite interpolation, chains of projectors, triangular elements, approximation order.

If P is projector on space X then the range space of projector P is denoted by

$$\mathcal{R}(P) = \{Pf|\ f \in X\} \tag{1}$$

The set of interpolation points of projector P is denoted by $\mathcal{P}(P)$.

Proposition 1.1. *If P, Q are comutative projectors*

$$\begin{aligned} &1) \quad \mathcal{R}(P \oplus Q) = \mathcal{R}(P) + \mathcal{R}(Q) \\ &2) \quad \mathcal{P}(P \oplus Q) = \mathcal{P}(P) \cup \mathcal{P}(Q). \end{aligned} \tag{2}$$

If P_1 and P_2 are projectors on space X, we define relation "$\leq$"

$$P_1 \leq P_2 \Leftrightarrow P_1 P_2 = P_1$$

Let be $f \in C(X \times Y)$ and $x \in X$. We define $f^x \in C(Y)$ by

$$f^x(t) = f(x,t), \quad t \in Y$$

For $y \in Y$ we define ${}^y f \in C(X)$ by

$${}^y f(s) = f(s,y), \quad s \in X$$

Let P be a liniar and bounded operator on $C(X)$. The parametric extension P' of P is defined by

$$(P'f)(x,y) = (P^y f)(x) \tag{3}$$

If P is a liniar and bounded operator on $C(Y)$ then the parametric extension Q'' of Q is defined by

$$(Q''f)(x,y) = (Qf^x)(y) \tag{4}$$

Proposition 1.2. *Let $r \in \mathbb{N}$, $P_1, \ldots, P_r$ be univariate interpolation projectors on $C(X)$ and $Q_1, \ldots, Q_r$ univariate interpolation projectors on $C(Y)$. Let $P_1', \ldots, P_r', Q_1'', \ldots, Q_r''$ be the corresponding parametric extension. We assume that*

$$P_1 \leq P_2 \leq \cdots \leq P_r, \quad Q_1 \leq Q_2 \leq \cdots \leq Q_r \tag{5}$$

Then

$$B_r = P_1' Q_r'' \oplus P_2' Q_{r-1}'' \oplus \cdots \oplus P_r' Q_1'' \tag{6}$$

is projector and it has representation

$$B_r = \sum_{m=1}^{r} P'_m Q''_{r+1-m} - \sum_{m=1}^{r-1} P'_m Q''_{r-m} \tag{7}$$

Moreover, we have

$$B_r^c = P_r'^c + P_{r-1}'^c Q_1''^c + \cdots + P_1'^c Q_{r-1}''^c + Q_r''^c - (P_r'^c Q_1''^c + \cdots + P_1'^c Q_r''^c) \tag{8}$$

where $P^c = I - P$, I *the identity operator.*

2. Biermann interpolation

In this section we present the Biermann interpolation operator and some of this properties from [5].

Let be the univariate projectors of polynomial interpolation

$$P_1, \ldots, P_r, Q_1, \ldots, Q_r$$

given by the folowing relations

$$(P_m f)(x) = \sum_{i=1}^{k_m} f(x_i)\phi_{i,m}(x)$$

$$(Q_n g)(y) = \sum_{j=1}^{l_n} g(y_j)\psi_{j,n}(y)$$

The sets of interpolation points are

$$\{x_1, \ldots, x_{k_m}\} \subseteq [a,b], \quad \{y_1, \ldots, y_{l_n}\} \subseteq [c,d]$$

with

$$1 \le k_1 < k_2 < \cdots < k_r, \quad 1 \le l_1 < l_2 < \cdots < l_r \tag{9}$$

The cardinal functions are given by

$$\phi_{i,m}(x) = \prod_{\substack{k=1\\k\neq i}}^{k_m} \frac{x - x_k}{x_i - x_k}, \quad 1 \le i \le k_m$$

$$\psi_{j,n}(y) = \prod_{\substack{l=1\\l\neq j}}^{l_n} \frac{y - y_l}{y_j - y_l}, \quad 1 \le j \le l_n$$

Then we have

$$\mathcal{R}(P_m) = \langle 1, x, \ldots, x^{k_m - 1}\rangle = \Pi_{k_m - 1}$$
$$\mathcal{R}(Q_n) = \langle 1, y, \ldots, y^{l_n - 1}\rangle = \Pi_{l_n - 1} \tag{10}$$

From (9) and (10) we have that parametric extensions

$$P'_1, \ldots, P'_r, Q''_1, \ldots, Q''_r$$

are bivariate interpolation projectors which form chains.

$$P'_1 \le \cdots \le P'_r, \quad Q''_1 \le \cdots \le Q''_r \tag{11}$$

Moreover we have

$$P'_m Q''_n = Q''_n P'_m, \quad 1 \le m, n \le r$$

$P'_m Q''_n$ is the tensor product projector of bivariate polynomial interpolation.

We have the representation

$$(P'_m Q''_n f)(x, y) = \sum_{l=1}^{k_m} \sum_{j=1}^{l_n} f(x_i, y_j)\phi_{i,m}(x)\psi_{j,n}(y)$$

$P'_m Q''_n$ has the interpolation properties

$$(P'_m Q''_n f)(x_i, y_j) = f(x_i, y_j), \quad 1 \le i \le k_m,\ 1 \le j \le l_n.$$

The range space defined by $P'_m Q''_n$ is

$$\mathcal{R}(P'_m Q''_n) = \Pi_{k_m - 1} \otimes \Pi_{l_n - 1}.$$

The projectors

$$P'_1, \ldots, P'_r, Q''_1, \ldots, Q''_r$$

generate a distributive lattice ξ of interpolation projectors on $C([a,b] \times [c,d])$. The interpolation projector B_r is defined by relation

$$B_r = P'_1 Q''_r \oplus P'_2 Q''_{r-1} \oplus \cdots \oplus P'_r Q''_1$$

where $r \in \mathbb{N}$. The projectors $B_1, \ldots, B_r$ form a finit chain

$$B_1 \leq B_2 \leq \cdots \leq B_r$$

Proposition 2.1. *The range space of Biermann interpolation projector is*

$$\mathcal{R}(B_r) = \Pi_{k_1-1} \otimes \Pi_{l_r-1} + \cdots + \Pi_{k_r-1} \otimes \Pi_{l_1-1} \tag{12}$$

Proposition 2.2. *The Biermann interpolation projector satisfies the interpolation properties*

$$(B_r f)(x_i, y_j) = f(x_i, y_j), \quad 1 \leq i \leq k_m,\ 1 \leq j \leq l_{r+1-m},\ 1 \leq m \leq r \tag{13}$$

The set of interpolation points possesses a disjoint representation

$$P(B_r) = \bigcup_{m=1}^{r} \bigcup_{n=0}^{r-m} \{(x_i, y_j):\ k_{m-1} < l \leq k_m,\ l_{r-m-n} < j \leq l_{r+1-m-n}\} \tag{14}$$

with $k_0 := 0,\ l_0 := 0$.

Using disjoint representation (14) of interpolation set $\mathcal{P}(B_r)$ of Biermann interpolation projector we obtain Lagrange representation of Biermann interpolant

$$B_r f = \sum_{m=1}^{r} \sum_{n=0}^{r-m} \sum_{i=1+k_{m-1}}^{k_m} \sum_{j=1+l_{r-m-n}}^{l_{r+1-m-n}} f(x_i, y_j) \Phi_{ij} \tag{15}$$

Proposition 2.3. *The cardinal function of Biermann interpolation is given by*

$$\Phi_{i,j}(x,y) = \sum_{s=m}^{m+n} \phi_{i,s}(x) \psi_{j,r+1-s}(y) - \sum_{s=m}^{m+n-1} \phi_{i,s}(x) \psi_{j,r-s}(y), \tag{16}$$

$$k_{m-1} < i \leq k_m,\ l_{r-m-n} < j \leq l_{r+1-m-n},\ 0 \leq n \leq r-m,\ 1 \leq m \leq r$$

Proposition 2.4. *The Cauchy form of remainder formula in bivariate Biermann interpolation is*

$$\begin{aligned} & f(x,y) - (B_r f)(x,y) \qquad (17) \\ &= (x-x_1)\ldots(x-x_{k_r}) \frac{f^{(k_r,0)}(\xi_r, y)}{k_r!} + (y-y_1)\ldots(y-y_{l_r}) \frac{f^{(0,l_r)}(x, \eta_r)}{l_r!} \\ &+ \sum_{m=1}^{r-1} \prod_{l=1}^{k_{r-m}} (x-x_i) \prod_{j=1}^{l_m} (y-y_j) \frac{f^{(k_{r-m}, l_m)}(\xi_{r-m}, \eta_m)}{k_{r-m}! l_m!} \end{aligned}$$

$$-\sum_{m=1}^{r}\prod_{l=1}^{k_{r+1-m}}(x-x_i)\prod_{j=1}^{l_m}(y-y_j)\frac{f^{(k_{r+1-m},l_m)}(\sigma_{r+1-m},\tau_m)}{k_{r+1-m}!l_m!}$$

where $\xi_i,\sigma_i\in[a,b]$, $\eta_i,\tau_i\in[c,d]$ *with* $1\le i\le r$.

Proposition 2.5. *Let be* $q=\min\{k_{r-m}+l_m:\ 0\le m\le r\}$ *with* $k_0=0$, $l_0=0$. *Then* q *is the approximation order in bivariate Biermann interpolation, i.e.*

$$f(x,y)-(B_rf)(x,y)=O(h^q),\quad h\to 0 \tag{18}$$

Example Let be r=3 and triangular elements

$$x_i=\frac{ih}{3},y_j=\frac{jh}{3},1\le i,j\le 3,h>0$$

$$k_m=m,l_n=n,1\le m,n\le 3$$

.

The cardinal functions are

$$\Phi_{13}(x,y)=\phi_{11}(x)\psi_{33}(y)$$

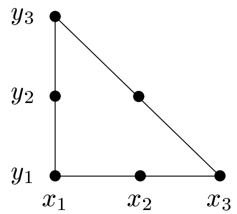
$$\Phi_{22}(x,y)=\phi_{22}(x)\psi_{22}(y)$$

$$\Phi_{31}(x,y)=\phi_{33}(x)\psi_{11}(y)$$

$$\Phi_{12}(x,y)=\phi_{11}(x)\psi_{23}(y)+\phi_{12}(x)\psi_{22}(y)-\phi_{11}(x)\psi_{22}(y)$$

$$\Phi_{21}(x,y)=\phi_{22}(x)\psi_{12}(y)+\phi_{23}(x)\psi_{11}(y)-\phi_{22}(x)\psi_{11}(y)$$

$$\Phi_{11}(x,y)=\phi_{11}(x)\psi_{13}(y)+\phi_{12}(x)\psi_{12}(y)+\phi_{13}(x)\psi_{11}(y)-\phi_{11}(x)\psi_{12}(y)-\phi_{12}(x)\psi_{11}(y)$$



The order of aproximation is 3.

3. Main result

Our goal is to construct Hermite projectors which form the chains and with their aid the Biermann operator of Hermite type.

Let be the univariate projectors of Hermite interpolation

$$P_1, \ldots, P_r, Q_1, \ldots, Q_r$$

given by relations

$$(P_m f)(x) = \sum_{i=1}^{k_m} \sum_{p=0}^{u_{i,m}} f^{(p)}(x_i) h_{ip}^m(x), \quad 1 \le m \le r$$

$$(Q_n g)(y) = \sum_{j=1}^{l_n} \sum_{q=0}^{v_{j,n}} g^{(q)}(y_j) \widetilde{h}_{jq}^n(y), \quad 1 \le n \le r$$

Assume that

$$\{x_1, \ldots, x_{k_m}\} \subseteq [a, b]$$

$$\{y_1, \ldots, y_{l_n}\} \subseteq [c, d]$$

with

$$\begin{gathered} 1 \le k_1 < k_2 < \cdots < k_r \\ 1 \le l_1 < l_2 < \cdots < l_r \end{gathered} \tag{19}$$

and

$$\begin{gathered} u_{i,m} \le u_{i,m+1}, \ i = \overline{1, k_m}, \ m = \overline{1, r-1} \\ v_{j,n} \le v_{j,n+1}, \ i = \overline{1, l_n}, \ n = \overline{1, r-1} \end{gathered} \tag{20}$$

The cardinal functions h_{ip}^m, $m = \overline{1, r}$ and $\widetilde{h}_{jq}^n$, $n = \overline{1, r}$ satisfy

$$\begin{cases} h_{ip}^{m\ (j)}(x_\nu) = 0, \ \nu \ne i, \ j = \overline{0, u_{\nu,m}} \\ h_{ip}^{m\ (j)}(x_i) = \delta_{jp}, \ j = \overline{0, u_{i,m}} \end{cases}$$

for $p = \overline{0, u_{i,m}}$, $\nu, i = \overline{1, k_m}$ and respective

$$\begin{cases} \widetilde{h}_{jq}^{n\ (i)}(y_\nu) = 0, \ \nu \ne j, \ i = \overline{0, v_{\nu,n}} \\ \widetilde{h}_{jq}^{n\ (i)}(y_j) = \delta_{iq}, \ i = \overline{0, v_{j,n}} \end{cases}$$

for $q = \overline{0, v_{j,n}}$, $\nu, j = \overline{1, l_n}$.

Theorem. 1. *The parametric extensions*

$$P'_1, \dots, P'_r, Q''_1, \dots, Q''_r$$

are bivariate interpolation projectors which form the chains

$$P'_1 \leq \dots \leq P'_r, \quad Q''_1 \leq \dots \leq Q''_r \tag{21}$$

Proof. Let be $1 \leq m_1 \leq m_2 \leq r$. Then

$$k_{m_1} \leq k_{m_2},$$

$$u_{i,m_1} \leq u_{i,m_2}, i \leq k_{m_1} \tag{22}$$

We have that

$$(P'_{m_1} P'_{m_2} f)(x, y) = \sum_{i_1=1}^{k_{m_1}} \sum_{p_1=0}^{u_{i_1 m_1}} \left(\sum_{i_2=1}^{k_{m_2}} \sum_{p_2=0}^{u_{i_2 m_2}} f^{(p_2,0)}(x_{i_2}, y) h_{i_2 p_2}^{m_2\ (p_1)}(x_{i_1}) \right) h_{i_1 p_1}^{m_1}(x) \tag{23}$$

But

$$h_{i_2 p_2}^{m_2\ (p_1)}(x_{i_1}) = \delta_{i_1 i_2} \delta_{p_1 p_2} \tag{24}$$

From (22), (23) and (24) we have

$$(P'_{m_1} P'_{m_2} f)(x, y) = \sum_{i_1=1}^{k_{m_1}} \sum_{p_1=0}^{u_{i_1 m_1}} f^{(p_1,0)}(x_{i_1}, y) h_{i_1 p_1}^{m_1}(x) = (P'_{m_1} f)(x, y)$$

i.e.

$$P'_{m_1} \leq P'_{m_2}$$

Thus $P'_1, P'_2, ..., P'_r$ form a chain. Analogous $Q''_1, Q''_2, ..., Q''_r$ are projectors which form a chain. □

Moreover we have

$$P'_m Q''_n = Q''_n P'_m, \quad 1 \leq m, n \leq r$$

The tensor product projector $P'_m Q''_n$ of bivariate interpolation has representation

$$(P'_m Q''_n f)(x, y) = \sum_{i=1}^{k_m} \sum_{p=0}^{u_{i,m}} \sum_{j=1}^{l_n} \sum_{q=0}^{v_{j,n}} f^{(p,q)}(x_i, y_j) h_{ip}^m(x) \widetilde{h}_{jq}^n(y)$$

and it has the interpolation properties

$$(P'_m Q''_n f)^{(p,q)}(x_i, y_j) = f^{(p,q)}(x_i, y_j)$$

$$1 \le i \le k_m, \quad 1 \le j \le l_n, \quad 0 \le p \le u_{i,m}, \quad 0 \le q \le v_{j,n}$$

The projectors $P'_1, \ldots, P'_r, Q''_1, \ldots, Q''_r$ generate a distributive lattice ξ of projectors on $C([a,b] \times [c,d])$. A special element in this lattice is

$$B_r^H = P'_1 Q''_r \oplus P'_2 Q''_{r-1} \oplus \cdots \oplus P'_r Q''_1, \quad r \in \mathbb{N} \tag{25}$$

called Biermann projector of Hermite type

Let be $\alpha_i = u_{1,i} + \cdots + u_{k_i,i} + k_i, \quad \beta_i = v_{1,i} + \cdots + v_{l_i,i} + l_i, \quad 1 \le i \le r.$

Proposition 3.1. *The range space of projector B_r^H is given by*

$$\mathcal{R}(B_r^H) = \Pi_{\alpha_1 - 1} \otimes \Pi_{\beta_r - 1} + \cdots + \Pi_{\alpha_r - 1} \otimes \Pi_{\beta_1 - 1} \tag{26}$$

Proof. Using Proposition 1.1. we have

$$\mathcal{R}(B_r^H) = \mathcal{R}(P'_1 Q''_r) \cup \ldots \cup \mathcal{R}(P'_1 Q''_r)$$

Taking into account that

$$\mathcal{R}(P_m) = \Pi_{\alpha_m - 1}, 1 \le m \le r$$

$$\mathcal{R}(Q_n) = \Pi_{\beta_n - 1}, 1 \le m \le r$$

we get (26).□

Proposition 3.2. *Biermann projector B_r^H has the interpolation properties*

$$(B_r^H f)^{(p,q)}(x_i, y_j) = f^{(p,q)}(x_i, y_j) \tag{27}$$

$$1 \le i \le k_m, \ 1 \le j \le l_{r+1-m}, 1 \le m \le r,$$

$$u_{i,m-1} < p \le u_{i,m}, \quad 0 \le q \le v_{j,r-m+1},$$

where $u_{i,m-1} = -1$, $k_{m-1} < i \le k_m$, $1 \le m \le r$, with $k_0 = 0$

Proof. Using Proposition 1.1. we have

$$\mathcal{P}(B_r^H) = \mathcal{P}(P_1'Q_r'') \cup ... \cup \mathcal{P}(P_1'Q_r'')$$

If we denote

$$\mathcal{I}(P) = \{f^{(p,q)}(x_i.y_j) | (Pf^{(p,q)})(x_i.y_j) = f^{(p,q)}(x_i.y_j)\}$$

then

$$\mathcal{I}(B_r^H) = \mathcal{I}(P_1'Q_r'') \cup ... \cup \mathcal{I}(P_1'Q_r'')$$

We have

$$\begin{aligned}\mathcal{I}(P_1'Q_r'') &= \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_1},\, j = \overline{1,l_r},\, p = \overline{0,u_{i1}}, q = \overline{0,v_{jr}}\ \} \\ &= \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_1},\, j = \overline{1,l_r},\, p = \overline{u_{i0}+1,u_{i1}}, q = \overline{0,v_{jr}}\ \} \\ &= C_1\end{aligned}$$

with $u_{i0} = -1,\, i = \overline{k_0+1,k_1}$

For $m = \overline{2,r}$ we have

$$\begin{aligned}&\mathcal{I}(P_m'Q_{r+1-m}'') = \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_m},\, j = \overline{1,l_{r+1-m}},\, p = \overline{0,u_{im}}, q = \overline{0,v_{j,r+1-m}}\} \\ &= \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_{m-1}},\, j = \overline{1,l_{r+1-m}},\, p = \overline{0,u_{i,m-1}}, q = \overline{0,v_{j,r+1-m}}\} \\ &\cup \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_{m-1}},\, j = \overline{1,l_{r+1-m}},\, p = \overline{u_{i,m-1}+1,u_{i,m}}, q = \overline{0,v_{j,r+1-m}}\} \\ &\cup \{f^{(p,q)}(x_i.y_j) | i = \overline{k_{m-1}+1,k_m},\, j = \overline{1,l_{r+1-m}},\, p = \overline{0,u_{i,m}}, q = \overline{0,v_{j,r+1-m}}\} \\ &= \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_{m-1}},\, j = \overline{1,l_{r+1-m}},\, p = \overline{0,u_{i,m-1}}, q = \overline{0,v_{j,r+1-m}}\} \\ &\cup \{f^{(p,q)}(x_i.y_j) | i = \overline{1,k_m},\, j = \overline{1,l_{r+1-m}},\, p = \overline{u_{i,m-1}+1,u_{i,m}}, q = \overline{0,v_{j,r+1-m}}\} \\ &= A_m \cup C_m\end{aligned}$$

with $u_{im} = -1,\, i = \overline{k_{m-1}+1,k_m}$.

