Anul L **2005**

S T U D I A
UNIVERSITATIS BABEŞ-BOLYAI

MATHEMATICA
2

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

Telefon: 405300

SUMAR – CONTENTS – SOMMAIRE

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

QUANTITATIVE APPROXIMATIONS BY USING SCALING TYPE FUNCTIONS

OCTAVIAN AGRATINI

Abstract. The focus of the paper is to study a class of linear positive operators constructed by using a quasi-scaling type function. Jackson type inequalities are established in the framework of different function spaces.

1. **Introduction**

In Approximation Theory an interesting tool with a rich mathematical content and great potential for applications, is materialized by sequences of linear positive operators generated by a scaling type function.

The aim of the present note is to investigate a general class $(L_k)_{k\in\mathbb{Z}}$ of linear positive operators of wavelet type. Our paper is designed as follows. Following [1], in Section 2 we recall the construction of L_k, $k \in \mathbb{Z}$, operators and we indicate the main notations and results which will be used in the sequel. Further on, in Section 3 we use this class to approximate smooth real valued signals, more precisely, functions which possess derivatives of high order. We establish both pointwise and global estimates of the rate of convergence of our operators. Under additional assumptions, we prove that each $\gamma^{-2}\delta^{-1}L_k$ operator has the degree of exactness equal to 1. The last section is devoted to estimate the approximation of bounded functions by L_k, with the help of a Lipschitz-type maximal function.

Clearly, the research along a certain line can be developed by different angles. We point out that our approach is made by using tools and methods which characterize

Received by the editors: 11.02.2005.

2000 *Mathematics Subject Classification.* 41A36, 41A25.

Key words and phrases. Quasi-scaling type function, linear positive operator, modulus of smoothness, rate of convergence.

the approximation of functions by linear positive operators. Similar results are quite familiar in the Littlewood-Paley and wavelets literature. A good illustration of this can be found in Daubechies' book [4; *Section 6.5*] and elsewhere.

2. **Background and preliminaries**

Setting $\mathbb{N}_0 := \mathbb{N}\cup\{0\}$, we consider a bi-dimensional net $(a_k, \delta j)$, $(k,j) \in \mathbb{Z}\times\mathbb{Z}$, $\delta \in]0,\infty[$, and

$$a_{-k} = a_k^{-1},\ 0 < a_k < a_{k+1},\ \text{for every } k \in \mathbb{N}_0. \tag{1}$$

Clearly, $a_{-(p+1)} < a_{-p}$ for each $p \in \mathbb{N}_0$ and $a_0 = 1$. We point out that the above net generalizes the couples $(2^k, j)$, $(k,j) \in \mathbb{Z}\times\mathbb{Z}$, broad used in the construction of many wavelet type discrete operators. Operating both on the sequence $(a_k)_{k\in\mathbb{N}_0}$ and on the ratio δ, we are able to transform the net in accordance with the problem data and, therefore, it is more flexible then the previous one.

Let $L_{1,loc}(\mathbb{R})$ be the vector space of the real-valued functions defined on $\mathbb{R}$ and locally integrable, i.e. integrable on any compact interval of the real line. We make the following informal definition.

Definition 2.1. *Let $\delta > 0$ be fixed. A function $\varphi : \mathbb{R} \to [0,\infty[$ satisfying the following conditions:*

(i) φ is a bounded function belonging to $L_{1,loc}(\mathbb{R})$,

(ii) a positive constant α exists such that $\operatorname{supp}(\varphi) \subset [-\alpha, \alpha]$, (2)

(iii) a positive constant γ exists with the property

$$\sum_{j=-\infty}^{\infty} \varphi(x+\delta j) = \gamma,\ \textit{for every } x \in \mathbb{R}, \tag{3}$$

is called a scaling function of (δ,γ) type.

Using the sequence $(a_k)_{k\in\mathbb{Z}}$ defined by (1) and a scaling function φ of (δ,γ) type we generate the functions

$$\varphi_{k,j}(x) := \sqrt{a_k}\,\varphi(a_k x + \delta j), \quad x \in \mathbb{R},\ (k,j) \in \mathbb{Z}\times\mathbb{Z}. \tag{4}$$

As usual in wavelet transforms, k is named the dilation index and j is named the translation index. Dilation by larger k compresses the function φ on the x-axis.

Altering j has the effect of sliding the function φ along the x-axis. We mention that condition (3) has nothing to do with the property of an orthogonal scaling function of a multiresolution analysis (MRA), intensively used in the signals theory.

At this point we are in position to introduce the announced sequence of operators.

For every $k \in \mathbb{Z}$ and $f \in L_{1,loc}(\mathbb{R})$ we define the operator L_k as follows

$$(L_k f)(x) := \sum_{j=-\infty}^{\infty} (f, \varphi_{k,j})\varphi_{k,j}(x), \quad x \in \mathbb{R}, \tag{5}$$

where the functions $\varphi_{k,j}$ are given by (4) and $(f, \varphi_{k,j}) = \int_{\mathbb{R}} f(t)\varphi_{k,j}(t)dt$.

As usual, we denote by $C(\mathbb{R})$ ($B(\mathbb{R})$, respectively) the space of all continuous (bounded, respectively) real valued functions on $\mathbb{R}$. The spaces $B(\mathbb{R})$ and $B(\mathbb{R}) \cap C(\mathbb{R})$ can be equipped with the norm $\|\cdot\|_\infty$ of the uniform convergence (briefly, the sup-norm). Also $L_p(\mathbb{R})$, $p \geq 1$, stands for the vector space of all real valued Lebesgue integrable functions defined on $\mathbb{R}$ endowed with the usual norm $\|\cdot\|_{L_p(\mathbb{R})}$. In the Hilbert space of square integrable functions, the inner product is denoted by $(\cdot, \cdot)$.

Examining Definition 2.1 we deduce that φ belongs to the Lebesgue space $L_2(\mathbb{R})$. The same statement is true for $\varphi_{k,j}$. Also, for each $(k, j) \in \mathbb{Z} \times \mathbb{Z}$ the coefficient $(f, \varphi_{k,j})$ exists and is finite. Because of the function φ has bounded support, for any real x the summation in (5) involves only a finite number of terms and, consequently, $(L_k f)(x)$ is well-defined on $\mathbb{R}$.

A more explicit look of $L_k f$ is the following

$$(L_k f)(x) = \sum_{j=-\infty}^{\infty} \varphi(a_k x + \delta j) \int\limits_{\mathrm{supp}(\varphi)} \varphi(u) f\left(\frac{u - \delta j}{a_k}\right) du. \tag{6}$$

The construction of $L_k f$ guarantees that L_k is a positive linear operator.

In the particular case $a_k = 2^k$ this operator becomes the operator A_k studied in [3]. The authors have used a scaling function φ of (1,1) type.

As regards L_k operator, a result presented in [1; *Theorem 1*] will be read as follows.

Proposition 2.1. *Let L_k, $k \in \mathbb{Z}$, be defined by (5). For every function $f \in C(\mathbb{R})$ the following inequality*

$$|(L_k f)(x) - \gamma^2 \delta f(x)| \leq \gamma^2 \delta \omega(f; 2\alpha a_{-k}), \quad k \in \mathbb{Z}, \ x \in \mathbb{R},$$

holds true, where α is given at (2) and $\omega(f;\cdot)$ represents the modulus of continuity associated to f.

Further on we collect some direct properties of the functions φ and $\varphi_{k,j}$.

Lemma 2.2. *If φ is a scaling function of (δ, γ) type then one has*

(i) $\|\varphi\|_{L_1(\mathbb{R})} = \int_{\mathbb{R}} \varphi(x)dx = \int_{\mathbb{R}} \varphi(x+\delta j)dx = \gamma\delta, \ j \in \mathbb{Z};$ (7)

(ii) $\dfrac{\gamma\delta}{\sqrt{2\alpha}} \leq \|\varphi\|_{L_2(\mathbb{R})} \leq \sqrt{\gamma\delta \sup_{x\in\mathbb{R}} \varphi(x)};$

(iii) $\|\varphi_{k,j}\|_{L_1(\mathbb{R})} = \sqrt{a_{-k}}\gamma\delta, \ \|\varphi_{k,j}\|_{L_2(\mathbb{R})} = \|\varphi\|_{L_2(\mathbb{R})}.$

Since the proof is based on simple computations, we omit it.

We end this section proving that L_k enjoys the self-adjointness property on the Hilbert space $L_2(\mathbb{R})$.

Lemma 2.3. *For every f and g belonging to $L_2(\mathbb{R})$, the operator L_k defined by (5) verifies $(L_k f, g) = (f, L_k g)$.*

Proof. We can write successively

$$\begin{aligned}(L_k f, g) &= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} \varphi_{k,j}(t) f(t) dt \Big) \varphi_{k,j}(x) g(x) dx \\ &= \sum_{j\in\mathbb{Z}} \Big(\int_{\mathbb{R}} \varphi_{k,j}(t) f(t) dt \Big) (g, \varphi_{k,j}) = (f, L_k g).\end{aligned}$$

□

3. Estimates for high order differentiable functions

In most practical problems, the functions possess some degree of smoothness.

Letting $C^n(\mathbb{R})$, $n \in \mathbb{N}$, the space of n-times continuously differentiable real valued functions defined on $\mathbb{R}$, we are concerned to give bounds for the approximation error $|L_k f - f|$, where $f \in C^n(\mathbb{R})$.

At first step we recall the Taylor formula. If $f \in C^n(\mathbb{R})$ then the following identity

$$f(y) = \sum_{i=0}^{n} \frac{f^{(i)}(x)}{i!}(y-x)^i + \frac{1}{(n-1)!}\int_x^y (f^{(n)}(t) - f^{(n)}(x))(y-t)^{n-1}dt \qquad (8)$$

holds true for every $(x, y) \in \mathbb{R} \times \mathbb{R}$.

At second step we need a technical result useful in the proof of Theorem 3.2.

Lemma 3.1. *Let φ be a scaling function of (δ, γ) type. Let fix $(x, a_k) \in \mathbb{R}\times]0, \infty[$ and define*

$$\begin{cases} J_{k,x} := \{j \in \mathbb{Z} |\ a_k x + j\delta \in [-\alpha, \alpha]\}, \\ r_n(f; a_k, x, t) := \left| \int_x^{t/a_k} (f^{(n)}(u) - f^{(n)}(x)) \left(\frac{t}{a_k} - u\right)^{n-1} du \right|, \ f \in C^n(\mathbb{R}) \end{cases} \qquad (9)$$

For each $j_0 \in J_{k,x}$ and each $t \in [-\alpha - j_0\delta, \alpha - j_0\delta]$, the following inequalities hold

$$\text{(i)} \left|\frac{t}{a_k} - x\right| \leq \frac{2\alpha}{a_k}, \quad \text{(ii)}\ r_n(f; a_k, x, t) \leq \frac{1}{n}\left(\frac{2\alpha}{a_k}\right)^n \omega\left(f^{(n)}; \frac{2\alpha}{a_k}\right). \qquad (10)$$

Proof. Since $-2\alpha + a_k x \leq -\alpha - j_0\delta \leq t \leq \alpha - j_0\delta \leq 2\alpha + a_k x$, the first relation is evident. In order to prove the second inequality, we shall analyze 2 cases taking in view the first inequality of this Lemma.

Case 1. $x < t/a_k$. We have

$$\begin{aligned} r_n(f; a_k, x, t) &\leq \int_x^{t/a_k} |f^{(n)}(u) - f^{(n)}(x)| \left(\frac{t}{a_k} - u\right)^{n-1} du \\ &\leq \int_x^{t/a_k} \omega(f^{(n)}; |u - x|) \left(\frac{t}{a_k} - u\right)^{n-1} du \\ &\leq \int_x^{t/a_k} \left(\frac{t}{a_k} - u\right)^{n-1} du\, \omega\left(f^{(n)}; \left|\frac{t}{a_k} - x\right|\right) \\ &\leq \frac{1}{n}\left(\frac{t}{a_k} - x\right)^n \omega\left(f^{(n)}; \frac{2\alpha}{a_k}\right). \end{aligned}$$

Case 2. $x > t/a_k$. Following the same line, we get

$$r_n(f; a_k, x, t) \leq \int_{t/a_k}^{x} |f^{(n)}(u) - f^{(n)}(x)| \left(u - \frac{t}{a_k}\right)^{n-1} du$$

$$\le \int_{t/a_k}^{x} \omega(f^{(n)}; |u-x|) \left(u - \frac{t}{a_k}\right)^{n-1} du$$

$$\le \int_{t/a_k}^{x} \left(u - \frac{t}{a_k}\right)^{n-1} du\, \omega\left(f^{(n)}; \left|\frac{t}{a_k} - x\right|\right) \le \frac{1}{n}\left(x - \frac{t}{a_k}\right)^{n} \omega\left(f^{(n)}; \frac{2\alpha}{a_k}\right).$$

In both cases, taking again the advantage of the first inequality we obtain the desired result. The case $x = t/a_k$ is trivial and the proof is complete. □

We present the main result of this section.

Theorem 3.2. *Let $f \in C^n(\mathbb{R})$. For every $x \in \mathbb{R}$ the operators L_k, $k \in \mathbb{Z}$, defined by (5) verify*

$$|(L_k f)(x) - \gamma^2 \delta f(x)|$$
$$\le \gamma^2 \delta \left(\sum_{i=1}^{n} \frac{|f^{(i)}(x)|}{i!} \left(\frac{2\alpha}{a_k}\right)^{i} + \frac{1}{n!}\left(\frac{2\alpha}{a_k}\right)^{n} \omega\left(f^{(n)}; \frac{2\alpha}{a_k}\right)\right). \tag{11}$$

Proof. Let fix $x \in \mathbb{R}$ and $k \in \mathbb{Z}$. Successively based on (5), (3), (4) and (9) we get

$$|(L_k f)(x) - \gamma^2 \delta f(x)| = \Big| \sum_{j\in\mathbb{Z}} (f, \varphi_{k,j})\varphi_{k,j}(x) - \gamma\delta f(x) \sum_{j\in\mathbb{Z}} \varphi(a_k x + j\delta) \Big|$$

$$\le \sqrt{a_k} \sum_{j \in J_{k,x}} |(f, \varphi_{k,j}) - \gamma\delta\sqrt{a_{-k}} f(x)| \varphi(a_k x + j\delta). \tag{12}$$

In the above we also used (1). It is obvious that, in what follows, we are interested only on the indices j belonging to $J_{k,x}$.

With the help of relations (4) and (7) we can write

$$|(f, \varphi_{k,j}) - \gamma\delta\sqrt{a_{-k}} f(x)|$$
$$= \Big| \sqrt{a_{-k}} \int_{\mathbb{R}} f\left(\frac{t}{a_k}\right) \varphi(t + \delta j) dt - \sqrt{a_{-k}} f(x) \int_{\mathbb{R}} \varphi(t + \delta j) dt \Big|$$
$$\le \sqrt{a_{-k}} \int_{\mathbb{R}} \left| f\left(\frac{t}{a_k}\right) - f(x) \right| \varphi(t + \delta j) dt. \tag{13}$$

Choosing in (8) $y := t/a_k$, $t \in [-\alpha - j\delta, \alpha - j\delta]$, and using both (9) and (10) we have

$$\left| f\left(\frac{t}{a_k}\right) - f(x) \right| \le \sum_{i=1}^{n} \frac{|f^{(i)}(x)|}{i!} \left|\frac{t}{a_k} - x\right|^{i} + \frac{1}{(n-1)!} r_n(f; a_k, x, t/a_k)$$

$$\le \sum_{i=1}^{n} \frac{|f^{(i)}(x)|}{i!} \left(\frac{2\alpha}{a_k}\right)^{i} + \frac{1}{n!}\left(\frac{2\alpha}{a_k}\right)^{n} \omega\left(f^{(n)}; \frac{2\alpha}{a_k}\right).$$

Returning on (13) and further on (12), we obtain the claimed result. □

Letting $C_b^n(\mathbb{R}) := \{f \in C^n(\mathbb{R}) |\ f^{(i)} \in B(\mathbb{R}),\ 0 \le i \le n\}$, relation (11) leads us to the following global estimate of the error.

Theorem 3.3. *For every $f \in C_b^n(\mathbb{R})$, the operators L_k, $k \in \mathbb{Z}$, defined by (5) verify*

$$\|L_k f - \gamma^2 \delta f\|_\infty \le \gamma^2 \delta \left(\sum_{i=1}^{n} \frac{\beta_k^i}{i!} \|f^{(i)}\|_\infty + \frac{\beta_k^n}{n!} \omega(f^{(n)}; \beta_k) \right), \tag{14}$$

where $\beta_k := 2\alpha a_{-k}$.

In the above, under the hypothesis $\lim_{k\to\infty} a_k = \infty$, one has $\beta_k < 1$ for sufficiently large k. Considering the semi-norm $|\cdot|_{C_b^n(\mathbb{R})}$ of the vector space $C_b^n(\mathbb{R})$ defined by $|h|_{C_b^n(\mathbb{R})} := \sum_{i=1}^{n} \|h^{(i)}\|_\infty$, relation (14) implies

$$\left\| \frac{1}{\gamma^2 \delta} L_k f - f \right\|_\infty \le \left(\frac{2\alpha}{a_k} \right) \left(|f|_{C_b^n(\mathbb{R})} + \omega \left(f^{(n)}; \frac{2\alpha}{a_k} \right) \right),$$

for every $f \in C_b^n(\mathbb{R})$ and sufficiently large k.

4. On the degree of exactness

In what follows, for any integer $s \ge 0$ we denote by e_s the test function defined by $e_s(x) = x^s$, $x \in \mathbb{R}$.

Under an additional assumption, we prove that the operator $(1/\gamma^2\delta)L_k$ reproduces the affine functions, in other words it has the degree of exactness equal to 1. We assume that the scaling function φ of (δ, γ) type has the following property

$$\sum_{j=-\infty}^{\infty} j\varphi(x + \delta j) = -\frac{\gamma}{\delta} x, \quad x \in \mathbb{R}. \tag{15}$$

Lemma 4.1. *Let φ be a scaling function of (δ, γ) type such that condition (15) is fulfilled. One has*

$$\int_{\mathbb{R}} u\varphi(u) du = 0. \tag{16}$$

Proof. We observe that

$$\int_{\mathbb{R}} u\varphi(u)du = \sum_{j\in\mathbb{Z}} \int_{\delta j}^{\delta(j+1)} u\varphi(u)du = \sum_{j\in\mathbb{Z}} \int_0^{\delta} (x+\delta j)\varphi(x+\delta j)dx$$
$$= \int_0^{\delta} x\Big(\sum_{j\in\mathbb{Z}} \varphi(x+\delta j)\Big)dx + \delta \int_0^{\delta} \Big(\sum_{j\in\mathbb{Z}} j\varphi(x+\delta j)\Big)dx.$$

Taking into account identities (3) and (15), the proof is finished. □

We come now to the main result of the section.

Theorem 4.2. *Let L_k, $k \in \mathbb{Z}$, be defined by (5) such that (15) is fulfilled. For every real-valued polynomial p of degree less or equal to 1, one has $L_k p = \gamma^2 \delta p$.*

Proof. Obviously, it is enough to verify the claimed identity only for the monomials e_0 and e_1. For computations we use the formula given at (6).

Based on (7) and (3) one gets $L_k e_0 = \gamma^2 \delta e_0$. The same quoted relations together with (16) guarantee that $L_k e_1 = \gamma^2 \delta e_1$. The conclusion follows. □

At this moment, the idea to present $L_k e_2$ comes out. In order to achieve it, we introduce the function θ given by

$$\theta(x) = \sum_{j=-\infty}^{\infty} j^2 \varphi(x+\delta j), \quad x \in \mathbb{R}. \tag{17}$$

Since (2) takes place, the above sum is finite and θ is well-defined. Moreover, θ is non-negative and belongs to $L_{1,loc}(\mathbb{R})$.

Theorem 4.3. *Let L_k, $k \in \mathbb{Z}$, be defined by (5) such that (15) is fulfilled. If θ is given by (17) then the following identities hold true*

(i) $(L_k e_2)(x) = \dfrac{\gamma}{a_k^2}(\|e_2\varphi\|_{L_1(\mathbb{R})} + \delta^3 \theta(a_k x))$, $x \in \mathbb{R}$, (18)

(ii) $\|e_2\varphi\|_{L_1(\mathbb{R})} = \delta^2 \displaystyle\int_0^{\delta} \theta(t)dt - \frac{\gamma}{3}\delta^3$.

Proof. (i) Clearly, $e_2\varphi \in L_1(\mathbb{R})$. Resorting to (6) we can write

$$(L_k e_2)(x) = \frac{1}{a_k^2} \sum_{j\in\mathbb{Z}} \varphi(a_k x + \delta j)\Big(\|e_2\varphi\|_{L_1(\mathbb{R})} - 2\delta j \int_{\mathbb{R}} u\varphi(u)du + \delta^2 j^2 \|\varphi\|_{L_1(\mathbb{R})}\Big).$$

Taking in view relations (16), (7), (3) and (17) we obtain (18).

(ii) Since $\|e_2\varphi\|_{L_1(\mathbb{R})} = \sum_{j\in\mathbb{Z}} \int_{\delta j}^{\delta(j+1)} u^2\varphi(u)du = \sum_{j\in\mathbb{Z}} \int_0^{\delta} (t+\delta j)^2\varphi(t+\delta j)dt$, with the help of (3), (15) and (17), our statement is proved. □

5. Estimates for bounded functions

Based on (6) and (3) we can remark in passing that

$$\|L_k f\|_\infty \le \gamma^2\delta\|f\|_\infty, \text{ for every } f \in B(\mathbb{R}) \cap L_{1,loc}(\mathbb{R}).$$

Consequently, for $\gamma^2\delta < 1$ each operator L_k is a contraction.

The aim of this section is to give bounds for error approximation by using a Lipschitz-type function introduced by Lenze [5; *Eq. (1.5)*]. We recall this map we will have to deal with. Let $J \subset \mathbb{R}$ be an interval. Let $f \in \mathbb{R}^J$ be bounded and $\mu \in]0,1]$. The Lipschitz-type maximal function of order μ associated to f is defined as

$$f_\mu^\sim(x) := \sup_{\substack{t\neq x\\ t\in J}} \frac{|f(t)-f(x)|}{|t-x|^\mu}, \quad x \in J. \tag{19}$$

The local behaviour of function f can be measured by $f_\mu^\sim$. The finiteness of $f_\mu^\sim$ gives a local control for the smoothness of f. Roughly speaking, the boundedness of $f_\mu^\sim$ is equivalent to $f \in \text{Lip}\mu$ on J.

Theorem 5.1. *Let L_k, $k \in \mathbb{Z}$, be defined by (5) such that (15) is fulfilled. For every $\mu \in]0,1]$ and $f \in B(\mathbb{R}) \cap L_{1,loc}(\mathbb{R})$ the following inequality*

$$|(L_k f)(x) - \gamma^2\delta f(x)| \le M_\varphi \left(\frac{\alpha^2}{3a_k^2} + \frac{\delta^2}{\gamma a_k^2}\theta(a_k x) - x^2\right)^{\mu/2} f_\mu^\sim(x), \quad x \in \mathbb{R},$$

holds true, where $M_\varphi = \gamma(2\alpha)^{\mu/2}\|\varphi\|_{L_p([-\alpha,\alpha])}$, $p = 2/(2-\mu)$ and θ is given at (17).

Proof. Let fix $x \in \mathbb{R}$ and $k \in \mathbb{Z}$. In what follows, for the sake of simplicity, we set

$$(I_{k,j}\varphi)(x) := \int_{\mathbb{R}} \left|\frac{t}{a_k} - x\right|^\mu \varphi(t+\delta j)dt,$$

$$c_{k,j}(x) := \int_{\text{supp}(\varphi)} \left(\frac{t-\delta j}{a_k} - x\right)^2 dt, \ (j \in \mathbb{Z}),$$

and

$$s_k(x) := \sum_{j\in\mathbb{Z}} c_{k,j}(x)\varphi(a_k x + \delta j).$$

In concordance with formula (19) we write

$$|f(t/a_k) - f(x)| \le f^{\sim}_{\mu}(x)|t/a_k - x|^{\mu}.$$

Taking the advantage of relations (12) and (13), one obtains

$$|(L_k f)(x) - \gamma^2\delta f(x)| \le \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} \left| f\left(\frac{t}{a_k}\right) - f(x)\right| \varphi(t+\delta j)dt\varphi(a_k x + \delta j)$$
$$\le f^{\sim}_{\mu}(x) \sum_{j\in\mathbb{Z}} (I_{k,j}\varphi)(x)\varphi(a_k x + \delta j).$$

By using Hölder's integral inequality with parameters $q := 2/\mu$ and $p := 2/(2-\mu)$, we deduce

$$(I_{k,j}\varphi)(x)$$
$$= \int\limits_{\mathrm{supp}(\varphi)} \left|\frac{u-\delta j}{a_k} - x\right|^{\mu} \varphi(u)du \le c^{\mu/2}_{k,j}(x)\Big(\int\limits_{\mathrm{supp}(\varphi)} \varphi^{2/(2-\mu)}(u)du\Big)^{(2-\mu)/2}. \qquad (21)$$

The last quantity represents $\|\varphi\|_{L_p([-\alpha,\alpha])}$, see (2).