As $A_m \subset \mathcal{I}(P_{m-1}'Q_{r+2-m}'')$, $m = \overline{2,r}$ it folows that

$$\mathcal{I}(B_r^H) = \mathcal{I}(P_1'Q_r'') \cup ... \cup \mathcal{I}(P_1'Q_r'') = C_1 \cup ... \cup C_r$$

q.e.d.□

Remark. 2. *The sets $C_i, i = \overline{1,r}$ are dizjoncts.*

From (7) we get the folowing representation for the projector B_r^H

$$(B_r^H f)(x,y) = \sum_{m=1}^{r}\sum_{i=1}^{k_m}\sum_{p=0}^{u_{i,m}}\sum_{j=1}^{l_{r+1-m}}\sum_{q=0}^{v_{j,r+1-m}} h_{ip}^m(x)\widetilde{h}_{jq}^{r+1-m}(y) f^{(p,q)}(x_i,y_j) \tag{28}$$

$$-\sum_{m=1}^{r-1}\sum_{i=1}^{k_m}\sum_{p=0}^{u_{i,m}}\sum_{j=1}^{l_{r-m}}\sum_{q=0}^{v_{j,r-m}} h_{ip}^m(x)\widetilde{h}_{jq}^{r-m}(y) f^{(p,q)}(x_i,y_j)$$

Taking into account (27) we obtain

$$(B_r^H f)(x,y) = \sum_{m=1}^{r}\sum_{i=1}^{k_m}\sum_{p=u_{i,m-1}+1}^{u_{im}}\sum_{j=1}^{l_{r-m+1}}\sum_{q=0}^{v_{j,r-m+1}} f^{(p,q)}(x_i,y_i)\Phi_{ij}^{pq}(x,y) \tag{29}$$

Proposition 3.3. *The cardinal functions* Φ_{ij}^{pq} *given by*

$$\Phi_{ij}^{pq}(x,y) = \sum_{m\in A_{ij}^{pq}} h_{ip}^m(x)\widetilde{h}_{jq}^{r-m+1}(y) - \sum_{m\in B_{ij}^{pq}} h_{ip}^m(x)\widetilde{h}_{jq}^{r-m}(y) \tag{30}$$

$$1\le i\le k_m,\ \ 1\le j\le l_{r+1-m}, 1\le m\le r,$$

$$u_{i,m-1} < p \le u_{i,m},\quad 0\le q\le v_{j,r-m+1},$$

where

$A_{ij}^{pq} = \{m\in\{1,...,r\} | i\in X_m, p\le u_{i,m}, j\in Y_{r+1-m}, q\le v_{j,r+1-m}\}$
$B_{ij}^{pq} = \{m\in\{1,...,r-1\} | i\in X_m, p\le u_{i,m}, j\in Y_{r-m}, q\le v_{j,r-m}\}$
$X_m = \{1,...,k_m\}, Y_n = \{1,...,l_n\}$

Proof. For the function

$$f(x,y) = h_{ip}^r(x)\widetilde{h}_{jq}^r(y)$$

we have

$$B_r f = \Phi_{ij}^{pq}$$

Taking into accoumt relation (7) we get

$$\begin{aligned}\Phi_{ij}^{pq} &= \sum_{m=1}^{r} P'_m(h_{ip}^r)\otimes Q''_{r+1-m}(\widetilde{h}_{jq}^r) - \sum_{m=1}^{r-1} P'_m(h_{ip}^r)\otimes Q''_{r-m}(\widetilde{h}_{jq}^r)\\
&= \sum_{\substack{m\in\{1,\dots,r\}\\ i\in X_m,\, p\geq u_{i,m}}} h_{ip}^m \otimes Q''_{r+1-m}(\widetilde{h}_{jq}^r) - \sum_{\substack{m\in\{1,\dots,r-1\}\\ i\in X_m,\, p\geq u_{i,m}}} h_{ip}^m \otimes Q''_{r-m}(\widetilde{h}_{jq}^r)\\
&= \sum_{\substack{m\in\{1,\dots,r\}\\ i\in X_m,\, p\geq u_{i,m}\\ j\in Y_{r+1-m},\, q\geq v_{j,r+1-m}}} h_{ip}^m \otimes \widetilde{h}_{jq}^{r+1-m} - \sum_{\substack{m\in\{1,\dots,r\}\\ i\in X_m,\, p\geq u_{i,m}\\ j\in Y_{r-m},\, q\geq v_{j,r-m}}} h_{ip}^m \otimes \widetilde{h}_{jq}^{r-m}\end{aligned}$$

Proposition 3.4. *If $f\in C^{k_r,l_r}([a,b]\times[c,d])$ we have Cauchy form of remainder*

$$f(x,y)-(B_r^H f)(x,y) \tag{31}$$

$$\begin{aligned}&= (x-x_1)^{u_{1,r}+1}\dots(x-x_{k_r})^{u_{k_r,r}+1}\frac{f^{(\alpha_r,0)}(\xi_r,y)}{\alpha_r!}\\
&+(y-y_1)^{v_{1,r}+1}\dots(y-y_{l-r})^{v_{l_r,r}+1}\frac{f^{(0,\beta_r)}(x,\eta_r)}{\beta_r!}\\
&+\sum_{m=1}^{r-1}\prod_{i=1}^{k_{r-m}}(x-x_i)^{u_{i,r-m}+1}\prod_{j=1}^{l_m}(y-y_j)^{v_{j,m}+1}\frac{f^{(\alpha_{r-m},\beta_m)}(\xi_{r-m},\eta_m)}{\alpha_{r-m}!\beta_m!}\\
&-\sum_{m=1}^{r}\prod_{i=1}^{k_{r+1-m}}(x-x_i)^{u_{i,r+1-m}+1}\prod_{j=1}^{l_m}(y-y_j)^{v_{j,m}+1}\frac{f^{(\alpha_{r+1-m},\beta_m)}(\sigma_{r+1-m},\tau_m)}{\alpha_{r+1-m}!\beta_m}\end{aligned}$$

$$\xi_i,\sigma_i\in[a,b],\quad \eta_i,\tau_i\in[c,d], 1\leq i\leq r$$

Proof. We have that

$$(P_m^c f)(x)=f(x)-(P_m f)(x)=\prod_{i=1}^{k_m}(x-x_i)^{u_{i,m}+1}\frac{f^{(\alpha_m)}(\xi_m)}{\alpha_m!},\xi_m\in[a,b], f\in C^{k_m}([a,b])$$

$$(Q_n^c g)(y)=g(y)-(Q_m f)(y)=\prod_{j=1}^{l_n}(y-y_j)^{v_{j,n}+1}\frac{g^{(\beta_n)}(\eta_n)}{\beta_n!},\eta_n\in[c,d], g\in C^{l_n}([c,d])$$

$$1\leq m,n\leq r$$

Taking into account relation (8) we get (31).□

Let be $h=b-a=d-c$ and $q=\min\{\alpha_{r-m}+\beta_m,\ 0\leq m\leq r\}$ with $\alpha_0=0$, $\beta_0=0$. Then we have

$$f(x,y)-(B_r^H f)(x,y)=O(h^q),\quad h\to 0. \tag{32}$$

Example. We determine the order of approximation of the Biermann interpolation projector B_r^H for triangular elements.We choice r=3 and

$$x_i = \frac{ih}{3}, y_j = \frac{jh}{3}, 1 \leq i, j \leq 3, h > 0$$

$$k_m = m, l_n = n, 1 \leq m, n \leq 3$$

Let be the Hermite interpolation projectors

$$P_1 = H^x \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \ P_2 = H^x \begin{pmatrix} x_1 & x_2 \\ 1 & 0 \end{pmatrix}, \ P_3 = H^x \begin{pmatrix} x_1 & x_2 & x_3 \\ 1 & 0 & 1 \end{pmatrix}$$

$$Q_1 = H^y \begin{pmatrix} y_1 \\ 1 \end{pmatrix}, \ Q_2 = H^y \begin{pmatrix} y_1 & y_2 \\ 1 & 1 \end{pmatrix}, \ Q_3 = H^y \begin{pmatrix} y_1 & y_2 & y_3 \\ 1 & 1 & 0 \end{pmatrix}$$

i.e.

$u_{11} = 0;$

$u_{12} = 1;\ u_{22} = 0;$

$u_{13} = 1;\ u_{23} = 0;\ u_{33} = 1;$

$v_{11} = 1;$

$v_{12} = 1;\ v_{22} = 1;$

$v_{13} = 1;\ v_{23} = 1;\ v_{33} = 0;$

It folows that parametric extension form the chains

$$P_1' \leq P_2' \leq P_3' \quad Q_1'' \leq Q_2'' \leq Q_3''$$

and we can define the Biermann operator of Hermite type

$$B_3^H = P_1'Q_3'' \oplus P_2'Q_2'' \oplus P_3'Q_2''.$$

The operator B_3^H has the interpolation properties

$(B_3^H f)^{(p,q)}(x_1, y_1) = f^{(p,q)}(x_1, y_1), (p, q) \in \{(0,0), (0,1), (1,0), (1,1)\}$

$(B_3^H f)^{(p,q)}(x_1, y_2) = f^{(p,q)}(x_1, y_2), (p, q) \in \{(0,0), (0,1), (1,0), (1,1)\}$

$(B_3^H f)^{(p,q)}(x_1, y_3) = f^{(p,q)}(x_1, y_3), (p, q) = (0,0)$

$(B_3^H f)^{(p,q)}(x_2, y_1) = f^{(p,q)}(x_2, y_1), (p, q) \in \{(0,0), (0,1)\}$

$(B_3^H f)^{(p,q)}(x_2, y_2) = f^{(p,q)}(x_2, y_2), (p, q) \in \{(0,0), (0,1)\}$

$(B_3^H f)^{(p,q)}(x_3, y_1) = f^{(p,q)}(x_3, y_1), (p, q) \in \{(0,0), (0,1), (1,0), (1,1)\}$

The cardinal functions of B_3^H are

$$\Phi_{11}^{pq}(x,y) = h_{1p}^1(x)\widetilde{h}_{1q}^3(y) + h_{1p}^2(x)\widetilde{h}_{1q}^2(y) + h_{1p}^3(x)\widetilde{h}_{1q}^1(y) - h_{1p}^1(x)\widetilde{h}_{1q}^2(y) - h_{1p}^2(x)\widetilde{h}_{1q}^1(y),\ (p,q) \in \{(0,0),(0,1)\}$$

$$\Phi_{11}^{pq}(x,y) = h_{1p}^2(x)\widetilde{h}_{1q}^2(y) + h_{1p}^3(x)\widetilde{h}_{1q}^1(y) - h_{1p}^2(x)\widetilde{h}_{1q}^1(y),\ (p,q) \in \{(1,0),(1,1)\}$$

$$\Phi_{12}^{pq}(x,y) = h_{1p}^1(x)\widetilde{h}_{2q}^3(y) + h_{1p}^2(x)\widetilde{h}_{2q}^2(y) - h_{1p}^1(x)\widetilde{h}_{2q}^2(y),\ (p,q) \in \{(0,0),(0,1)\}$$

$$\Phi_{12}^{pq}(x,y) = h_{1p}^2(x)\widetilde{h}_{2q}^2(y), (p,q) \in \{(1,0),(1,1)\}$$

$$\Phi_{13}^{pq}(x,y) = h_{1p}^1(x)\widetilde{h}_{3q}^3(y),\ (p,q) = (0,0)$$

$$\Phi_{21}^{pq}(x,y) = h_{2p}^2(x)\widetilde{h}_{1q}^2(y) + h_{2p}^3(x)\widetilde{h}_{1q}^1(y) - h_{2p}^2(x)\widetilde{h}_{1q}^1(y),\ (p,q) \in \{(0,0),(0,1)\}$$

$$\Phi_{22}^{pq}(x,y) = h_{2p}^2(x)\widetilde{h}_{2q}^2(y), (p,q) \in \{(0,0),(0,1)\}$$

$$\Phi_{31}^{pq}(x,y) = h_{3p}^3(x)\widetilde{h}_{1q}^1(y),\ (p,q) \in \{(0,0),(0,1),(1,0),(1,1)\}$$

We compute the approximation order

$\alpha_1 = u_{11} + 1 = 1;$

$\alpha_2 = u_{12} + u_{22} + 2 = 3;$

$\alpha_3 = u_{13} + u_{23} + u_{33} + 3 = 5;$

$\beta_1 = v_{11} + 1 = 2;$

$\beta_2 = v_{12} + v_{22} + 2 = 4;$

$\beta_3 = v_{13} + v_{23} + v_{33} + 3 = 5;$

$q = \min\{\alpha_3,\ \alpha_2 + \beta_1,\ \alpha_1 + \beta_2,\ \beta_3\} = 5;$

The order of approximation in this case is 5.

References

[1] Bărbosu, A.D., *Aproximarea functiilor de mai multe variabile prin sume booleene de operatori liniari de tip interpolator*, Ed Risoprint, Cluj-Napoca, 2002

[2] Biermann, O., Uber naherungsweise Cubaturen, Monatshefte fur Mathematik und Physic **14**(1903), 211-225.

[3] Coman, Gh. *Analiză numerică*, Ed libris, Cluj, 1995

[4] Coman, Gh., *Multivariate approximation schemes and the approximation of linear functionals*, Rev. d'Analyse Numerique et la Theory de l'Approximation, Mathemathica, **16** (1974), 229-249.

[5] Delvos, F.-J.; Schempp, W. *Boolean methods in interpolation and approximation.* Pitman Research Notes in Mathematics, Series 230, New York 1989.

[6] Delvos, F.-J.; Posdorf, H. *Generalized Biermann interpolation.* Resultate Math. 5 (1982), no. 1, 6–18.

[7] Gordon, William J. *Distributive lattices and the approximation of multivariate functions.* 1969 Approximations with Special Emphasis on Spline Functions (Proc. Sympos. Univ. of Wisconsin, Madison, Wis., 1969) pp. 223–277 Academic Press, New York

[8] Gordon, William J.; Hall, Charles A. *Transfinite element methods: blending-function interpolation over arbitrary curved element domains.* Numer. Math. **21** (1973/74), 109–129.

[9] Stancu, D. D. *The remainder of certain linear approximation formulas in two variables.* J. SIAM Ser. B Numer. Anal. **1** (1964), 137–163.

Babeş-Bolyai University, Faculty of Mathematics
and Computer Science, Str. Kogălniceanu, 1
Cluj-Napoca, Romania
E-mail address: mbirou@math.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

TREE WAYS OF DEFINING THE BIVARIATE SHEPARD OPERATOR OF LIDSTONE TYPE

TEODORA CĂTINAŞ

Abstract. In this paper they are given three possible definitions of the bivariate Shepard operator of Lidstone type. Also, they are given error estimations for the corresponding interpolation formulas.

1. First variant of the Shepard operator of Lidstone type

Let f be a real-valued function defined on $X \subset \mathbb{R}^2$, $(x_i, y_i) \in X,\ i = 0, ..., N$ some distinct points and $r_i(x, y)$, the distances between a given point $(x, y) \in X$ and the points $(x_i, y_i),\ i = 0, 1, ..., N$.

First, we consider the original bivariate operator introduced by Shepard in 1968. This operator is defined by:

$$(S_{N,\mu} f)(x, y) = \sum_{i=0}^{N} A_i(x, y) f(x_i, y_i), \tag{1}$$

where

$$A_i(x, y) = \frac{\prod\limits_{\substack{j=0 \\ j \neq i}}^{N} r_j^{\mu}(x, y)}{\sum\limits_{k=0}^{N} \prod\limits_{\substack{j=0 \\ j \neq k}}^{N} r_j^{\mu}(x, y)}, \tag{2}$$

with $\mu \in \mathbb{R}_+$.

The functions $A_i,\ i = 1, \ldots, N$ have the cardinality properties:

$$A_i(x_\nu, y_\nu) = \delta_{i\nu}, \quad i, \nu = 1, \ldots, N,$$

Received by the editors: 20.09.2005.

2000 *Mathematics Subject Classification.* 41A05, 41A80, 41A25.

Key words and phrases. Shepard operator, Lidstone operator, interpolation, error estimation.

and

$$\sum_{i=0}^{N} A_i\left(x,y\right)=1. \tag{3}$$

The main properties of $S_{N,\mu}$ are:

1. the interpolation property:

$$\left(S_{N,\mu}f\right)\left(x_i,y_i\right)=f\left(x_i,y_i\right), \quad i=0,1,\ldots,N$$

2. the degree of exactness is:

$$\operatorname{dex}\left(S_{N,\mu}\right)=0.$$

Consider $a,b,c,d\in\mathbb{R}$, $a<b$ and $c<d$ and let $\Delta : a=x_0<x_1<\ldots<x_{M+1}=b$ and $\Delta' : c=y_0<y_1<\ldots<y_{N+1}=d$ denote uniform partitions of $[a,b]$ and $[c,d]$ with stepsizes $h=(b-a)/(M+1)$ and $l=(d-c)/(N+1)$, respectively. Further, let $\rho=\Delta\times\Delta'$ be a rectangular partition of $[a,b]\times[c,d]$.

In [4] it was introduced the bivariate Shepard operator of Lidstone type, using the classical definition of the Shepard operator (1).

For a function $f\in C^{2m-2}[a,b]$, according to [1], the Lidstone interpolant uniquely exists and it is of the form

$$(L_m^{\Delta}f)(x)=\sum_{i=0}^{M+1}\sum_{\mu=0}^{m-1} r_{m,i,\mu}(x)f^{(2\mu)}(x_i), \tag{4}$$

where $r_{m,i,j}$, $0\le i\le M+1$, $0\le j\le m-1$ are satisfying

$$D^{2\upsilon}r_{m,i,j}(x_\mu)=\delta_{i\mu}\delta_{2\upsilon,j}, \quad 0\le\mu\le N+1, 0\le\upsilon\le m-1. \tag{5}$$

On the subinterval $[x_i,x_{i+1}]$, $0\le i\le M$, the polynomial $L_m^{\Delta}f$ can be explicitly expressed as

$$\begin{aligned}(L_m^{\Delta,i}f)(x) &:= (L_m^{\Delta}f)|_{[x_i,x_{i+1}]}(x)= \\ &=\sum_{k=0}^{m-1}\left[\Lambda_k\left(\tfrac{x_{i+1}-x}{h}\right)f^{(2k)}(x_i)+\Lambda_k\left(\tfrac{x-x_i}{h}\right)f^{(2k)}(x_{i+1})\right]h^{2k},\end{aligned} \tag{6}$$

where Λ_k is the Lidstone polynomial of degree $2k+1$, $k\in\mathbb{N}$. In analogous way it is obtained the expression of $L_m^{\Delta',i}f$, corresponding to Δ'.

For a function $f \in C^{2m-2,2m-2}([a,b]\times[c,d])$, the bivariate Lidstone interpolant $L_m^\rho f$ uniquely exists and can be explicitly expressed as

$$(L_m^\rho f)(x,y) = \sum_{i=0}^{M+1}\sum_{\mu=0}^{m-1}\sum_{j=0}^{N+1}\sum_{\nu=0}^{m-1} r_{m,i,\mu}(x) r_{m,j,\nu}(y) f^{(2\mu,2\nu)}(x_i,y_j), \tag{7}$$

with $r_{m,i,j}$, $0 \le i \le M+1$, $0 \le j \le m-1$ given by (5).

Lemma 1.1. [1] *If $f \in C^{2m-2,2m-2}([a,b]\times[c,d])$ then*

$$(L_m^\rho f)(x,y) = (L_m^\Delta L_m^{\Delta'} f)(x,y) = (L_m^{\Delta'} L_m^\Delta f)(x,y).$$

Corollary 1.1. [1] *For a function $f \in C^{2m-2,2m-2}([a,b]\times[c,d])$, from Lemma 1.1, we have that*

$$\begin{aligned} f - L_m^\rho f &= (f - L_m^\Delta f) + L_m^\Delta(f - L_m^{\Delta'} f) \\ &= (f - L_m^\Delta f) + [L_m^\Delta(f - L_m^{\Delta'} f) - (f - L_m^{\Delta'} f)] + (f - L_m^{\Delta'} f). \end{aligned} \tag{8}$$

We recall that the $k-$th modulus of smoothness of $f \in L_p[a,b]$, $0 < p < \infty$, or of $f \in C[a,b]$, if $p = \infty$, is defined by (see, e.g., [11]):

$$\omega_k(f;t)_p = \sup_{0<h\le t} \left\| \Delta_h^k f(x) \right\|_p,$$

where

$$\Delta_h^k f(x) = \sum_{i=0}^{k} (-1)^{k+i} \binom{k}{i} f(x+ih).$$

In what follows $\|\cdot\|$ detones the uniform norm over the corresponding interval.

We have some error bound for the bivariate Lidstone interpolation, that is useful in what follows. It is obtained based on some results from [5].

Theorem 1.2. *If $f \in C^{2m-2,2m-2}([a,b]\times[c,d])$ then*

$$\begin{aligned} \|f - L_m^\rho f\| \le & \left(1 + \left\|L_m^\Delta\right\|\right) W_{2m} \max_{y\in[c,d]} \omega_{2m}\left(f(\cdot,y); \tfrac{b-a}{2m}\right) \\ & + \left(1 + \left\|L_m^\Delta\right\|\right) W_{2m} \max_{y\in[c,d]} \omega_{2m}\left((f - L_m^{\Delta'} f)(\cdot,y); \tfrac{b-a}{2m}\right) \\ & + \left(1 + \|L_m^{\Delta'}\|\right) W_{2m} \max_{x\in[a,b]} \omega_{2m}\left(f(x,\cdot); \tfrac{d-c}{2m}\right), \end{aligned} \tag{9}$$

where W_k is Whitney's constant.

The bivariate Shepard operator of Lidstone type is given by:

$$(S^{Li}f)(x,y) = \sum_{i=0}^{N} A_i(x,y)(L_m^{\rho,i}f)(x,y), \tag{10}$$

where $L_m^{\rho,i}f$ is the restriction of $L_m^{\rho}f$, given by (7), to the subrectangle $[x_i, x_{i+1}] \times [y_i, y_{i+1}]$, $0 \leq i \leq N$.

We have the bivariate Shepard-Lidstone interpolation formula,

$$f = S^{Li}f + R^{Li}f, \tag{11}$$

where $R^{Li}f$ is the remainder term.

Estimations of the remainder of this interpolation formula were obtained by us in [4] and [5].

2. **Second variant of the Shepard operator of Lidstone type**

For a function $f : [0,1] \times [0,1] \to R$ we consider the bivariate Shepard operator as a tensor product [13]:

$$(S_{M,N}f)(x,y) = \sum_{i=0}^{M}\sum_{j=0}^{N} s_{i,\lambda}(x)s_{j,\mu}(y)f\left(\tfrac{i}{M}, \tfrac{j}{N}\right), \tag{12}$$

where $\lambda, \mu > 1$ and

$$s_{i,\lambda}(x) = \frac{\left|x - \frac{i}{M}\right|^{-\lambda}}{\sum\limits_{k=0}^{M} \left|x - \frac{k}{M}\right|^{-\lambda}},$$

$$s_{j,\mu}(y) = \frac{\left|y - \frac{j}{N}\right|^{-\mu}}{\sum\limits_{k=0}^{N} \left|y - \frac{k}{N}\right|^{-\mu}}.$$

If we denote as in [13], by $S_{M,\lambda}(f,\cdot)$ the univariate Shepard operator regarding a univariate function f we have that

$$(S_{M,N}f)(x,y) = S_{M,\lambda}(f,x)S_{N,\mu}(f,y).$$

For a function $f \in C^{2m-2,2m-2}(D)$, $m \in \mathbb{N}$, the Shepard operator of Lidstone type corresponding to (12), is defined by

$$(S_{M,N}^{Li} f)(x,y) = \sum_{i=0}^{M} \sum_{j=0}^{N} s_{i,\lambda}(x) s_{j,\mu}(y) (L_m^{\rho,i,j} f) \left(\tfrac{i}{M}, \tfrac{j}{N} \right), \tag{13}$$

where $L_m^{\rho,i,j} f$ is the restriction of $L_m^{\rho} f$, given by (7), to the subrectangle $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$, $0 \le i \le M$, $0 \le j \le N$.

The corresponding interpolation formula is

$$f = S_{M,N}^{Li} f + R_{M,N}^{Li} f,$$

where $R_{M,N}^{Li} f$ denotes the remainder.

Further we give some error bounds for this interpolation procedure. First, we recall some known results.

Theorem 2.1. [5] *If $f \in C^{2m-2}[a,b]$ then*

$$\| f - S_{M,\lambda}(f,x) \| \le (1 + \left\| L_m^{\Delta} \right\|) W_{2m} \omega_{2m} \left(f; \tfrac{b-a}{2m} \right). \tag{14}$$

Theorem 2.2. *For any $f \in C^{2m-2,2m-2}(D)$ and $\mu > 2$ we have*

$$\begin{aligned} \left\| f - S_{M,N}^{Li} f \right\| \le & \left(1 + \left\| L_m^{\Delta} \right\| \right) W_{2m} \max_{y \in [0,1]} \omega_{2m} \left(f(\cdot, y); \tfrac{1}{2m} \right) \\ & + \left(1 + \left\| L_m^{\Delta} \right\| \right) W_{2m} \max_{y \in [0,1]} \omega_{2m} \left((f - L_m^{\Delta'} f)(\cdot, y); \tfrac{1}{2m} \right) \\ & + \left(1 + \| L_m^{\Delta'} \| \right) W_{2m} \max_{x \in [0,1]} \omega_{2m} \left(f(x, \cdot); \tfrac{1}{2m} \right). \end{aligned} \tag{15}$$

Proof. By Corollary 1.1 and taking into account (13) it follows that

$$\begin{aligned} \left\| f - S_{M,N}^{Li} f \right\|_{C[0,1]} = & \left\| f - S_{M,\lambda} f \right\|_{C[0,1]} \\ & + \left\| (f - L_m^{\Delta',i} f) - S_{M,\lambda} (f - L_m^{\Delta',i} f) \right\|_{C[0,1]} \\ & + \left\| f - S_{N,\mu} f \right\|_{C[0,1]}. \end{aligned}$$

and from (14) we obtain (15).

3. Third variant of the Shepard operator of Lidstone type

In [13] was introduce another type of the bivariate Shepard operator which has good approximation properties and better global smoothness preservation properties then that defined by (1). It is defined by

$$S_{M,N}(f;x,y) = \frac{T_{M,N}(f;x,y)}{T_{M,N}(1;x,y)},$$

with $\mu > 0$, $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^2$, $D = [0,1] \times [0,1]$, $x_i = i/M$, $i = 0,...,M$; $y_i = j/N$, $j = 0,...,N$ and

$$T_{M,N}(f;x,y) = \sum_{i=0}^{M}\sum_{j=0}^{N} \frac{f(x_i,y_j)}{[(x-x_i)^2+(y-y_j)^2]^\mu}.$$

For a function $f \in C^{2m-2,2m-2}(D)$, $m \in \mathbb{N}$, the corresponding Shepard operator of Lidstone type is given by

$$(S_{M,N}^{Li} f)(x,y) = \frac{T_{M,N}^{Li}(f;x,y)}{T_{M,N}(1;x,y)}, \tag{16}$$

with

$$T_{M,N}^{Li}(f;x,y) = \sum_{i=0}^{M}\sum_{j=0}^{N} \frac{(L_m^{\rho,i,j} f)(x_i,y_j)}{[(x-x_i)^2+(y-y_j)^2]^\mu},$$

where $L_m^{\rho,i,j} f$ is the restriction of $L_m^\rho f$, given by (7), to the subrectangle $[x_i,x_{i+1}] \times [y_j,y_{j+1}]$, $0 \le i \le M$, $0 \le j \le N$.

Theorem 3.1. [13] *For any $f \in C(D)$ and $\mu > 3/2$ we have*

$$\|f - S_{M,N}(f)\| \le c\omega(f; \tfrac{1}{M}, \tfrac{1}{N}), \tag{17}$$

where

$$\omega(f;\delta,\eta) = \sup\{|f(x+h,y+k) - f(x,y)| : 0 \le h \le \delta, 0 \le k \le \eta\}.$$

Using Theorem 3.1 we can give some error bounds.

Theorem 3.2. *For any* $f \in C^{2m-2,2m-2}(D)$ *and* $\mu > 3/2$ *we have*

$$\begin{aligned}\left\|f - S^{Li}_{M,N}f\right\| \leq & c\omega(L^{\rho}_m f; \tfrac{1}{M}, \tfrac{1}{N}) + \left(1 + \left\|L^{\Delta}_m\right\|\right) W_{2m} \max_{y\in[0,1]} \omega_{2m}\left(f(\cdot, y); \tfrac{1}{2m}\right) \\ & + \left(1 + \left\|L^{\Delta}_m\right\|\right) W_{2m} \max_{y\in[0,1]} \omega_{2m}\left((f - L^{\Delta'}_m f)(\cdot, y); \tfrac{1}{2m}\right) \\ & + \left(1 + \|L^{\Delta'}_m\|\right) W_{2m} \max_{x\in[0,1]} \omega_{2m}\left(f(x, \cdot); \tfrac{1}{2m}\right).\end{aligned} \tag{18}$$

Proof. We have

$$\left\|f - S^{Li}_{M,N}f\right\| \leq \|f - L^{\rho}_m f\| + \left\|L^{\rho}_m f - S^{Li}_{M,N}f\right\|$$

and by (16) and (17) we obtain

$$\left\|f - S^{Li}_{M,N}f\right\| \leq \|f - L^{\rho}_m f\| + c\omega(L^{\rho}_m f; \tfrac{1}{M}, \tfrac{1}{N}).$$

By (9) it follows (18).

References

[1] R. Agarwal, P.J.Y. Wong, *Error Inequalities in Polynomial Interpolation and their Applications,* Kluwer Academic Publishers, Dordrecht, (1993).

[2] T. Cătinaş, *The combined Shepard-Abel-Goncharov univariate operator,* Rev. Anal. Numér. Théor. Approx., **32** (2003) no. 1, pp. 11–20.

[3] T. Cătinaş, *The combined Shepard-Lidstone univariate operator,* "T. Popoviciu" Itinerant Seminar of Functional Equations, Approx. and Convexity, Cluj-Napoca, May 21–25, 2003, pp. 3–15.

[4] T. Cătinaş, *The combined Shepard-Lidstone bivariate operator,* Trends and Applications in Constructive Approximation, (Eds. M.G. de Bruin, D.H. Mache, J. Szabados), International Series of Numerical Mathematics, Vol. 151, 2005, Springer Group-Birkhäuser Verlag, pp. 77-89.

[5] T. Cătinaş, *Estimation of the remainder for the Shepard-Lidstone bivariate operator,* Rev. Anal. Numér. Théor. Approx., **36** (2005), no. 1, to appear.

[6] T. Cătinaş, *The Lidstone interpolation on tetrahedron*, J. Appl. Funct. Anal., 2006, no. 1, Nova Science Publishers, Inc., New York, to appear.

[7] Gh. Coman, *The remainder of certain Shepard type interpolation formulas,* Studia Univ. "Babeş-Bolyai", Mathematica, **32** (1987) no. 4, pp. 24–32.

[8] Gh. Coman, *Shepard operators of Birkhoff type,* Calcolo, **35** (1998), pp. 197–203.

[9] Gh. Coman and R. Trîmbiţaş, *Combined Shepard univariate operators,* East Jurnal on Approximations, **7** (2001) no. 4, pp. 471–483.

[10] Gh. Coman and R. Trîmbiţaş, *Univariate Shepard-Birkhoff interpolation,* Rev. Anal. Numér. Théor. Approx., **30** (2001) no. 1, pp. 15–24.

[11] Z. Ditzian and V. Totik, *Moduli of Smoothness*, Springer-Verlag, Berlin-Heidelberg-New York, Series in Computational Mathematics, vol. 9, 1987.

[12] S.G. Gal, J. Szabados, *On the preservation of global smoothness by some interpolation operators,* Studia Sci. Math. Hung., **35** (1999), no. 3-4, pp. 397-414.

[13] S.G. Gal, J. Szabados, *Global smoothness preservation by bivariate interpolation operators,* Analysis in Theory and Applications, **19** (2003), no. 3, pp. 193-208.

[14] J. Szabados, *Direct and converse approximation theorems for the Shepard operator,* Approximation Theory and its Applications, **7** (1991), no. 3, pp. 63-76.

[15] C. Zuppa, *Error estimates for modified local Shepard's interpolation formula*, Appl. Numer. Math., 49 (2004), pp. 245-259.

Babeş-Bolyai University, Faculty of Mathematics
and Computer Science, Str. Kogălniceanu, 1
Cluj-Napoca, Romania
E-mail address: tcatinas@math.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

CONTINUATION METHODS FOR INTEGRAL EQUATIONS IN LOCALLY CONVEX SPACES

A. CHIŞ

Abstract. The continuation method is used to investigate the existence of solutions to integral equations in locally convex spaces.

1. **Introduction**

In this article we study the problem of the existence of solutions for the Fredholm integral equation

$$x(t) = \int_0^1 K(t, s, x(s))ds, \qquad t \in [0, 1]. \tag{1.1}$$

and the Volterra integral equation

$$x(t) = \int_0^t K(t, s, x(s))ds, \qquad t \in [0, 1] \tag{1.2}$$

where the functions x, K have values in a locally convex space.

In paper [2] the above equations are studied using fixed point theorems for self-maps. Our approach is based on the continuation method.

The results presented in this paper extend and complement those in [2]-[5].

We finish this section by stating the main result from [1] which will be used in the next section.

Received by the editors: 8.09.2005.

2000 *Mathematics Subject Classification.* 45B05, 45D05.

Key words and phrases. Integral equations, locally convex spaces, Banach spaces, Volterra integral equations, Fredholm integral equations.

For a map $H : D \times [0,1] \to X$, where $D \subset X$, we will use the following notations:

$$\begin{aligned} &\Sigma = \{(x,\lambda) \in D \times [0,1] : H(x,\lambda) = x\}, \\ &S = \{x \in D : H(x,\lambda) = x \text{ for some } \lambda \in [0,1]\}, \\ &\Lambda = \{\lambda \in [0,1] : H(x,\lambda) = x \text{ for some } x \in D\}. \end{aligned} \tag{1.3}$$

Theorem 1.1. *Let X be a set endowed with the separating gauge structures $\mathcal{P} = \{p_\alpha\}_{\alpha \in A}$ and $\mathcal{Q}^\lambda = \{q_\beta^\lambda\}_{\beta \in B}$ for $\lambda \in [0,1]$. Let $D \subset X$ be $\mathcal{P}$-sequentially closed, $H : D \times [0,1] \to X$ a map, and assume that the following conditions are satisfies:*

(i) for each $\lambda \in [0,1]$, there exists a function $\varphi_\lambda : B \to B$ and $a^\lambda \in [0,1)^B$, $a^\lambda = \{a_\beta^\lambda\}_{\beta \in B}$ such that

$$q_\beta^\lambda(H(x,\lambda), H(y,\lambda)) \le a_\beta^\lambda q_{\varphi_\lambda(\beta)}^\lambda(x,y), \tag{1.4}$$

$$\sum_{n=1}^{\infty} a_\beta^\lambda a_{\varphi_\lambda(\beta)}^\lambda a_{\varphi_\lambda^2(\beta)}^\lambda ... a_{\varphi_\lambda^{n-1}(\beta)}^\lambda q_{\varphi_\lambda^n(\beta)}^\lambda(x,y) < \infty \tag{1.5}$$

for every $\beta \in B$ and $x, y \in D$;

(ii) there exists $\rho > 0$ such that for each $(x,\lambda) \in \Sigma$, there is a $\beta \in B$ with

$$\inf\{q_\beta^\lambda(x,y) : y \in X \backslash D\} > \rho; \tag{1.6}$$

(iii) for each $\lambda \in [0,1]$, there is a function $\psi : A \to B$ and $c \in (0,\infty)^A$, $c = \{c_\alpha\}_{\alpha \in A}$ such that

$$p_\alpha(x,y) \le c_\alpha q_{\psi(\alpha)}^\lambda(x,y) \qquad \text{for all } \alpha \in A \text{ and } x,y \in X; \tag{1.7}$$

(iv) $(X,\mathcal{P})$ is a sequentially complete gauge space;

(v) if $\lambda \in [0,1]$, $x_0 \in D$, $x_n = H(x_{n-1},\lambda)$ for $n = 1,2,...$, and $\mathcal{P}$-$\lim_{n\to\infty} x_n = x$, then $H(x,\lambda) = x$;

(vi) for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ with

$$q_{\varphi_\lambda^n(\beta)}^\lambda(x, H(x,\lambda)) \le (1 - a_{\varphi_\lambda^n(\beta)}^\lambda)\varepsilon$$

for $(x,\mu) \in \Sigma$, $|\lambda - \mu| \le \delta$, all $\beta \in B$, and $n \in \mathbb{N}$.