Further on, based on Hölder's discrete inequality with the same parameters q, p, we have

$$\sum_{j\in\mathbb{Z}} c^{\mu/2}_{k,j}(x)\varphi(a_k x + \delta j) = \sum_{j\in\mathbb{Z}} (c_{k,j}(x)\varphi(a_k x+\delta j))^{\mu/2}\varphi^{1-\mu/2}(a_k x + \delta j)$$

$$\le \Big(\sum_{j\in\mathbb{Z}} c_{k,j}(x)\varphi(a_k x + \delta j)\Big)^{\mu/2}\Big(\sum_{j\in\mathbb{Z}} \varphi(a_k x + \delta j)\Big)^{(2-\mu)/2} = \gamma^{(2-\mu)/2} s^{\mu/2}_k(x). \qquad (22)$$

In order to evaluate $s_k(x)$, we shall use (3), (15), (18) and (2).

$$s_k(x) = \frac{1}{a_k^2} \int\limits_{\mathrm{supp}(\varphi)} \sum_{j\in\mathbb{Z}} (t - \delta j - a_k x)^2 \varphi(a_k x + \delta j)dt$$

$$= \frac{1}{a_k^2} \int\limits_{\mathrm{supp}(\varphi)} (t^2\gamma + \delta^2\theta(a_k x) - a_k^2\gamma x^2)dt \le \frac{2\alpha\gamma}{a_k^2}\left(\frac{\alpha^2}{3} + \frac{\delta^2}{\gamma}\theta(a_k x) - a_k^2 x^2\right). \qquad (23)$$

Collecting (23), (22), (21) and substituting in (20) we finish the proof. □

Remarks. (i) In the particular case $\gamma=\delta=1$, $a_k=2^k$, L_k turns into Anastassiou's original operator A_k, [2; §*6.1*]. As far as we know, Theorem 5.1 establishes a new result for A_k which involves Lenze's function.

(ii) For comparison, it is a standard fact in Littlewood-Paley theory that if f and φ are both Hölder continuous of order $\mu>0$ and if φ has compact support then the μ-Hölder norm of L_kf-f decays like $1/a_k$.

(iii) Regarding this note, we mention that a similar approach could be made by considering the following $(L_a)_{a>0}$ net of operators, $L_af:=\sum_{j\in\mathbb{Z}}(f,\varphi_{j,a})\varphi_{j,a}$ with $\varphi_{j,a}(x)=\sqrt{a}\varphi(ax+j\delta)$. The estimates would be exactly the same.

References

[1] Agratini, O., Operators generated by a quasi-scaling type function, *Revista de la Unión Matemática Argentina*, **44**(2003), 2, 21-30.

[2] Anastassiou, G., *Quantitative Approximations*, Chapman & Hall/CRC, Boca Raton, 2001.

[3] Anastassiou, G., Yu, X. M., Monotone and probabilistic wavelet approximation, *Stochastic Anal. Appl.*, **10**(1992), 251-264.

[4] Daubechies, I., *Ten Lectures on Wavelets*, NSF-CBMS Regional Conference Series in Applied Math. **61**, SIAM Publications, Philadelphia, 1992.

[5] Lenze, B., Bernstein-Baskakov-Kantorovič operators and Lipschitz-type maximal functions, *Colloquia Mathematica Societatis János Bolyai 58, Approximation Theory*, Kecskemèt, Hungary (J. Szabados, K. Tandori, eds.), pp. 469-496, North-Holland Publishing Company, Amsterdam – Oxford – New York, 1990.

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

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

ON THE EXISTENCE OF VIABLE SOLUTIONS FOR A CLASS OF NONAUTONOMOUS NONCONVEX DIFFERENTIAL INCLUSIONS

AURELIAN CERNEA

Abstract. We prove the existence of viable solutions to the Cauchy problem $x' \in F(t,x), x(0) = x_0$ in M, where F is a multifunction and M is a convex locally compact set of a Hilbert space that satisfy $F(t,x) \cap K_xM \cap \partial V(x) \neq \emptyset$, with K_xM the contingent cone to M at x and ∂V is the subdifferential of a convex function V.

1. **Introduction**

Consider H a real Hilbert space and $F : M \subset H \to \mathcal{P}(H)$ a multifunction that defines the Cauchy problem

$$x' \in F(x), \quad x(0) = x_0, \tag{1.1}$$

In the theory of differential inclusions the viability problem consists in proving the existence of viable solutions, i.e. $\forall t$, $x(t) \in M$, to the Cauchy problem (1.1).

Under the assumptions that $H = R^n$, F is an upper semicontinuous nonempty convex compact valued multifunction and M is locally compact, in [5] Haddad proved that a necessary and sufficient condition for the existence of viable trajectories starting from $x_0 \in M$ of problem (1.1) is the tangential condition

$$\forall x \in M \quad F(x) \cap K_xM \neq \emptyset, \tag{1.2}$$

where K_xM is the contingent cone to M at $x \in M$.

Received by the editors: 19.09.2004.
2000 *Mathematics Subject Classification.* 34A60.
Key words and phrases. Viable solutions, convex functions, differential inclusions.

Rossi, in [7], proved the existence of viable solutions to problem (1.1) replacing the convexity conditions on the images on F with

$$F(x) \subset \partial V(x) \quad \forall x \in M, \tag{1.3}$$

where ∂V is the subdifferential, in the sense of Convex Analysis, of a proper convex function V. In [4] condition (1.3) is improved in the sense that instead of (1.2) and (1.3) we assume that $F(.)$ verifies

$$F(x) \cap K_x M \cap \partial V(x) \neq \emptyset \quad \forall x \in M, \tag{1.4}$$

with V as in [7], provided M is convex.

The aim of the present paper is to extend the result in [4] to the case of nonautonomous problems

$$x' \in F(t,x), \quad x(0) = x_0. \tag{1.5}$$

We note that in [6] a similar type of result is proved for a function V that is assumed to be lower regular, i.e. a locally Lipschitz continuous function whose upper Dini directional derivative coincides with the Clarke directional derivative.

The idea of the proof of our result is to use the regularizing technique in [6] and to apply the known result for autonomous problems in [4].

The paper is organized as follows: in Section 2 we recall some preliminary facts that we need in the sequel and in Section 3 we prove our main result.

2. **Preliminaries**

Let H be a real separable Hilbert space and $\Omega \subset H$ a given set. By $\mathcal{P}(H)$ we denote the family of all subsets of H. A multifunction $F : \Omega \to \mathcal{P}(H)$ is called (Hausdorff) upper semicontinuous at $x_0 \in \Omega$, $\forall \epsilon > 0$ there exists $\delta > 0$ such that $||x - x_0|| < \delta$ implies $F(x) \subset F(x_0) + \epsilon B$, where B is the unit ball in H. For $\epsilon > 0$ we put $B(x, \epsilon) = \{y \in H; ||y - x|| < \epsilon\}$.

Let $V : H \to R \cup \{+\infty\}$ be a function with domain $D(V) = \{x \in H; V(x) < +\infty\}$. If $D(V) \neq \emptyset$, then f is called *proper*. We recall that the *subdifferential* (in the sense of Convex Analysis) of the convex function V is the multifunction $\partial V : H \to \mathcal{P}(H)$ defined by

$$\partial V(x) = \{y \in H; \quad V(z) - V(x) \geq < y, z - x > \quad \forall z \in H\}.$$

In what follows we assume:

Hypothesis 2.1. i) $F : [0, \infty) \times M \subset H \to \mathcal{P}(H)$ is a bounded set valued map, measurable in t, upper semicontinuous with respect to x, with nonempty closed values.

ii) There exists a proper lower semicontinuous convex function $V : H \to R \cup \{+\infty\}$ such that

$$F(t, x) \cap K_x M \cap \partial V(x) \neq \emptyset \quad \forall x \in M, \ a.e. \ t \in [0, \infty), \tag{2.1}$$

where $K_x M = \{v \in H; \quad \liminf_{h \to 0+} \frac{1}{h} d(x + hv, M) = 0\}$ is the contingent cone to M at $x \in M$.

3. **The main result**

Our main result is the following.

Theorem 3.1. *Let $M \subset H$ be a convex and locally compact set and let $F : [0, \infty) \times M \subset H \to \mathcal{P}(H)$ be a set-valued map satisfying Hypothesis 2.1.*

Then for every $x_0 \in M$ there exists $T > 0$ such that problem (1.5) admits a solution on $[0, T]$ satisfying $x(t) \in M$, $\forall t \in [0, T]$.

Proof. Let $x_0 \in M$. Since $M \subset H$ is locally compact, there exists $r > 0$ such that $M_0 := M \cap B(x_0, r)$ is compact. Consider $L := \sup_{(t,x) \in [0,\infty) \times M} ||F(t, x)||$, define $T := \frac{r}{L+1}$ and take $n \in N$ such that $\frac{1}{n} < T$.

By regularizing the set valued map F on the right hand side of the Cauchy problem (1.5) we reduce the nonautonomous problem to the autonomous case ([6]). We can find a countable collection of disjoint subintervals $(a_j, b_j) \subset [0, T]$, $j = 1, 2, ...$ such that their total length is less then $\frac{1}{n}$ and a set valued map F_n defined on $D := ([0, T] \setminus \cup_{j=1}^{\infty} (a_j, b_j)) \times M$ that is jointly upper semicontinuous and $F_n(t, x) \subset F(t, x)$

for each $(t,x) \in D$. Moreover, if $u(.)$ and $v(.)$ are measurable functions on $[0,T]$ such that $u(t) \in F(t,v(t))$ a.e. $t \in [0,T]$ then for a.e. $t \in ([0,T] \setminus \cup_{j=1}^{\infty} (a_j,b_j))$ we have $u(t) \in F_n(t,v(t))$ (we refer to [8] for this Scorza Dragoni type theorem). It is obvious that all trajectories of F are also trajectories of F_n. We extend F_n to the whole $[0,T] \times M$. We define

$$\tilde{F}_n(t,x) = \begin{cases} F_n(t,x) & \text{if} \quad t \in [0,T] \setminus \cup_{j=1}^{\infty} (a_j,b_j) \\ F_n(a_j,x) & \text{if} \quad a_j < t < \frac{a_j+b_j}{2} \\ F_n(b_j,x) & \text{if} \quad \frac{a_j+b_j}{2} < t < b_j \\ F_n(a_j,x) \cup F_n(b_j,x) & \text{if} \quad t = \frac{a_j+b_j}{2}. \end{cases}$$

It is easy to see that $\tilde{F}_n(.,.)$ still satisfies the tangential condition (2.1). On the other hand, according to Lemma 4 in [6], $\tilde{F}_n(.,.)$ is upper semicontinuous on $[0,T] \times M$.

By extending the state space from H to $R \times H$ we can reduce our problem to the autonomous case. For every $(t,x) \in [0,T] \times M$ we define

$$\tilde{V}(t,x) = t + V(x).$$

Obviously, $\tilde{V}(.,.)$ is a proper lower semicontinuous convex function and $(1,v) \in \partial\tilde{V}(t,x)$ if and only if $v \in \partial V(x)$ for all $(t,x) \in [0,T] \times M$. At the same time, standard arguments show that $(1,v) \in K_{(t,x)}([0,T] \times M)$ if and only if $v \in K_x M$.

Therefore, the tangential condition (2.1) implies that

$$(1,\tilde{F}_n(t,x)) \cap K_{(t,x)}([0,T] \times M) \cap \partial\tilde{V}(t,x) \neq \emptyset \quad \forall (t,x) \in [0,T] \times M. \tag{3.1}$$

Thus applying Theorem 3.1 in [4] we obtain the existence of an absolutely continuous function $x_n(.) : [0,T] \to H$ that satisfies

$$(1,x_n'(t)) \in (1,\tilde{F}_n(t,x_n(t))) \cap \partial\tilde{V}(t,x_n(t)) \quad a.e.\ [0,T], \quad x_n(0) = x_0$$

and

$$(t,x_n(t)) \in [0,T] \times M \quad \forall t \in [0,T].$$

It follows that $x_n(.)$ verifies

$$x_n'(t) \in F_n(t, x_n(t)) \cap \partial V(x_n(t)) \quad a.e.\ [0,T], \quad x_n(0) = x_0 \tag{3.2}$$

and

$$x_n(t) \in M \quad \forall t \in [0,T]. \tag{3.3}$$

Therefore from (3.2) we have

$$||x_n'(t)|| \leq L. \tag{3.4}$$

On the other hand, from (3.3) $graph(x_n(.))$ is contained in $[0,T] \times M$ and $x_n(.)$ is also a solution to the inclusion (1.5) except for a set (say) E_n of measure not exceeding $\frac{1}{n}$ for each $n \in N$. Hence, from (3.4) and Theorem III. 27 in [3] there exists a subsequence (again denoted by $x_n(.)$) and an absolutely continuous function $x(.) : [0,T] \to H$ such that

$$x_n(.) \quad \text{converges uniformly to} \quad x(.),$$

$$x_n'(.) \quad \text{converges weakly in } L^2([0,T], H) \text{ to} \quad x'(.).$$

Since $V(.)$ is lower semicontinuous, it follows that $graph(\partial V)$ is closed and thus, by (3.2), one has

$$x'(t) \in \partial V(x(t)) \quad a.e.\ [0,T]. \tag{3.5}$$

We apply Lemma 3.3 in [2] and by (3.5) we obtain

$$(V(x(t)))' =< x'(t), x'(t) >= ||x'(t)||^2 \quad a.e.\ [0,T];$$

and thus, $V(x(T)) - V(x_0) = \int_0^T ||x'(t)||^2 dt$.

On the other hand, from (3.2) we deduce that

$$\int_0^T ||x_n'(t)||^2 dt = \int_0^T (V \circ x_n)'(t) dt = V(x_n(T)) - V(x_0).$$

Hence, by the lower semicontinuity of V, we get

$$\lim_{n\to\infty} \int_0^T ||x_n'(t)||^2 dt = V(x(T)) - V(x_0) = \int_0^T ||x'(t)||^2 dt$$

and so $\{x_n'(.)\}$ converges strongly in $L^2([0,T], H)$.

Hence, there exists a subsequence (still denoted) $x'_n(.)$ which converges pointwise almost everywhere to $x'(.)$. From (3.2) and the fact that $graph(F)$ is closed we have

$$x'(t) \in F(t, x(t)) \quad a.e. \ [0, T]$$

and from (3.3) we obtain that $\forall t \in [0, T]$, $x(t) \in M$.

References

[1] Aubin, J. P., Cellina, A., *Differential Inclusions*, Springer, Berlin, 1984.

[2] Brezis, H., *Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North Holland, Amsterdam, 1973.

[3] Brezis, H., *Analyse fonctionelle, theorie et applications*, Masson, Paris, 1983.

[4] Cernea, A., *Viable solutions for a class of differential inclusions without convexity*, Anale. Univ. "Al. I. Cuza" Iaşi, Mat., submitted.

[5] Haddad, G., *Monotone trajectories of differential inclusions and functional inclusions with memory*, Israel J. Math., **39**(1981), 83-100.

[6] Kannai, Z., Tallos, P., *Viable solutions to nonautonomous inclusions without convexity*, Central European J. Operational Research, **11**(2003), 47-55.

[7] Rossi, P., *Viability for upper semicontinuous inclusions without convexity*, Diff. Integral Eqs., **5**(1992), 455-459.

[8] Rzezuchowski, T., *Scorza-Dragoni type theorems for upper semicontinuous multivalued functions*, Bull. Acad. Pol. Math., **28**(1980), 61-66.

Faculty of Mathematics and Informatics, University of Bucharest
Academiei 14, 010014 Bucharest, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

MAXIMUM PRINCIPLES FOR A CLASS OF SECOND ORDER ELLIPTIC SYSTEMS IN DIVERGENCE FORM

CRISTIAN CHIFU-OROS

Abstract. The purpose of this paper is to prove a maximum principle for weakly coupled second order elliptic systems in divergence form using an iterative estimation technique. A similar method was used by A.W. Tursky in 1992 to prove a maximum principle for elliptic equations in divergence form.

1. **Introduction**

Let us consider the following second order elliptic system:

$$\sum_{i,j=1}^{N}(a_{ij}(x)(u_p)_{x_j})_{x_i}+\sum_{j=1}^{N}b_j(x)(u_p)_{x_j}+\sum_{l=1}^{n}c_{pl}(x)u_l+f_p(x)=0,\ p=\overline{1,n} \tag{1}$$

in Ω, where $\Omega\subset\mathbb{R}^N$ is a smooth bounded domain. Assume that:

(i) $a_{ij}, b_j \in C^1(\overline{\Omega})\ , i,j=\overline{1,N},\ c_{pl}\in C(\overline{\Omega}), p,l=\overline{1,n},$

$$\sum_{i,j=1}^{N}a_{ij}(x)\xi_i\xi_j\geq 0, \forall x\in\Omega, \forall\xi\in\mathbb{R}^N;$$

(ii) f_p is continuous in $\overline{\Omega}, p=\overline{1,n}$.

Let

$$\|u\|=\max\{\|u_p\| : p=\overline{1,n}\},$$

where:

$$\|u_p\|=\inf\{M : |u_p(x)|\leq M,\ a.e.\ x\in\Omega\}.$$

Received by the editors: 26.04.2005.

2000 *Mathematics Subject Classification.* 35B50, 35J45, 35J55.

Key words and phrases. Second order elliptic systems in divergence form, maximum principles.

Denote

$$M_p = \sup_{\Omega} \|f_p\|, M = \max_p M_p,$$

$$m_p = \sup_{\partial\Omega} |u_p(x)|, m = \max_p m_p.$$

In the proof of the main result we shall need the following Lemma given in [6]:

Lemma 1. *If*

$$G_a(x) = (1 + x^a)[1 - x(1 + x^a)^{-\frac{1}{a}}], a \in \mathbb{Z}, x > 0,$$

then

$$G_a(x) \geq \frac{1}{a}, \forall x > 0. \tag{2}$$

In what follows we shall give the main result of the paper. We shall show that the method which was used by A. W. Turski in [6] for second order elliptic equations in divergence form works for weakly coupled second order elliptic systems in divergence form. Our result is similar with other results as in [3],[4] or [5] but the method of proving is different. On the other hand the technique may be used to study weak solutions of the second order elliptic systems.

2. Main result

Theorem 1. *Assume that there exit two real constants h, β with $0 \leq \beta < h$ such that:*

1. $c_{pp} \leq -h < 0, p = \overline{1,n}$;
2. $0 \leq c_{pl} \leq \frac{\beta}{n}, p, l = \overline{1,n}, p \neq l$.

Then for each $u \in C^2(\Omega, \mathbb{R}^n) \cap C^0(\overline{\Omega}, \mathbb{R}^n)$ satisfying (1),

$$\|u\| \leq \max\left\{m, \frac{M}{h-\beta}\right\}.$$

Proof: Because the derivatives of u which occurs in system (1) are continuous just in Ω, to be able to use the continuity of these on the boundary, for every $q \in \{1, 2, 3, ...\}$, we shall define $\Omega_q = \{x \in \Omega : d(x, \partial\Omega) > \frac{1}{q}\}$. Our assumptions still hold in $\overline{\Omega_q}$. If we multiply system (1) by $u_p^{2^k-1}$ and integrate on Ω_q with q fixed, we obtain:

$$\int\limits_{\Omega_q}\sum_{i,j=1}^{N}(a_{ij}(x)(u_p)_{x_j})_{x_i}u_p^{2^k-1}dx+\int\limits_{\Omega_q}\sum_{j=1}^{N}b_j(x)(u_p)_{x_j}u_p^{2^k-1}dx+$$

$$+\int\limits_{\Omega_q}\sum_{l=1}^{n}c_{pl}(x)u_l u_p^{2^k-1}dx+\int\limits_{\Omega_q}f_p(x)u_p^{2^k-1}dx=0. \tag{3}$$

We have

$$\int\limits_{\Omega_q}\sum_{i,j=1}^{N}(a_{ij}(x)(u_p)_{x_j})_{x_i}u_p^{2^k-1}dx=\int\limits_{\partial\Omega_q}\sum_{i,j=1}^{N}a_{ij}(x)(u_p)_{x_j}u_p^{2^k-1}\cos(N,x_i)ds-$$

$$-(2^k-1)\int\limits_{\Omega_q}\sum_{i,j=1}^{N}a_{ij}(x)(u_p)_{x_j}(u_p)_{x_i}u_p^{2^k-2}dx.$$

Denote

$$m_{p_q}=\sup_{x\in\partial\Omega_q}|u_p(x)|,$$

$$H_{1_q}:=\int\limits_{\partial\Omega_q}\left|\sum_{i,j=1}^{N}a_{ij}(x)(u_p)_{x_j}\cos(N,x_i)\right|ds.$$

Because of the assumption (i) we obtain

$$\int\limits_{\Omega_q}\sum_{i,j=1}^{N}(a_{ij}(x)(u_p)_{x_j})_{x_i}u_p^{2^k-1}dx\le m_{p_q}^{2^k-1}H_{1_q}. \tag{4}$$

Integrating by parts we have

$$\int\limits_{\Omega_q}\sum_{j=1}^{N}b_j(x)(u_p)_{x_j}u_p^{2^k-1}dx=$$

$$=2^{-k}\int\limits_{\partial\Omega_q}\sum_{j=1}^{N}b_j(x)u_p^{2^k}\cos(N,x_j)ds-2^{-k}\int\limits_{\Omega_q}\sum_{j=1}^{N}(b_j(x))_{x_j}u_p^{2^k}dx.$$

Denote

$$H_{2_q}:=\int\limits_{\partial\Omega_q}\left|\sum_{j=1}^{N}b_j(x)\cos(N,x_j)\right|ds.$$

Because $b_j \in C^1(\overline{\Omega_q}\ )$, there exists $B_q > 0$ such that

$$\left| \sum_{j=1}^{N} (b_j(x))_{x_j} \right| \le B_q.$$

In this way

$$\int_{\Omega_q} \sum_{j=1}^{N} b_j(x)(u_p)_{x_j} u_p^{2^k-1} dx \le 2^{-k} m_{p_q}^{2^k} H_{2_q} + 2^{-k} B_q \int_{\Omega_q} u_p^{2^k} dx,$$

$$\int_{\Omega_q} \sum_{j=1}^{N} b_j(x)(u_p)_{x_j} u_p^{2^k-1} dx \le 2^{-k} m_{p_q}^{2^k} H_{2_q} + 2^{-k} B_q \left\| u_p \right\|^{2^k}. \tag{5}$$

On the other hand

$$\int_{\Omega_q} \sum_{l=1}^{n} c_{pl}(x) u_l u_p^{2^k-1} dx = \sum_{l=1, l\neq p}^{n} \int_{\Omega_q} c_{pl}(x) u_l u_p^{2^k-1} dx + \int_{\Omega_q} c_{pp}(x) u_p^{2^k} dx.$$

Using Hölder inequality with the exponents 2^k and $\frac{2^k}{2^k-1}$ and the hypothesis of theorem we'll obtain

$$\int_{\Omega_q} \sum_{l=1}^{n} c_{pl}(x) u_l u_p^{2^k-1} dx \le \frac{\beta}{n} \sum_{l=1, l\neq p}^{n} \left(\int_{\Omega_q} u_l^{2^k} \right)^{2^{-k}} \left(\int_{\Omega_q} u_p^{2^k} \right)^{1-2^{-k}} - h \int_{\Omega_q} u_p^{2^k} dx.$$

In this way

$$\int_{\Omega_q} \sum_{l=1}^{n} c_{pl}(x) u_l u_p^{2^k-1} dx \le \frac{\beta}{n} \sum_{l=1, l\neq p}^{n} \left\| u_l \right\| \left\| u_p \right\|^{2^k-1} - h \left\| u_p \right\|^{2^k}. \tag{6}$$

Also

$$\int_{\Omega_q} f_p(x) u_p^{2^k-1} dx \le \left(\int_{\Omega_q} f_p^{2^k} \right)^{2^{-k}} \left(\int_{\Omega_q} u_p^{2^k} \right)^{1-2^{-k}}.$$

Let $|\Omega_q|$ be the Lebesque measure of Ω_q. Because of the assumption for f_p,

$$\int_{\Omega_q} f_p(x) u_p^{2^k-1} dx \le M_{p_q} \left| \Omega_q \right|^{2^{-k}} (\left\| u_p \right\|^{2^k})^{1-2^{-k}}. \tag{7}$$

From $(4), (5), (6)$ and (7), (3) becomes

$$m_{p_q}^{2^k-1} H_{1_q} + 2^{-k} m_{p_q}^{2^k} H_{2_q} + (2^{-k} B_q - h) \left\| u_p \right\|^{2^k} + \frac{\beta}{n} \sum_{l=1, l\neq p}^{n} \left\| u_l \right\| \left\| u_p \right\|^{2^k-1} +$$

$$+M_{p_q}\left|\Omega_q\right|^{2^{-k}}\left(\|u_p\|^{2^k}\right)^{1-2^{-k}} \geq 0.$$