In addition, assume that $H_0 := H(.,0)$ has a fixed point. Then, for each $\lambda \in [0,1]$, the map $H_\lambda := H(.,\lambda)$ has a unique fixed point.

Remark 1.2. *Notice that, by condition (ii) we have: for each* $(x, \lambda) \in \Sigma$*, there is a* $\beta \in B$ *such that the set*

$$B(x, \lambda, \beta) = \{y \in X : q^{\lambda}_{\varphi^n_{\lambda}(\beta)}(x, y) \leq \rho, \ \forall n \in \mathbb{N}\} \subset D. \tag{1.8}$$

The proof of Theorem 1.1, in [1]*, shows that the contraction condition (1.4) given on* D*, can be asked only on sets of the form (1.8), more exactely for* $(x, \lambda) \in \Sigma$ *and* $y \in B(x, \lambda, \beta)$*.*

2. Existence Results

This section contains existence results for the equations (1.1) and (1.2).

Theorem 2.1. *Let* E *be a locally convex space, Hausdorff separated, complete by sequences, with the topology defined by the saturated and sufficient set of semi-norms* $\{|.|_{\alpha}, \alpha \in A\}$ *and let* $\delta > 0$ *be a fixed number. Assume that the following conditions are satisfied:*

(1) $K : [0,1]^2 \times E \to E$ *is continuous;*

(2) there exists $r = \{r_{\alpha}\}_{\alpha \in A}$ *such that, any solution* x *of the equation*

$$x(t) = \lambda \int_0^1 K(t, s, x(s))ds, \qquad t \in [0,1], \tag{2.9}$$

for some $\lambda \in [0,1]$ *satisfies* $|x(t)|_{\alpha} \leq r_{\alpha}$*, for all* $t \in [0,1]$ *and* $\alpha \in A$*;*

(3) there exists $\{L_{\alpha}\}_{\alpha \in A} \in [0,1)^A$ *such that*

$$|K(t, s, x) - K(t, s, y)|_{\alpha} \leq L_{\alpha} |x - y|_{f(\alpha)} \tag{2.10}$$

whenever $\alpha \in A$*,for all* $t, s \in [0,1]$ *and* $x, y \in E_r$ *where* $E_r = \{x \in E :$ *there exists* $\alpha \in A$ *such that* $|x|_{\alpha} \leq r_{\alpha} + \delta\}$*;*

(4)

$$\sum_{n=0}^{\infty} L_{\alpha} L_{f(\alpha)} ... L_{f^n(\alpha)} < \infty \tag{2.11}$$

for every $\alpha \in A$*;*

(5) for every $\alpha \in A$ *and for each continuous function* $g : [0,1] \to E$ *one has*

$$\sup\{|g(t)|_{f^n(\alpha)} : t \in [0,1], \ n = 0, 1, 2, ...\} < \infty;$$

(6) there exists C with $0 < C \leq \dfrac{1 - L_{f^n(\alpha)}}{M_{f^n(\alpha)}}$ for all $\alpha \in A$ and $n \in \mathbb{N}$, where $M_\alpha := \sup\limits_{\substack{t,s\in[0,1],\\ |x|_{f(\alpha)} \leq r_{f(\alpha)}}} |K(t,s,x)|_\alpha$.

Then problem (1.1) has a solution.

Notice that $M_\alpha < \infty$. Indeed, from (2.10) we have

$$\begin{aligned} |K(t,s,x)|_\alpha &\leq |K(t,s,x) - K(t,s,0)|_\alpha + |K(t,s,0)|_\alpha \\ &\leq L_\alpha r_{f(\alpha)} + \max_{t,s\in[0,1]} |K(t,s,0)|_\alpha < \infty \end{aligned}$$

for all $t,s \in [0,1]$ and $x \in E$ with $|x|_{f(\alpha)} \leq r_{f(\alpha)}$.

Proof. We shall apply Theorem 1.1. Let $X = C([0,1],E)$. For each $\alpha \in A$ we define the map $d_\alpha : X \times X \to \mathbb{R}_+$, by

$$d_\alpha(x,y) = \max_{t\in[0,1]} |x(t) - y(t)|_\alpha .$$

It is easy to show that d_α is a pseudo-metric on X and the family $\{d_\alpha\}_{\alpha\in A}$ defines on X a gauge structure, separated and complete by sequences.

Here $\mathcal{P} = \mathcal{Q}^\lambda = \{d_\alpha\}_{\alpha\in A}$ for every $\lambda \in [0,1]$. Let D be the closure in X of the set

$$\{x \in X : \ d_\alpha(x,0) \leq r_\alpha + \delta \text{ for some } \alpha \in A\}.$$

We define $H : D \times [0,1] \to X$, by $H(x,\lambda) = \lambda A(x)$, where

$$A(x)(t) = \int_0^1 K(t,s,x(s))ds.$$

In what follows we shall check conditions (i)-(vi) in Theorem 1.1. We shall start with condition (ii) by technical reason.

Condition (ii) becomes: there exists $\rho > 0$ such that for each solution $(x,\lambda) \in D \times [0,1]$, of $x = H(x,\lambda)$, there is an $\alpha \in A$ with

$$\inf\{d_\alpha(x,y) : y \in X\backslash D\} > \rho.$$

To prove this, let us note that if $y \in X \backslash D$, one has $d_\alpha(y,0) > r_\alpha + \delta$ for every $\alpha \in A$. Consequently, for at least one $t \in [0,1]$,

$$|x(t) - y(t)|_\alpha \geq |y(t)|_\alpha - |x(t)|_\alpha > r_\alpha + \delta - r_\alpha = \delta.$$

Then $d_\alpha(x,y) > \delta$. Hence (ii) holds for any $\rho \in (0,\delta)$.

Condition (i) becomes: for each $\alpha \in A$ there exists $f(\alpha) \in A$ and $L_\alpha \in [0,1)$ such that

$$d_\alpha(H(x,\lambda), H(y,\lambda)) \leq L_\alpha d_{f(\alpha)}(x,y), \tag{2.12}$$

$$\sum_{n=1}^{\infty} L_\alpha L_{f(\alpha)} ... L_{f^{n-1}(\alpha)} d_{f^n(\alpha)}(x,y) < \infty, \tag{2.13}$$

for all $x, y \in D$.

According to Remark 1.2, it suffices to have (2.12) on sets of the form (1.8). Let $(x,\lambda) \in D \times [0,1]$, such that $H(x,\lambda) = x$, and let $\beta \in A$. The set $B(x,\lambda,\beta) := \{y \in X : d_{f^n(\beta)}(x,y) \leq \rho, \forall\, n \in \mathbb{N}\}$ is included in D. From the fact that $H(x,\lambda) = x$ it follows that $|x(t)|_\alpha \leq r_\alpha$, for every $t \in [0,1]$ and $\alpha \in A$; from $y \in B(x,\lambda,\beta)$ it follows that $|y(t)|_\beta \leq r_\beta + \delta$, for every $t \in [0,1]$.

Then for x with $H(x,\lambda) = x$ and $y \in B(x,\lambda,\beta)$ we have

$$\begin{aligned}
|H(x,\lambda)(t) - H(y,\lambda)(t)|_\alpha &= \lambda \left| \int_0^1 (K(t,s,x(s)) - K(t,s,y(s)))\, ds \right|_\alpha \\
&\leq \lambda \int_0^1 |K(t,s,x(s)) - K(t,s,y(s))|_\alpha\, ds \\
&\leq \lambda \int_0^1 L_\alpha |x(s) - y(s)|_{f(\alpha)}\, ds \\
&\leq \lambda L_\alpha \max_{t\in[0,1]} |x(s) - y(s)|_{f(\alpha)} \\
&= \lambda L_\alpha d_{f(\alpha)}(x,y) \\
&\leq L_\alpha d_{f(\alpha)}(x,y).
\end{aligned}$$

Then $\max_{t\in[0,1]} |H(x,\lambda)(t) - H(y,\lambda)(t)|_\alpha \leq L_\alpha d_{f(\alpha)}(x,y)$, that is (2.12).

Now (2.13) follows from (4) and (5).

Condition (iii) is trivial since $\mathcal{P} = \mathcal{Q}^\lambda$.

Condition (iv): $(X, \{d_\alpha\}_{\alpha\in A})$ is a sequentially complete gauge space since E is complete by sequences.

Condition (v): Let $\lambda \in [0,1],\ x_0 \in D,\ x_n = H(x_{n-1}, \lambda)$ for $n = 1, 2, ...$ and assume $\mathcal{P}\text{-}\lim\limits_{n\to\infty} x_n = x$. We wish to obtain that $H(x, \lambda) = x$.

We have

$$\begin{aligned}
|H(x,\lambda)(t) - x(t)|_\alpha &= |H(x,\lambda)(t) - x_n(t) + x_n(t) - x(t)|_\alpha \\
&\le |H(x,\lambda)(t) - x_n(t)|_\alpha + |x_n(t) - x(t)|_\alpha \\
&= |H(x,\lambda)(t) - H(x_{n-1},\lambda)(t)|_\alpha + |x_n(t) - x(t)|_\alpha \\
&\le \int_0^1 L_\alpha |x(s) - x_{n-1}(s)|_{f(\alpha)}\, ds + |x_n(t) - x(t)|_\alpha \\
&\le L_\alpha \max_{s\in[0,1]} |x(s) - x_{n-1}(s)|_{f(\alpha)} + \sup_{t\in[0,1]} |x_n(t) - x(t)|_\alpha \\
&= L_\alpha d_{f(\alpha)}(x_{n-1}, x) + d_\alpha(x_n, x).
\end{aligned}$$

Passing to the supremum we obtain

$$d_\alpha(H(x,\lambda), x) \le L_\alpha d_{f(\alpha)}(x_{n-1}, x) + d_\alpha(x_n, x).$$

Letting $n \to \infty$, we deduce $d_\alpha(H(x,\lambda), x) = 0$.Since this equality is true for all $\alpha \in A$ and $\{d_\alpha\}_{\alpha\in A}$ is separated, we have $H(x, \lambda) = x$ as we wished.

Condition (vi) becomes: for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$d_{f^n(\alpha)}(x, H(x,\lambda)) \le (1 - L_{f^n(\alpha)})\varepsilon$$

whenever $(x, \mu) \in D \times [0,1]$, $H(x, \mu) = x$, $|\lambda - \mu| \le \delta$, $\alpha \in A$ and $n \in \mathbb{N}$.

Indeed, using (2) and (6) we obtain

$$\begin{aligned}
|x(t)-H(x,\lambda)(t)|_{f^n(\alpha)} &= |H(x,\mu)(t)-H(x,\lambda)(t)|_{f^n(\alpha)} \\
&= |\mu-\lambda| \left| \int_0^1 K(t,s,x(s))ds \right|_{f^n(\alpha)} \\
&\le |\mu-\lambda| \int_0^1 |K(t,s,x(s))|_{f^n(\alpha)}\, ds \\
&\le |\mu-\lambda|\, M_{f^n(\alpha)} \\
&\le |\mu-\lambda| \frac{1-L_{f^n(\alpha)}}{C}.
\end{aligned}$$

So condition (vi) is true with $\delta(\varepsilon)=C\varepsilon$.

In addition $H(.,0)=0\cdot A(.)=0$. Hence $H(.,0)$ has a fixed point.

Thus all the assumptions of Theorem 1.1 are satisfied and the proof is completed. □

In Banach space, Theorem 2.1 becomes the following well-known result.

Corollary 2.2. *Let $(E,|.|)$ be a Banach space. Assume that the following conditions are satisfied:*

(1) $K:[0,1]^2\times E\to E$ is continuous;

(2) there exists $r>0$ such that, any solution x of the equation

$$x(t)=\lambda\int_0^1 K(t,s,x(s))ds, \qquad t\in[0,1], \tag{2.14}$$

for some $\lambda\in[0,1]$ satisfies $|x(t)|<r$, for all $t\in[0,1]$ and any $\lambda\in[0,1]$;

(3) there exists $L\in[0,1)$ such that

$$|K(t,s,x)-K(t,s,y)|\le L\,|x-y| \tag{2.15}$$

for all $t,s\in[0,1]$ and $x,y\in E$ with $|x|,|y|\le r$.

Then problem (1.1) has a solution.

Notice that an analogue result is true for Volterra integral equation (1.2).

In particular, we obtain an existence principle for the initial value problem

$$\begin{cases} x'(t) = K(t, x(t)) & t \in [0,1] \\ x(0) = 0 \end{cases} \tag{2.16}$$

is equivalent to the integral equation

$$x(t) = \int_0^t K(s, x(s))ds, \qquad t \in [0,1] \tag{2.17}$$

for which the following result holds.

Theorem 2.3. *Let E be a locally convex space, Hausdorff separated, complete by the sequences, with the topology defined by the saturated and sufficient set of semi-norms $\{|.|_\alpha, \alpha \in A\}$ and let $\delta > 0$ be a fixed number. Assume that the following conditions are satisfied:*

(1) $K : [0,1] \times E \to E$ is continuous;

(2) there exists $r = \{r_\alpha\}_{\alpha \in A}$ such that, any solution x of the equation

$$x(t) = \lambda \int_0^t K(s, x(s))ds \qquad t \in [0,1]$$

for some $\lambda \in [0,1]$ satisfies $|x(t)|_\alpha \leq r_\alpha$, for all $t \in [0,1]$ and $\alpha \in A$;

(3) there exists $\{L_\alpha\}_{\alpha \in A} \in [0,1)^A$ such that

$$|K(t,x) - K(t,y)|_\alpha \leq L_\alpha |x - y|_{f(\alpha)}$$

whenever $\alpha \in A$, for all $t \in [0,1]$ and $x, y \in E_r$;

(4) $\sum_{n=0}^{\infty} L_\alpha L_{f(\alpha)} ... L_{f^n(\alpha)} < \infty$, for every $\alpha \in A$;

(5)for every $\alpha \in A$ and for each continuous function $g : [0,1] \to E$, one has

$$\sup\{|g(t)|_{f^n(\alpha)} : t \in [0,1],\ n = 0,1,2,...\} < \infty;$$

(6) there exists C with $0 < C \leq \dfrac{1 - L_{f^n(\alpha)}}{M_{f^n(\alpha)}}$, for all $\alpha \in A$ and $n \in \mathbb{N}$, where $M_\alpha := \sup\limits_{\substack{t \in [0,1], \\ |x|_{f(\alpha)} \leq r_{f(\alpha)}}} |K(t,x)|_\alpha$.

Then, the problem (2.17) has a solution.

The next theorem is concerning with the "a priori" boundedness condition (2).

Theorem 2.4. *Assume $K : [0,1] \times E \to E$ is continuous. In addition assume that for each $\alpha \in A$, there exists $\beta_\alpha \in C([0,1], \mathbb{R}_+)$ and $\psi_\alpha : \mathbb{R}_+ \to (0,\infty)$ nondecreasing with $\frac{1}{\psi_\alpha} \in L^1_{loc}(\mathbb{R}_+)$ such that*

$$|K(t,x)|_\alpha \leq \beta_\alpha(t)\psi_\alpha(|x|_\alpha), \text{ for } x \in E, t \in [0,1] \tag{2.18}$$

and

$$\int_0^\infty \frac{d\tau}{\psi_\alpha(\tau)} > \int_0^1 \beta_\alpha(s)ds. \tag{2.19}$$

Then condition (2) in Theorem 2.3 is satisfied..

Proof. Let x be any solution of the problem

$$\begin{cases} x'(t) = \lambda K(t, x(t)), & t \in [0,1] \\ x(0) = 0 \end{cases}$$

for some $\lambda \in [0,1]$, and let $\alpha \in A$ by arbitrary. Then

$$x(t) = \lambda \int_0^t K(s, x(s))ds, \qquad t \in [0,1]$$

and so

$$|x(t)|_\alpha \leq \lambda \int_0^t |K(s,x(s))|_\alpha \, ds = \lambda \int_0^t |x'(s)|_\alpha \, ds.$$

Let $w_\alpha(t) = \int_0^t |x'(s)|_\alpha \, ds$. Then $|x(t)|_\alpha \leq w_\alpha(t)$ on $[0,1]$. Using (2.18) we obtain

$$w'_\alpha(t) = |x'(t)|_\alpha = \lambda \, |K(t,x(t))|_\alpha \leq \lambda\beta_\alpha(t)\psi_\alpha(|x(t)|_\alpha) \leq \lambda\beta_\alpha(t)\psi_\alpha(w_\alpha(t))$$

on $[0,1]$. Next

$$\frac{w'_\alpha(t)}{\psi_\alpha(w_\alpha(t))} \leq \lambda\beta_\alpha(t) \leq \beta_\alpha(t)$$

and

$$\int_0^t \frac{w'_\alpha(s)}{\psi_\alpha(w_\alpha(s))} ds \leq \int_0^t \beta_\alpha(s)ds \leq \int_0^1 \beta_\alpha(s)ds.$$

Make the following change of variable $w_\alpha(s) = \tau$ and use (2.19) to derive

$$\int_0^{w_\alpha(t)} \frac{d\tau}{\psi_a(\tau)} \leq \int_0^1 \beta_\alpha(s)ds < \int_0^\infty \frac{d\tau}{\psi_\alpha(\tau)}.$$

The last inequality implies that there exists $r_\alpha < \infty$ such that $w_\alpha(t) \leq r_\alpha$ for every $t \in [0,1]$. Hence $|x(t)|_\alpha \leq r_\alpha$, for every $t \in [0,1]$. Therefore (2) holds. □

A better existence result is true for the Volterra integral equation

$$x(t)=\int_0^t K(t,s,x(s))ds, \qquad t\in[0,1]. \tag{2.20}$$

Theorem 2.5. *Let E be a locally convex space, Hausdorff separated, complete by the sequences, with the topology defined by the saturated and sufficient set of semi-norms $\{|.|_\alpha : \alpha\in A\}$ and let $\delta>0$ a fixed number. Assume that the following conditions are satisfied:*

(1) $K:[0,1]^2\times E\to E$ is continuous;

(2) there exists $r=\{r_\alpha\}_{\alpha\in A}$ such that each solution x of the equation

$$x(t)=\lambda\int_0^t K(t,s,x(s))ds, \qquad t\in[0,1]$$

for some $\lambda\in[0,1]$ satisfies $|x(t)|_\alpha\le r_\alpha$, for all $t\in[0,1]$ and $\alpha\in A$;

(3) there exists $\{L_\alpha\}_{\alpha\in A}\in(0,\infty)^A$ such that

$$|K(t,s,x)-K(t,s,y)|_\beta\le L_\alpha\,|x-y|_{f(\beta)} \qquad \text{for every } \beta\in\mathcal{O}_\alpha$$

whenever $\alpha\in A$; $t,s\in[0,1]$ and $x,y\in E_r$; here $\mathcal{O}_\alpha:=\{\alpha,f(\alpha),f^2(\alpha),...\}$;

(4) for every $\alpha\in A$ and for each continuos function $g:[0,1]\to E$ one has

$$\sup\{|g(t)|_{f^n(\alpha)} : t\in[0,1],\ n=0,1,2,...\}<\infty;$$

(5) $\sup\limits_n M_{f^n(\alpha)}<\infty$, for every $\alpha\in A$.

Then problem (2.20) has a solution.

Proof. We also apply Theorem 1.1. Let $X=C([0,1],E)$. We define the applications $\|.\|_\alpha : X\to\mathbb{R}_+$ by

$$\|x\|_\alpha=\max_{t\in[0,1]}\left(|x(t)|_\alpha\, e^{-\theta_\alpha t}\right)$$

where $\theta_\alpha>0$ will we precised in what follows. This applications are semi-norms on the linear space X, and the family $\{\|.\|_\alpha\}_{\alpha\in A}$ defines on X a structure of a locally convex space, separated, complete by sequences.

Let $a<1$. For each $\alpha\in A$ and $\theta_\alpha>0$, we define the pseudo-metric d_α : $X\times X\to\mathbb{R}_+$, by

$$d_\alpha(x,y)=\|x-y\|_\alpha\,.$$

Here again $\mathcal{P} = \mathcal{Q}^\lambda = \{d_\alpha\}_{\alpha \in A}$ for all $\lambda \in [0,1]$. Let D be the closure of

$$\{x \in X : \text{there is } \alpha \in A \text{ with } d_\alpha(x,0) \leq r_\alpha + \delta\}.$$

We define $H : D \times [0,1] \to X$, by $H(x,\lambda) = \lambda A(x)$, where

$$A(x)(t) = \int_0^t K(t,s,x(s))ds.$$

Now we check conditions (i)-(vi) from Theorem 1.1.

First we check *condition (ii)*: For any $y \in X \backslash D$ one has $d_\alpha(y,0) > r_\alpha + \delta$ for every $\alpha \in A$. Then for at least one $t \in [0,1]$, we have

$$\begin{aligned} |x(t) - y(t)|_\alpha e^{-\theta_\alpha t} &\geq (|y(t)|_\alpha - |x(t)|_\alpha) e^{-\theta_\alpha t} \\ &= |y(t)|_\alpha e^{-\theta_\alpha t} - |x(t)|_\alpha e^{-\theta_\alpha t} \\ &\geq d_\alpha(y,0) - d_\alpha(x,0) \\ &> r_\alpha + \delta - r_\alpha = \delta. \end{aligned}$$

Then $d_\alpha(x,y) > \delta$ for all $y \in X \backslash D$. So $\inf\{d_\alpha(x,y) : y \in X \backslash D\} > \rho$ for any $\rho \in (0,\delta)$.

Condition (i): Using the statements made in Remark 1.2,we will check the condition (1.4) on sets of the form (1.8). Let $(x,\lambda) \in D \times [0,1]$, such that $H(x,\lambda) = x$, and let $\beta \in A$. The set $B(x,\lambda,\beta) := \{y \in X : d_{f^n(\beta)}(x,y) \leq \rho, \ \forall \ n \in \mathbb{N}\}$ is included in D. From the fact that $H(x,\lambda) = x$ it follows that $|x(t)|_\alpha e^{-\theta_\alpha t} \leq r_\alpha$, for every $t \in [0,1]$, every $\alpha \in A$ and $\theta_\alpha > 0$; from $y \in B(x,\lambda,\beta)$ it follows that $|y(t)|_\beta e^{-\theta_\beta t} \leq r_\beta + \delta$, for every $t \in [0,1]$ and $\theta_\beta > 0$.

Let x with $H(x,\lambda) = x$ and $y \in B(x,\lambda,\beta)$. Then for $\gamma \in \mathcal{O}_\beta$ we have

$$\begin{aligned}
|H(x,\lambda)(t) - H(y,\lambda)(t)|_\gamma &= \lambda \left| \int_0^t (K(t,s,x(s)) - K(t,s,y(s)))\, ds \right|_\gamma \\
&\leq \lambda \int_0^t |K(t,s,x(s)) - K(t,s,y(s))|_\gamma \, ds \\
&\leq \lambda \int_0^t L_\beta \, |x(s) - y(s)|_{f(\gamma)} \, e^{-\theta_\beta s} e^{\theta_\beta s} ds \\
&\leq \lambda L_\beta \max_{t\in[0,1]} \left(|x(s) - y(s)|_{f(\gamma)} \, e^{-\theta_\beta s} \right) \int_0^t e^{\theta_\beta s} ds \\
&= \lambda L_\beta d_{f(\gamma)}(x,y) \int_0^t e^{\theta_\beta s} ds \\
&\leq \frac{L_\beta}{\theta_\beta} d_{f(\gamma)}(x,y) e^{\theta_\beta t}.
\end{aligned}$$

So we have

$$|H(x,\lambda)(t) - H(y,\lambda)(t)|_\beta \, e^{-\theta_\beta t} \leq \frac{L_\beta}{\theta_\beta} d_{f(\gamma)}(x,y).$$

Consequently

$$d_\gamma(H(x,\lambda), H(y,\lambda)) \leq \frac{L_\beta}{\theta} d_{f(\gamma)}(x,y).$$

We choose $\theta_\alpha > 0$ large enough that

$$\frac{L_\alpha}{\theta_\alpha} \leq a$$

and

$$L_\alpha + \sup_n M_{f^n(\alpha)} \leq \theta_\alpha \tag{2.21}$$

for all $\alpha \in A$.

For each $\alpha \in A$ series (1.5) is dominated by the convergent series $\sum\limits_{n=0}^{\infty} a^n$ which obviously is convergent. This together with condition (4) guarantees condition (i) from Theorem 1.1.

For *condition (iii)* and *condition (iv)* see the proff of Theorem 2.1.

Condition (v): We have

$$\begin{aligned}
|H(x,\lambda)(t)-x(t)|_\alpha &= |H(x,\lambda)(t)-x_n(t)+x_n(t)-x(t)|_\alpha \\
&\le |H(x,\lambda)(t)-x_n(t)|_\alpha + |x_n(t)-x(t)|_\alpha \\
&= |H(x,\lambda)(t)-H(x_{n-1},\lambda)(t)|_\alpha + |x_n(t)-x(t)|_\alpha \\
&\le \int_0^t L_\alpha |x(s)-x_{n-1}(s)|_{f(\alpha)} e^{-\theta_\alpha s} e^{\theta_\alpha s} ds + |x_n(t)-x(t)|_\alpha \\
&\le L_\alpha \max_{s\in[0,1]} \left(|x(s)-x_{n-1}(s)|_{f(\alpha)} e^{-\theta_\alpha s} \right) \int_0^t e^{\theta_\alpha s} ds + \\
&+ |x_n(t)-x(t)|_\alpha \le \frac{L_\alpha}{\theta_\alpha} d_{f(\alpha)}(x_{n-1},x) e^{\theta_\alpha t} + |x_n(t)-x(t)|_\alpha .
\end{aligned}$$

Hence

$$|H(x,\lambda)(t)-x(t)|_\alpha \le \frac{L_\alpha}{\theta_\alpha} d_{f(\alpha)}(x_{n-1},x) e^{\theta_\alpha t} + |x_n(t)-x(t)|_\alpha .$$

If we multiply by $e^{-\theta_\alpha t}$, we obtain

$$|H(x,\lambda)(t)-x(t)|_\alpha e^{-\theta_\alpha t} \le d_{f(\alpha)}(x_{n-1},x) + |x_n(t)-x(t)|_\alpha e^{-\theta_\alpha t}.$$

Taking the supremum into the above inequality, we obtain

$$d_\alpha(H(x,\lambda),x) \le d_{f(\alpha)}(x_{n-1},x) + d_\alpha(x_n,x).$$

Letting $n\to\infty$, we deduce that $d_\alpha(H(x,\lambda),x)=0$ and so $H(x,\lambda)=x$.

Condition (vi) From

$$\begin{aligned}
|x(t)-H(x,\lambda)(t)|_{f^n(\alpha)} &= |H(x,\mu)(t)-H(x,\lambda)(t)|_{f^n(\alpha)} \\
&= |\mu-\lambda| \left| \int_0^t K(t,s,x(s))ds \right|_{f^n(\alpha)} \\
&\le |\mu-\lambda| \int_0^t |K(t,s,x(s))|_{f^n(\alpha)} e^{-\theta_\alpha s} e^{\theta_\alpha s} ds \\
&\le |\mu-\lambda| M_{f^n(\alpha)} \int_0^t e^{\theta_\alpha s} ds.
\end{aligned}$$

we obtain

$$|x(t)-H(x,\lambda)(t)|_{f^n(\alpha)} \le |\mu-\lambda| \frac{M_{f^n(\alpha)}}{\theta_\alpha} e^{\theta_\alpha t},$$

and using (2.21) we deduce

$$|x(t) - H(x,\lambda)(t)|_{f^n(\alpha)} e^{-\theta_\alpha t} \leq |\mu - \lambda| \frac{M_{f^n(\alpha)}}{\theta_\alpha} \leq |\mu - \lambda| \left(1 - \frac{L_\alpha}{\theta_\alpha}\right).$$

So condition(vi) is true for $\delta = \varepsilon$.

In addition $H(.,0) = 0 \cdot A(.) = 0$. Hence $H(.,0)$ has a fixed point. Thus Theorem (1.1), applies. □

In case that $f : A \to A$ is the identity map, Theorem 2.5 reduces to the following result.

Theorem 2.6. *Let E be a locally convex space, Hausdorff separated ,complete by the sequences, with the topology defined by the saturated and sufficient set of semi-norms $\{|.|_\alpha, \alpha \in A\}$ and $\delta > 0$ a fixed number. Assume that the following conditions are satisfied:*

(1) $K : [0,1]^2 \times E \to E$ is continuous;

(2) there exists $r = \{r_\alpha\}_{\alpha \in A}$ such that, each solution x of the problems

$$x(t) = \lambda \int_0^t K(t,s,x(s))ds$$

has the property $|x(t)|_\alpha \leq r_\alpha$, for all $t \in [0,1], \alpha \in A$ and every $\lambda \in [0,1]$;

(3)there exists $L_\alpha > 0$ such that

$$|K(t,s,x) - K(t,s,y)|_\alpha \leq L_\alpha |x - y|_{f\alpha}$$

whenever $\alpha \in A$, for all $t,s \in [0,1]$, and $x,y \in E_r$;

Then, the problem (2.20) has a solution.

References

[1] A. Chis and R. Precup, *Continuation theory for general contractions in gauge spaces*, Fixed Point Theory and Applications **2004:**3 (2004), 173-185.

[2] N. Gheorghiu and M. Turinici , *Equation intégrales dans les espaces localment convexes,* Rev. Roumaine Math Pures Appl. **23** (1978), no. 1, 33-40 (French).

[3] M. Frigon, *Fixed point results for generalized contractions in gauge space and applications,* Proc. Amer. Math. Soc. **128** (2000), 2957-2965.

[4] M. Frigon et A. Granas, *Resultats du type de Leray-Schauder pour des contractions multivoques,* Topological Methods Nonlinear Anal. **4** (1994), 197-208.

[5] D. O'Regan and R. Precup, *Theorems of Leray-Schauder Type and Applications*, Gordon and Breach Science Publishers, Amsterdam, 2001.

Department of Mathematics
Technical University
Cluj-Napoca, Romania
E-mail address: Adela.Chis@math.utcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

MULTI-CLASS INFERENCE WITH GAUSSIAN PROCESSES

BOTOND CSEKE   LEHEL CSATÓ

Abstract. A Bayesian probabilistic framework for multi-class classification is presented. We employ Gaussian processes as latent variable models for each of the classes and present a Bayesian inference scheme. The problem is not analytically tractable and we present approximation schemes and assess the approximation on different problems.

1. Introduction

The problem of "recognizing" patterns mathematically is formulated as the assignment of labels to specific inputs $\mathbf{x}$. The set of labels has finite cardinality, therefore the problem of label assignment is one of classification where the number of classes equals the cardinality of labels.

Binary classification is thoroughly studied and well understood for several problem domains [**?**]. It is easier to *model* the binary classification since it reduces to assigning *the sign* of a function to either of the classes. For the multi-class case it is a more difficult problem: more than two classes require to have an indicator for each class. To avoid the multiplication of these indicators, several alternative models have been proposed, all of them transform the *single* multi-class classification into *several* binary classification problems and then combine the results of the binary classifications into a single "output" [**?**, **?**]. In this article we model the multi-class classification. We use a probabilistic modelling and latent variables to model the class-conditional densities. A flexible modelling strategy is the use of random functions, namely the stochastic Gaussian processes as latent variables associated to each class.

Received by the editors: 15.09.2005.

2000 *Mathematics Subject Classification.* 68T05,68T10.

Key words and phrases. machine learning, pattern recognition, graphical models, Gaussian processes.

We present the general framework of modelling with latent variables (section 1), the models using Gaussian processes (sections 2,3) and the modelling of the multi-class classification (section 4). The article ends with the discussion of further research points worth to be carried out.

1.1. **The Classification Problem.** Let be given a set of data $D = \{(\mathbf{x}_i, y_i) : \mathbf{x}_i \in \mathcal{X}, y_i \in \{1, \ldots, C\}, i = 1, \ldots, n\}$ sampled independently from an unknown distribution $P(\mathbf{x}, y)$ our task is to build a *classifier* – a function $\mathbf{x} \to y$ where $y \in \{1, \ldots, C\}$ – which produces a reasonably small generalization error i.e. for a given new input $\mathbf{x}_*$, it gives a relatively good approximation for $P(y_*|\mathbf{x}_*, D)$ where y_* is in $\{1, \ldots, C\}$.

The set D is usually called *training set*, the process of finding the model is called *training* or *learning process*. In most cases there is also a set S called *test set* on which we measure the performance of the model. We point out that $\mathbf{x}_*$ may be any point of the input space and the train–test method described above is just a common technique and we attempt to solve a supervised learning problem – to provide prediction for arbitrary input points $\mathbf{x}_*$– not a transductive one – to provide predictions for a fixed set of input points.

It is desirable to build a classifier which produces low errors both on training and test sets. A too low error on training set usually leads to weak prediction – high error on the test set – performance since the model is fitted "too tight" to the training data. This effect is known as *over-fitting*. Usually we expect that inputs close to each other belong to the same class – the modelled classifier is smooth in some sense – so it is plausible to penalize overly complex candidates which usually produce low errors on the training set.

1.2. **Probabilistic models and Bayesian inference.** When building model for data one usually postulates some structure for the hidden mechanism that supposedly produced it – depending on the nature of the problem at hand. This assumption leads to the introduction of *hidden* or *latent variables* – u from now on – and the assumption that the outputs y and the inputs $\mathbf{x}$ are conditionally independent given u. There may be various practical motivations for modelling with hidden variables like: it is easier to introduce *smoothness* criteria and it is easier to model $P(y|u)$ and $P(u|\mathbf{x})$ independently than modelling the relation $P(y|\mathbf{x})$ – the relation between $\mathbf{x}$ and y – directly. This assumption can be written in the form:

$$P(y|\mathbf{x}) = \int P(y|u)P(u|\mathbf{x})du.$$

One postulates the model by specifying the distributions $P(y|u)$ and $P(u|\mathbf{x})$.

The distribution $P(y|u)$ is usually called *noise* distribution/model or *random component* and it expresses our belief about how the *hidden variables* produce the output y. There are two ways of defining the *prior distribution* $P(u|\mathbf{x})$ of the *hidden variables*: (1) in a *parametric* manner: $u(\mathbf{x}; w)$ is a parametric function and we place some prior distribution $P(w)$ on parameters w which usually have a low dimensionality not depending on the cardinality of the data sets (2) in a *non-parametric* manner: we place a prior $P(u|\mathbf{x})$ directly on $u(\mathbf{x})$. If $u(\cdot)$ is a Gaussian Process then it is specified by its mean value and covariance function.

In some cases $P(y|u)$ and $P(u|\mathbf{x})$ depend on further parameters θ called *hyperparameters*, they control characteristics like parameters of the distribution $P(w)$ or parameters of functions which define $u(\cdot)$. Smoothness criteria – mentioned a few paragraph earlier – may be expressed by the priors placed on $u(\cdot)$ or $u(\cdot; w)$ (a prior on w) assigning low probability to models leading to overly complex functions. We shall see later – sections 2.1 and 3.1 – how these probabilities control the smoothness of the model.

In this text we are concerned with *non-parametric* models and in the following we shall present how the Bayesian machinery – repeated application of Bayes' rule (see for example Soós [**?**]) in a hierarchy – can be put in work in such cases.

In order to make predictions one has to calculate the posterior probability of the hidden variables at training input locations. Denoting them by $u(X_D)$ we have:

$$P(u(X_D)|D) \propto P(D|u(X_D))P(u(X_D))$$

and using the assumption that y and x are conditionally independent w.r.t u one gets:

$$P(u_*|\mathbf{x}_*, D) = \int P(u_*|u(X_D))P(u(X_D)|D)du(X_D). \tag{1}$$

and

$$P(y_*|\mathbf{x}_*, D) = \int P(y_*|u_*)P(u_*|\mathbf{x}_*, D)du_*.$$

where we have denoted by $\mathbf{x}_*$ the input location where the prediction needs to be done, $u_* = u(\mathbf{x}_*)$ the value of the latent variable at this location and by y_* the predicted output variable.

When *hyperparameters* are involved, one uses a second level inference. One has to weight the prediction distributions $P(y_*|\mathbf{x}_*, D, \theta)$ with the suitability of the model with parameter θ. This suitability is usually measured by the posterior distributions of θ:

$$\begin{aligned} P(\theta|D) &\propto P(\theta)P(D|\theta) \\ &\propto P(\theta)\int P(D|u(X_D),\theta)P(u(X_D)|\theta)du(X_D). \end{aligned} \tag{2}$$

When the cardinality of the training data is sufficiently large, then $P(\theta|D)$ is highly peaked around its mode $\hat{\theta}$ .This means that the posterior is unimodal, therefore it is a common practice to substitute it by $\delta_{\hat{\theta},\theta}$. This method is called *maximum likelihood II* and the prediction we obtain using this method is called *maximum a-posteriori (MAP)* approximation. Using the Bayesian approach one has to sum over all possible parameters and gets:

$$P(y_*|x_*, D) = \int P(y_*|x_*, \theta)P(\theta|D)d\theta.$$

The process of learning is realized trough Bayesian estimations i.e. it means the updating of model parameters from the *priors* $P(\theta)$ to *posteriors* $P(\theta|D)$.