If we take the maximum after $p=\overline{1,n}$ then

$$m_q^{2^k-1}H_{1q}+2^{-k}m_q^{2^k}H_{2_q}+(2^{-k}B_q+\beta-h)\left\|u\right\|^{2^k}+$$

$$+\,M_q\left|\Omega_q\right|^{2^{-k}}\left(\|u\|^{2^k}\right)^{1-2^{-k}} \geq 0.$$

Denoting

$$y_k := \|u\|^{2^k},$$

we have

$$m_q^{2^k-1}H_{1_q}+2^{-k}m_q^{2^k}H_{2_q}+(2^{-k}B_q+\beta-h)y_k+M_q\left|\Omega\right|^{2^{-k}}y_k^{1-2^{-k}} \geq 0. \tag{8}$$

Let

$$\alpha_k := h-\beta-2^{-k}B_q>0,$$

for k large enough,

$$\beta_k := M_q\left|\Omega_q\right|^{2^{-k}},$$

$$\gamma_k := 2^k m_q^{2^k-1}H_{1_q}+m_q^{2^k}H_{2_q}.$$

With these notations (8) becomes

$$2^{-k}\gamma_k-\alpha_k y_k+\beta_k y_k^{1-2^{-k}} \geq 0,$$

$$\alpha_k y_k-\beta_k y_k^{1-2^{-k}} \leq 2^{-k}\gamma_k. \tag{9}$$

Let

$$F(y)=\alpha_k y-\beta_k y^{1-2^{-k}}.$$

We observe that if $y \geq \left(\frac{\beta_k}{\alpha_k}\right)^{2^k}$, then $F'(y) \geq 0$, which means that on $I := \left[\left(\frac{\beta_k}{\alpha_k}\right)^{2^k},\infty\right)$ the function F is increasing. On the other hand we have

$$F(y_k) \leq 2^{-k}\gamma_k. \tag{10}$$

Let

$$y^*=\left(\frac{\beta_k}{\alpha_k}\right)^{2^k}+\frac{\gamma_k}{\alpha_k} \in I.$$

We shall prove that $F(y^*) \geq 2^{-k}\gamma_k$ which, because of (10) and the monotonicity of F, will imply that $y_k \leq y^*$.

$$F(y^*) = \alpha_k \left(\frac{\beta_k^{2^k}}{\alpha_k^{2^k}} + \frac{\gamma_k}{\alpha_k} \right) - \beta_k \left(\frac{\beta_k^{2^k}}{\alpha_k^{2^k}} + \frac{\gamma_k}{\alpha_k} \right)^{1-2^{-k}},$$

$$F(y^*) = \gamma_k \left(\frac{\beta_k^{2^k}}{\alpha_k^{2^k-1}\gamma_k} + 1 \right) \left[1 - \frac{\beta_k}{\alpha_k^{1-2^{-k}}\gamma_k^{2^{-k}}} \left(\frac{\beta_k^{2^k}}{\alpha_k^{2^k-1}\gamma_k} + 1 \right)^{2^{-k}} \right].$$

If we denote by $\Phi := \frac{\beta_k}{\alpha_k^{1-2^{-k}}\gamma_k^{2^{-k}}}$ then $\Phi^{2^k} = \frac{\beta_k^{2^k}}{\alpha_k^{2^k-1}\gamma_k}$ and in this way we obtain

$$F(y^*) = \gamma_k \left(1 + \Phi^{2^k}\right) \left[1 - \Phi(1 + \Phi^{2^k})^{-2^{-k}}\right] = \gamma_k G_{2^k}(\Phi).$$

From Lemma 1 we have

$$F(y^*) \geq 2^{-k}\gamma_k. \tag{11}$$

From (10) and (11) it follows that $F(y_k) \leq F(y^*)$ and hence $y_k \leq y^*$.

$$y_k \leq \left(\frac{\beta_k}{\alpha_k}\right)^{2^k} + \frac{\gamma_k}{\alpha_k} \Rightarrow y_k^{2^{-k}} \leq \left[\left(\frac{\beta_k}{\alpha_k}\right)^{2^k} + \frac{\gamma_k}{\alpha_k}\right]^{2^{-k}} \leq 2^{2^{-k}}\ max\left\{\frac{\beta_k}{\alpha_k}, \left(\frac{\gamma_k}{\alpha_k}\right)^{2^{-k}}\right\},$$

$$\|u\| \leq 2^{2^{-k}}\ max\left\{\frac{\beta_k}{\alpha_k}, \left(\frac{\gamma_k}{\alpha_k}\right)^{2^{-k}}\right\}.$$

For $k \to \infty$ we obtain

$$\|u\| \leq max\left\{\lim_{k\to\infty}\frac{\beta_k}{\alpha_k}, \lim_{k\to\infty}\left(\frac{\gamma_k}{\alpha_k}\right)^{2^{-k}}\right\},$$

$$\frac{\beta_k}{\alpha_k} = \frac{M_q\,|\Omega_q|^{2^{-k}}}{h - \beta - 2^{-k}B_q} \to \frac{M_q}{h-\beta},$$

$$\left(\frac{\gamma_k}{\alpha_k}\right)^{2^{-k}} = m_q\left(2^k\frac{2^kH_{1_q} + m_qH_{2_q}}{[2^k(h-\beta) - B_q]m_q}\right)^{2^{-k}} \to m_q.$$

Hence

$$\|u\| \leq \max\left\{m_q, \frac{M_q}{h-\beta}\right\}.$$

For $q \to \infty$ we obtain (see [7])

$$\|u\| \leq \max\left\{m, \frac{M}{h-\beta}\right\}.$$

Remark 1. *If $n = 1$ and $\beta = 0$ then we obtain the result given in* [6].

Theorem 2. *Under the assumptions of Theorem 1, if* $u \in C^2(\Omega, \mathbb{R}^n) \cap C^0(\overline{\Omega}, \mathbb{R}^n)$ *satisfies* (1), *with* $f \equiv 0$, *then*

$$\|u\| \leq \sup_{\partial\Omega} |u(x)| .$$

3. Application

Let us consider the following boundary value problem:

$$\begin{cases} -\Delta u_p - \sum_{l=1}^{n} c_{pl} u_l = f_p, \; in\, \Omega \\ u_p = 0, \; on\, \partial\Omega \end{cases}, \; p = \overline{1,n}. \tag{12}$$

Assume that $\Omega \subset \mathbb{R}^N$ is a smooth bounded domain and f_p is continuous in $\overline{\Omega}$.

Theorem 3. *Under the assumptions of Theorem 1, the boundary value problem (12) has at most one solution in* $C^2(\Omega, \mathbb{R}^n) \cap C^0(\overline{\Omega}, \mathbb{R}^n)$.

Proof: Suppose that problem (12) has two solutions $u_1, u_2 \in C^2(\Omega, \mathbb{R}^n) \cap C^0(\overline{\Omega}, \mathbb{R}^n), u_1 \neq u_2$. Then $u = u_1 - u_2 \in C^2(\Omega, \mathbb{R}^n) \cap C^0(\overline{\Omega}, \mathbb{R}^n)$ is a solution of the following boundary value problem

$$\begin{cases} \Delta u_p + \sum_{l=1}^{n} c_{pl} u_l = 0, \; in\, \Omega \\ u_p = 0, \; on\, \partial\Omega \end{cases}, \; p = \overline{1,n}. \tag{13}$$

Because of the assumptions of the theorem we have

$$\|u\| \leq \max\left\{m, \frac{M}{h - \beta}\right\}.$$

In this case $m = 0, M = 0$, hence

$$\|u\| \leq 0.$$

This will imply that $u \equiv 0$ which is a contradiction. In conclusion problem (12) has at most one solution.

References

[1] Dlotko, T., *Remarks on global bounds of solutions of parabolic equations in divergence form*, Rocky Mountain J. Math., 17(3) (1987), 499-510.

[2] Gilbarg, D., Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order*, Springer Verlag, Berlin, Heidelberg, New York, 1977.

[3] Rus, I. A., *A maximum principle for an elliptic system of partial differential equations*, Proceedings of the Conference on Differential Equations and their Applications, Iaşi, Octombrie 1973, 77-80, Ed. Academiei Romane, Bucuresti, 1977.

[4] Rus, I. A., *Maximum principle for elliptic systems*, International Series of Numerical Mathematics, vol. 107, Birkhauser Verlag, Basel, 1992.

[5] Protter, M. H., Weinberger, H. F., *Maximum principles in differential equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1967.

[6] Turski, A. W., *An iteration proof of the maximum principle*, Demonstratio Mathematica, XXV, 1-2(1992), 147-152.

[7] Yosida, K., *Functional Analysis*, Spinger Verlag, Berlin, Heidelberg, New York, 1978.

Babeş-Bolyai University, Faculty of Business, Department of Business
7^{th}, Horea Street, 400174, Cluj-Napoca, Romania
E-mail address: cochifu@tbs.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume L, Number 2, June 2005

MIXED FUNCTIONAL DIFFERENTIAL EQUATION WITH PARAMETER

VERONICA ANA DÂRZU ILEA

Abstract. In this paper we study some special functional differential equation of mixed type. First we take an equation with some initial condition:

$$-\lambda x\prime(t,\lambda) = f(t, x(t,\lambda), x(t-h_1,\lambda), x(t+h_2,\lambda)), t \in R \tag{1}$$

$$x(t,\lambda) = \varphi(t,\lambda), \quad t \in [t_0 - h_1, t_0 + h_2], \tag{2}$$

where $h_1,\ h_2 > 0,\ \varphi \in C^1[t_0 - h_1, t_0 + h_2],\ \lambda \in R$ parameter.

The problem that we are interested in is the convergence of the solution of the problem (1)+(2) in the case $\lambda \neq 0$ to the solution of the same problem in case $\lambda = 0$. As an example of this problem we give the linear case of functional differential equation of mixed type. At the end of the article we discuss some inequalities between the solution of equation (1), inequalities that depend on λ.

1. **Introduction**

In this section we'll discuss the linear case of the mixed functional differential equation (MFDE) with parameter .

Let us consider the problem:

$$-\lambda x\prime(t,\lambda) = \alpha x(t,\lambda) + \beta x(t-h_1,\lambda) + \gamma x(t+h_2,\lambda), \quad t \in R \tag{3}$$

$$x(t,\lambda) = \varphi(t,\lambda), \quad t \in [t_0 - h_1, t_0 + h_2],\ \varphi \in C^1[t_0 - h_1, t_0 + h_2] \tag{4}$$

Received by the editors: 26.10.2004.

2000 *Mathematics Subject Classification.* 34K07.

Key words and phrases. Functional differential equation.

This paper was presented at International Conference on Nonlinear Operators, Differential Equations and Applications held in Cluj-Napoca (Romania) from August 24 to August 27, 2004.

where $h_1 \neq h_2,\ \beta, \gamma \neq 0,\ \lambda \in R$.

By a solution of equation (3) we mean a function $x \in C^1(R^2)$ which satisfy the relation (4) for any $t \in R,\ \lambda \in R$.

Remark. If $\lambda = 0$, the equation becomes a difference equation. Both the differential equation case $\lambda \neq 0$ and the difference equation case $\lambda = 0$, will be of interest.

In the first part of this section we assume that $\lambda \neq 0$. By the method of steps we'll find the solution for problem (3) + (4).

Let $t \in [t_0, t_0 + h_2]$. We have

$$-\lambda\varphi'(t,\lambda) = \alpha\varphi(t,\lambda) + \beta\varphi(t - h_1, \lambda) + \gamma x(t + h_2, \lambda) \tag{5}$$

It follows that

$$x(t) := x_1(t) = -\frac{1}{\gamma}[\alpha\varphi(t - h_2, \lambda) + \beta\varphi(t - h_1 - h_2, \lambda) + \lambda\varphi\prime(t - h_2, \lambda)], \tag{6}$$

$t \in [t_0 + h_2, t_0 + 2h_2]$.

Let $t \in [t_0 + h_2, t_0 + 2h_2]$. We have

$$-\lambda x_1'(t,\lambda) = \alpha x_1(t,\lambda) + \beta y_1(t - h_1, \lambda) + \gamma x(t + h_2, \lambda) \tag{7}$$

where

$$y_1(t - h_1, \lambda) := \begin{cases} \varphi(t - h_1, \lambda), & h_1 > h_2 \\ \varphi(t - h_1, \lambda), & h_1 < h_2,\ t \in [t_0 + h_2, t_0 + h_1 + h_2] \\ x_1(t - h_1, \lambda), & h_1 < h_2,\ t \in [t_0 + h_1 + h_2, t_0 + 2h_2] \end{cases} \tag{8}$$

It follows that

$$x(t) := x_2(t) = -\frac{1}{\gamma}[\alpha x_1(t - h_2, \lambda) + \beta y_1(t - h_1 - h_2, \lambda) + \lambda x_1'(t - h_2, \lambda)], \tag{9}$$

$t \in [t_0 + 2h_2, t_0 + 3h_2]$.

The same way, it follows that

$$x_n(t) = -\frac{1}{\gamma}[\alpha x_{n-1}(t - h_2, \lambda) + \beta y_{n-1}(t - h_1 - h_2, \lambda) + \lambda x_{n-1}'(t - h_2, \lambda)], \tag{10}$$

$t \in [t_0 + nh_2, t_0 + (n+1)h_2]$, where

$$y_{n-1}(t - h_1 - h_2, \lambda) :=$$

$$\begin{cases} \varphi(t - h_1 - h_2, \lambda), & h_1 > (n-1)h_2 \\ x_1(t - h_1 - h_2, \lambda), & (n-1)h_2 > h_1 > (n-2)h_2, \\ \dots & \\ x_{n-2}(t - h_1 - h_2, \lambda), & 2h_2 > h_1 > h_2, \\ x_{n-2}(t - h_1 - h_2, \lambda), & h_1 < h_2,\ t \in [t_0 + nh_2, t_0 + nh_2 + h_1), \\ x_{n-1}(t - h_1 - h_2, \lambda), & h_1 < h_2,\ t \in [t_0 + nh_2 + h_1, t_0 + (n+1)h_2] \end{cases} \tag{11}$$

We apply the same method on the left of t_0 and it follows

$$x_{-n}(t) = -\frac{1}{\beta}[\alpha x_{-(n-1)}(t+h_1, \lambda) + \gamma y_{-(n-1)}(t+h_1+h_2, \lambda) + \lambda x'_{-(n-1)}(t+h_1, \lambda)], \tag{12}$$

$t \in [t_0 - (n+1)h_1, t_0 - nh_1]$, where

$$y_{-(n-1)}(t + h_1 + h_2, \lambda) :=$$

$$\begin{cases} \varphi(t + h_1 + h_2, \lambda), & h_2 > (n-1)h_1 \\ x_{-1}(t + h_1 + h_2, \lambda), & (n-1)h_1 > h_2 > (n-2)h_1 \\ \dots & \\ x_{-(n-2)}(t + h_1 + h_2, \lambda), & 2h_1 > h_2 > h_1 \\ x_{-(n-2)}(t + h_1 + h_2, \lambda), & h_2 < h_1,\ t \in [t_0 - nh_1, t_0 - nh_1 + h_2) \\ x_{-(n-1)}(t + h_1 + h_2, \lambda), & h_2 < h_1,\ t \in [t_0 - nh_1 + h_2, t_0 + (n-1)h_1] \end{cases} \tag{13}$$

Next we have to find a condition for unique of the solution.

Let $\varphi \in C^\infty[t_0 - h_1, t_0 + h_2]$.

Let $x \in C^\infty(R)$ a solution of the problem (3)+(4).

We have

$$-\lambda x^{(k+1)}(t, \lambda) = \alpha x^{(k)}(t, \lambda) + \beta x^{(k)}(t - h_1, \lambda) + \gamma x^{(k)}(t + h_2, \lambda),\ k \in 0, 1, \cdots, n. \tag{14}$$

For $t = t_0$ we have

$$-\lambda \varphi^{(k+1)}(t, \lambda) = \alpha \varphi^{(k)}(t_0, \lambda) + \beta \varphi^{(k)}(t_0 - h_1, \lambda) + \gamma \varphi^{(k)}(t + h_2, \lambda),\ k \in 0, 1, \cdots, n. \tag{15}$$

Theorem 1.1. *The problem (3)+(4) have a unique solution if the relation (15) is done and the solution have the form*

$$x(t,\lambda)=\begin{cases} x_{-k}(t,\lambda), & t\in[t_0-(k+1)h_1,t_0-kh_1],\ k=1,2,\cdots,n,\ n\to\infty \\ \varphi(t,\lambda), & t\in[t_0-h_1,t_0+h_2] \\ x_k(t,\lambda), & t\in[t_0+kh_2,t_0+(k+1)h_2] \end{cases} \tag{16}$$

where

$$x_k(t,\lambda)=-\frac{1}{\gamma}[\alpha x_{k-1}(t-h_2,\lambda)+\beta y_{k-1}(t-h_1-h_2,\lambda)+\lambda x'_{k-1}(t-h_2,\lambda)], \tag{17}$$

$t\in[t_0+kh_2,t_0+(k+1)h_2]$

$$y_{k-1}(t-h_1-h_2,\lambda):=$$

$$\begin{cases} \varphi(t-h_1-h_2,\lambda), & h_1>(k-1)h_2 \\ x_1(t-h_1-h_2,\lambda), & (k-1)h_2>h_1>(k-2)h_2, \\ \cdots & \\ x_{k-2}(t-h_1-h_2,\lambda), & 2h_2>h_1>h_2, \\ x_{k-2}(t-h_1-h_2,\lambda), & h_1<h_2,\ t\in[t_0+kh_2,t_0+kh_2+h_1), \\ x_{k-1}(t-h_1-h_2,\lambda), & h_1<h_2, t\in[t_0+kh_2+h_1,t_0+(k+1)h_2] \end{cases} \tag{18}$$

$$x_{-k}(t,\lambda)=-\frac{1}{\beta}[\alpha x_{-(k-1)}(t+h_1,\lambda)+\gamma y_{-(k-1)}(t+h_1+h_2,\lambda)+\lambda x'_{-(k-1)}(t+h_1,\lambda)], \tag{19}$$

$t\in[t_0-(k+1)h_1,t_0-kh_1]$,

$$y_{-(k-1)}(t+h_1+h_2,\lambda):=$$

$$\begin{cases} \varphi(t+h_1+h_2,\lambda), & h_2>(k-1)h_1 \\ x_{-1}(t+h_1+h_2,\lambda), & (n-1)h_1>h_2>(n-2)h_1 \\ \cdots & \\ x_{-(k-2)}(t+h_1+h_2,\lambda), & 2h_1>h_2>h_1 \\ x_{-(k-2)}(t+h_1+h_2,\lambda), & h_2<h_1,\ t\in[t_0-kh_1,t_0-kh_1+h_2) \\ x_{-(k-1)}(t+h_1+h_2,\lambda), & h_2<h_1,\ t\in[t_0-kh_1+h_2,t_0+(k-1)h_1] \end{cases} \tag{20}$$

Let now $\lambda = 0$. The problem becomes:

$$0 = \alpha x(t,0) + \beta x(t-h_1,0) + \gamma x(t+h_2,0),\ t \in R \tag{21}$$

$$x(t,0) = \varphi(t,0),\ t \in [t_0-h_1, t_0+h_2] \tag{22}$$

If we apply in the same way the method of steps it follows

$$x_{-k}(t,0) = -\frac{1}{\beta}[\alpha x_{-(k-1)}(t+h_1,0) + \gamma y_{-(k-1)}(t+h_1+h_2,0)], \tag{23}$$

$t \in [t_0-(k+1)h_1, t_0-kh_1]$,

$$y_{-(k-1)}(t+h_1+h_2,0) :=$$

$$\begin{cases} \varphi(t+h_1+h_2,0), & h_2 > (k-1)h_1 \\ x_{-1}(t+h_1+h_2,0), & (n-1)h_1 > h_2 > (n-2)h_1 \\ \cdots & \\ x_{-(k-2)}(t+h_1+h_2,0), & 2h_1 > h_2 > h_1 \\ x_{-(k-2)}(t+h_1+h_2,0), & h_2 < h_1,\ t \in [t_0-kh_1, t_0-kh_1+h_2) \\ x_{-(k-1)}(t+h_1+h_2,0), & h_2 < h_1,\ t \in [t_0-kh_1+h_2, t_0+(k-1)h_1] \end{cases} \tag{24}$$

$$x_k(t,0) = -\frac{1}{\gamma}[\alpha x_{k-1}(t-h_2,0)+\beta y_{k-1}(t-h_1-h_2,0)],\ t \in [t_0+kh_2, t_0+(k+1)h_2] \tag{25}$$

$$y_{k-1}(t-h_1-h_2,0) :=$$

$$\begin{cases} \varphi(t-h_1-h_2,0), & h_1 > (k-1)h_2 \\ x_1(t-h_1-h_2,0), & (k-1)h_2 > h_1 > (k-2)h_2, \\ \cdots & \\ x_{k-2}(t-h_1-h_2,0), & 2h_2 > h_1 > h_2, \\ x_{k-2}(t-h_1-h_2,0), & h_1 < h_2,\ t \in [t_0+kh_2, t_0+kh_2+h_1), \\ x_{k-1}(t-h_1-h_2,0), & h_1 < h_2,\ t \in [t_0+kh_2+h_1, t_0+(k+1)h_2] \end{cases} \tag{26}$$

and

$$x(t,0) = \begin{cases} x_{-k}(t,0), & t \in [t_0-(k+1)h_1, t_0-kh_1],\ k = 1,2,\cdots,n,\ n \to \infty \\ \varphi(t,0), & t \in [t_0-h_1, t_0+h_2] \\ x_k(t,0), & t \in [t_0+kh_2, t_0+(k+1)h_2] \end{cases} \tag{27}$$

Theorem 1.2. *The problem (21) + (22) has a unique solution if and only if we have*

$$0 = \alpha\varphi^{(k)}(t_0, 0) + \beta\varphi^{(k)}(t_0 - h_1, 0) + \gamma\varphi^{(k)}(t_0 + h_2, 0),\ k \in 0, 1, \cdots, n. \quad (28)$$

We can give the following theorem

Theorem 1.3. *Let the problem (3) + (4) with the solution given by (16). If we put a limit on (16) when* $\lambda \to 0$ *we obtain the relation (27), which is the unique solution of the problem (21) + (22).*

2. Main results

In this section we consider the problem

$$-\lambda x\prime(t, \lambda) = f(t, x(t, \lambda), x(t - h_1, \lambda), x(t + h_2, \lambda)),\ t \in R \quad (29)$$

$$x(t, \lambda) = \varphi(t, \lambda),\ t \in [t_0 - h_1, t_0 + h_2],\ \varphi \in C^1[t_0 - h_1, t_0 + h_2] \quad (30)$$

where

$$\frac{\partial f(t, u)}{\partial u_j} \neq 0,\ j = 2, 3,\ u = (u_1, u_2, u_3),\ f \in C^\infty(R^4). \quad (31)$$

We first suppose that $\lambda \neq 0$.

Let the following conditions:

$(C1)$ For any $u_1,\ u_2,\ u_4,\ u_5 \in R$, there exist a unique $u_3 \in R$, $u_3 = f_1(u_1, u_2, u_4, u_5)$, $f_1 \in C^\infty(R^4)$, so that $u_5 = f(u_1, u_2, u_3, u_4)$.

$(C2)$ For any $u_1,\ u_2,\ u_3,\ u_5 \in R$, there exist a unique $u_4 \in R$, $u_4 = f_2(u_1, u_2, u_3, u_5)$, $f_2 \in C^\infty(R^4)$, so that $u_5 = f(u_1, u_2, u_3, u_4)$.

Remark. If $f \in C^\infty(R^4)$ and $x \in C^1(R^2)$ is a solution of the equation (29), then $x \in C^\infty(R^2)$.

Theorem 2.1. *Let* $f \in C^\infty(R^4)$ *satisfy the conditions* $(C1)$, $(C2)$.

If $\varphi \in C^\infty[t_0 - h_1, t_0 + h_2]$ *then the problem (29) + (30) has a unique solution if and only if the following relation takes place:*

$$-\lambda\varphi^{(k+1)}(t_0, \lambda) = [f(t, \varphi(t, \lambda), \varphi(t - h_1, \lambda), \varphi(t + h_2, \lambda))]^{(k)}_{t=t_0},\ k = 0, 1, \cdots, n. \quad (32)$$

Proof. By the method of steps we built the solution of the problem (29) + (30) as follows:

Let $t \in [t_0, t_0 + h_2]$,

$$-\lambda\varphi\prime(t,\lambda) = f(t, \phi(t,\lambda), \phi(t-h_1,\lambda), x(t+h_2,\lambda)). \tag{33}$$

From $(C2)$ we have

$$x(t,\lambda) := x_1(t,\lambda) = f_2(t-h_2, \varphi(t-h_2,\lambda), \varphi(t-h_1-h_2,\lambda), \lambda\varphi'(t-h_2,\lambda)), \tag{34}$$

$t \in [t_0 + h_2, t_0 + 2h_2]$.

By the same way we have:

$$x_k(t,\lambda) = f_2(t-h_2, x_{k-1}(t-h_2,\lambda), y_{k-1}(t-h_1-h_2,\lambda), \lambda x'_{k-1}(t-h_2,\lambda)), \tag{35}$$

$t \in [t_0 + kh_2, t_0 + (k+1)h_2]$, where

$$y_{k-1}(t-h_1-h_2,\lambda) :=$$

$$\begin{cases} \varphi(t-h_1-h_2,\lambda), & h_1 > (k-1)h_2 \\ x_1(t-h_1-h_2,\lambda), & (k-1)h_2 > h_1 > (k-2)h_2, \\ \dots & \\ x_{k-2}(t-h_1-h_2,\lambda), & 2h_2 > h_1 > h_2, \\ x_{k-2}(t-h_1-h_2,\lambda), & h_1 < h_2,\ t \in [t_0+kh_2, t_0+kh_2+h_1), \\ x_{k-1}(t-h_1-h_2,\lambda), & h_1 < h_2,\ t \in [t_0+kh_2+h_1, t_0+(k+1)h_2] \end{cases} \tag{36}$$

On the left of t_0 we obtain:

$$x_{-k}(t,\lambda) = f_1(t+h_1, x_{-(k-1)}(t+h_1,\lambda), y_{-(k-1)}(t+h_1+h_2,\lambda), \lambda x'_{-(k-1)}(t+h_1,\lambda)), \tag{37}$$

$t \in [t_0 - (k+1)h_1, t_0 - kh_1]$, where

$$y_{-(k-1)}(t+h_1+h_2,\lambda) :=$$

$$\begin{cases} \varphi(t+h_1+h_2,\lambda), & h_2 > (k-1)h_1 \\ x_{-1}(t+h_1+h_2,\lambda), & (n-1)h_1 > h_2 > (n-2)h_1 \\ \dots & \\ x_{-(k-2)}(t+h_1+h_2,\lambda), & 2h_1 > h_2 > h_1 \\ x_{-(k-2)}(t+h_1+h_2,\lambda), & h_2 < h_1,\ t \in [t_0-kh_1, t_0-kh_1+h_2) \\ x_{-(k-1)}(t+h_1+h_2,\lambda), & h_2 < h_1,\ t \in [t_0-kh_1+h_2, t_0+(k-1)h_1] \end{cases} \tag{38}$$

For the regularity of the solution we have the conditions:

$$\begin{cases} \varphi(t_0+h_2,\lambda) = x_1(t_0+h_2,\lambda) \\ x_p(t_0+(p+1)h_2,\lambda) = x_{p+1}(t_0+(p+1)h_2,\lambda),\ p \geq 1 \\ \varphi(t_0-h_1,\lambda) = x_{-1}(t_0-h_1,\lambda) \\ x_{-p}(t_0-(p+1)h_1,\lambda) = x_{-(p+1)}(t_0-(p+1)h_1,\lambda),\ p \geq 1 \end{cases} \tag{39}$$

So the solution is

$$x(t,\lambda) = \begin{cases} x_{-k}(t,\lambda), & t \in [t_0-(k+1)h_1, t_0-kh_1],\ k = 1,2,\cdots,n,\ n \to \infty \\ \varphi(t,\lambda), & t \in [t_0-h_1, t_0+h_2] \\ x_k(t,\lambda), & t \in [t_0+kh_2, t_0+(k+1)h_2] \end{cases} \tag{40}$$

We have to prove the necessity of the (42). Let $x \in C^1(R)$ solution of the problem (29)+(30). Then $x \in C^\infty(R)$ is a solution. We have

$$-\lambda x^{(k+1)}(t,\lambda) = [f(t,x(t,\lambda),x(t-h_1,\lambda),x(t+h_2,\lambda))]^{(k)},\ t \in R,\ k = 0,1,\cdots,n \tag{41}$$

For $t = t_0$ it follows

$$-\lambda\varphi^{(k+1)}(t_0,\lambda) = [f(t,\varphi(t,\lambda),\varphi(t-h_1,\lambda),\varphi(t+h_2,\lambda))]^{(k)}_{t=t_0},\ k = 0,1,\cdots,n. \tag{42}$$

Let now $\lambda = 0$

The problem becomes

$$0 = f(t,x(t,0),x(t-h_1,0),x(t+h_2,0)),\ t \in R \tag{43}$$

$$x(t,0) = \varphi(t,0),\ t \in [t_0-h_1,t_0+h_2],\ \varphi \in C^1[t_0-h_1,t_0+h_2] \tag{44}$$

Theorem 2.2. *Let $f \in C^{\infty}(R^4)$ satisfy the conditions $(C1)$, $(C2)$.*

If $\varphi \in C^{\infty}[t_0-h_1, t_0+h_2]$ then the problem (43) + (44) has a unique solution if and only if the following relation takes place:

$$-\lambda\varphi^{(k+1)}(t_0,0) = [f(t,\varphi(t,0),\varphi(t-h_1,0),\varphi(t+h_2,0))]^{(k)}_{t=t_0},\ k=0,1,\cdots,n. \quad (45)$$

Proof. The proof is similar with the proof of the last theorem. We find by the method of steps that the unique solution of the problem (43) + (44) has the form

$$x(t,0) = \begin{cases} x_{-k}(t,0), & t\in[t_0-(k+1)h_1, t_0-kh_1],\ k=1,2,\cdots,n,\ n\to\infty \\ \varphi(t,0), & t\in[t_0-h_1,t_0+h_2] \\ x_k(t,0), & t\in[t_0+kh_2,t_0+(k+1)h_2] \end{cases} \quad (46)$$

where

$$x_k(t,0) = f_2(t-h_2, x_{k-1}(t-h_2,0), y_{k-1}(t-h_1-h_2,0)), t\in[t_0+kh_2, t_0+(k+1)h_2], \quad (47)$$

$$y_{k-1}(t-h_1-h_2,0) :=$$

$$\begin{cases} \varphi(t-h_1-h_2,0), & h_1 > (k-1)h_2 \\ x_1(t-h_1-h_2,0), & (k-1)h_2 > h_1 > (k-2)h_2, \\ \cdots & \\ x_{k-2}(t-h_1-h_2,0), & 2h_2 > h_1 > h_2, \\ x_{k-2}(t-h_1-h_2,0), & h_1 < h_2,\ t\in[t_0+kh_2, t_0+kh_2+h_1), \\ x_{k-1}(t-h_1-h_2,0), & h_1 < h_2,\ t\in[t_0+kh_2+h_1, t_0+(k+1)h_2] \end{cases} \quad (48)$$

$$x_{-k}(t,0) = f_1(t+h_1, x_{-(k-1)}(t+h_1,0), y_{-(k-1)}(t+h_1+h_2,0)), \quad (49)$$

$t\in[t_0-(k+1)h_1, t_0-kh_1]$,

$$y_{-(k-1)}(t+h_1+h_2,0) :=$$

$$\begin{cases} \varphi(t+h_1+h_2,0), & h_2 > (k-1)h_1 \\ x_{-1}(t+h_1+h_2,0), & (n-1)h_1 > h_2 > (n-2)h_1 \\ \dots & \\ x_{-(k-2)}(t+h_1+h_2,0), & 2h_1 > h_2 > h_1 \\ x_{-(k-2)}(t+h_1+h_2,0), & h_2 < h_1,\ t \in [t_0-kh_1, t_0-kh_1+h_2) \\ x_{-(k-1)}(t+h_1+h_2,0), & h_2 < h_1,\ t \in [t_0-kh_1+h_2, t_0+(k-1)h_1] \end{cases} \tag{50}$$

It is natural that the solution of problem (29) + (30) converge, when $\lambda \to 0$, to the solution of the problem (43) + (44).

So we can give the following theorem:

Theorem 2.3. *Let $f \in C(R^4), \lambda$ parameter. Let $x(t,\lambda)$ given by (40) the solution of the problem (29) + (30). If we put $\lambda \to 0$ then $x(t,\lambda)$ given by (40) converge punctually to $x(t,0)$ given by (46).*

3. **The comparing of the solutions**

Let the equation

$$-\lambda x\prime(t,\lambda) = f(t, x(t,\lambda), x(t-h_1,\lambda), x(t+h_2,\lambda)),\ t \in R \tag{51}$$

where

(i) $h_1 \neq h_2$;

(ii) $f: R^4 \to R,\ f(t,u)$ - continuous, local Lipschitz on u, $u = (u_1, u_2, u_3)$;

(iii) for any $t \in R$: $\frac{\partial f(t,u)}{\partial u_j} \geq 0$, $u \in R^3$, $j = 2, 3$, $u = (u_1, u_2, u_3)$.

Lemma 3.1. *Assume that (ii) and (iii) above hold. Let $x_j : R \to R$, for $j = 1, 2$, be two solutions of equation (51) at some nonzero parameter value $\lambda \in R^*$. Assume that*

$$x_1(t,\lambda) \geq x_2(t,\lambda),\ t \in R. \tag{52}$$

Then if $x_1(\tau,\lambda) = x_2(\tau,\lambda)$ at some $\tau \in R$, we have that

$x_1(\xi,\lambda) = x_2(\xi,\lambda)$ for all $\xi \geq \tau$, in case $\lambda > 0$, and that

$x_1(\xi,\lambda) = x_2(\xi,\lambda)$ for all $\xi \leq \tau$, in case $\lambda < 0$.

Proof. Let $y(t,\lambda) := x_1(t,\lambda) - x_2(t,\lambda) \geq 0$, and we assume that the inequality (52) is an equality at some $\tau \in R$.

It means that $y(\tau,\lambda) = 0$, so that $x_1(\tau,\lambda) = x_2(\tau,\lambda)$.

We prove for $c > 0$; the proof when $c < 0$ being similar.

Let the function

$$H(\xi,u) := -\lambda^{-1}[G(\xi, u + x_2(\xi,\lambda), x_1(\xi - h_1,\lambda), x_1(\xi + h_2,\lambda))$$

$$-G(\xi, x_2(\xi,\lambda), x_2(\xi - h_1,\lambda), x_2(\xi + h_2,\lambda)]. \tag{53}$$

We observe that $u = y(\xi,\lambda)$ satisfies $u' = H(\xi,u)$.

From (52) and (iii) we have that $H(\xi,0) \leq 0$, for every $\xi \in R$.

It follows that $y(\xi,\lambda) \leq 0$ for all $\xi \geq \tau$ by a standard differential inequality. So that $y(\xi,\lambda) = 0$ for all $\xi \geq \tau$.

Thus $x_1(\xi,\lambda) = x_2(\xi,\lambda)$ for all $\xi \geq \tau$, in case $\lambda > 0$ and $x_1(\xi,\lambda) = x_2(\xi,\lambda)$ for all $\xi \leq \tau$, in case $\lambda < 0$.

Lemma 3.2. *Assume that the conditions of lemma (3.1) hold, except that solutions x_1, x_2 satisfy equation (51) at different values of λ, with $\lambda_1 > \lambda_2$, and where either $\lambda_1 = 0$ or $\lambda_2 = 0$ are permitted. Suppose that $x_1(\tau,\lambda_1) = x_2(\tau,\lambda_2)$ at some $\tau \in R$. Then if*

$$x_2(\xi,\lambda_2) \text{ is monotone increasing in } \xi \in R$$

and $\lambda_1 > 0$ we have that $x_1(\xi,\lambda_1) = x_2(\xi,\lambda_2)$ is constant for all $\xi \geq \tau + h_1$; while if

$$x_1(\xi,\lambda_1) \text{ is monotone increasing in } \xi \in R$$

and $\lambda_2 > 0$ we have that $x_1(\xi,\lambda_1) = x_2(\xi,\lambda_2)$ is constant for all $\xi \leq \tau - h_2$;

Proof. The proof is very similar to that of Lemma (3.1) except for the choice of the function H. Several cases must be considered, based on the signs of λ_1 and λ_2.

First, we suppose that (3.2) holds and $\lambda_2 \neq 0$.

Let

$$H(\xi,u) := -\lambda_1^{-1}G(\xi, u + x_2(\xi,\lambda), x_1(\xi - h_1,\lambda), x_1(\xi + h_2,\lambda))$$

$$+\lambda_2^{-1}G(\xi, x_2(\xi\lambda), x_2(\xi - h_1,\lambda), x_2(\xi + h_2,\lambda). \tag{54}$$

Then $u = x_1(\xi, \lambda_1) - x_2(\xi, \lambda_1)$ satisfy $u' = H(\xi, u)$.

By replacing x_1 with x_2 in the formula of H and using the inequality (52) and (iii), form (3.2) follows that

$H(\xi, 0) \leq (\lambda_2\lambda_1^{-1} - 1)x_2'(\xi, \lambda_2) \leq 0.$

Thus $x_1(\xi, \lambda_1) - x_2(\xi, \lambda_2) \leq 0$ for all $\xi \geq \tau$.

So that $x_1(\xi, \lambda_1) = x_2(\xi, \lambda_2)$, $\xi \geq \tau$.

From (51) we have

$\lambda_1 x_1'(\xi, \lambda_1) = \lambda_2 x_2'(\xi, \lambda_2)$ for all $\xi \geq \tau + h_1$.

If $\lambda_1 \neq \lambda_2$ we have $x_1'(\xi, \lambda_1) = x_2'(\xi, \lambda_2) = 0$, $\xi \geq \tau + h_1$; so $x_1(\xi, \lambda_1) = x_2(\xi, \lambda_2) = const$, $\xi \geq \tau + h_1$.

Now suppose that (3.2) holds and $\lambda_2 = 0$.

Let

$H(\xi, u) := -\lambda_1^{-1} G(\xi, u, x_1(\xi - h_1, \lambda_1), x_1(\xi + h_2, \lambda_1))$

$$\widetilde{H}(\xi, u) = \begin{cases} H(\xi, x_2(\xi, \lambda_2)), & u \geq x_2(\xi, \lambda_2) \\ H(\xi, u), & u \leq x_2(\xi, \lambda_2) \end{cases} \tag{55}$$

H can be easily checks that $H(\xi, x_2(\xi, \lambda_2)) \leq 0$, so that $\widetilde{H}(\xi, u) \leq 0$. From inequality, form the fact that x_2 is monotone increasing and that $\widetilde{H}$ satisfies the standard Caratheodory and Lipschitz conditions, we have that the unique solution $u = x_3(\xi, \lambda_2)$ satisfies $u' = H(\xi, u)$, for any $\xi \geq \tau$.

But $u = x_1(\xi, \lambda_1)$ satisfies $u' = H(\xi, u)$ and $x_3(\tau) = x_2(\tau) = x_1(\tau)$.

So that $x_1(\xi, \lambda_1) \leq x_2(\xi, \lambda_2)$ for any $\xi \geq \tau$.

Thus $x_1(\xi, \lambda_1) = x_2(\xi, \lambda_2)$ for any $\xi \geq \tau$.

So $x_1'(\xi) = x_2'(\xi) = 0$ for any $\xi \geq \tau + h_1$.

It follows that $x_1(\xi, \lambda_1) = x_2(\xi, \lambda_2) = const$ for any $\xi \geq \tau + h_1$.

References

[1] Darzu, V. A., *Some Existence Results for Differential Equations of Mixed Type,* Itinerant Seminar Tiberiu Popoviciu, mai 2001, 43-46.

[2] Darzu, V. A., *Weeler-Feynman problem on a compact interval*, Semin. Fixed Point Theory Cluj-Napoca 3(2002), 389-392.

[3] Darzu, V.A., *Functional Differential Equation of Mixed Type, via Weakly Picard Operators*, Proc. 6th Conf. of the Romanian Math. Soc., 276-284.

[4] Darzu, V. A., *Data dependence for functionall differential equations of mixed type*, Mathematica (Cluj) (to appear).

[5] Wheeler, J. A., Feynman, R. P., *Classical electordynamics in terms of direct interparticles action*, Rev. Mod. Phys. 21(1949), 425-433.

[6] Mallet-Paret, J., *The global structure of travelling waves in spatially discrete dynamical systems*, J. Dynam. Diff. Eq., vol. 11, no. 1(1999), 49-127.

[7] Mallet-Paret, J., *Exponential dichotonomies and Wiener-Hopf factorizations for mixed-type functional differential equations*, 2001, 1-86.

[8] Mallet-Paret, J., *The Fredholm alternative for functional differential equations of mixed type*, J. Dynam. Diff. Eq., vol 11, no.1(1999), 1-46.

[9] Precup, R., *Some Existence Results for Differential Equations with Both Retarded and Advanced Arguments*, Mathematica, Cluj-Napoca, 2002, to appear.

[10] Rus, I. A., *Ecuaţii diferenţiale. Ecuaţii integrale şi sisteme dinamice*, Transilvnia Press, Cluj-Napoca, 1996.

[11] Rus, I. A., Iancu, C., *Wheeler-Feynman problem for mixed first order functional differential equations*, Seminar Itinerant Tiberiu Popoviciu, Cluj-Napoca, mai 2000, 197-200.

[12] Rus, I. A., *Principii şi aplicaţii ale teoriei punctului fix*, Editura Dacia, Cluj-Napoca, 1979.

Department of Applied Mathematics, Babeş-Bolyai University,
Str. M. Kogălniceanu 1, Cluj-Napoca, Romania
E-mail address: vdarzu@math.ubbcluj.ro

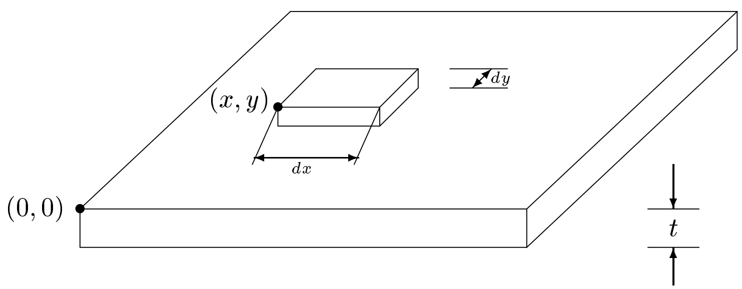
STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

ON SOME DISTRIBUTION PROBLEM OF THE TEMPERATURE IN A METALLIC PLATE

CĂTĂLIN MITRAN

Abstract. Many problems of technical, industrial, economical type have can be simulated using the differential equations or the partial equations derivatives. But many times the determining of an analytical solution is a difficult, even impossible problem. This is the reason for which the numerical approximation, generally, and in this case the finite differences are a good solution for solving the mentioned problems.

We shall consider a plate made uniformly which has a thickness of t and who contain an element of measure $dx \times dy$. We shall take u the independent variable who represent the temperature into the element. The location of the element is in (x, y), situated in the left position of the plate. We shall consider that the high tide cross the element on the positive x axis direction and also cross the element in the y direction too as it can be seen in the next figure:



Received by the editors: 13.10.2004.
2000 *Mathematics Subject Classification.* 35Q05, 35Q72.
Key words and phrases. Partial differential equations and applications.

The ratio in which the heat divides the element in the x direction is given by

$$-kA\frac{\partial u}{\partial x} = -k(tdy)\frac{\partial u}{\partial x} \tag{1}$$

and, similar, in the y direction is given by

$$-kA\frac{\partial u}{\partial y} = -k(tdx)\frac{\partial u}{\partial y} \tag{2}$$

where k is the conductivity and A is the area.

We have that the ratio of the high tide who cross in must be equal with the ratio of the high tide who cross out. The high tide who cross in is given by

$$-kA\frac{\partial u}{\partial x} - kA\frac{\partial u}{\partial y} = -k(tdy)\frac{\partial u}{\partial x} - k(tdx)\frac{\partial u}{\partial y} \tag{3}$$

and the high tide who cross out is given by

$$-k(tdy)\left[\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}dx\right] - k(tdx)\left[\frac{\partial u}{\partial y} + \frac{\partial^2 u}{\partial y^2}\right] + Q(dxdy) \tag{4}$$

We shall obtain that

$$kt\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)(dxdy) = Q(dxdy) \tag{5}$$

or

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{Q}{kt} \tag{6}$$

where Q is the heat.

If the object is considered as being in the space the relation (6) will become

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{Q}{kt} \tag{7}$$

that means

$$\Delta u = \frac{Q}{kt}. \tag{8}$$

If the thickness of the plate is variable with x and y the relation (7) become

$$t\Delta^2 u + \frac{\partial t}{\partial x}\left(\frac{\partial u}{\partial x}\right) + \frac{\partial t}{\partial y}\left(\frac{\partial u}{\partial y}\right) = \frac{Q}{k} \tag{9}$$

If the thickness and the thermic conductivity of the plate are variable the relation (8) will become

$$kt\Delta^2 u + \left(k\frac{\partial t}{\partial x} + t\frac{\partial k}{\partial x}\right)\left(\frac{\partial u}{\partial x}\right) + \left(k\frac{\partial t}{\partial y} + t\frac{\partial k}{\partial y}\right)\left(\frac{\partial u}{\partial y}\right) = Q \tag{10}$$

We shall have the next approximation:

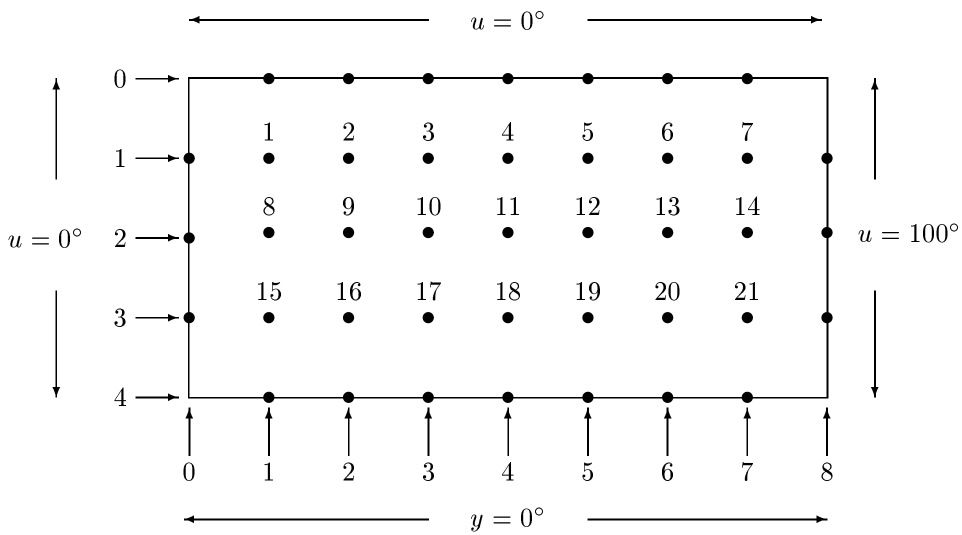
$$\frac{\partial^2 u}{\partial x^2} = \frac{u_L - 2u_0 + u_R}{(\Delta x)^2}$$
$$\frac{\partial^2 u}{\partial y^2} = \frac{u_U - 2u_0 + u_D}{(\Delta y)^2} \tag{11}$$

where u_L and u_R are the temperatures at the left and right and u_U and u_D are the temperatures up and down of the considered knot.

Usual, we have that $\Delta x = \Delta y = h$ and from this we shall obtain that

$$\Delta^2 u = \frac{u_L + u_R + u_U + u_D - 4u_0}{h^2} \tag{12}$$

We shall present an example who will show how can be applied the presented formulas with finite differences. Let us consider the next problem where the dates are presented in the next figure:



So, we have a plate of dimensions 20 cm × 10 cm and the space between the knots is about 2.5 cm. We have 21 interior knots. On three sides of the plate the value of u is 0°. On the last side the value of u is 100°. If we suppose that $Q = 0$ the equation (6) will be

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

If we take care of (10) and (11) we shall have

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_L - 2u_0 + u_R}{2.5^2}$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_U - 2u_0 + u_R}{2.5^2}$$

respectively

$$\frac{u_L + u_R + u_U + u_D - 4u_0}{2.5^2} = 0.$$

We shall write the expression (12) in the case of a lot of knots. We shall have:

1) for the knot no. 1: $-4u_1 + u_2 + u_8 = 0$;

2) for the knot no. 7: $u_6 - 4u_7 + u_{14} = -100$;

3) for the knot no. 9: $u_2 + u_8 - 4u_9 + u_{10} + u_{16} = 0$;

4) for the knot no. 19: $u_{18} + u_{12} - 4u_{19} + u_{20} = 0$;

5) for the knot no. 21: $u_{20} + u_{14} - 4u_{21} = -100$.