2. Modelling with Gaussian Processes

When modelling with Gaussian Processes (GPs) we place a Gaussian process prior on the random function u, thus the *hidden variable/function* has the property that for any collection of possible different inputs $X = \{\mathbf{x}'_1, \ldots, \mathbf{x}'_m\} \subset \mathcal{X}$ the random variable $u(X) = (u(\mathbf{x}'_1), \ldots, u(\mathbf{x}'_m))^T$ is a Gaussian random vector. The process is determined by its mean value function $m(\mathbf{x}) = \mathrm{E}[u(\mathbf{x})]$ and covariance function $K(\mathbf{x}.\mathbf{x}') = \mathrm{E}[(u(\mathbf{x}) - m(\mathbf{x}))(u(\mathbf{x}') - m(\mathbf{x}'))]$ which is a positive definite symmetric function, thus "producing" valid covariance matrices for any finite dimensional distribution.

Gaussian processes have long been studied in probability and statistics and used for various problems in nonparametric estimation but they have been "rediscovered" by the *ML* community only a decade ago when Neal [**?**], Williams and Rasmussen [**?**] showed that the output distribution of a simple two layer Bayesian Neural Network with increasing number of hidden units converges to a Gaussian process. Their nonparametric nature makes them relatively insensible to data dimensionality and reasonably complex models can be built with a few number of *hyperparameters* only. Being a nonparametric method smoothness conditions can be imposed by the choice of covariance function as it was pointed out by Parzen [**?**] then later on by Kimeldorf and Wahba [**?**], therefore Gaussian Processes are a tempting device for attacking *ML* problems.

Let X_S with $X_S \cap X_D = \emptyset$ be a set of test "locations" where estimations needs to be done and let us denote $X_D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ the set of training input locations and $\mathbf{u} = u(X_D) = (u(\mathbf{x}_1), \ldots, u(\mathbf{x}_n))^T$ the hidden variable vector at the training input points.

Applying the Bayesian model presented above one gets:

$$P(u(X_S)|X_S, D) = \int P(u(X_S)|\mathbf{u})P(\mathbf{u}|D)d\mathbf{u}.$$

For notational simplicity we shall use from now on the notation $P(u_*|D)$ for $P(u(X_S)|X_S, D)$, – whenever it is unambiguous – expressing that X_S is arbitrary and that we do not consider modelling $P(\mathbf{x})$. The process resulting form the finite dimensional distributions $P(u(X_S)|X_S, D)$ is called *posterior predictive process* and we shall sometimes call it simply: *posterior process.*

2.1. **Gaussian process regression with Gaussian noise.** The simplest probabilistic model using Gaussian process as hidden function is the regression problem with zero-mean Gaussian noise $P(y|u) \sim N(y; u, \sigma^2)$, which has an analytically easily tractable formalism. We assume that the "hidden function" is a Gaussian random function having a priori zero mean and a covariance K. For notational simplicity let us denote $u_i = u(\mathbf{x}_i)$, $\mathbf{y} = (y_i)_i$, $\mathbf{k}(\mathbf{x}) = (K(\mathbf{x}, \mathbf{x}_i))_i$, $k_* = K(\mathbf{x}_*, \mathbf{x}_*)$ and the covariance matrix at training input locations $\mathbf{K} = (K(\mathbf{x}_i, \mathbf{x}_j))_{i,j}$. Employing again the Bayesian formalism presented above one gets the posterior distribution

$$\begin{aligned} P(\mathbf{u}|D) &\propto \prod_i P(y_i|u_i)P(\mathbf{u}) \\ &= N\left(\mathbf{K}(\sigma^2\mathbf{I}+\mathbf{K})^{-1}\mathbf{y}, \sigma^2(\sigma^2\mathbf{I}+\mathbf{K})^{-1}\mathbf{K}\right) \end{aligned}$$

and thus obtaining the predictive distribution:

$$P(u_*|D) = \int P(u_*|\mathbf{u})P(\mathbf{u}|D)d\mathbf{u} \tag{3}$$

$$= N\left(u_*|\mathbf{k}^T(\mathbf{x}_*)(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y}, k_* - \mathbf{k}^T(\mathbf{x}_*)(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{k}(\mathbf{x}_*)\right). \tag{4}$$

leading to a Gaussian *predictive process* $u(\cdot)|D$ with

$$\begin{aligned} \mathrm{E}[u(\mathbf{x})|D] &= \mathbf{k}^T(\mathbf{x})(\sigma^2\mathbf{I}+\mathbf{K})^{-1}\mathbf{y} \\ \mathrm{Cov}[u(\mathbf{x}), u(\mathbf{x}')|D] &= K(\mathbf{x}, \mathbf{x}') - \mathbf{k}^T(\mathbf{x})(\sigma^2\mathbf{I}+\mathbf{K})^{-1}\mathbf{k}(\mathbf{x}'). \end{aligned}$$

We remark that denoting $\mathbf{w} = (\sigma^2\mathbf{I}+\mathbf{K})^{-1}\mathbf{y}$ – which is independent of $\mathbf{x}_*$ – we have:

$$\mathrm{E}[u(\mathbf{x})|D] = \mathbf{w}^T\mathbf{k}(\mathbf{x}) = \sum w_i K(\mathbf{x}, \mathbf{x}_i).$$

Analyzing equation (5) we notice that the point-wise predictive variance is smaller than the prior variance.

Using an arbitrary noise model – changing the noise distribution $P(y_i|u(\mathbf{x}_i))$ – we might generalize this Gaussian process regression model but $P(u(\cdot)|D)$ is not Gaussian anymore, not is the *point-wise predictive distribution* and *posterior predictive process*. Using a fixed covariance function is not generally useful for practical purposes, because it's nature affects the "quality" of the approximation and prediction we obtain. Since the posterior mean value is a linear combination of functions

$K(\mathbf{x}, \mathbf{x}_i)$, choosing fast or slow decaying covariances may lead to poor approximations. With parameterized covariance function we may get better control over the flexibility of the functions/processes in issue. Due to ease in identifying its parameters role the square exponential

$$K(s,t) = b + a\exp(-\frac{1}{2}\sum_i v_i(s^i - t^i)^2) \tag{5}$$

is one of the most often used covariance functions in machine learning GP models. Figure 2.1 shows how the choice of the so called scale parameters v_i control the posterior mean value. (We may use $\theta_0 = (\log(b), \log(a), (\log v_i)_i)$ as covariance function parameter.)

Because $P(\theta|D)$ is not a Gaussian, the predictive distribution given in equation (2) is not analytically tractable and instead, finding a $MLII$ value or sampling methods (ex. Markov Chain Monte Carlo, Hybrid Monte Carlo) – from $P(\theta|D)$ – must be employed for carrying out the integration numerically. Both of these may be done by using the log–likelihood:

$$\log P(D|\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left|\mathbf{K} + \sigma^2\mathbf{I}\right| - \frac{1}{2}\mathbf{y}^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{y}. \tag{6}$$

The key of the relative ease in formalism of the regression problems is the closeness property of the Gaussian distributions function regarding multiplication and division. Gaussian likelihoods are able to model only a small proportion of real word problems but Gaussian processes can model a large variety of functions – for example the class of posterior mean value functions when a parametrized exponential covariance is used – therefore it worths to keep the function class and develop methods for a wider or arbitrary class of likelihoods, see Csató [**?**].

3. **Approximate inference**

As we have pointed out in the previous section non-Gaussian noise distributions are the ones which “break” the analytical tractability of the Bayesian model for GPs. Taking account that in cases when the log-likelihood $\log P(D|u)$ is concave

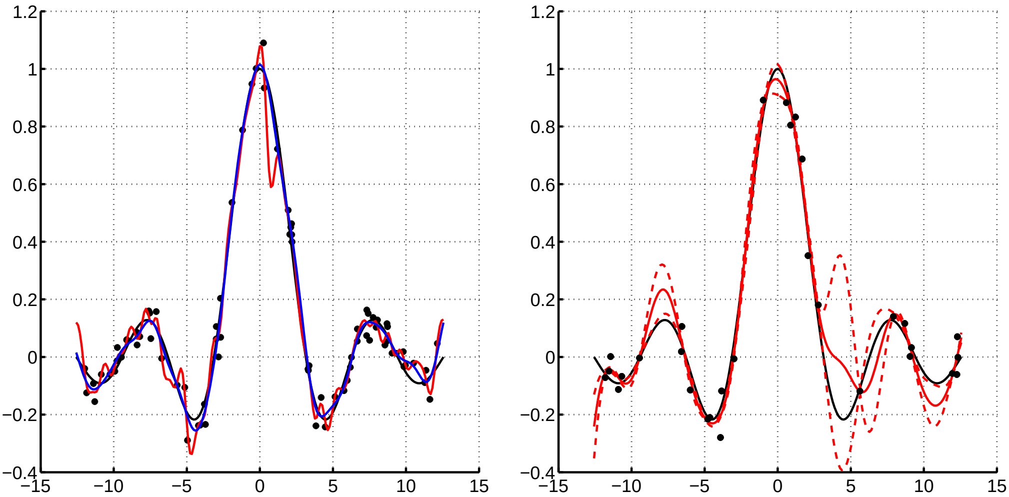


FIGURE 1. An illustration of overfitting and smooth fitting on noisy data generated by the *sinc* function (left) as well as the GP regression posterior mean and variance (right).

the posterior is unimodal it seems a good solution to approximate non-Gaussian posteriors with Gaussian ones. In the followings we shall present a few variants of this approach.

3.1. **Binary classification using Laplace approximation.** The main idea is to transform the classification problem into a regression one and the interpret the obtained results. Let $y_i \in \{0, 1\}$ and assume that if $\mathbf{x}_i$ belongs to class C_0 then $y_i = 0$. We model the problem in the following way: we shall transform the output i.e. the process u to the interval $[0, 1]$ with a suitable function σ and we shall interpret $\sigma(u(\mathbf{x}))$ as $P(\mathbf{x} \in C_1)$ - the probability of $\mathbf{x}$ belonging to the class C_1 denoted by the 1 values of y_i-s. We use the function $\sigma(x) = e^x/(1 + e^x)$, thus our goal is to approximate $P(\sigma(u(\mathbf{x}))|D)$ at a fixed point $\mathbf{x}$. For notational simplicity we demote $\pi(\mathbf{x}) = \sigma(u(\mathbf{x}))$. To apply the Bayesian treatment presented in the previous section one must postulate the corresponding conditional densities $P(y_i|u(\mathbf{x}_i))$. Assuming $\pi(\mathbf{x})$-s are probability of success for the Bernoulli random variables y and the samples $(\mathbf{x}_i, y_i)$ are independent:

$$P(D|\mathbf{u}) = \prod_i \pi(\mathbf{x}_i)^{y_i}(1-\pi(\mathbf{x}_i))^{(1-y_i)}.$$

Since in this case the posterior process is not a Gaussian, Williams and Barber [?] propose a Gaussian approximation for the posterior process. Using Laplace's method they approximate $P(u(\mathbf{x}_*), \mathbf{u}|D)$ by a Gaussian at it's mode, then they marginalize and obtain $P(u(\mathbf{x}_*)|D)$ and so the last step remains the calculation of $P(\pi(\mathbf{x}_*)|D)$. Applying Bayes' rule one gets:

$$\begin{aligned}\log P(u_*, \mathbf{u}|D) &= \log P(\mathbf{y}|\mathbf{u}_+) + \log P(\mathbf{u}_+) - \log P(D). \\ &= \mathbf{y}^T\mathbf{u} - \sum_{i=1}^{n} \log(1+e^{u_i}) - \frac{1}{2}\mathbf{u}_+{}^T\mathbf{K}_+^{-1}\mathbf{u}_+ - \frac{1}{2}\log|\mathbf{K}_+| + c\end{aligned}$$

where $\mathbf{K}_+$ is the extended – with $\mathbf{k}(\mathbf{x}_*)$ and $K(\mathbf{x}_*, \mathbf{x}_*)$ – covariance matrix and $\mathbf{u}_+$ is the extended hidden variable at inputs $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_*$. One may easily verify that at the maximum – $\nabla_{\mathbf{u}_+} \log P(\mathbf{u}_+|D) = 0$ – we have $(u_*)_{max} = \mathbf{k}^T(\mathbf{x}_*)\mathbf{K}^{-1}\mathbf{u}_{max}$ where $\mathbf{u}_{max} = \text{argmax}_{\mathbf{u}} P(\mathbf{u}|D)$.

Denoting by $\bar{u}|D$ the approximating Gaussian process of the posterior process where approximation is understood in the sense presented above: Laplace's method – one gets:

$$\bar{u}(\mathbf{x})|D \sim N(\mathbf{k}^T(\mathbf{x})\mathbf{K}^{-1}\mathbf{u}_{max}, K((\mathbf{x}),(\mathbf{x})) - \mathbf{k}^T(\mathbf{x})(\mathbf{I}+\mathbf{W}\mathbf{K})^{-1}\mathbf{W}\mathbf{k}(\mathbf{x}))$$

and we have used the notation $W = -(\nabla\nabla^T)_{\mathbf{u}} \log P(D|\mathbf{u})|_{\mathbf{u}=\mathbf{u}_{max}}$

When parameterized covariance function is used we obtain the likelihood approximation

$$\log P(D|\theta) \simeq \log P(\mathbf{u}_{max}|D) - \frac{1}{2}\log|\mathbf{K}^{-1} + \mathbf{W}(\mathbf{u}_{max})| + c.$$

This can be used for sampling from $P(\theta|D)$ to carry out the Bayesian averaging or hyperparameter optimization.

3.2. **Expectation–propagation.** When using non-Gaussian likelihoods the posterior process is not Gaussian which makes the estimation of predictions analytically intractable. The finite dimensional distributions of the posterior process may be written similarly to equation (1), meaning that, in order to get a Gaussian posterior we should approximate the non-Gaussian $P(\mathbf{u}|D)$ with a Gaussian $Q(\mathbf{u})$ and define the Gaussian approximation of the posterior process defined by the finite dimensional distributions

$$Q(u_*) = \int P(u_*|\mathbf{u})Q(\mathbf{u})d\mathbf{u}$$

In section 3.1 Laplace approximation has been applied to approximate the posterior process for the binary classification problem. Another plausible way to approximate the posterior is to find the Gaussian which minimizes the KL distance (see for example Cover and Thomas [?]) defined by

$$\mathrm{D}[P(\mathbf{u}|D)||Q(\mathbf{u})] = \int \ln\left[\frac{P(\mathbf{u}|D)}{Q(\mathbf{u})}\right] P(\mathbf{u}|D)d\mathbf{u}.$$

Because

$$\mathrm{D}[P(u_*, \mathbf{u}|D)||Q(u_*, \mathbf{u})] = \mathrm{D}[P(\mathbf{u}|D)||Q(\mathbf{u})]$$

and the minimization boils down to second and first order matching between $P(\mathbf{u}|D)$ and $Q(\mathbf{u})$, see Opper [?].

The *Expectation Propagation (EP)* method developed by Minka [?] proves to be an efficient method for doing KL-type approximate inference in probabilistic models using factorizing likelihoods, because the properties of KL distance endow *EP* with a particularly important local property in cases when the factors depend only on a few components of the hidden variable vector (a few number of hidden variables).

We shall present this method in the context of GP models. A complete and general exposition is found in Minka [?]. The main idea of *EP* consists in a novel interpretation of the Assumed Density Filtering (ADF) method. Supposing a factorizing likelihood $P(D|\mathbf{u}) = \prod_i t_i(\mathbf{u})$ – $t_i(\mathbf{u})$ standing for $P(y_i|\mathbf{u})$ – and a prior $P(\mathbf{u})$, *EP* approximates the posterior $P(\mathbf{u}|D)$ by a distribution $Q(\mathbf{u}) = P(\mathbf{u})\prod_i \bar{t}_i(\mathbf{u})$, thus approximating each "component" of the likelihood – also called sites – by an "easy to handle" distributions – in our case low dimensional Gaussians. It does this in the following way:

(1) the algorithm usually starts assuming $Q(\mathbf{u}) := P(\mathbf{u})$ thus setting $\bar{t}_i := 1$
(2) at each step a chosen – arbitrary or by other way – $\bar{t}_i(\mathbf{u})$ is removed from $Q(\mathbf{u})$ resulting:

$$Q^{\setminus i}(\mathbf{u}) \sim P(\mathbf{u}) \prod_{j \neq i} \bar{t}_j(\mathbf{u})$$

(3) it infers $\hat{P}(\mathbf{u}) \sim Q^{\setminus i}(\mathbf{u}) t_i(\mathbf{u})$ and
(4) approximates $\hat{P}(\mathbf{u})$ by $Q^{new}(\mathbf{u})$ such that $\bar{t}_i(\mathbf{u}) \sim Q^{new}(\mathbf{u})/Q^{\setminus i}(\mathbf{u})$ belongs to the assumed family.

Following steps (2-4), *EP* updates the influence of site $t_i(\mathbf{u})$.The repetitions of these steps lead to good approximations although convergence is not guaranteed.

Taking in consideration the closeness properties of the Gaussian family *EP* seems a well suited method for approximating $P(\mathbf{u}|D)$ because we only have to choose $\bar{t}_i(\mathbf{u})$ and $Q^{new}(\mathbf{u})$ as Gaussians densities to "make" this method "work".

Now suppose t_i–s depends only on a subset of parameters say $I_i \subset \{1, \ldots, n\}$ and let $R_i = \{1, \ldots, n\} \setminus I_i$ – both ordered – then $t_i(\mathbf{u}) = t_i(\mathbf{u}_{I_i})$. When updating $\bar{t}_i$ one has

$$\begin{aligned} \hat{P}(\mathbf{u}) &\propto Q^{\setminus i}(\mathbf{u}) t_i(\mathbf{u}_{I_i}) \\ &\propto Q^{\setminus i}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i}) Q^{\setminus i}(\mathbf{u}_{I_i}) t_i(\mathbf{u}_{I_i}) \\ &\propto Q^{\setminus i}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i}) \hat{P}(\mathbf{u}_{I_i}) \end{aligned}$$

and has to minimize

$$\begin{aligned} \mathrm{D}[\hat{P}(\mathbf{u})||Q^{new}(\mathbf{u})] &= \mathrm{D}[\hat{P}(\mathbf{u}_{I_i})||Q^{new}(\mathbf{u}_{I_i})] \\ &\quad + \mathrm{E}_{\mathbf{u}_{I_i} \sim \hat{P}(\mathbf{u}_{I_i})} \left[\mathrm{D}[Q^{\setminus i}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i})||Q^{new}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i})] \right] \end{aligned}$$

As the minimum of the positive second term of the left hand side is 0 when $Q^{\setminus i}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i}) = Q^{new}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i})$ and this puts no constraint criteria on the first term the minimizations means finding the moments up to second order of $\hat{P}(\mathbf{u}_{I_i}) \propto$

$Q^{\setminus i}(\mathbf{u}_{I_i})t_u(\mathbf{u}_{I_i})$ leading to $Q^{new}(\mathbf{u}) = Q^{\setminus i}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i})Q^{new}(\mathbf{u}_{I_i})$ and so the site approximation $\bar{t}_i(\mathbf{u}) = \bar{t}_i(\mathbf{u}_{I_i}) \propto Q^{new}(\mathbf{u}_{I_i})/Q^{\setminus i}(\mathbf{u}_{I_i})$ depends on the same set of hidden variables as its corresponding site. Now its is easy to show that $Q^{new}(\mathbf{u}_{R_i}|\mathbf{u}_{I_i}) = Q(\mathbf{u}_{R_i}|\mathbf{u}_{I_i})$ and thus the update step has a local nature.