Finally, we shall obtain a system of 21 equations with a number of 21 unknowns with the solution given by the next table:

Line	Column No. 1	Column No. 2	Column No. 3	Column No. 4	Column No. 5	Column No. 6	Column No. 7
1	0.3530	0.9132	2.0103	4.2957	9.1531	19.6631	43.2101
1	0.4988	1.2894	2.8323	7.0193	12.6537	27.2893	53.1774
3	0.3530	0.9132	2.0103	4.2957	9.1531	19.6631	43.2101

Many other techniques of solving the above mentioned equations use iterative methods in which a knot is written depending of other knots. For example we shall have the relations:

$$u_{i,j} = \frac{i_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}}{4}$$

or

$$u_{i,j} = \frac{u_{i-1,j-1} + 4u_{i,j-1} + u_{i+1,j} + 4u_{i,j-1} - 20u_{i,j} + 4u_{i,j+1} + u_{i+1,j-1} + 4u_{i+1,j} + u_{i+1,j+1}}{6}.$$

Definition 1. An equation of type:

$$\Delta u = R$$

where $R = R(x, y)$ is a function defined on the same domain with u will be named just a *Poisson equation*.

For example, if we have to solve the next equation:

$$\Delta^2 u = -2$$

we shall use the next approximation

$$u_{i,j} = \frac{u_{i,j-1} + u_{i,j+1} + u_{i-1,j} + u_{i+1,j} + 2}{4}.$$

Definition 2. An equation of type $\Delta u = 0$ and which satisfies a condition of the next type:

$$Au + B = Cu' \tag{13}$$

where A, B, C are constants will be named an *equation of Neumann type.*

Remark 1.

$$-ku' = H(u - u_s) \tag{14}$$

is also a condition of Neumann type if we take care of (12) and take $A = H$, $B = -Hu_s$, $C = -k$.

We shall present a new model with a plate having the dimensions 5 cm × 9 cm and a thickness of 0.5 cm. We will take the next values:

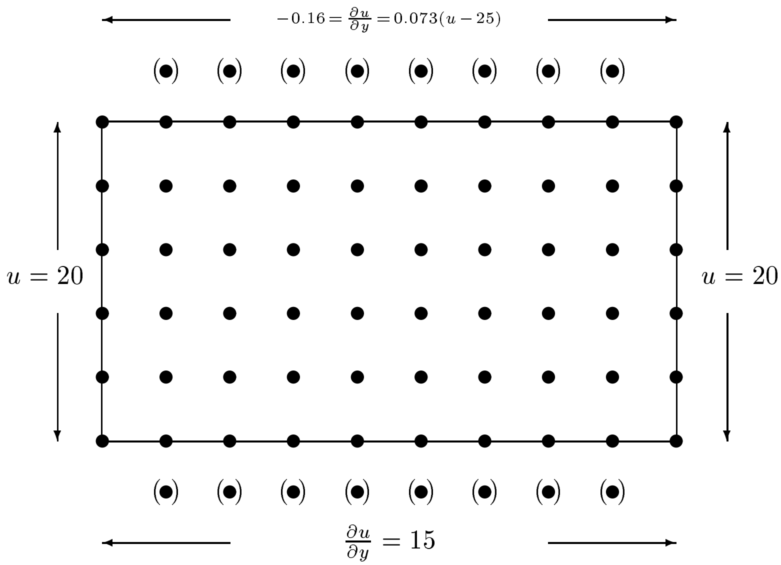
1) $h = 1$ cm;

2) $Q = 0.6 cal/s/cm^3$;

3) $k = 0.16$ is the thermic conductivity;

4) $H = 0.073$ is the coefficient of the heat transfer.

The frontier conditions are given by the next figure:



Solving the problem using the mentioned approximations we shall obtain the next values included in the next table:

20.000	73.510	107.915	128.859	138.826	138.826	128.859	107.915	73.510	20.000
20.000	99.195	137.476	167.733	180.743	180.743	167.733	137.476	99.195	20.000
20.000	99.793	155.061	189.855	207.669	207.669	189.855	155.061	99.793	20.000
20.000	103.918	163.119	200.956	219.410	219.410	200.956	163.119	103.918	20.000
20.000	102.762	162.539	201.442	220.604	220.604	201.442	162.539	102.762	20.000
20.000	94.589	152.834	191.669	210.959	210.959	191.669	152.834	94.589	20.000

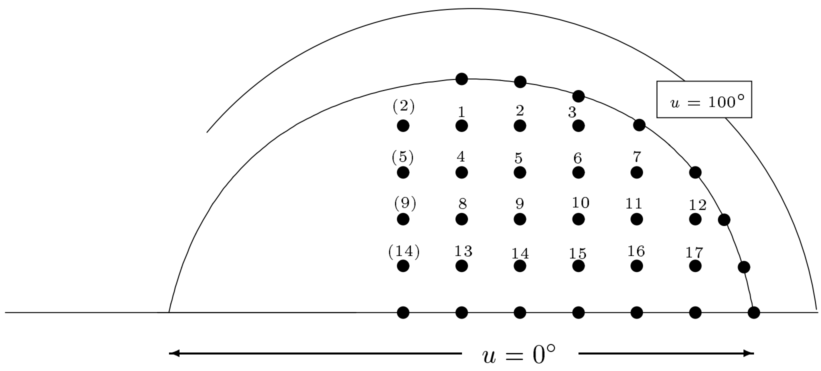
If we have a circular domain or, generally, an irregular domain it is recommended to use the polar coordinates:

$$\Delta^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} \tag{15}$$

and, from here, the approximation formula

$$\Delta^2 u = \frac{u_L - 2u_0 + u_R}{(\Delta r)^2} + \left(\frac{1}{r}\right)\left(\frac{u_R - u_L}{2\Delta r}\right) + \frac{1}{r^2}\left(\frac{u_U - 2u_0 + u_D}{(\Delta\theta)^2}\right) \tag{16}$$

Using (16) we shall solve the problem whose dates are given in the next figure:



We shall obtain the results presented in the next table:

Knot	Calculated value	Analytical value
1	86.053	85.906
2	87.548	87.417
3	92.124	92.094
4	69.116	68.807
5	70.733	70.482
6	75.994	765.772
7	85.471	85.405
8	48.864	48.448
9	50.436	50.000
10	55.606	55.151
11	65.891	65.593
12	84.189	84.195
13	25.466	25.133
14	27.501	27.109
15	30.102	29.527
16	38.300	37.436
17	57.206	57.006

References

[1] Heatley, G., *Numerical analysis with applications*, Addison Wesley Longman, Inc., U.S.A., 1999.

[2] Dragotă, I., Petrehuş, V., *Metode numerice pentru ecuaţii diferenţiale*, Editura Orizonturi Universitare, Timişoara, 2002.

[3] Arieh Iserles, *Numerical analysis of differential equations*, Cambridge Univ. Press, USA, 1996.

Department of Mathematics, University of Petroşani

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

COMPARISON BETWEEN DIFFERENT HARVESTING MODELS FOR NON-LINEAR AGE STRUCTURED FISH POPULATIONS

MOSTAFA K. S. MOHAMED AND S. M. KHALED

Abstract. The role of harvesting in discrete nonlinear age-structured fish population models has been studied. The overcompensatory Ricker recruitment function is considered in our model. We show numerically that the maximum sustainable yield (MSY) in harvesting with nets is different very little from (MSY) in selective harvesting. Our models contain a large number of parameters such as mortality, Von Bertalanffy growth parameter and recruitment parameters. The influence of mortality has been studied. The age structured matrix model (general Leslie model) for description of harvesting population dynamics has been used because most marine fish exhibit a clear yearly cycle of spawning, recruitment, migration and growth.

1. Introduction

Our basic model is a nonlinear discrete age-structured population model:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}_{t+1} = \begin{pmatrix} 0 & \alpha f_2 r(P) & \cdots & \alpha f_m r(P) \\ \tau_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \tau_{m-1} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}_t \tag{1}$$

The model is based on general biological principles and contains a large number of parameters and functions. Concrete data and further informations will be used to reduce the number of parameters and functions and determine the range of critical

Received by the editors: 01.02.2005.
2000 *Mathematics Subject Classification.* 92B10.
Key words and phrases. Age structured fish population, harvesting model, Ricker recruitment function, nonlinear Leslie matrix model.

parameters. Properties of this model class will be investigated and compared with observations. All models are deterministic. The ultimate goal of these models is to obtain precise informations about the state of the species. These informations can then be used to generate recommendations on catches, quotas and equipments.

The models are of the Leslie type and the only nonlinearity is the recruitment function, which we choose to be of Ricker type. Actual date give little support for the precise form of the recruitment function. Otherwise the parameters and functions used are chosen as to reflect concrete marine fish species. In the classical paper by Levin and Goodyear [2], model (1) was used in order to investigate the dynamics of the striped bass in the Hudson river. Such models have been described in many articles, cf. Cushing [1], Getz and Haight [3] and a classical paper by Leslie [4]. In such a model, time is considered as a discrete variable, measured in years. It is most sensible to identify the beginning of the year with spawning. Since selective harvesting is an unrealistic idealization, we concentrate on net harvesting and study in particular the role of the mesh width of fishing nets.

The goal of this paper is to compare between selective harvesting and harvesting with nets. Also the influence mortality on our models is investigated. The plane of the paper is as follows: In Section 2 we present the selective harvesting model while harvesting with nets is presented in Section 3. In Section 4 we give the results of our investigations by using haddock as a numerical example and finally in Section 5, we state the conclusions.

2. **The selective harvesting model**

Let $x_i(t)$, be the i-th age class of a fish population at time t. Denote the corresponding fecundity by f_i . Next, we let each age class i be exposed to harvesting with constant harvest rate $h_i, i=$ 2,3,...,m, i.e., there is no harvesting in the first class, where m is the maximum age class. So, the model after harvesting has the matrix form:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}_{t+1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-h_2 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \cdots & \tau_{m-1} & 1-h_m \end{pmatrix} \cdot$$

$$\begin{pmatrix} 0 & \alpha f_2 r(P) & \cdots & \alpha f_m r(P) \\ \tau_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \tau_{m-1} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}_t \tag{2}$$

or in vector form, $\underline{x}(t+1) = (I-H)A\underline{x}(t)$, where $P(t)$ is the number of recruits (new borns) of fish from all age classes at time t i.e.,

$$P(t) = \sum_{i=2}^{m} f_i x_i(t)$$

$\tau_i = 1 - \mu_i$ is the density independent probability of survival from age class i to age class $i+1$, where μ_i is the mortality rate of age class i. Finally $r(P)$ describes the recruitment process. $r(P)$ is is a non-negative monotonically decreasing function and α is the productivity parameter, representing the probability of survival of eggs at low densities.

Under the assumptions above, the components of equilibrium vector of (2) are:

$$x_1^* = \frac{L_1 r^{-1}\left(\frac{1}{n(h_j)}\right)}{n(h_j)}, \qquad x_2^* = \frac{L_2 H_2 r^{-1}\left(\frac{1}{n(h_j)}\right)}{n(h_j)}$$

and generally,

$$x_i^* = \frac{L_i H_i r^{-1}\left(\frac{1}{n(h_j)}\right)}{n(h_j)}, \qquad i = 1, 2, \ldots, m \tag{3}$$

where

$L_i = \prod_{j=1}^{i} \tau_{j-1}$ is the survival probability from age class 1 to age class i,
$(L_1 = 1), H_1 = 1, H_i = (1-h_2)(1-h_3)\cdots(1-h_i), \qquad i = 2, 3, \cdots, m$

and

$$n(h_i) = \sum_{i=2}^{m} \alpha f_i L_i H_i$$

is called "*the net reproductive number*" because biologically it gives the expected number of offspring per individual over its life time, cf., Cushing and Yicang [5]. Finally the yield is defined by:

$$Y(h_i) = \sum_{i=2}^{m} \frac{w_i h_i x_i^*}{(1-h_i)},$$

where w_i is the growth weight of adult fish and it is described by *Von Bertalanffy growth equation* [8].

$$w(t) = w_\infty (1 - e^{-K(t-\widetilde{t})^3}$$

where K is the rate at which growth weight tends towards its asymptotic value and $\widetilde{t}$ is the age at which growth weight starts.

So from equation (3), we get that the yield in selective harvesting is:

$$Y(h_i) = \frac{r^{-1}\left(\frac{1}{n(h_i)}\right)}{n(h_i)} \sum_{i=2}^{m} w_i h_i L_i H_{i-1}, \tag{4}$$

The maximum sustainable yield is now

$$\max_h Y(h) = Y_{\max}.$$

Reed [6] showed that for selective harvesting, the optimal policy is of the "two-age" type. This means that if we define $j(t)$ recursively by:

$$j(1) = \arg_j \max \frac{w_j L_j}{\sum_{i=j}^{m} f_i L_i}$$

and

$$j(t+1) = \arg_j \max \frac{w_j L_j - w_{j(t)} L_{j(t)}}{\sum_{i=j}^{j(t)-1} f_i L_i}, \ t = 1, 2, \cdots$$

There is a partial harvest at age $j(t+1)$ and a total harvest at age $j(t)$ where $j(t) > j(t+1)$. If we consider that the fecundity is proportional to the weight and assume the mortality is constant, one gets:

$$j(t) = m - (t-1), t = 1, 2, \cdots .$$

If fecundity is proportional to the weight and the mortality is increasing with age, then $j(1) = m$ and $j(t+1), t = 1, 2, \cdots$ are generally smaller than $j(t+1)$ in constant mortality, c.f. Mostafa K. S. Mohamed [7].

The maximum sustainable yield for selective harvesting is:

$$\begin{aligned} Y(h_{j(t+1)}) \quad = \quad & \frac{r^{-1}\left(\frac{1}{n(h_{j(t+1)})}\right)}{n(h_{j(t+1)})} \prod_{j=1}^{j(t+1)} \tau_{j-1} \cdot \\ & \left\{ w_{j(t)}(1-h_{j(t+1)}) \prod_{j=j(t+1)+1}^{j(t)} \tau_{j-1} + w_{j(t+1)} h_{j(t+1)} \right\} \end{aligned} \tag{5}$$

where $r^{-1}(x)$ is given from the stock-recruitment relationships and $h_{j(t+1)}$ is partial harvesting intensity. It can not determined analytically.

We will use the Ricker recruitment relationship as an example. Thus

$$r(x) = e^{-\beta x}, \qquad r^{-1}\left(\frac{1}{n(h)}\right) = \frac{Ln(n(h))}{\beta}$$

Note that in this case, we use a normalized formula for $r(P)$ i.e., $r(0) = 1$.

3. Harvesting with nets

If we want to model fishing with nets, we have to translate the width of fishing nets into this model. This will be done as follows. We write $H = h$ $diag(0, 0, \cdots, \gamma, 1, \cdots, 1)$ to describe the situation where all fish from class $k+1$ or more are caught, while all fish of class $k-1, k-2, \cdots$ can escape and fish of class k only a fraction $0 \leq \gamma \leq 1$ is retained. By this we mean that the mesh width is too small for fish from class $k+1$. With the term γ, we can model the fact that the mesh width is a continuous variable.

Now we can use the formulae from selective harvesting with

$$h_i = \begin{cases} 0, & 1 \le i \le k-1 \\ \gamma h, & i = k \\ h, & k \le i \le m \end{cases} \tag{6}$$

The components of equilibrium vector are:

$$\overline{x_i}(h) = \frac{L_i(h)}{n(h)}.r^{-1}\left(\frac{1}{n(h_i)}\right), \qquad i = 1, 2, \cdots, m$$

where

$$n(h) = \sum_{i=1}^{m} \alpha f_i.L_i(h)$$

and

$$L_i(h) = \begin{cases} \prod_{l=1}^{i} \tau_{l-1} & 1 \le i \le k-1 \\ \prod_{l=1}^{k-1} \tau_{l-1}\tau_{k-1}(1-\gamma h) & i = k \\ \prod_{l=1}^{k-1} \tau_{l-1}\tau_{k-1}(1-\gamma h) \prod_{j=k+1}^{i} \tau_{j-1}(1-h)^{j-k}, & k < i \le m. \end{cases}$$

The yield in net harvesting is then

$$Y(h) = \frac{C\ h\ r^{-1}\left(\frac{1}{n(h)}\right)}{n(h)} \left\{ w_k\ \gamma + \sum_{i=k+1}^{m} w_i\ (1-\gamma\ h) \prod_{j=k+1}^{i} \tau_{j-1}\ (1-h)^{j-k-1} \right\} \tag{7}$$

where

$$C = \prod_{l=1}^{k-1} \tau_l$$

We will consider for simplicity that $\gamma = 0$ only for all numerical computations, so equation (7) will be:

$$Y(h) = \frac{h\ r^{-1}\left(\frac{1}{n(h)}\right)}{n(h)} \left\{ w_{k+1}\ l_{k+1} + \sum_{i=k+2}^{m} w_i\ l_i(1-h)^{i-k-1} \right\}$$

where,

$$n(h) = \sum_{i=1}^{k} \alpha f_i l_i + \sum_{i=k+1}^{m} \alpha f_i l_i (1-h)^{i-k}$$

and

$$l_{i+1} = \prod_{j=0}^{i} \tau_j, \qquad l_1 = \tau_o = 1$$

One can show that the function Y has a unique maximum by deriving:

$$\frac{d^2Y}{dh^2} < 0.$$

4. **A numerical example**

Now, we will study the optimal harvesting for haddock using the Leslie model with $\gamma = 0$. The aim is to find a relation between optimal harvesting and beginning class of harvesting " k " for those fish species, k is a discrete parameter measuring the width of the meshes. Large k corresponding to large width. Also we will compare between selective harvesting and harvesting with nets. We showed in section (2) that partial harvesting $h_{j(t+1)}$ can't be determined analytically, so we will determine it numerically.

From Beverton-Holt [8], the maximum age of haddock is $m = 20$ years and it has a constant natural mortality of about 0.2 per year. In order to study the influence of the mortality on our models, we assume that the mortality is increasing as an example in the form

$$\mu(i) = \begin{cases} 0.2, & i \leq \dfrac{m}{2} \\ \dfrac{0.4 \times i}{m}, & i > \dfrac{m}{2} \end{cases}$$

The weight of haddock is determined from the formula

$$w(t) = 1.34 \times (1 - e^{-0.26(t+0.75)})^3 \qquad \text{kg}$$

and the fecundity which proportional to the weight, is determined from

$$f(t) = w(t) \cdot 10^5$$

Ricker stock-recruitment parameters for haddock are

$$\beta = \frac{1}{61.4}, \qquad \alpha = 1.53 \times 10^{-8}$$

4.1. **Influence of mortality.**

4.1.1. *Selective harvesting.* The influence of parameters in selective and net harvesting are determined from equations (5)and (7) respectively

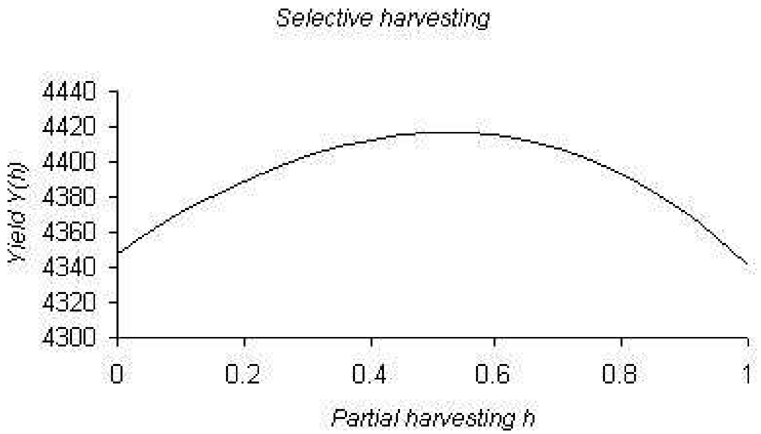


FIGURE 1. Maximum harvesting $h_{j(t+1)} = 0.7$ when constant mortality=0.2, $j(t+1) = 7$ years and MSY= 4407 gm

4.1.2. *Net harvesting.* mortality: MSY= 4562.81 gm, variable mortality: MSY= 4562.81 gm

4.2. **Influence of Von Bertlanffy parameter K.**

4.2.1. *Selective harvesting.*

4.2.2. *Net harvesting.* when $K = 0.2$, optimal mesh width $k = 9$ and MSY = 1874 gm,

when $K = 0.3$, optimal mesh width $k = 6$ and MSY = 4447 gm,

when $K = 0.4$, optimal mesh width $k = 5$ and MSY = 6809 gm

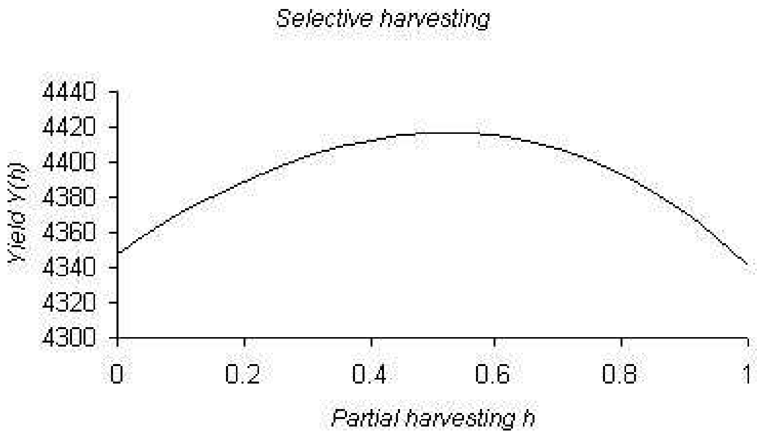


FIGURE 2. Maximum harvesting $h_{j(t+1)} = 0.7$, variable mortality, $j(t+1) = 7$ years and MSY= 4407 gm

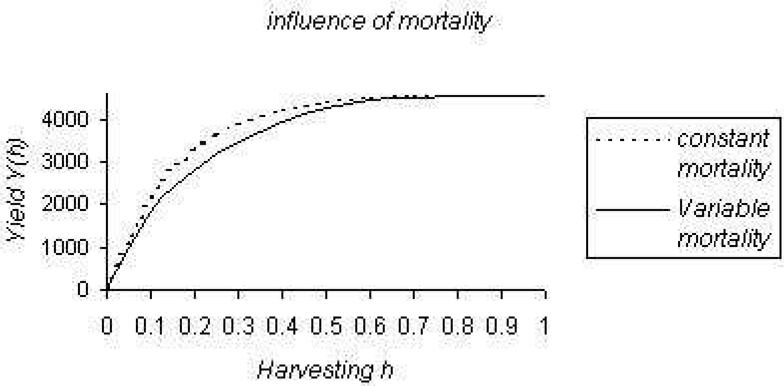
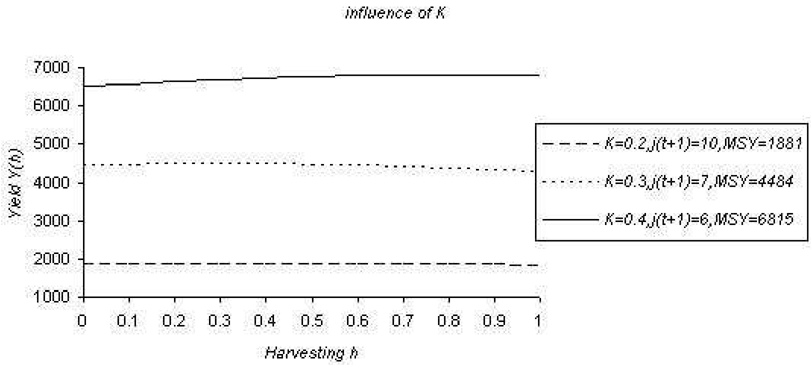


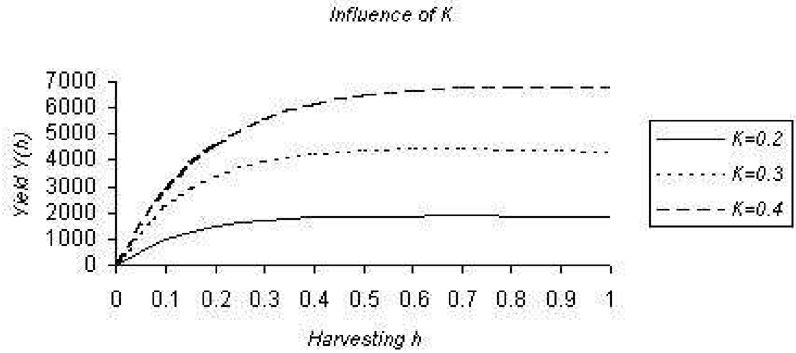
FIGURE 3. Optimal mesh width $k = 6$, $h_{\max} = 0.9$, constant mortality: MSY= 4562.81 gm, variable mortality: MSY= 4562.81 gm

5. **Conclusion and results**

The purpose of this paper is to compare between two cases of harvesting from a discrete nonlinear age-structured fish population with Ricker stock recruitment function (cf. the classical paper by Ricker [10] and Ricker [9]) and the influence



m



m

of paremetrs in our model. In our models, there are many parameters acting on the results such as mortality μ and von-Bertalanffy parameter K. The influence of mortality is that, in general, increasing mortality means the numbers of individual at high age classes are decreasing and MSY in this case is decreasing because the probability of dying is increasing. In our particular example, Figures 1 and 2 indicate that the influence of mortality in selective harvesting is ignored because $j(t+1)$ is less than $m/2$ and mortality is constant in this case. In Figure 3, the influence of mortality in net harvesting is that the values of MSY in variable mortality is slightly smaller than in constant mortality because when mortality is increasing, the survival probability L_i is decreasing i.e., the number of fish which arrive to fishable age is also decreasing. So the influence of mortality parameter on our models is small and we can use a constant mortality as a simplification of models. The influence of the Von-Bertalanffy growth parameter$K$ is that when K increases, the growth function arises to its asymptotic value more quickly, this means that weight is increasing more quickly too and since the heavier fish are more catchable, so the values of $j(t+1)$and optimal mesh width k are decreasing and the value of MSY is increasing as shown in Figures 4 and 5.