Assuming that $Q(\mathbf{u}) = N(\mathbf{u}; \mathbf{h}, \mathbf{A})$ the a approximation if the first two moments of $t_i(\mathbf{u}_{I_i})Q^{\setminus i}(\mathbf{u}_{I_i})$ cannot be done analytically and we employed Gauss-Hermite quadrature method – see for example Coman [**?**]. Applying change of variables one can factorize the weight function $N(\mathbf{u}_{I_i}; [\mathbf{h}^{\setminus i}]_{I_i}, [\mathbf{A}^{\setminus i}]_{I_i})$ in order to get a tractable quadrature formula. Let $\mathbf{x}$ and $\mathbf{w}$ be the d-th order Gauss-Hermite nodes and weights of $N(\cdot; 0, 1)$ and σ an element of the Descartes product $\{1, \dots, d\}^{|I_i|}$. Using a Cholesky decomposition $[\mathbf{A}^{\setminus i}]_{I_i} = \mathbf{L}\mathbf{L}^T$, and denoting $\mathbf{m} = [\mathbf{h}^{\setminus i}]_{I_i}$, the approximation formula for the normalization constant of $t_i(\mathbf{u}_{I_i})N(\mathbf{u}_{I_i}; [\mathbf{h}^{\setminus i}]_{I_i}, [\mathbf{A}^{\setminus i}]_{I_i})$ is given by:

$$Z_i \simeq \sum_{\sigma} \prod_{j} w_{\sigma}^{j}\, t_i(\mathbf{L}\mathbf{x}_{\sigma} + \mathbf{m})$$

and thus the approximation of first and second moments is straightforward. Unfortunately this quadrature method scales exponentially in $|I_i|$ which makes is practical only for very small values only.

Although we have used the minimization of the KL distance as approximation method, in order to avoid calculation of moments we can use a "hybrid" method: to do the last step of local approximation with a Laplace-type approximation which is plausible if $t_i(\mathbf{u}_{I_i})Q^{\setminus i}(\mathbf{u}_{I_i})$ is strictly log-concave in $\mathbf{u}_{I_i}$:

$$\begin{aligned}
\hat{\mathbf{m}} &\simeq \operatorname{argmax}_{\mathbf{u}_{I_i}} \left\{ \log(t_i(\mathbf{u}_{I_i})) - \frac{1}{2}(\mathbf{u}_{I_i} - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{u}_{I_i} - \mathbf{m}) \right\} \\
\hat{\mathbf{V}} &\simeq \left(\hat{\mathbf{W}} + \mathbf{V}^{-1} \right)^{-1} \\
\log Z_i &\simeq \log(t_i(\hat{\mathbf{m}})) - \frac{1}{2}(\hat{\mathbf{m}} - \mathbf{m})^T \mathbf{V}^{-1} (\hat{\mathbf{m}} - \mathbf{m}) - \frac{1}{2}\log|\mathbf{I} + \hat{\mathbf{W}}\mathbf{V}|
\end{aligned}$$

where $\hat{\mathbf{W}} = -(\nabla\nabla^T)_{\mathbf{u}_{I_i}} t_i(\mathbf{u}_{I_i})|_{\mathbf{u}_{I_i} = \hat{\mathbf{m}}}$. The approximation can be done with a Newton-Raphson method and it takes roughly $O(|I_i|^3)$ time.

4. An approach to multi-class problem

In the followings we present the extension of binary classification problem to multiple classes. We use the representation of Barber and Williams [?] to build the model.

In order to avoid dealing with a great number of data resulting from a multi-output Gausssian process one can model the C-class case by choosing C independent priors and taking account only their posterior cross-correlations which is realized trough the coupling(s) in the likelihood terms. In order to avoid confusions we settle first some notational conventions: the subscript indexing is used for referring to input locations while the upper indexing is be used for referring to class type indexing – $\mathbf{u}_i = (u_i^1, u_i^2, \ldots, u_i^C)$ and $\mathbf{u}^{(c)} = (u_1^{(c)}, u_2^{(c)}, \ldots, u_n^{(c)})$, we use $\mathbf{u} = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(C)})$ and the corresponding prior covariance matrix $\mathbf{K} = \text{diag}(\mathbf{K}^{(c)})_c$.

As we have seen the two-class case, finding a likelihood term for finite number of outputs is not an easy task and one of the easiest ways to solve the problem is to turn it into a regression-like one. A multi-dimensional extension of the logistic function, the *softmax* function $\pi(\mathbf{u}) = \exp(\mathbf{u})/\mathbf{1}^T \exp(\mathbf{u})$ is used to model the class probabilities. Modelling y_i belonging to class c by $y_j = \mathbf{e}_c$ one defines the likelihood:

$$P(y|\mathbf{u}) = y^T \pi(\mathbf{u}).$$

The likelihood $\log P(y|\mathbf{u})$ is not strictly log-concave – due to $1^T\pi(\mathbf{u}) = 1$ it is strictly log-concave only on $\{\mathbf{1}\}^{\perp}$. This could constitute a risk to the approximation in cases when the local marginals of the likelihood approximations collapse or are close to it since both of the approximations processes presented above – minimization of KL distance and Laplace method – rely on this property.

In order to apply the EP procedure presented earlier we set the (site) likelihood:

$$t_i(\mathbf{u}_i) = P(y_i|\mathbf{u}_i) = (\pi(\mathbf{u}_i))_c$$

where c is the nonzero element of y_i and the use site approximations $\bar{t}_i(\mathbf{u}_i) \sim N(\mathbf{u}_i;\mathbf{m}_i,\mathbf{V}_i)$. Using the notation $\mathbf{V} = (\mathrm{diag}(\mathbf{V}_i^{(s,t)})_i)_{t,s=1,\ldots,C}$ the Gaussian approximation of the posterior $P(\mathbf{u}|D)$ has the covariance $\mathbf{K} - \mathbf{K}(\mathbf{K}+\mathbf{V})^{-1}\mathbf{K}$ and mean value $(\mathbf{I}+\mathbf{V}\mathbf{K}^{-1})^{-1}(\mathbf{m}^{(c)})_c$, thus the the prediction $\mathbf{u}_*$ is normal with:

$$\begin{aligned}\mathrm{E}[\mathbf{u}_*|D] &= \mathbf{K}_*^T(\mathbf{V}+\mathbf{K})^{-1}(\mathbf{m}^{(c)})_c\\ \mathrm{Cov}[\mathbf{u}_*|D] &= \mathrm{diag}\left(K^{(c)}(\mathbf{x}_*,\mathbf{x}_*)\right)_c - \mathbf{K}_*^T(\mathbf{V}+\mathbf{K})^{-1}\mathbf{K}_*\end{aligned}$$

where we used the notation $\mathbf{K}_* = \mathrm{diag}\left(\mathbf{k}^{(c)}(\mathbf{x}_*)\right)_c$. In order to calculate the intractable $P(y_* = c|\mathbf{x}_*, D)$ one has to apply once again numerical quadrature formulas.

5. **Experiments**

We implemented the two local approximation methods described in section 3. We built upon the *OGP toolbox* developed by Csató (see, [?]) which was implemented based on Csató and Opper [?] and is publicly available with full documentation. The *OGP toolbox* provides *a sparse approximation* method for a variety of likelihoods – user defined ones as well. This makes the toolbox easily applicable for artificial or real world problems that employ *Gaussian Process models.* The Gauss-Hermite Quadrature rule was ran using 7 nodes. Increasing its order did not lead to a significant change in accuracy. In fact fewer nodes proved to be suitable, the reason for it might be the good convergence properties of the Taylor expansion of the likelihood function. Heuristics like employing the symmetry property of nodes and weights when using Gauss-Hermite quadrature formulas – as it was pointed out by Seeger and Jordan [?] – can be used but significant improvement in time-performance cannot be achieved without further making use of the likelihood structure. However, factorizing likelihood sites, or likelihoods that speed up the Gaussian quadrature routine might lead to worse performance because of weak couplings in the variables of the likelihood.

The "local" Laplace method was implemented using Newton-Raphson method for $\mathbf{a} = \mathbf{V}^{-1}(\mathbf{s}-\mathbf{m})$ with the update equation

$$\mathbf{a}^{new} = (\mathbf{I}+\mathbf{W}\mathbf{V})^{-1}\left(\mathbf{W}\mathbf{V}\mathbf{a} + \pi(\mathbf{s}) - \mathbf{e}_c\right)$$

where we have used the notations from section 3. In each step we have used BiCG (implemented in Matlab) to solve the linear system. The method takes around $O(n_{bicg}n_{nr}C^2)$ time where n_{bicg} and n_{nr} are denoting the number of BiCG resp. Newton-Raphson steps.

The plots on figure 5 show the error rates we achieved on a 3-class data set using a spherical (the scaling parameters v_i are equal) square exponential kernel from equation 5. We implemented hard classification rules i.e. an item belongs to the class which has the greatest probability – this was done in order to compare our results with the benchmark *5NN* classifier. For real world problems however, one could exploit the multi-output probabilistic outputs returned by the system – the freedom of choice is significantly more.

The basis for comparison was the *5NN* rule. The data was generated in the following way: we generated from 9 Gaussians with mean values chosen randomly from $[0,\ldots,10]^2$ and labelled these randomly. The resulting data were preprocessed: we whitened the data. We used 1000 samples splitting them in 1/2 train/test ratio. Figure 5 shows hard-classification boundaries on a data set of 250 examples.

6. Conclusions/Further research

The presented methods outperformed *5NN* in cases when the scaling parameters did not have extreme values. We aim to further develop and implement *hyper-parameter optimization* methods, as well as approximations for *posterior probabilities* to be used with *Markov Chain Monte Carlo, Hybrid Monte Carlo* type methods for a "complete" Bayesian inference like in section 1 to integrate out the posterior process and the hyperparameters from the model. Our interest in developing further local approximation methods is still active.

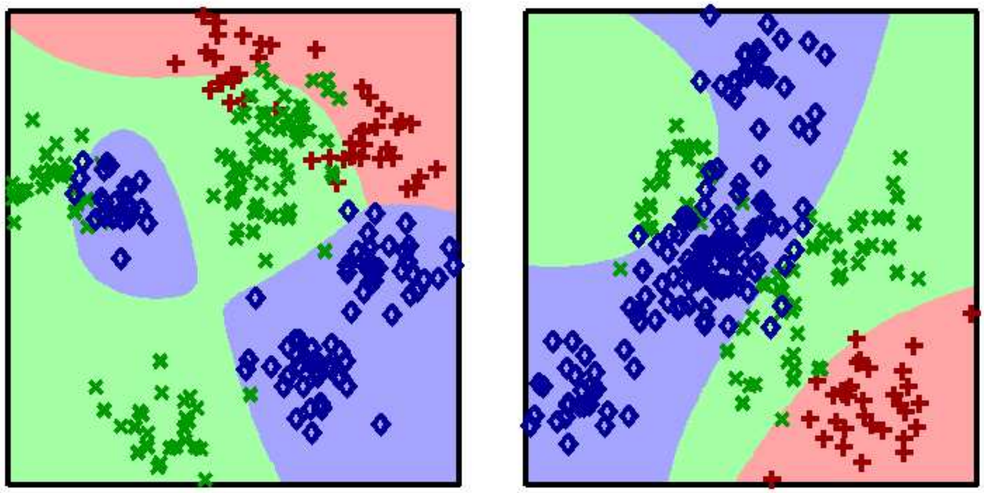


FIGURE 2. Hard classification discriminant curves for a 3-class case with 250 samples.

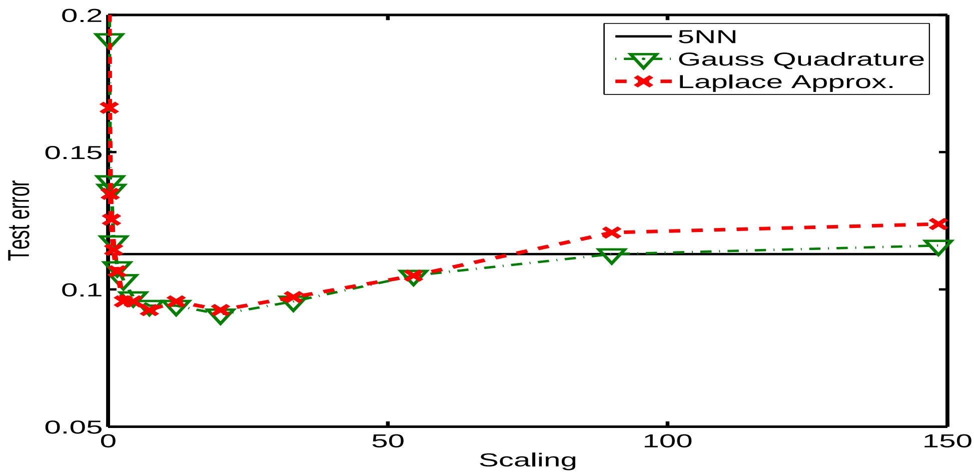


FIGURE 3. Performance of a multi-class GP methods with spherical square exponential kernel on data set with 3 classes.

BABEŞ-BOLYAI UNIVERSITY, FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, STR. KOGĂLNICEANU, 1 CLUJ-NAPOCA, ROMANIA

E-mail address: `csekeb@math.ubbcluj.ro`, `lehel.csato@cs.ubbcluj.ro`

DATA DEPENDENCE FOR SOME INTEGRAL EQUATIONS VIA WEAKLY PICARD OPERATORS

ION MARIAN OLARU

Abstract. In this paper we study data dependence for the following integral equation:

$$u(x) = h(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K(x, s, u(\theta_1 s_1, \cdots, \theta_m s_m))ds,$$

$$x \in \prod_{i=1}^{m} [0, b_i], \theta_i \in (0, 1), (\forall) i = \overline{1, m}$$

by using c-WPOs.

1. **Introduction**

Let (X, d) be a metric space and $A : X \to X$ an operator. We shall use the following notations:

$F_A := \{x \in X \mid A(x) = x\}$ the fixed points set of A.

$I(A) := \{Y \in P(X) \mid A(Y) \subset Y\}$ the family of the nonempty invariant subsets of A.

$A^{n+1} = A \circ A^n, A^0 = 1_X, A^1 = A, n \in N.$

Definition 1.1. [1] *An operator A is weakly Picard operator (WPO) if the sequence*

$$(A^n(x))_{n \in N}$$

converges, for all $x \in X$ and the limit (which depend on x) is a fixed point of A.

Definition 1.2. [1] *If the operator A is WPO and $F_A = \{x^*\}$ then by definition A is Picard operator.*

Received by the editors: 10.09.2005.

2000 *Mathematics Subject Classification.* 34K10, 47H10.

Key words and phrases. Picard operators, weakly Picard operators, fixed points, data dependence.

Definition 1.3. [1] *If A is WPO, then we consider the operator*

$$A^\infty : X \to X, A^\infty(x) = \lim_{n\to\infty} A^n(x).$$

We remark that $A^\infty(X) = F_A$.

Definition 1.4. [1] *Let be A an WPO and $c > 0$. The operator A is c-WPO if*

$$d(x, A^\infty(x)) \leq c \cdot d(x, A(x)).$$

We have the following characterization of the WPOs:

Theorem 1.1. [1]*Let (X, d) be a metric space and $A : X \to X$ an operator. The operator A is WPO (c-WPO) if and only if there exists a partition of X,*

$$X = \bigcup_{\lambda\in\Lambda} X_\lambda$$

such that

(a) $X_\lambda \in I(A)$

(b) $A \mid X_\lambda : X_\lambda \to X_\lambda$ *is a Picard (c-Picard) operator, for all* $\lambda \in \Lambda$.

For the class of c-WPOs we have the following data dependence result:

Theorem 1.2. [1] *Let (X, d) be a metric space and $A_i : X \to X, i = \overline{1,2}$ operators. We suppose that:*

(i) *the operator A_i is $c_i - WPO$, $i = \overline{1,2}$.*

(ii) *there exists $\eta > 0$ such that*

$$d(A_1(x), A_2(x)) \leq \eta, (\forall)x \in X.$$

Then

$$H(F_{A_1}, F_{A_2}) \leq \eta\ max\{c_1, c_2\}.$$

Here stands for Hausdorff-Pompeiu functional.

We have:

Lemma 1.1. [1], [3] *Let* $(X, d, \leq)$ *be an ordered metric space and* $A : X \to X$ *an operator such that:*

a) A *is monotone increasing.*

b) A *is WPO.*

Then the operator A^∞ *is monotone increasing.*

Lemma 1.2. [1], [3] *Let* $(X, d, \leq)$ *be an ordered metric space and* $A, B, C : X \to X$ *such that :*

(i) $A \leq B \leq C$.

(ii) the operators A,B,C are WPOs.

(iii) the operator B is monotone increasing.

Then

$$x \leq y \leq z \Longrightarrow A^\infty(x) \leq B^\infty(y) \leq C^\infty(z).$$

2. **Main results**

Data dependence for functional integral equations was studied [1], [2], [3]. In what follows we consider the integral equation

$$u(x) = h(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K(x, s, u(\theta_1 s, \cdots, \theta_m s))ds, \tag{1}$$

where

$$x \in \prod_{i=1}^{m}[0, b_i], \theta_i \in (0, 1), (\forall) i = \overline{1, m}.$$

We denote $D = \prod_{i=1}^{m}[0, b_i]$.

Theorem 2.1. *We suppose that:*

(i) $h \in C(D \times R)$ *and* $K \in C(D \times D \times R)$.

(ii) $h(0, \alpha) = \alpha, (\forall)\alpha \in R$.

(iii) there exists $L_K > 0$ *such that*

$$|K(x, s, u_1) - K(x, s, u_2)| \leq L_K |u_1 - u_2|,$$

for all $x, s \in D$ *and* $u_1, u_2 \in R$.

In these conditions the equation (1) has in C(D) an infinity of solutions.

Moreover if

(iv) $h(x, \cdot)$ *and* $K(x, s, \cdot)$ *are monotone increasing for all* $x, s \in D$

then if u *and* v *are solutions of the equation (1) such that* $u(0) \leq v(0)$ *we have* $u \leq v$.

Proof. Consider the operator

$$A : (C(D), \|\cdot\|_B) \to (C(D), \|\cdot\|_B),$$

$$A(u)(x) := h(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K(x, s, u(\theta_1 s, \cdots, \theta_m s))ds.$$

Here $\|u\|_B = \max_{x \in D} |u(x)| e^{-\sum_{i=1}^{m} x_i}$.