The main conclusion is that the MSY in net harvesting is slightly smaller than the MSY in selective harvesting. This is because in selective harvesting, the MSY is over a cube with $m-1$ dimension (the values of $h_i, h_1 = 0$) but in harvesting with nets, the MSY is over a subsets of that cube. These subsets are lines of diagonal of that cube.

Acknowledgments. I would to thank Prof. Dr. Horst Behncke and Dr. Ali Ben Amor for their suggestions.

Research partially supported by the Hungarian National Foundation for Scientific Research.

References

[1] Cushing, J.M., *An introduction to structured population dynamics.* Siam Philadelphia 1998.

[2] Levin, S.A., Goodyear, P.H., *Analysis of an age-structured fishery model*, J. Math. Biol. 9, 245-274 (1980).

[3] Getz, W.M., Haight, R.G., *Population Harvesting*, Princeton Press, 1989.

[4] Leslie, P.H., *On the use of matrices in certain population mathematics*, Biometrika 33, 183-212 (1945).

[5] Cushing, J.M., and Yicang, Z., *The net reproductive value and stability in matrix population models*, Natural Resource Modeling,8(4), 297-333, (1994).

[6] Reed, W.J., *Optimum age-specific harvesting in nonlinear population model*, Biometrics. 36, 579-593 (1980)

[7] Mostafa, K.S. Mohamed, *Harvesting of an age-structured fish population*, Dr. Scient. thesis, University of Osnabrueck, Osnabrueck, Germany, (2004).

[8] Beverton, J.H.R., and Holt, J. S., *On the dynamics of Exploited Fish populations*, Fish and Fisheries Series 11, 1993.

[9] Ricker ,W.E., *Computation and Interpretation of Biological Statistics of fish populations*, Fisheries Research Board of Canada, Bulletin 191. (1975)

[10] Ricker , W.E., *Stock and Recruitment*, J. Fish. Res. Board of Canada, 11, 559-623 (1954).

Department of Mathematics and Computer science
University of Osnabrueck, Germany
E-mail address: mostafaks2002@yahoo.com

Institute of Nuclear Techniques
Budapest University of Technology
Budapest, Hungary

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

NOTE ON A TWO-POINT BOUNDARY VALUE PROBLEM UNDER NONRESONANCE CONDITION

DEZIDERIU MUZSI

Abstract. The nonresonance method of Mawhin and Ward Jr. is used to discuss the existence of solutions to two point boundary value problems for second order functional-differential equations.

1. **Introduction**

In this paper we present existence results for the two point boundary value problem

$$\begin{cases} -u''(t) = cu(t) + F(u)(t), \quad t \in (0,1) \\ u(0) = u(1) = 0 \end{cases} \tag{1}$$

under the assumption that the constant c is not an eigenvalue of the operator $-u''$ (nonresonance condition) and the growth of $F(u)$ on u is at most linear. More exactly, we will apply the fixed point theorems of Banach, Schauder and the Leray-Schauder principle in order to obtain weak solutions to (1), that is a function $u \in H_0^1(0,1)$ with

$$\int_0^1 u'(t)v'(t)dt = \int_0^1 (cu(t) + F(u)(t))v(t)dt, \quad \text{for all } v \in H_0^1(0,1).$$

The method we use was introduced by J. Mawhin and J. Ward Jr. in [2]. See also [3], [4], [5] for its applications to differential equations. This paper was inspired by [7] and [6], chapter 6. The novelty in this note is that the term $F(u)$ is given by a general operator F from $L^2(0,1)$ to $L^2(0,1)$. In particular, F can be the usual superposition operator $f(t, u(t))$ as in[6] and [7], or a delay operator $f(t, u(t-\tau))$.

Received by the editors: 10.05.2005.

2000 *Mathematics Subject Classification.* 34L30.

Key words and phrases. Boundary value problems, functional-differential equations, nonresonance.

1.1. **Fixed point formulation of problem (1).** We consider $F : L^2(0,1) \to L^2(0,1)$ to be a continuous operator and we define

$$L : H^2(0,1) \cap H_0^1(0,1) \to L^2(0,1), \quad Lu = -u'' - cu$$

Let $L^{-1} : L^2(0,1) \to H^2(0,1) \subset L^2(0,1)$ be the inverse of L. If we look a priori for a solution u of the form $u = L^{-1}v$ with $v \in L^2(0,1)$, then we have to solve the fixed point problem on $L^2(0,1)$:

$$(F \circ L^{-1})(v) = v \tag{2}$$

Throughout this paper we denote:

$$\langle u,v\rangle_{L^2} = \int_0^1 uvdx, \;\; \|u\|_{L^2} = \left(\int_0^1 u^2dx\right)^{1/2}, \;\; \|v\|_{H_0^1} = \left(\int_0^1 (v')^2\, dx\right)^{1/2}$$

1.2. **An auxiliarly result.** We present first an auxiliarly result given in [7]. Let $(\lambda_k)_{k\geq 1}$ be the sequence of all eigenvalues of $-u''$ with respect to the boundary condition $u(0) = u(1) = 0$, and let $(\phi_k)_{k\geq 1}$ be the corresponding eigenfunctions, with $\|\phi_k\|_{L^2} = 1$.

Lemma 1. *Let c be any constant with $c \neq \lambda_k$ for $k = 1, 2, \ldots$. For each $v \in L^2(0,1)$, there exists a unique weak solution $u \in H_0^1(0,1)$ to the problem*

$$\begin{cases} -u'' - cu = v, & \text{on } (0,1) \\ u(0) = u(1) = 0 \end{cases}$$

denoted by $L^{-1}v$, and the following eigenfunction expansion holds

$$L^{-1}v = \sum_{k=1}^{\infty} (\lambda_k - c)^{-1} \langle v, \phi_k\rangle_{L^2} \tag{3}$$

where the series converges in $H_0^1(0,1)$. In addition,

$$\left\|L^{-1}v\right\|_{L^2} \leq \mu_c \|v\|_{L^2} \;\; \text{for all } v \in L^2(0,1) \tag{4}$$

where

$$\mu_c = \max\left\{|\lambda_k - c|^{-1}; k = 1, 2, \ldots\right\}.$$

2. Existence results

We first show how the fixed point theorems of Banach and Schauder can be used to obtain existence results for problem (1).

Theorem 2. *Suppose*

$$\lambda_j < c < \lambda_{j+1} \text{ for some } j \in \mathbb{N}, j \geq 1, \text{ or } 0 \leq c < \lambda_1 \tag{5}$$

Also assume that

$$\|F(v_1) - F(v_2)\|_{L^2} \leq a \|v_1 - v_2\|_{L^2} \tag{6}$$

for all $v_1, v_2 \in L^2(0,1)$, where a is a nonnegative constant such that

$$a\mu_c < 1. \tag{7}$$

Then (1) has a unique solution $u \in H_0^1(0,1) \cap H^2(0,1)$. In addition

$$(F \circ L^{-1})^n(v_0) \to v \text{ in } L^2(0,1) \text{ as } n \to \infty$$

for any $v_0 \in L^2(0,1)$, where $v = Lu$.

Proof. We will show that $F \circ L^{-1}$ is a contraction on $L^2(0,1)$. For this, let $v_1, v_2 \in L^2(0,1)$. Using (6) and (4) we have

$$\left\|F(L^{-1}(v_1)) - F(L^{-1}(v_2))\right\|_{L^2} \leq a \left\|L^{-1}(v_1 - v_2)\right\|_{L^2} \leq a\mu_c \|v_1 - v_2\|_{L^2}.$$

This together with (7) shows that $F \circ L^{-1}$ is a contraction. The conclusion follows from Banach's fixed point theorem.□

Theorem 3. *Suppose that (5) holds, F is continuous and satisfies the growth condition*

$$\|F(u)\|_{L^2} \leq a \|u\|_{L^2} + h \tag{8}$$

for all $u \in L^2(0,1)$, where $h \in \mathbb{R}_+$ and $a \in \mathbb{R}_+$ is as in (7). Then (1) has at least one solution $u \in H^2(0,1) \cap H_0^1(0,1)$.

Proof. We have $F \circ L^{-1} = F \circ J \circ L_0^{-1}$ where

$$\begin{cases} L_0^{-1} : L^2(0,1) \to H^2(0,1), \ L_0^{-1}u = L^{-1}u \text{ and} \\ \\ J : H_0^1(0,1) \to L^2(0,1), \ Ju = u. \end{cases}$$

Recall that F is continuous and by (8) is bounded. Next, by Rellich-Kondrachov theorem (see [1]), the imbedding of $H_0^1(0,1)$ into $L^2(0,1)$ is completely continuous. Thus, $F \circ L^{-1}$ is a completely continuous operator. On the other hand, from (8) and (4) we have

$$\left\| F(L^{-1}(v)) \right\|_{L^2} \leq a \left\| L^{-1}(v) \right\|_{L^2} + h \leq a\mu_c \left\| v \right\|_{L^2} + h.$$

Now (7) guarantees that $F \circ L^{-1}$ is a self-map of a sufficiently large closed ball of $L^2(0,1)$. Thus we may apply Schauder's fixed point theorem.□

Better results can be obtained if we use the Leray-Schauder principle (see [6]).

Theorem 4. *Suppose that F is continuous and has the decomposition*

$$F(u) = G(u)u + F_0(u) + F_1(u)$$

Also assume that

$$\left\| F_0(u) \right\|_{L^2} \leq a \left\| u \right\|_{L^2} + h_0 \tag{9}$$

$$\left\| F_1(u) \right\|_{L^2} \leq b \left\| u \right\|_{L^2} + h_1 \tag{10}$$

$$\langle u, F_1(u) \rangle_{L^2} \leq 0 \tag{11}$$

$$-M \leq G(u)(t) + c \leq \beta < \lambda_1 \tag{12}$$

for all $u \in L^2(0,1)$, where $a, b, h_0, h_1, M, \beta \in \mathbb{R}_+$. In addition assume that $0 \leq c \leq \beta$ and

$$a/\lambda_1 < 1 - \beta/\lambda_1. \tag{13}$$

Then (1) has at least one solution $u \in H^2(0,1) \cap H_0^1(0,1)$.

Proof. We look for a fixed point $v \in L^2(0,1)$ of $F \circ L^{-1}$. As above, $F \circ L^{-1}$ is a completely continuous operator. We will show that the set of all solutions to

$$v = \lambda (F \circ L^{-1})(v) \tag{14}$$

when $\lambda \in [0,1]$ is bounded in $L^2(0,1)$. Let $v \in L^2(0,1)$ be any solution of (14). Let $u = L^{-1}v$. It is clear that u solves

$$\begin{cases} -u''(t) - cu(t) = \lambda F(u)(t), & t \in (0,1) \\ u(0) = u(1) = 0. \end{cases} \tag{15}$$

Since u is a weak solution of (15), we have

$$\|u\|_{H_0^1}^2 = \langle cu + \lambda F(u), u \rangle_{L^2}.$$

It is easy to check that

$$\langle cu + \lambda G(u)u, u \rangle_{L^2} \leq \beta \|u\|_2^2. \tag{16}$$

We define

$$R(u) := \|u\|_{H_0^1}^2 - \beta \|u\|_2^2 \tag{17}$$

and using (11), (16) and $c \leq \beta$, we obtain

$$R(u) \leq \|u\|_{H_0^1}^2 - \langle cu + \lambda G(u)u, u \rangle_{L^2} \leq |\langle F_0(u), u \rangle_{L^2}|.$$

On the other hand, if we denote $c_k = \langle u, \phi_k \rangle_{L^2} = \langle u, \phi_k \rangle_{H_0^1} / \lambda_k$, we see that

$$\begin{aligned} R(u) &= \sum_{k=1}^{\infty} (\lambda_k - \beta) c_k^2 \geq \sum_{k=1}^{\infty} \lambda_k (1 - \beta/\lambda_1) c_k^2 \\ &\geq (1 - \beta/\lambda_1) \|u\|_{H_0^1}^2. \end{aligned} \tag{18}$$

Recall that

$$\lambda_1 = \inf \left\{ \|u\|_{H_0^1}^2 / \|u\|_2^2 \,; u \in H_0^1(0,1) \setminus \{0\} \right\}$$

and using (18), (17), (9) and Holder's inequality we obtain

$$\begin{aligned} (1 - \beta/\lambda_1) \|u\|_{H_0^1}^2 &\leq |\langle F_0(u), u \rangle_{L^2}| \leq \|F_0(u)\|_{L^2} \|u\|_{L^2} \leq a \|u\|_{L^2}^2 + h_0 \|u\|_{L^2} \\ &\leq \frac{a}{\lambda_1} \|u\|_{H_0^1}^2 + C \|u\|_{H_0^1} \end{aligned}$$

for some constant $C > 0$. Thus (13) guarantees that there is a constant $r > 0$ independent of λ with $\|u\|_{H_0^1} \leq r$. Finally, a bound for $\|v\|_{L^2}$ can be immediately derived from $u = L^{-1}v$. The colnclusion now follows from the Leray-Schauder principle. □

3. **Particular cases**

Particular case 1. Let $F(u)$ be the usual superposition operator, $F(u)(t) = f(t, u(t))$. Then for the problem

$$\begin{cases} -u''(t) = cu(t) + f(t, u(t)),\ t \in (0,1) \\ u(0) = u(1) = 0 \end{cases} \tag{19}$$

we have the following existence result given in [7]:

Theorem 5. *Assume that $f : (0,1) \times \mathbb{R} \to \mathbb{R}$ is a Caratheodory function, $f(\cdot, 0) \in L^2(0,1)$ and that f satisfies the Lipschitz condition*

$$|f(t, v_1) - f(t, v_2)| \leq a\,|v_1 - v_2| \tag{20}$$

for every $v_1, v_2 \in \mathbb{R}$, $t \in (0,1)$ and some $a \geq 0$. Also assume that the conditions (5) and (7) from Thorem 2 are satisfied.

Then (19) has a unique solution $u \in H_0^1(0,1) \cap H^2(0,1)$.

Proof. Using (20) we deduce

$$|f(t,u)| \leq |f(t,u) - f(t,0)| + |f(t,0)| \leq a\,|u| + |f(t,0)|$$

for every $u \in \mathbb{R}$ and $t \in (0,1)$. Moreover, f beeing a Caratheodory function, we have that the Nemitskii operator

$$u \longmapsto f(\cdot, u(\cdot))$$

is well defined, bounded and continuous from $L^2(0,1)$ into $L^2(0,1)$. Using again (20) we obtain

$$\int_0^1 |f(t, v_1(t)) - f(t, v_2(t))|^2\, dt \leq a^2 \int_0^1 |v_1(t) - v_2(t)|^2\, dt$$

so

$$\|F(v_1) - F(v_2)\|_{L^2} \leq a\,\|v_1 - v_2\|_{L^2}.$$

The conclusion follows now by applying Theorem 2. □

Particular case 2. Let $0 < \tau < 1$ and let F be defined by

$$F(u)(t) = \begin{cases} f(t, u(t-\tau)), & \tau < t < 1 \\ g(t), & 0 < t < \tau. \end{cases} \tag{21}$$

Theorem 6. *Assume that $f : (\tau, 1) \times \mathbb{R} \to \mathbb{R}$ is a Caratheodory function, $f(\cdot, 0) \in L^2(\tau, 1)$ and that f satisfies the Lipschitz condition*

$$|f(t, v_1) - f(t, v_2)| \leq a\,|v_1 - v_2| \tag{22}$$

for all $v_1, v_2 \in \mathbb{R}$, $t \in (\tau, 1)$ and some $a > 0$. Also assume that $g \in L^2(0, \tau)$ and that the conditions (5) and (7) from Theorem 2 are satisfied.

Then (1) with F defined by (21) has a unique solution $u \in H_0^1(0,1) \cap H^2(0,1)$.

Proof. Let $u \in L^2(0,1)$. Then $u(\cdot - \tau) \in L^2(\tau, 1)$. Hence, $f(\cdot, u(\cdot - \tau)) \in L^2(\tau, 1)$. Moreover, since $g \in L^2(0, \tau)$ we have $F(u) \in L^2(0,1)$ is well defined as operator from $L^2(0,1)$ into $L^2(0,1)$.

Let (u_k) be a sequence wich converges to u in $L^2(0,1)$. Let $v_k(t) = u_k(t-\tau)$ and $v(t) = u(t-\tau)$. Then

$$\begin{aligned} \int_\tau^1 (v_k(t) - v(t))^2 dt &= \int_\tau^1 (u_k(t-\tau) - u(t-\tau))^2 dt \\ &= \int_0^{1-\tau} (u_k(t) - u(t))^2 dt \longrightarrow 0, \text{ as } k \to \infty, \end{aligned}$$

so $v_k \to v$ in $L^2(\tau, 1)$ as $k \to \infty$. Consequently, $f(\cdot, v_k(\cdot)) \longrightarrow f(\cdot, v(\cdot))$ in $L^2(\tau, 1)$ and by the definition of F it follows that $F(u_k) \to F(u)$ in $L^2(0,1)$. Using (22) we deduce

$$\begin{aligned} \int_0^1 (F(v_1)(t) - F(v_2)(t))^2 dt &\leq \int_\tau^1 (f(t, v_1(t-\tau)) - f(t, v_2(t-\tau)))^2 dt \\ &\leq a^2 \int_\tau^1 (v_1(t-\tau) - v_2(t-\tau))^2 dt \\ &\leq a^2 \int_0^{1-\tau} (v_1(s) - v_2(s))^2 ds \\ &\leq a^2 \int_0^1 (v_1(s) - v_2(s))^2 ds \end{aligned}$$

and finally

$$\|F(v_1) - F(v_2)\|_{L^2} \leq a \|v_1 - v_2\|_{L^2} .$$

The conclusion follows now by applying Theorem 2. □

References

[1] Brezis, H., *Analyse Foctionelle. Theorie et applications*,Dunod,Paris,1983

[2] Mawhin, J., Ward Jr., J., *Nonresonance and existence for nonlinear elliptic boundary value problems*, Nonlinear Anal. **6** (1981), 677-684.

[3] Ntouyas, S. K., Sficas, Y. G., Tsamatos, P. Ch., *Boundary Value Problems for Functional Differential Equations,* J. Math. Anal. Appl. **199** (1996), 213-230.

[4] O'Regan, D., *Nonresonant nonlinear singular problems in the limit circle case,* J. Math. Anal. Applic., **197**(1996), 708-725.

[5] O'Regan, D., *Caratheodory theory of nonresonant second order boundary value problems*, Differential Equations and Dynamical Systems, **4** (1996), 57-77.

[6] O'Regan,. D., Precup, R., *Theorems of Leray-Schauder Type and Applications,* Gordon and Breach Science Publishers, Amsterdam, 2001.

[7] Precup, R., *Existence results for nonlinear boundary value problems under nonresonance conditions,* in: Qualitative Problems for Differential Equations and Control Theory, C. Corduneanu (ed.), World Scientific, Singapore, 1995, 263-273.

Babeş-Bolyai University, Department of Applied Mathematics
Str. M. Kogălniceanu 1, 400084 Cluj-Napoca, Romania
E-mail address: dmuzsi@math.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume L, Number 2, June 2005

ON A FIRST-ORDER NONLINEAR DIFFERENTIAL SUBORDINATION II

GH. OROS AND GEORGIA IRINA OROS

Abstract. We find conditions on the complex-valued functions A, B, C, D defined in the unit disc U and the positive constants M and N such that

$$|A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)| < M$$

implies $|p(z)| < N$, where p is analytic in U, with $p(0) = 0$.

1. **Introduction and preliminaries**

In [1] chapter IV, the authors have analyzed a first-order linear differential subordination

$$B(z)zp'(z) + C(z)p(z) + D(z) \prec h(z), \tag{1}$$

where B, C, D and h are complex-valued functions.

A more general version of (1) is given by

$$B(z)zp'(z) + C(z)p(z) + D(z) \in \Omega, \tag{2}$$

where $\Omega \subseteq \mathbb{C}$.

In this paper we shall extend this problem by considering a first-order nonlinear differential subordination given by

$$A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z) \prec h(z). \tag{3}$$

A more general version of (3) is given by:

$$A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z) \in \Omega, \tag{4}$$

Received by the editors: 06.05.2005.
2000 *Mathematics Subject Classification.* 30C80.
Key words and phrases. Differential subordination, dominant.

where $\Omega \subseteq \mathbb{C}$.

The general problem is to find conditions on the complex-valued functions A, B, C, D and h such that the differential subordination given by (3) or (4) will have dominants and even best dominant.

We let U denote the class of holomorphic functions in the unit disc

$$U = \{z \in \mathbb{C};\ |z| < 1\}, \quad \overline{U} = \{z \in \mathbb{C};\ |z| \le 1\}.$$

For $a \in \mathbb{C}$ and $n \in \mathbb{N}^*$ we let

$$\mathcal{H}[a,n] = \{f \in U,\ f(z) = a + a_n z^n + a_{n+1} z^{n+1} + \dots,\ z \in U\}$$

and

$$A_n = \{f \in U,\ f(z) = z + a_{n+1} z^{n+1} + \dots,\ z \in U\}$$

and $A_1 = A$.

We let Q denote the class of functions q that are holomorphic and injective in $\overline{U} \setminus E(q)$, where

$$E(q) = \left\{\zeta \in \partial U :\ \lim_{z \to \zeta} q(z) = \infty\right\}$$

and furthermore $q'(\zeta) \neq 0$ for $\zeta \in \partial U \setminus E(q)$, where $E(q)$ is called exception set.

In order to prove the new results we shall use the following:

Lemma A. [1] (Lemma 2.2.d p.24) *Let $q \in Q$, with $q(0) = a$, and let*

$$p(z) = a + a_n z^n + a_{n+1} z^{n+1} + \dots$$

be analytic in U with $p(z) \not\equiv a$, and $n \ge 1$.

If p is not subordinate to q, then there exist points $z_0 = r_0 e^{i\theta_0} \in U$, $r_0 < 1$ and $\zeta \in \partial U \setminus E(q)$, and an $m \ge n \ge 1$ for which $p(U_{r_0}) \subset q(U)$

(i) $p(z_0) = q(\zeta)$

(ii) $z_0 p'(z_0) = m\zeta q'(\zeta)$, and

(iii) $\mathrm{Re}\, \dfrac{z_0 p''(z_0)}{p'(z_0)} + 1 \ge m \mathrm{Re}\, \left[\dfrac{\zeta q''(\zeta)}{q'(\zeta)} + 1\right]$.

In this paper we consider the first-order nonlinear differential subordination (4) in which $\Omega = \{w;\ |w| < M\}$. Given the functions A, B, C, D and the constant

M, our problem is to find a constant N such that, for $p \in \mathcal{H}[0,n]$, the differential inequality

$$|A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)| < M$$

implies

$$|p(z)| < N.$$

If $D(0) = 0$, then this result can be written in terms of the differential subordination as

$$A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z) \prec Mz$$

implies $p(z) \prec Nz$.

2. Main results

In this paper we improve the results obtained in [2].

Theorem 1. *Let $M > 0$, $N > 0$ and let n be a positive integer. Suppose that the functions $A, B, C, D : U \to \mathbb{C}$ satisfy*

$$n|A(z)| - |C(z)| \geq \frac{M + N^2|B(z)| + |D(z)|}{N}. \tag{5}$$

If $p \in \mathcal{H}[0,n]$ and

$$|A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)| < M \tag{6}$$

then

$$|p(z)| < N, \quad z \in U.$$

Proof. If we let

$$w(z) = A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z),$$

then from (6) we obtain

$$|w(z)| = |A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)|. \tag{7}$$

From (7) and (6) we have

$$|w(z)| < M, \quad z \in U. \tag{8}$$

Assume that $|p(z)| \not< N$, which is equivalent with $p(z) \not\prec Nz = q(z)$.

According to Lemma A, with $q(z) = Nz$, there exist $z_0 \in U$, $z_0 = r_0 e^{i\theta_0}$, $r_0 < 1$, $\theta_0 \in [0, 2\pi)$, $\zeta \in \partial U$, $|\zeta| = 1$ and $m \geq n$, such that $p(z_0) = N\zeta$ and $z_0 p'(z_0) = mN\zeta$.