Let $\lambda \in R$ and $X_\lambda = \{u \in C(D) \mid u(0) = \lambda\}$. Then

$$C(D) = \bigcup_{\lambda \in R} X_\lambda.$$

is a partition of $C(D)$ and $X_\lambda \in I(A)$, for all $\lambda \in R$.

For all $u, v \in X_\lambda$, we have have

$$|A(u)(x) - A(v)(x)| \leq \frac{L_K}{\tau^m \theta_1 \cdots \theta_m} e^{\tau \sum_{i=1}^{m} x_i} \|u - v\|_B .$$

So the restriction of the operator A on X_λ is a c-Picard operator with $c = (1 - \frac{L_K}{\tau^m \theta_1 \cdots \theta_m})^{-1}$, for a suitable choices of τ such that $\frac{L_K}{\tau^m \theta_1 \cdots \theta_m} < 1$.
If $u \in R$ then we denote by $\widetilde{u}$ the constant operator

$$\widetilde{u} : C(D) \to C(D)$$

defined by

$$\widetilde{u}(t) = u.$$

If $u, v \in C(D)$ are the solutions of (1) with $u(0) \leq v(0)$ then $\widetilde{u(0)} \in X_{u(0)}, \widetilde{v(0)} \in X_{v(0)}$.

By lema 1.1 we have that

$$\widetilde{u(0)} \leq \widetilde{v(0)} \Longrightarrow A^\infty(\widetilde{u(0)}) \leq A^\infty(\widetilde{v(0)}).$$

But

$$u = A^\infty(\widetilde{u(0)}), v = A^\infty(\widetilde{v(0)}).$$

So, $u \leq v$.

Theorem 2.2. *Let $h_i \in C(D \times R)$ and $K_i \in C(D \times D \times R)$, $i = \overline{1,3}$ satisfy the conditions (i), (ii), (iii) from the Theorem 2.1. We suppose that*

(a) $h_2(x, \cdot)$ and $K_2(x, s, \cdot)$ are monotone increasing, for all $x, s \in D$.

(b) $h_1 \leq h_2 \leq h_3$ and $K_1 \leq K_2 \leq K_3$.

Let u_i be a solution of the equation (1) corresponding to h_i and K_i.

Then

$$u_1(0) \leq u_2(0) \leq u_3(0) \ \ \textit{imply} \ \ u_1 \leq u_2 \leq u_3.$$

Proof. The proof follows from Lemma 1.2.

For studding of data dependence we consider the following equations:

$$u(x) = h_1(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K_1(x, s, u(\theta_1 s_1, \cdots, \theta_m s_m))ds \tag{2}$$

$$u(x) = h_2(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K_2(x, s, u(\theta_1 s_1, \cdots, \theta_m s_m))ds \tag{3}$$

Theorem 2.3. *We consider (2), (3) under the following conditions:*

(i) $h_i \in C(D \times R)$ and $K_i \in C(D \times D \times R)$, $i = \overline{1,2}$.

(ii) $h_i(0, \alpha) = \alpha, (\forall)\alpha \in R$, $i = \overline{1,2}$.

(iii) there exists $L_{K_i} > 0$, $i = \overline{1,2}$ such that

$$|K_i(x, s, u_1) - K_i(x, s, u_2)| \leq L_{K_i}|u_1 - u_2|, \quad i = \overline{1,2}$$

for all $x, s \in D$ *and* $u_1, u_2 \in R$.

(iv) $(\exists)\eta_1, \eta_2 > 0$ *such that*

$$|h_1(x,u) - h_2(x,u)| \leq \eta_1,$$

$$|K_1(x,s,u) - K_2(x,s,u)| \leq \eta_2,$$

$(\forall)x, s \in D, u \in R$.

If $S_1,\ S_2$ *are the solutions sets of the equations (2), (3), then we have:*

$$H(S_1, S_2) \leq (\eta_1 + \eta_2 \prod_{i=1}^{m} b_i) \max_{i=\overline{1,2}} \left\{ \frac{1}{1 - \dfrac{L_{K_i}}{\tau^m \theta_1 \cdots \theta_m}} \right\},$$

for $\tau > \max\limits_{i=\overline{1,2}} \left\{ \sqrt[m]{\dfrac{L_{K_i}}{\theta_1 \cdots \theta_m}} \right\}$.

Proof. We consider the following operators:

$$A_i : (C(D), \|\cdot\|_B) \to (C(D), \|\cdot\|_B)),$$

$$A_i u(x) := h_i(x, u(0)) + \int_0^{x_1} \cdots \int_0^{x_m} K_i(x, s, u(\theta_1 s, \cdots, \theta_m s)) ds,\ \ i = \overline{1,2}$$

From:

$$|A_1(u)(x) - A_2(u)(x)| \leq |h_1(x, u(0)) - h_2(x, u(0))| +$$

$$\int_0^{x_1} \cdots \int_0^{x_m} \|K_1(x, s, u(\theta_1 s \cdot \theta_m s)) - K_2(x, s, u(\theta_1 s, \cdots \theta_m s))\| ds \leq$$

$$\leq \eta_1 + \eta_2 \prod_{i=1}^{m} b_i.$$

we have that $\|A(u) - A(v)\|_B \leq \eta_1 + \eta_2 \prod\limits_{i=1}^{m} b_i$

Like in the proof of Theorem 1.2 we obtain that the operators $A_i, i = \overline{1,2}$ are c_i-WPOs with $c_i = \left(1 - \dfrac{L_{K_i}}{\tau^m \theta_1 \cdots \theta_m}\right)^{-1}$, $\tau > \max\limits_{i=\overline{1,2}} \left\{ \sqrt[m]{\dfrac{L_{K_i}}{\theta_1 \cdots \theta_m}} \right\}$.

From this and by Theorem 1.2. we have conclusion.

References

[1] I. A. Rus, *Weakly Picard operators and applications*, Seminar on Fixed Point Theory, Cluj-Napoca, vol. 2, 2001, 41-57.

[2] I.A.Rus, *Generalized contractions*, Seminar on Fixed Point Theory, No. 3, 1983, 1-130.

[3] I. A. Rus, *Functional-Differential equation of mixed type, via weakly Picard operator*, Seminar on Fixed Point Theory Cluj-Napoca, vol. 3, 2002, 335-345.

Department of Mathematics ,
University "Lucian Blaga " ,
Sibiu, Romania
E-mail address: olaruim@yahoo.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 3, September 2005

CYCLIC REPRESENTATIONS AND PERIODIC POINTS

GABRIELA PETRUŞEL

Abstract. The purpose of this note is to give some existence results of periodic points for some classes of single-valued operators. The fixed point structures technique and an abstract periodic point lemma given by I. A. Rus are used.

1. **Introduction**

Throughout this paper, we will use the notations and terminologies in [4], [5].

Let (X, d) be a metric space and $f : X \to X$ an operator. By $F_f := \{x \in X |\ x = f(x)\}$ we will denote the fixed point set of the operator f.

We will also use the following symbols:

$P(X) := \{Y \subseteq X | Y \neq \emptyset\}$, $P_{cl}(X) := \{Y \in P(X) |\ Y \text{ is closed}\}$, $P_{cp}(X) := \{Y \in P(X) |\ Y \text{ is compact}\}$ and $P_b(X) := \{Y \in P(X) |\ Y \text{ is bounded }\}$.

Let X, Y be nonempty sets. We will denote by $\mathbb{M}(X, Y)$ the set of all single-valued operators from $f : X \to Y$. If $X = Y$ then $\mathbb{M}(Y) := \mathbb{M}(Y, Y)$.

Definition 1.1. Let X be a nonempty set. By definition (see [4]), the triple $(X, S(X), M)$ is a fixed point structure (briefly f. p. s.) if:

(i) $S(X) \subset P(X), S(X) \neq \emptyset$

(ii) $M : P(X) \to \bigcup_{Y \in P(X)} \mathbb{M}(Y)$ is a selection operator, such that if $Z \subset Y, Z \neq \emptyset$ then $M(Z) \supset \{f_{|_Z} | f \in M(Y), Z \in I(f)\}$

(iii) for each $Y \in S(X)$ and $f \in M(Y)$ we have that $F_f \neq \emptyset$.

Received by the editors: 12.09.2005.

2000 *Mathematics Subject Classification.* 47H10, 54H25.

Key words and phrases. Periodic point, fixed point, cyclic representation, fixed point structure.

Definition 1.2. (I. A. Rus [5]) Let X be a nonempty set and $f : X \to X$ an operator. By definition, $X = \bigcup_{i=1}^{m} X_i$ (where $X_i \subset X$, for each $i \in \{1, 2, \cdots, m\}$) is a cyclic representation of X with respect to f if $f(X_1) \subset X_2, \cdots, f(X_{m-1}) \subset X_m, f(X_m) \subset X_1$.

In [3], W. A. Kirk, P. S. Srinivasan, P. Veeramani proved some fixed point theorems for single-valued operators satisfying some cyclical contractive assumptions. Then, I. A. Rus generalize these results in terms of the fixed point structures (see [5]).

Also, in Rus [5], the following periodic points lemma is given:

Lemma 1.3. *Let $(X, S(X), M)$ be a fixed point structure, where X is a nonempty set. Let $A_i \in P(X)$, for each $i \in \{1, 2, \cdots, m\}$. Denote $Y := \bigcup_{i=1}^{m} A_i$ and consider $f : Y \to Y$. Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ is a cyclic representation of Y with respect to f;

(ii) $A_i \in S(X)$ for some $i \in \{1, 2, \cdots, m\}$;

(iii) $g_1, g_2 \in M(Y)$ implies $g_1 \circ g_2 \in M(Y)$.

Then $F_{f^m} \neq \emptyset$.

The purpose of this paper is to give some applications of the previous lemma.

2. Periodic points for Knaster-Tarski type operators

Let $(X, \leq)$ be an ordered set,

$S(X) := \{Y \in P(X) | (Y, \leq)$ is a complete latice$\}$ and

$M(Y) := \{f : Y \to Y |$ f is increasing $\}$. Then $(X, S(X), M)$ is a f. p. s. (Knaster-Tarski, see [1]).

Then, by applying Lemma 1.3., one obtains:

Theorem 2.1. *Let $(X, \leq)$ be an ordered set, $A_i \in P(X)$, for $i \in \{1, 2, \cdots, m\}$, such that there is $i_0 \in \{1, 2, \cdots, m\}$ with A_{i_0} a complete lattice. Denote $Y := \bigcup_{i=1}^{m} A_i$ and consider $f : Y \to Y$. Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ is a cyclic representation of Y with respect to f;

(ii) $f(x_1) \leq f(x_2)$*, for each* $x_1 \in A_i$ *and each* $x_2 \in A_{i+1}$*,* $(i \in \{1, 2, \cdots, m\})$ *with* $x_1 \leq x_2$ *(where* $A_{m+1} = A_1$*).*

Then $F_{f^m} \neq \emptyset$*.*

Proof. Let us remark that the fixed point structure of Knaster-Tarski satisfies the conditions (i)-(iii) in Lemma 1.3. □

3. Periodic points for generalized contractions

Let (X, d) be a complete metric space. Then the operator $f : X \to X$ is called a φ-contraction if there exists a comparison function φ (i. e. $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is non-decreasing and $(\varphi^n(t))_{n\in\mathbb{N}} \to 0$, as $n \to +\infty$, for each $t > 0$) such that

$$d(f(x_1), f(x_2)) \leq \varphi(d(x_1, x_2)), \text{ for all } x_1, x_2 \in X.$$

If we consider $S(X) := P_{cl}(X)$ and one define
$M(Y) := \{f : Y \to Y |\text{exists a comparison function } \varphi \text{ such that f is a } \varphi\text{-contraction}\}$, then $(X, S(X), M)$ is a f. p. s. (see Rus [4])).

The following result follow now from Lemma 1. 3.

Theorem 3.1. *Let* (X, d) *be a complete metric space,* $A_i \in P(X)$*, for* $i \in \{1, 2, \cdots, m\}$*, such that there is* $i_0 \in \{1, 2, \cdots, m\}$ *with* $A_{i_0} \in P_{cl}(X)$*. Denote* $Y := \bigcup_{i=1}^{m} A_i$ *and consider* $f : Y \to Y$*. Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ *is a cyclic representation of* Y *with respect to* f*;*

(ii) there exists a comparison function φ *such that*

$$d(f(x_1), f(x_2)) \leq \varphi(d(x_1, x_2)), \text{ for all } x_1 \in A_i, \text{ and } x_2 \in A_{i+1}, i \in \{1, 2, \cdots, m\},$$

where $A_{m+1} = A_1$*.*

Then $F_{f^m} \neq \emptyset$*.*

Proof. Let $g_1, g_2 \in M(Y)$. It follows that there exist the comparison functions φ_1, φ_2 such that g_i is a φ_i-contraction, for $i \in \{1, 2\}$. Since the composition of two comparison functions is a comparison function, it follows immediately that the condition (iii) in Lemma 1.3. holds. □

4. **Periodic points for contractive operators**

Let (X, d) be a metric space. Then the operator $f : X \to X$ is called contractive if $d(f(x_1), f(x_2)) < d(x_1, x_2)$, for all $x_1, x_2 \in X, x_1 \neq x_2$. If $S(X) := P_{cl}(X)$ and $M(Y) := \{f : Y \to Y|\text{ f is contractive }\}$. If (X, d) is a compact metric space, then $(X, S(X), M)$ is a f. p. s. (Nemytzki-Edelstein, see [4])).

From Lemma 1.3. we have:

Theorem 3.1. *Let (X, d) be a compact metric space, $A_i \in P(X)$, for $i \in \{1, 2, \cdots, m\}$, such that there is $i_0 \in \{1, 2, \cdots, m\}$ with $A_{i_0} \in P_{cl}(X)$. Denote $Y := \bigcup_{i=1}^{m} A_i$ and consider $f : Y \to Y$. Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ is a cyclic representation of Y with respect to f;

(ii) $d(f(x_1), f(x_2)) < d(x_1, x_2)$, for each $x_1 \in A_i$, and $x_2 \in A_{i+1}$, with $x_1 \neq x_2$, for $i \in \{1, 2, \cdots, m\}$, where $A_{m+1} = A_1$.

Then $F_{f^m} \neq \emptyset$.

Proof. Let g, h be contractive operators. Then, for any two elements x_1, x_2 from X, with $x_1 \neq x_2$, we have: $d((g \circ h)(x_1), (g \circ h)(x_2)) \leq d(h(x_1), h(x_2)) < d(x_1, x_2)$. hence all the conditions of Lemma 1.3. are satisfy. □

5. **Periodic points for nonexpansive operators**

Let (X, d) be an uniformly convex Banach space. Then the operator $f : X \to X$ is called nonexpansive if $d(f(x_1), f(x_2)) \leq d(x_1, x_2)$, for all $x_1, x_2 \in X$. If $S(X) := P_{b,cl,cv}(X)$ and $M(Y) := \{f : Y \to Y|\text{ f is nonexpansive }\}$. Then $(X, S(X), M)$ is a f. p. s.(Browder - Ghöde - Kirk, see [1], [2]).

For nonexpansive operators we have:

Theorem 4.1. *Let X be an uniformly convex Banach space, $A_i \in P(X)$, for $i \in \{1, 2, \cdots, m\}$, such that there is $i_0 \in \{1, 2, \cdots, m\}$ with $A_{i_0} \in P_{b,cl,cv}(X)$. Denote $Y := \bigcup_{i=1}^{m} A_i$ and consider $f : Y \to Y$. Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ is a cyclic representation of Y with respect to f;

(ii) $d(f(x_1), f(x_2)) \leq d(x_1, x_2)$, *for each* $x_1 \in A_i$, *and* $x_2 \in A_{i+1}$, *for* $i \in \{1, 2, \cdots, m\}$, *where* $A_{m+1} = A_1$.

Then $F_{f^m} \neq \emptyset$.

Proof. Since the composition of two nonexpansive operators is a nonexpansive operator, we remark that the condition (iii) in Lemma 1.3 holds. The conclusion follows now by Lemma 1.3. □

6. Periodic points for Perov type operators

Let (X, d) be a generalized metric space, in the sense that $d(x, y) \in \mathbb{R}^k$. The operator $f : X \to X$ is called a Perov type contraction (or S-contraction) if $S \in \mathcal{M}_{kk}(\mathbb{R})$, with $S^n \to 0$, as $n \to +\infty$, such that $d(f(x_1), f(x_2)) \leq S \cdot d(x_1, x_2))$, for all $x_1, x_2 \in X$. If $S(X) := P_{cl}(X)$ and $M(Y) := \{f : Y \to Y|$ f is a Perov contraction $\}$. Then $(X, S(X), M)$ is a f. p. s. (Perov, see [4])).

In the setting of the Perov's f. p. s., Lemma 1. 3. gives us:

Theorem 5. 1. *Let* (X, d) *be a complete generalized metric space,* $A_i \in P(X)$, *for* $i \in \{1, 2, \cdots, m\}$, *such that there is* $i_0 \in \{1, 2, \cdots, m\}$ *with* $A_{i_0} \in P_{cl}(X)$. *Denote* $Y := \bigcup_{i=1}^{m} A_i$ *and consider* $f : Y \to Y$. *Suppose that:*

(i) $Y := \bigcup_{i=1}^{m} A_i$ *is a cyclic representation of* Y *with respect to* f*;*

(ii) There exists a matrix $S \in \mathcal{M}_{kk}(\mathbb{R})$, *with* $S^n \to 0$, *as* $n \to +\infty$ *such that* $d(f(x_1), f(x_2)) \leq S \cdot d(x_1, x_2)$, *for each* $x_1 \in A_i$, *and* $x_2 \in A_{i+1}$, *for* $i \in \{1, 2, \cdots, m\}$, *where* $A_{m+1} = A_1$.

Then $F_{f^m} \neq \emptyset$.

References

[1] J. Dugundji, A. Granas, *Fixed point theory*, Springer Verlag, Berlin, 2003.

[2] W. A. Kirk, B. Sims (editors), *Handbook of metric fixed point theory*, Kluwer Acad. Publ., Dordrecht, 2001.

[3] W. A. Kirk, P. S. Srinivasan, P. Veeramani, *Fixed poinys for mappings satisfying cyclical contractive conditions*, Fixed Point Theory, **4**(2003), 79-89.

[4] I. A. Rus, *Generalized contractions and applications*, Cluj Univ. Press, 2001.

[5] I. A. Rus, *Cyclic representations and fixed points*, Annals of the Tiberiu Popoviciu Seminar, 2005, to appear.

[6] I. A. Rus, A. Petruşel, G. Petruşel, *Fixed Point Theory 1950-2000: Romanian Contributions*, House of the Book of Science, Cluj-Napoca, 2002.

Babeş-Bolyai University, Faculty of Mathematics
and Computer Science, Str. Kogălniceanu, 1
Cluj-Napoca, Romania
E-mail address: gabip@math.ubbcluj.ro