Using these conditions in (3) we obtain for $z = z_0$

$$|w(z_0)| = |A(z_0)z_0 p'(z_0) + B(z_0)p^2(z_0) + C(z_0)p(z_0) + D(z_0)| \tag{9}$$

$$= |A(z_0)mN\zeta + B(z_0)N^2\zeta^2 + C(z_0)N\zeta + D(z_0)|$$

$$\geq |A(z_0)mN + B(z_0)N^2\zeta + C(z_0)N| - |D(z_0)|$$

$$\geq N|A(z_0)m + C(z_0)| - N^2|B(z_0)| - |D(z_0)|$$

$$\geq mn|A(z_0)| - N|C(z_0)| - N^2|B(z_0)| - |D(z_0)|$$

$$\geq [n|A(z_0)| - |C(z_0)|]N - N^2|B(z_0)| - |D(z_0)| \geq M.$$

Since (9) contradicts (8) we obtain the desired results $|p(z)| < N$. □

Instead of prescribing the constant N in Theorem 1, in some cases we can use in (5) to determine an appropriate $N = N(M, n, A, B, C, D)$ so that (6) implies $|p(z)| < N$. This can be accomplished by solving (5) for N and by taking the supremum of the resulting function over U. The condition (5) is equivalent to

$$N^2|B(z)| - N[n|A(z)| - |C(z)|] + |D(z)| + M \leq 0. \tag{10}$$

Suppose $B(z) \neq 0$, the inequality (10) holds if

$$[n|A(z)| - |C(z)|]^2 \geq 4|B(z)|[|D(z)| + M]. \tag{11}$$

The roots of the trinomial in (10) are

$$N_{1,2} = \frac{n|A(z)| - |C(z)| \pm \sqrt{[n|A(z)| - |C(z)|]^2 - 4|B(z)|[|D(z)| + M]}}{2|B(z)|}.$$

Let

$$N = \sup_{|z|<1} \frac{n|A(z)| - |C(z)| - \sqrt{[n|A(z)| - |C(z)|]^2 - 4|B(z)|[|D(z)| + M]}}{2|B(z)|}$$

$$= \sup_{|z|<1} \frac{2[|D(z)| + M]}{n|A(z)| - |C(z)| + \sqrt{[n|A(z)| - |C(z)|]^2 - 4|B(z)|[|D(z)| + M]}}.$$

If this supremum is finite, we have the following version of the Theorem 1:

Theorem 2. *Let $M > 0$, $N > 0$ and n be a positive integer. Suppose that $p \in \mathcal{H}[0, n]$ and the functions $A, B, C, D : U \to \mathbb{C}$, with $B(z) \neq 0$, satisfy:*

$$[n|A(z)| - |C(z)|]^2 \geq 4|B(z)|[|D(z)| + M].$$

$$N = \sup_{|z|<1} \frac{2[|D(z)| + M]}{n|A(z)| - |C(z)| + \sqrt{[n|A(z)| - |C(z)|]^2 - 4|B(z)|[|D(z)| + M]}}$$

then

$$|A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)| < M$$

implies

$$|p(z)| < N, \quad z \in U.$$

If $D(z) \equiv 0$, the Theorem 1 can be rewritten as the following:

Corollary 1. *Let $M > 0$, $N > 0$ and n be a positive integer. Suppose that the functions $A, B, C : U \to \mathbb{C}$ satisfy*

$$n|A(z)| - |C(z)| \geq \frac{M + N^2|C(z)|}{N}.$$

If $p \in \mathcal{H}[0, n]$ and

$$|A(z)zp'(z) + B(z)p^2(z) + C(z)p(z) + D(z)| < M$$

then

$$|p(z)| < N, \quad z \in U.$$

References

[1] Miller, S. S., Mocanu, P. T., *Differential Subordinations. Theory and Applications*, Marcel Dekker Inc., New York, Basel, 2000.

[2] Oros, Gh., Oros, Georgia Irina, *On a first-order nonlinear differential subordination I*, Analele Univ. Oradea, Fasc. Matematică, Tome IX, 5-12, 2002, 65-70.

Department of Mathematics, University of Oradea
Str. Armatei Române, No. 5, 410087 Oradea, Romania
E-mail address: georgia_oros_ro@yahoo.co.uk

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

GENERATION OF NON-UNIFORM LOW-DISCREPANCY SEQUENCES IN QUASI-MONTE CARLO INTEGRATION

NATALIA ROŞCA

Abstract. We propose an inversion type method that can be used in Quasi-Monte Carlo integration to generate low-discrepancy sequences with an arbitrary distribution function G. The method is based on the approximation of the inverse of the distribution function by linear Lagrange interpolation or cubic Hermite interpolation. We also give bounds for the G-discrepancy of the generated sequences.

1. Discrepancy and error bounds

In quasi-Monte Carlo integration one approximates $\int_{[0,1]^s} f(x)dx$ by sums of the form $\frac{1}{N}\sum_{k=1}^{N} f(x_k)$, where $f : [0,1]^s \to \mathbb{R}$ and $(x_1, \ldots, x_N)$ is a sequence of deterministic points, with $x_k = (x_k^{(1)}, \ldots, x_k^{(s)}) \in [0,1]^s$, $k = 1, \ldots, N$. A well-known measure of the distribution properties of the sequences used in quasi-Monte Carlo integration is the discrepancy.

Definition 1 (discrepancy). *Let $P = (x_1, \ldots, x_N)$ be a sequence of points in $[0,1]^s$. The discrepancy of sequence P is defined as*

$$D_N(P) = \sup_{J \subseteq [0,1]^s} \left| \frac{1}{N} A_N(J,P) - \lambda(J) \right|,$$

where A_N counts the number of elements of sequence P, falling into the interval J, i.e.,

Received by the editors: 04.04.2005.

2000 *Mathematics Subject Classification.* 65C05, 65C10, 11K38.

Key words and phrases. Quasi-Monte Carlo integration, low-discrepancy sequences, inversion method.

$$A_N(J,P) = \sum_{k=1}^{N} 1_J(x_k).$$

1_J is the characteristic function of J and λ is the s-dimensional Lebesgue measure. The sequence P is called uniformly distributed if $D_N(P) \to 0$ when $N \to \infty$.

For $s = 1$, we may arrange the points $x_1, \ldots, x_N$ of a given point set in nondecreasing order. The following result is due to Niederreiter [10].

Theorem 2. *If $x_0 := 0 \leq x_1 \leq x_2 \leq \ldots \leq x_N \leq 1 =: x_{N+1}$, then we have the following expression for the discrepancy of sequence $(x_1, \ldots, x_N)$*

$$D_N(x_1, \ldots, x_N) = \frac{1}{N} + \max_{1 \leq n \leq N} \left(\frac{n}{N} - x_n \right) - \min_{1 \leq n \leq N} \left(\frac{n}{N} - x_n \right) = \max_{\substack{0 \leq i \leq N \\ 1 \leq j \leq N+1}} \left| \frac{1}{N} + r_i - r_j \right| \tag{1}$$

where $r_n = \frac{n}{N} - x_n$ for $0 \leq n \leq N+1$.

The monograph by Niederreiter [10] provides a comprehensive overview on discrepancy, low-discrepancy sequences and their properties. Halton [5], Faure [3], [4], Niederreiter [10] and others constructed famous low-discrepancy sequences.

A similar concept of discrepancy can be defined for non-uniformly distributed sequences.

Definition 3 (non-uniform discrepancy). *Consider an s-dimensional distribution on $[0,1]^s$, with distribution function G. Let λ_G be the probability measure corresponding to G. Let $P = (x_1, \ldots, x_N)$ be a sequence of points in $[0,1]^s$. The G-discrepancy of sequence $P = (x_1, \ldots, x_N)$ is defined as*

$$D_{N,G}(P) = \sup_{J \subseteq [0,1]^s} \left| \frac{1}{N} A_N(J,P) - \lambda_G(J) \right|.$$

The sequence P is called G-distributed, if $D_{N,G}(P) \to 0$ when $N \to \infty$.

If f is a function with finite variation in the sense of Hardy and Krause, $V_{HK}(f) < +\infty$ (see eg. Owen [12]), then an upper bound for the error of the approximation in quasi-Monte Carlo integration is given by the non-uniform Koksma-Hlawka inequality (see Chelson [1]).

Theorem 4 (non-uniform Koksma-Hlawka inequality). *Let* $f : [0,1]^s \to \mathbb{R}$ *be of bounded variation in the sense of Hardy and Krause. Moreover, let* G *be a distribution function with continuous density on* $[0,1]^s$ *and* $(x_1, \ldots, x_N)$ *a sequence on* $[0,1]^s$*. Then, for any* $N > 0$

$$\left| \frac{1}{N} \sum_{k=1}^{N} f(x_k) - \int_{[0,1]^s} f(x) dG(x) \right| \leq V_{HK}(f) D_{N,G}(x_1, \ldots, x_N),$$

where $V_{HK}(f)$ *is the variation of* f *in the sense of Hardy and Krause.*

2. **Inversion method**

The inversion method produces a random variable with desired distribution function by making use of the inverse of the distribution function.

Consider a distribution on $[0,1]$ with continuous density function g and distribution function $G(x) = \int_0^x g(t)dt$. Assume that there exists the inverse function G^{-1}. The inversion method is based on the following principle.

Theorem 5. *Let* U *be a random variable uniformly distributed on* $[0,1]$*. Then the distribution function of the random variable* $G^{-1}(U)$ *is* G*.*

Proof. We denote by $F_{G^{-1}(U)}$ the distribution function of $G^{-1}(U)$. We have

$$F_{G^{-1}(U)}(x) = P(G^{-1}(U) < x) = P(U < G(x)) = G(x).$$

Thus, the distribution function of the random variable $G^{-1}(U)$ is G. □

Such a transformation preserves the discrepancy in one dimension, as showed in the following theorem (see Okten [11]).

Theorem 6. *Let* $P = (x_1, \ldots, x_N)$ *be a sequence in* $[0,1]$ *and* G *a distribution function on* $[0,1]$*.*

Construct the sequence $(y_1, \ldots, y_N) = (G^{-1}(x_1), \ldots, G^{-1}(x_N))$*. Then the G-discrepancy of the constructed sequence is given by*

$$D_{N,G}(y_1, \ldots, y_N) = D_N(x_1, \ldots, x_N).$$

In other words,

$$D_{N,G}(G^{-1}(P)) = D_N(P).$$

As a consequence, for generating low-discrepancy sequences with an arbitrary distribution function G, we can transform uniformly distributed low-discrepancy sequences using the inverse function G^{-1}. In most cases, however, the inverse G^{-1} cannot be given analytically. In such cases, we may use the inversion technique with an approximation of the inverse function G^{-1}.

3. Existing inversion type methods

The inversion type transformations presented in this paper are designed for the one-dimensional case. They all propose different modalities of approximating the inverse G^{-1}.

3.1. Hlawka-Mück method. Hlawka [7] defines a transformation and bounds the G-discrepancy of the transformed sequence as follows.

Theorem 7. *Consider a continuous type distribution on* $[0,1]$*, with density* g *and distribution function* G*. Assume that the distribution function* G *is invertible and* $M = \sup_{x\in[0,1]} g(x) < \infty$*. Furthermore, let* $(x_1, x_2, \ldots, x_N)$ *be a sequence in* $[0,1]$*. Generate the point set* $(y_1, y_2, \ldots, y_N)$ *with*

$$y_k = \frac{1}{N}\sum_{r=1}^{N}\left[1 + x_k - G(x_r)\right] = \frac{1}{N}\sum_{r=1}^{N} 1_{[0,x_k]}(G(x_r)), \tag{2}$$

where $[a]$ *denotes the integer part of* a*. Then the generated sequence has a G-discrepancy of*

$$D_{N,G}(y_1, \ldots, y_N) \le (2 + 6M) D_N(x_1, \ldots, x_N). \tag{3}$$

This method is known in the literature as the Hlawka-Müch method and it is a generalization of an earlier version proposed by Hlawka and Mück in [8], [9]. The main disadvantage of the Hlawka-Müch method is that all the points of the sequence

$(y_1, y_2, \ldots, y_N)$ are of the form $i/N, (i = 0, \ldots, N)$. This implies that, when adding some points, all the other points have to be regenerated.

3.2. **Method proposed by Hartinger and Kainhofer.** They propose (see [6]) an inversion type transformation that is also shown to generate G-distributed low-discrepancy sequences.

Theorem 8. *(see [6]) Let $P = (x_1, x_2, \ldots, x_N)$ be a sequence in $[0, 1]$. Consider a continuous type distribution on $[0, 1]$, with density g and distribution function G. Assume that the distribution function G is invertible and $M = \sup_{x\in[0,1]} g(x) < \infty$. Define for $k = 1, \ldots, N$*

$$x_k^- = \max_{\mathcal{A}=\{x_i\in P | G(x_i)<x_k\}} x_i$$

$$x_k^+ = \min_{\mathcal{B}=\{x_i\in P | x_k\leq G(x_i)\}} x_i.$$

Set $x_k^- = 0$ if $\mathcal{A} = \emptyset$ and $x_k^+ = 1$ if $\mathcal{B} = \emptyset$.
Then the G-discrepancy of any transformed sequence $(y_1, y_2, \ldots, y_N)$ with the property that $y_k \in (x_k^-, x_k^+]$ for all $1 \leq k \leq N$ is bounded by

$$D_{N,G}(y_1, \ldots, y_N) \leq (1 + 2M) D_N(x_1, \ldots, x_N).$$

In the method proposed by Hartinger and Kainhofer, any value in the interval $(x_k^-, x_k^+]$ can be considered as $G^{-1}(x_k)$. They do not analyze the possibility of approximating G^{-1} using interpolation methods. Thus, in their method, the kind of interpolation is not relevant for the discrepancy bound and the smoothness of the interpolation is not taken into account.

4. Inversion method using linear Lagrange interpolation

Next, we propose an inversion type method that can be used to generate one-dimensional low-discrepancy sequences with an arbitrary distribution function G. The method is based on the approximation of the inverse of the distribution function by linear Lagrange interpolation. We also give bounds for the G-discrepancy of the generated sequence. Our method is based on the following idea:

Let $0 \leq x_1 < \ldots < x_N \leq 1$ be a one-dimensional sequence. Let G be an invertible distribution function. We define $x_0 = 0$ and $x_{N+1} = 1$. To approximate $G^{-1}(x_k)$, we proceed as follows. First, we determine the interval $(x_i, x_{i+1}]$, with $i \in \{0, 1, \ldots, N\}$ such that $G^{-1}(x_k) \in (x_i, x_{i+1}]$, based on the following equivalence:

$$G^{-1}(x_k) \in (x_i, x_{i+1}], \quad \text{iff} \quad G(x_i) < x_k \leq G(x_{i+1}).$$

Then, we approximate $G^{-1}(x_k)$ with a value y_k in the interval $(x_i, x_{i+1}]$, which is calculated using linear Lagrange interpolation of G^{-1}.

Before describing our method, we recall a lemma from Niederreiter [10].

Lemma 9. *Let $P_1 = (u_1, \ldots, u_N)$ and $P_2 = (v_1, \ldots, v_N)$ be two sequences in $[0,1]$. If, for all $1 \leq n \leq N$, the following condition takes place*

$$|u_n - v_n| \leq \varepsilon$$

then

$$|D_N(u_1, \ldots, u_N) - D_N(v_1, \ldots, v_N)| \leq 2\varepsilon. \tag{4}$$

To prove the main theorems of this paper, we formulate and prove the following results.

Proposition 10. *Let $(x_1, x_2, \ldots, x_N)$ be a one-dimensional sequence in $[0,1]$, with $x_0 := 0 \leq x_1 < x_2 < \ldots < x_N \leq 1 =: x_{N+1}$. The following inequality takes place*

$$|x_n - x_{n+1}| \leq D_N(x_1, \ldots, x_N), \qquad n = 0, \ldots, N. \tag{5}$$

Proof. We note that

$$|x_n - x_{n+1}| = \left| \frac{1}{N} + \left(\frac{n}{N} - x_n \right) - \left(\frac{n+1}{N} - x_{n+1} \right) \right| = \left| \frac{1}{N} + r_n - r_{n+1} \right|$$

where $r_n = \frac{n}{N} - x_n$, $n = 0, \ldots, N+1$.

It follows that

$$|x_n - x_{n+1}| \leq \max_{\substack{0 \leq i \leq N \\ 1 \leq j \leq N+1}} \left| \frac{1}{N} + r_i - r_j \right| = D_N(x_1, \ldots, x_N).$$

In the last equality, we used the result from Theorem 2. □

Lemma 11. *Consider a continuous type distribution on* $[0,1]$*, with density* g *and distribution function* G*. Assume that the distribution function* G *is invertible and* $g(t) \neq 0, \quad \forall t \in [0,1]$*. If* $G \in C^4([0,1])$ *then* $G^{-1} \in C^4([0,1])$ *and the derivatives of* G^{-1} *have the following expressions:*

$$\begin{aligned}
(G^{-1})' &= \frac{1}{g(G^{-1})} \\
(G^{-1})'' &= -\frac{g'(G^{-1})}{(g(G^{-1}))^3} \\
(G^{-1})^{(3)} &= -\frac{g''(G^{-1})g(G^{-1}) - 3(g'(G^{-1}))^2}{(g(G^{-1}))^5} \\
(G^{-1})^{(4)} &= -\frac{g'''(G^{-1})g^2(G^{-1}) - 10g''(G^{-1})g'(G^{-1})g(G^{-1}) + 15(g'(G^{-1}))^3}{(g(G^{-1}))^7}.
\end{aligned}$$

Their norms are given by:

$$||(G^{-1})''||_\infty = \left|\left| \frac{g'}{g^3} \right|\right|_\infty$$

$$||(G^{-1})^{(4)}||_\infty = \left|\left| \frac{g'''g^2 - 10g''g'g + 15g'^3}{g^7} \right|\right|_\infty.$$

Proof. The proof is immediately. □

Theorem 12 (Lagrange interpolated inversion method). *Let* $0 \leq x_1 < \ldots < x_N \leq 1$ *be a one-dimensional sequence. We consider a continuous type distribution on* $[0,1]$*, with density* g *and distribution function* G*. We assume that the distribution function* G *is invertible,* $\sup_{t\in[0,1]} g(t) \leq M < \infty$ *and* $g(t) \neq 0, \quad \forall t \in [0,1]$*. For each point* x_k*,* $k = 1, \ldots, N$*, we determine the interval* $(x_i, x_{i+1}]$ *such that*

$$G(x_i) < x_k \leq G(x_{i+1}).$$

We denote by $(x_k^-, x_k^+]$ *the determined interval* $(x_i, x_{i+1}]$ *.*

We set $x_k^- = 0$ *if* $x_k \leq G(x_1)$ *and* $x_k^+ = 1$ *if* $G(x_N) < x_k$.

We generate the sequence $(y_1, \ldots, y_N)$ *with*

$$y_k = \frac{x_k - G(x_k^+)}{G(x_k^-) - G(x_k^+)} x_k^- + \frac{x_k - G(x_k^-)}{G(x_k^+) - G(x_k^-)} x_k^+, \qquad k = 1, \ldots, N. \tag{6}$$

If $G \in C^2[0,1]$ *and* $\left\|\frac{g'}{g^3}\right\|_\infty \leq L$, *then the G-discrepancy of the sequence* $(y_1, y_2, \ldots, y_N)$ *is bounded by*

$$D_{N,G}(y_1, \ldots, y_N) \leq \big(1 + M^3 L\big) D_N(x_1, \ldots, x_N). \tag{7}$$

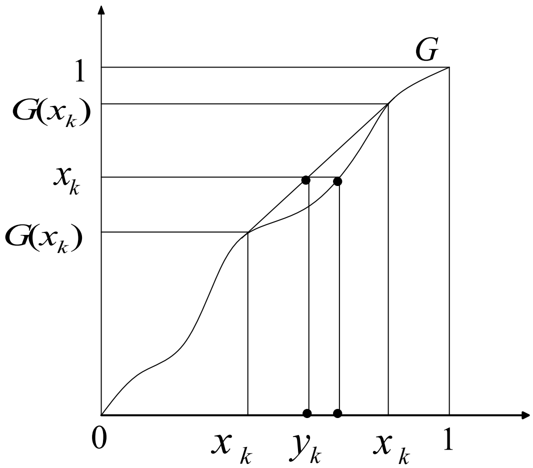


FIGURE 1. Inversion method.

Proof. First, we illustrate how we obtained the values y_k given by (6).

We consider a linear Lagrange interpolation of G^{-1}, with nodes $G\big(x_k^-\big)$ and $G\big(x_k^+\big)$. The values of G^{-1} at the nodes are $G^{-1}(G(x_k^-)) = x_k^-$ and $G^{-1}(G(x_k^+)) = x_k^+$. The interpolation formula is:

$$G^{-1} = L_1 G^{-1} + R_1 G^{-1},$$

where L_1G^{-1} is the Lagrange interpolation polynomial of degree 1 and R_1G^{-1} is the remainder. Using the expression of the Lagrange interpolation polynomial, we get

$$(L_1G^{-1})(x_k) = \frac{x_k - G(x_k^+)}{G(x_k^-) - G(x_k^+)}x_k^- + \frac{x_k - G(x_k^-)}{G(x_k^+) - G(x_k^-)}x_k^+ = y_k.$$

Next, we prove inequality (7). For this, we use the result from Theorem 6 and we obtain

$$\begin{aligned} D_{N,G}(y_1, \ldots, y_N) &= D_{N,G}\big(G^{-1}(G(y_1)), \ldots, G^{-1}(G(y_N))\big) \\ &= D_N\big(G(y_1), \ldots, G(y_N)\big). \end{aligned}$$

It follows that

$$\big|D_{N,G}(y_1, \ldots, y_N) - D_N(x_1, \ldots, x_N)\big| = \big|D_N(G(y_1), \ldots, G(y_N)) - D_N(x_1, \ldots, x_N)\big|. \tag{8}$$

Our intention is to apply Lemma 9 with $P_1 = (x_1, \ldots, x_N)$ and $P_2 = (G(y_1), \ldots, G(y_N))$. For this, we first estimate $|G(y_k) - x_k|$, for $1 \le k \le N$, as follows

$$\big|G(y_k) - x_k\big| = \big|G(y_k) - G(G^{-1}(x_k))\big| = \left|\int_{G^{-1}(x_k)}^{y_k} g(t)dt\right| \le M\big|G^{-1}(x_k) - y_k\big|. \tag{9}$$

We use the bound for the interpolation error (see [2])

$$\big|G^{-1}(x_k) - y_k\big| = \big|R_1(G^{-1})(x_k)\big| \le \frac{|u(x_k)|}{2!}\big|\big|(G^{-1})''\big|\big|_\infty \tag{10}$$

where

$$|u(x_k)| = \big|\big(x_k - G(x_k^-)\big)\big(x_k - G(x_k^+)\big)\big|. \tag{11}$$

Considering the fact that $G(x_k^-) < x_k \le G(x_k^+)$, we get

$$\big|x_k - G(x_k^-)\big| \le \big|G(x_k^+) - G(x_k^-)\big| = \left|\int_{x_k^-}^{x_k^+} g(t)dt\right| \le M\big|x_k^+ - x_k^-\big|. \tag{12}$$

Since $[x_k^-, x_k^+]$ is an interval of type $[x_i, x_{i+1}]$, we apply Proposition 10 and we get

$$|x_k^+ - x_k^-| = |x_{i+1} - x_i| \leq D_N(x_1, \ldots, x_N).$$

Relation (12) becomes

$$|x_k - G(x_k^-)| \leq MD_N(x_1, \ldots, x_N). \tag{13}$$

In a similar way, we obtain

$$\left|x_k - G(x_k^+)\right| \leq MD_N(x_1, \ldots, x_N). \tag{14}$$

From (11), (13), (14), it follows that

$$|u(x_k)| = \left|\left(x_k - G(x_k^-)\right)\left(x_k - G(x_k^+)\right)\right| \leq M^2 D_N^2(x_1, \ldots, x_N). \tag{15}$$

Using (15) and the result from Lemma 11, relation (10) becomes

$$\left|G^{-1}(x_k) - y_k\right| \leq \frac{M^2 D_N^2(x_1, \ldots, x_N)}{2}\left\|\frac{g^{'}}{g^3}\right\|_\infty \leq \frac{M^2 D_N^2(x_1, \ldots, x_N)}{2} L. \tag{16}$$

Replacing (16) into (9) we obtain

$$|G(y_k) - x_k| \leq \frac{M^3 D_N^2(x_1, \ldots, x_N)}{2} L, \qquad k = 1, \ldots, N.$$

Applying Lemma 9 with $P_1 = (x_1, \ldots, x_N)$, $P_2 = (G(y_1), \ldots, G(y_N))$, $\varepsilon = \frac{M^3 D_N^2(x_1, \ldots, x_N)}{2} L$ and using $D_N^2 \leq D_N$, as $D_N \leq 1$, we get

$$\left|D_N\left(G(y_1), \ldots, G(y_N)\right) - D_N(x_1, \ldots, x_N)\right| \leq 2\varepsilon \leq M^3 L D_N(x_1, \ldots, x_N). \tag{17}$$

From (8) and (17), we obtain

$$|D_{N,G}(y_1, \ldots, y_N) - D_N(x_1, \ldots, x_N)| \leq M^3 L D_N(x_1, \ldots, x_N).$$

The final result is

$$D_{N,G}(y_1, \ldots, y_N) \leq \big(1 + M^3 L\big) D_N(x_1, \ldots, x_N).$$

□

5. Inversion method using cubic Hermite interpolation

Next, we propose a transformation where the inverse of the distribution function G is approximated using cubic Hermite interpolation. We also give bounds for the G-discrepancy of the generated sequence.

Theorem 13 (Hermite interpolated inversion method). *On the same conditions as in Theorem 12, we consider the sequence $(y_1, \ldots, y_N)$ generated by*

$$y_k = h_{00}(x_k) x_k^- + h_{10}(x_k) x_k^+ + h_{01}(x_k) \frac{1}{g(x_k^-)} + h_{11}(x_k) \frac{1}{g(x_k^+)} \tag{18}$$

where

$$\begin{aligned} h_{00}(x_k) &= \frac{\big(x_k - G(x_k^+)\big)^2}{\big(G(x_k^-) - G(x_k^+)\big)^2} \left(1 - 2 \frac{x_k - G(x_k^-)}{G(x_k^-) - G(x_k^+)}\right) \\ h_{10}(x_k) &= \frac{\big(x_k - G(x_k^-)\big)^2}{\big(G(x_k^+) - G(x_k^-)\big)^2} \left(1 - 2 \frac{x_k - G(x_k^+)}{G(x_k^+) - G(x_k^-)}\right) \\ h_{01}(x_k) &= \frac{\big(x_k - G(x_k^-)\big)\big(x_k - G(x_k^+)\big)^2}{\big(G(x_k^-) - G(x_k^+)\big)^2} \\ h_{11}(x_k) &= \frac{\big(x_k - G(x_k^+)\big)\big(x_k - G(x_k^-)\big)^2}{\big(G(x_k^+) - G(x_k^-)\big)^2} \end{aligned}$$

for all $k = 1, \ldots, N$.
If $G \in C^4[0,1]$ and $\left\| \frac{g''' g^2 - 10 g'' g' g + 15 g'^3}{g^7} \right\|_\infty \leq L$, then the G-discrepancy of the sequence $(y_1, \ldots, y_N)$ is bounded by

$$D_{N,G}(y_1, \ldots, y_N) \leq \left(1 + \frac{M^5 L}{12}\right) D_N(x_1, \ldots, x_N). \tag{19}$$

Proof. First, we explain how we generated the values y_k given by (18).

We consider a cubic Hermite interpolation of G^{-1} with double nodes $G(x_k^-)$ and $G(x_k^+)$. The values of G^{-1} and $(G^{-1})'$ at the nodes are $G^{-1}(G(x_k^-)) = x_k^-$, $G^{-1}(G(x_k^+)) = x_k^+$, $(G^{-1})'(G(x_k^-)) = \frac{1}{g(x_k^-)}$ and $(G^{-1})'(G(x_k^+)) = \frac{1}{g(x_k^+)}$.

The Hermite interpolation formula is:

$$G^{-1} = H_3G^{-1} + R_3G^{-1},$$

where H_3G^{-1} is the Hermite interpolation polynomial of degree 3 and R_3G^{-1} is the remainder. Using the expression of the Hermite polynomial with double nodes (see [2]), it can be proved, after some calculus, that

$$(H_3G^{-1})(x_k) = y_k.$$

Next, we follow the same steps as in Theorem 12. We point out only the differences. The bound for the interpolation error (see [2]) is given by

$$\left|R_3(G^{-1})(x_k)\right| = \left|G^{-1}(x_k) - y_k\right| \leq \frac{|u(x_k)|}{4!}\left|\left|(G^{-1})^{(4)}\right|\right|_\infty \tag{20}$$

where

$$|u(x_k)| = \left|\left(x_k - G\left(x_k^-\right)\right)^2\left(x_k - G(x_k^+)\right)^2\right| \leq M^4D_N^4(x_1, \ldots, x_N).$$

We use the result from Lemma 11

$$\left|\left|(G^{-1})^{(4)}\right|\right|_\infty = \left|\left|\frac{g'''g^2 - 10g''g'g + 15g'^3}{g^7}\right|\right|_\infty \leq L.$$

Relation (20) becomes

$$|G^{-1}(x_k) - y_k| \leq \frac{M^4D_N^4(x_1, \ldots, x_N)}{4!}L.$$

Similar to Theorem 12, we obtain

$$|G(y_k) - x_k| \leq \frac{M^5L}{24}D_N^4(x_1, \ldots, x_N).$$

Applying Lemma 9 with $\varepsilon = \frac{M^5 L}{24} D_N^4(x_1, \ldots, x_N)$ and using $D_N^4 \leq D_N$, as $D_N \leq 1$, we get

$$|D_N(G(y_1), \ldots, G(y_N)) - D_N(x_1, \ldots, x_N)| \leq 2\varepsilon \leq \frac{M^5 L}{12} D_N(x_1, \ldots, x_N). \tag{21}$$

Similar to Theorem 12, this implies that

$$|D_{N,G}(y_1, \ldots, y_N) - D_N(x_1, \ldots, x_N)| \leq \frac{M^5 L}{12} D_N(x_1, \ldots, x_N). \tag{22}$$

The final result is

$$|D_{N,G}(y_1, \ldots, y_N)| \leq \left(1 + \frac{M^5 L}{12}\right) D_N(x_1, \ldots, x_N). \tag{23}$$

□

The inversion method using linear Lagrange interpolation or cubic Hermite interpolation can be used to G-distributed low discrepancy sequences. The generation of low-discrepancy sequences with an arbitrary distribution function G is described in the Algorithm 14.

Algorithm 14. *Inversion method using interpolation*

Input data: the uniformly distributed low-discrepancy sequence $(x_1, \ldots, x_N)$;

for $k = 1, \ldots, N$ **do**

  Find the values x_k^- and x_k^+;

  Calculate the point y_k;

end for

Output data: the G-distributed low-discrepancy sequence $(y_1, \ldots, y_N)$.

The method that we proposed uses some values of G, g and derivatives of g to approximate G^{-1}. This is an advantage, as for some distributions the expression of G^{-1} is not known. Note that, in our method, adding one point to the generated sequence would not change the other elements of the sequence, which is another advantage.

References

[1] Chelson, P., *Quasi-Random Techniques for Monte Carlo Methods*, Ph. D. Dissertation, The Claremont Graduate School, 1976.

[2] Coman, Gh., *Numerical Analysis*, Ed. Libris, Cluj-Napoca, 1995 (in Romanian).

[3] Faure, H., *Discrépances de suites associées à un système de numération* (en dimension un), Bull. Soc. Math. France, 109 (1981), 143-182.

[4] Faure, H., *Discrépances de suites associées à un système de numération* (en dimension s), Acta Arith., 41 (1982), 337-351.

[5] Halton, J. H., *On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals*, Numer. Math, 2 (1960), 84-90.

[6] Hartinger, J., Kainhofer, R., *Non-uniform low-discrepancy sequence generation and integration of singular integrands*, Submitted, 2004.

[7] Hlawka, E., *Gleichverteilung und Simulation*, Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II, 206 (1997), 183-216.

[8] Hlawka, E., Mück, R., *A transformation of equidistributed sequences*, In Applications of number theory to numerical analysis, Academic Press, New York, 1972, 371-388.

[9] Hlawka, E., Mück, R., *Über Eine Transformation von gleichverteilten Folgen*, II, Computing, 9 (1972), 127-138.

[10] Niederreiter, H., *Random number generation and Quasi-Monte Carlo methods*, volume 63 of SIAM Conf. Ser. Appl. Math. SIAM, Philadelphia, 1992.

[11] Okten, G., *Error Reduction Techniques in Quasi-Monte Carlo Integration*, Mathematical and Computer Modelling, Vol. 30, No 7-8 (1999), 61-69.

[12] Owen, A. B., *Multidimensional variation for Quasi-Monte Carlo*, Technical report, Standford University, 2004.

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

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

THE UNITARY TOTIENT MINIMUM AND MAXIMUM FUNCTIONS

JÓZSEF SÁNDOR

Abstract. The unitary totient function has been introduced by E. Cohen [1]. The Euler minimum function has been first studied by P. Moree and H. Roskam [2], and independently by the author [4], who introduced more general concepts (and duals). A particular case is obtained for the unitary totient. Basic properties for this minimum, as well as maximum functions are pointed out. These include inequalities, divisibility properties, and values taken at special arguments. The necessary exponential diophantine equations are treated by elementary arguments.

1. Introduction

A divisor d of n is called **unitary** if $\left(d, \frac{n}{d}\right) = 1$. Let $(k, n)_*$ denote the greatest divisor of k which is a unitary divisor of n. The arithmetical functions associated with unitary divisors have been introduced by E. Cohen [1]. The multiplicative function

$$\mu^*(n) = (-1)^{\omega(n)},$$

where $\omega(n)$ denotes the number of distinct prime factors of n, is the unitary analogue of the Möbius function $\mu(n)$ and we have

$$\sum_{d||n} \mu^*(d) = \begin{cases} 1, & n = 1 \\ 0, & n > 1 \end{cases}$$

Received by the editors: 01.07.2004.

2000 *Mathematics Subject Classification.* 11A25, 11D61, 26D15.

Key words and phrases. Euler totient, unitary divisor, unitary totient, Euler minimum and maximum functions, inequalities, diophantine equations.

where $d||n$ denotes that d is a unitary divisor of n. Let $\varphi^*(n)$ denote the unitary analogue of the Euler totient function, that is $\varphi^*(n)$ represents the number of positive integers $k \leq n$ with $(k,n)_* = 1$. Then, it is easy to see that

$$\varphi^*(n) = \sum_{d|n} d\mu^* \left(\frac{n}{d}\right),$$

so $\varphi^*(n)$ is multiplicative, being the unitary convolution of two multiplicative functions (see [1]), and for $n = p_1^{\alpha_1} \dots p_r^{\alpha_r} > 1$ (prime factorization of n) we have

$$\varphi^*(n) = (p_1^{\alpha_1} - 1) \dots (p_r^{\alpha_r} - 1) \tag{1}$$

Put $\varphi^*(1) = 1$.

Let $A \subset \mathbb{N}^* = \{1, 2, \dots\}$ be a given set, and $f, g : \mathbb{N}^* \to \mathbb{N}^*$ two given functions. In [4] and [5] we have introduced the functions $F_f^A(n)$, $G_g^A(n)$ by (if these are well-defined)

$$F_f^A(n) = \min\{k \in A : \ n|f(k)\}, \tag{2}$$

and its "dual" by

$$G_g^A(n) = \max\{k \in A : \ g(k)|n\} \tag{3}$$

For $A = \mathbb{N}^*$, $f(k) = g(k) = \varphi(n)$, one obtains the "Euler minimum" and "Euler maximum" functions, given by

$$E(n) = \min\{k \in \mathbb{N}^* : \ n|\varphi(k)\}, \tag{4}$$

$$E_*(n) = \max\{k \in \mathbb{N}^* : \ \varphi(k)|n\} \tag{5}$$

For properties of $E(n)$ given by (4) see [2] and [6], while function (5) appears for the first time in [4] and [6].

The author has considered also other particular cases of (2) and (3) for $f(k) = g(k) = \sigma(k)$ (sum of divisors of k), $f(k) = d(k)$ (number of divisors of k) [7], $f(k) = g(k) = T(k)$ (product of divisors of k) [9], $f(k) = g(k) = S(k)$ (Smarandache function) [8], $f(k) = g(k) = Z(k)$ (pseudo-Smarandache function) [11], $f(k) = \varphi_e(n)$ (exponential totient function) [10]. It is interesting to note that, for $g(k) = d(k)$ or $g(k) = \varphi_e(n)$ the analogues functions to (5) are not well-defined.

The aim of this paper is the introduction and the initial study of the particular cases $f(k) = g(k) = \varphi^*(k)$, the unitary totient function. In analogy with (4) and (5) define

$$E^*(n) = \min\{k \geq 1 : \ n|\varphi^*(k)\}, \tag{6}$$

$$E^*_*(n) = \max\{k \geq 1 : \ \varphi^*(k)|n\} \tag{7}$$

First note that the functions $E^*(n)$ and $E^*_*(n)$ are well-defined. Indeed, by Dirichlet's theorem on arithmetic progressions, for each $n \geq 1$ there exists $a \geq 1$ so that $k = an + 1$ is a prime (see e.g. [3]). Then, since by (1) $\varphi^*(k) = k - 1 = an$, which is a multiple of n, (6) is well-defined. On the other hand, remark that

$$\varphi(n) \leq \varphi^*(n), \tag{8}$$

with equality only for $n = 1$ and $n =$ squarefree (i.e. product of distinct primes). Since φ and φ^* are multiplicative, (8) follows from

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} \leq p^\alpha - 1 = \varphi^*(p^\alpha)$$

(p prime, $\alpha \geq 1$), where for $\alpha = 1$ there is equality.

Now, since $\varphi(k) > \sqrt{k}$ for $k > 6$ (see e.g. [3]) and from $\varphi^*(k)|n$ it follows $\varphi^*(k) \leq n$, so $\sqrt{k} < n$, implying $k < n^2$. Thus $E^*_*(n) \leq \max\{6, n^2\} < +\infty$, so this function is well-defined, too.

2. Main results

Lemma 1. *For all $n \geq 2$ one has*

$$P(n) - 1 \leq \varphi^*(n) \leq n - 1, \tag{9}$$

where $P(n)$ denotes the greatest prime factor of n.

Proof. The left side inequality follows by $(p_1^{\alpha_1} - 1) \dots (p_r^{\alpha_r} - 1) \geq p_r^{\alpha_r} - 1 \geq p_r - 1$, where $p_1 < p_2 < \cdots < p_r$ are the distinct prime factors of n. Then by (1), $\varphi^*(n) \geq p_r - 1 = P(n) - 1$.

For the right side of (9), apply the obvious relation $(1 + y_1) \dots (1 + y_r) \geq 1 + y_1 \dots y_r$ ($y_i > 0$ for $i = 1, 2, \dots, r$) to $y_i = p_i^{\alpha_i} - 1 > 0$. Then we get $p_1^{\alpha_1} \dots p_r^{\alpha_r} \geq$

$1+(p_1^{\alpha_1}-1)\dots(p_r^{\alpha_r}-1)$, so by (1), the required result follows. Since here there is equality only for $r=1$, the equality sign in right-side of (9) is attained only for $n=$ prime power. Clearly, for the left side of (9) there is equality for $n=$ prime.

Lemma 2. *Let $r=\omega(n)$ be the number of distinct prime factors of n. Then*

$$\varphi^*(n) \le (n^{\frac{1}{r}}-1)^r \text{ for all } n \ge 2 \tag{10}$$

Proof. Apply the Huyggens inequality

$$\sqrt[r]{(1+y_1)\dots(1+y_r)} \ge 1+\sqrt[r]{y_1\dots y_r} \quad (y_i>0)$$

to $y_i=p_i^{\alpha_i}-1$. Then by (1), inequality (10) follows.

Theorem 1.

$$\varphi^*(E_*^*(n))|n|\varphi^*(E^*(n)) \text{ for all } n \ge 1. \tag{11}$$

Particularly,

$$\varphi^*(E_*^*(n)) \le n, \tag{12}$$

$$\varphi^*(E^*(n)) \ge n. \tag{13}$$

Proof. Let $E^*(n)=k_0$. By Definition (6), $n|\varphi^*(k_0)$. This gives the right side of (11). If $E_*^*(n)=k_1$, then by (7), $\varphi^*(k_1)|n$, so the left side of (11) follows. Relations (12) and (13) are direct consequences of (11).

Corollary 1.

$$E^*(n) \ge (n^{\frac{1}{r}}+1)^r \ge n+1, \text{ for } n \ge 2, \tag{14}$$

where $r=\omega(E^*(n))$.

Proof. By (10), $\varphi^*(E^*(n)) \le (E^*(n)^{\frac{1}{r}}-1)^r$, so by (13) we get $n \le (E^*(n)^{\frac{1}{r}}-1)^r$. This gives the first relation of (14). The second one is a trivial consequence of $(a+b)^r \ge a^r+b^r$ $(a,b>0,\ r\ge 1)$, which follows e.g. by the binomial theorem.

Corollary 2.

$$P(E_*^*(n)) \le n+1, \quad n \ge 2, \tag{15}$$

where $P(m)$ denotes the greatest prime factor of m.

Proof. This is similar to the proof of (14), by applying (12) and the left side of (9).

Remark 1. The weaker inequality of (14), i.e. $E^*(n) \geq n+1$ for $n \geq 2$ follows also by (13) and the right side of (9). This inequality becomes an equality for many values of n, e.g. for $n = 2, 3, 4, 6, 7, 8, 10, 12, 15, 16, 18, 22, \ldots$. Particularly, we prove:

Theorem 2. *If $p \geq 3$ is a prime, then*

$$E^*(p-1) = p \tag{16}$$

Proof. Since $(p-1)|\varphi^*(p)$ (because of $\varphi^*(p) = p-1$), by definition (6) it follows $E^*(p-1) \leq p$. On the other hand, applying $E^*(n) \geq n+1$ for $n = p-1 \geq 2$ one gets $E^*(p-1) \geq p$, so (16) is proved.

Clearly, since $\varphi^*(p)|(p-1)$, too, by (7) and (15) we get

Theorem 3. *For all primes p,*

$$E_*^*(p-1) \geq p, \tag{17}$$

and

$$P(E_*^*(p-1)) \leq p \text{ for } p \geq 3 \tag{18}$$

Remark 2. The exact calculation of $E_*^*(p-1)$ seems difficult. However, the determination of $E_*^*(p)$ is given by the following

Theorem 4.

$$E_*^*(p) = \begin{cases} 6, & \text{if } \; p = 2, \\ 2, & \text{if } \; p \geq 3 \text{ is not a Mersenne prime}, \\ 2^n, & \text{if } \; p = 2^n - 1 \text{ is a Mersenne prime} \end{cases} \tag{19}$$

First we prove the following auxiliary result:

Lemma 3. *Let p be a prime. Then the equation*

$$\varphi^*(x) = p$$

is solvable if and only if $p = 2$ or p is a Mersenne prime (with a single solution).

Proof. If x is composite, $x = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, with $\omega(x) = r \geq 2$, then $\varphi^*(x) = (p_1^{\alpha_1} - 1) \dots (p_r^{\alpha_r} - 1)$ is always composite, so $\neq p$. If $r = 1$, i.e. $x = q^\alpha$, then $\varphi^*(x) = q^\alpha - 1 = p$ iff $q^\alpha = p + 1$. Now, if $p \geq 3$, then $p+1$ is even, so we must have $q = 2$, i.e. $p = 2^\alpha - 1 =$ Mersenne prime (see [3]). If $p = 2$, we get $q = 3$, $\alpha = 1$ so x is not composite.

If $x = q$ is a prime, then $\varphi^*(x) = q - 1 = p \Leftrightarrow q = p + 1$, and this is solvable only if $p = 2$, since for $p \geq 3$, $p + 1$ being even, cannot be a prime.

Now, for the proof of (19), let $\varphi^*(k)|p$. Then $\varphi^*(k) = 1$ (i.e. $k = 1$ or 2), or $\varphi^*(k) = p$. Since $p \geq 2$, always, the result follows from Lemma 3, by taking into account the form of the solution, when p is a Mersenne prime.

We now prove:

Lemma 4. *Let $k \geq 1$ be an integer. Then the equation*

$$\varphi^*(x) = 2^k$$

is always solvable, and its solutions are of the form $x = F$, or $x = 2F$, where $F = 9$; a Fermat prime; or the product of distinct Fermat primes.

Proof. Let $x = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, when $\varphi^*(x) = (p_1^{\alpha_1} - 1) \dots (p_r^{\alpha_r} - 1) = 2^k \Leftrightarrow p_1^{\alpha_1} - 1 = 2^{a_1}, \dots, p_r^{\alpha_r} - 1 = 2^{a_r}$, with $a_1 + \dots + a_r = k$. Thus $p_1^{\alpha_1} = 2^{a_1} + 1, \dots, p_r^{\alpha_r} = 2^{a_r} + 1$, so each p_i $(i = 1, 2, \dots, r)$ is odd, so x must be odd. Since we can have also the case $2^1 - 1 = 2^0$, x could be also of the form $x = 2F$, where F is odd. Therefore we must study an equation of type

$$p^\alpha = 2^a + 1 \quad (a \geq 1) \tag{20}$$

1) If $\alpha = 2m$ is even, then $(p^m - 1)(p^m + 1) = 2^a$ gives $p^m - 1 = 2^u$, $p^m + 1 = 2^v$ $(u + v = a)$, so $2^v - 2^u = 2$, i.e. $2^v = 2(1 + 2^{u-1})$, which is possible only if $u = 1$,

$v = 2$. Then $p^m - 1 = 2$, $p^m + 1 = 4$, giving $p = 3$, $m = 1$, $\alpha = 2$; so $a = 3$ and $x = p^\alpha = 9$.

2) If $\alpha = 2m + 1$ $(m \geq 0)$, for $m = 0$ we get $\alpha = 1$, so $p = 2^a + 1$ is a prime, so it is a Fermat prime (see [3]). Let $m \geq 1$. Then since $p^{2m+1} - 1 = (p-1)(p^{2m} + p^{2m-1} + \cdots + p + 1)$ and p is odd, the second term contains a number of $2m + 1$ odd terms, so it is odd. Thus (20) is impossible.

This finishes the proof of Lemma 4.

Theorem 5. *$E_*^*(2^t) = 2m$, where m is the greatest of the products $(2^{a_1} + 1) \dots (2^{a_r} + 1)$ of Fermat primes, where $a_1 + \cdots + a_r \leq t$.*

Proof. Let $\varphi^*(k) | 2^t$. Then $\varphi^*(k) = 2^a$, where $0 \leq a \leq t$. By Lemma 3, the greatest such k is $k = 2m$, where $m = (2^{a_1} + 1) \dots (2^{a_r} + 1)$, with $a_1 + \cdots + a_r = a \leq t$ and r is maximal (i.e. m is maximal if $a_1 + \cdots + a_r \leq t$).

Example. $E_*^*(8) = 30$.

Indeed, $8 = 2^3$, $a_1 + \cdots + a_r \leq 3 \iff r = 2$, since $2^1 + 1 = 3$, $2^2 + 1 = 5$ are Fermat primes and $1 + 2 = 3$. So $2m = 2 \cdot 3 \cdot 5 = 30$.

Lemma 5. *Let p be a prime. Then*

$$\varphi^*(x) = p^2$$

is solvable iff $p = 2$. The solutions are $x = 5, 10$.

Proof. 1) Let $x = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ be odd. Then $(p_1^{\alpha_1} - 1) \dots (p_r^{\alpha_r} - 1) = p^2$ iff

a) $p_1^{\alpha_1} - 1 = p^2$;

b) $p_1^{\alpha_1} - 1 = 1$, $p_2^{\alpha_2} - 1 = p^2$;

c) $p_1^{\alpha_1} - 1 = 1$, $p_2^{\alpha_2} - 1 = 1$, $p_3^{\alpha_3} - 1 = p^2$;

d) $p_1^{\alpha_1} - 1$, $p_2^{\alpha_2 - 1} = p$, $p_3^{\alpha_3} - 1 = p$.

Remark that cases b), c), d) are impossible, since then $p_1 = 2$ always, and this contradicts $x =$ odd. In case a), since p_1 is odd we must have $p =$ even, so $p = 2$. But in this case, $p_1 = 5$, $\alpha_1 = 1$, so $x = 5$.

2) If x is even, then $p_1 = 2$. In case a) we can write $2^{\alpha_1} = p^2 + 1$. For $p = 2, 3, 5$ this is impossible. If $p > 5$, then it is known (see e.g. [3]) that p must have

the forms $p = 6M \pm 1$, so $p^2 = 36k^2 \pm 12k + 1 = 12k(3k \pm 1) + 1 = 24M + 1$, so $p^2 + 1 = 2(12M + 1) \neq 2^{\alpha_1}$.

In case b) $\alpha_1 = 1$, $p_2^{\alpha_2} = 5$ ($p_2 = 5$, $\alpha_2 = 1$) is possible, implying $x = 2{\cdot}5 = 10$.

Cases c), d) cannot hold, since then e.g. in case c) $\alpha_1 = 1$, $\alpha_2 = 1$, $p_2 = p_3$ and this is a contradiction to $p_3 > p_2$. Similarly, in case d).

Theorem 6.

$$E_*^*(p^2) = \begin{cases} 10, & \textit{if} \quad p = 2 \\ 2, & \textit{if} \quad p \geq 3 \textit{ is not a Mersenne prime} \\ 2^k, & \textit{if} \quad p = 2^k - 1 \textit{ is a Mersenne prime} \end{cases}$$

Proof. $\varphi^*(k) | p^2 \iff \varphi^*(k) \in \{1, p, p^2\}$. Now, apply Lemmas 3, 5, and definition (7).

Lemma 6. *Let p be a prime, $k > 1$ an integer. Then the equation*

$$\varphi^*(x) = p^k$$

is solvable only for $p = 2$.

Proof. First we prove an auxiliary result:

Lemma 6'. *If $k > 1$ and p is a prime, $\alpha \geq 1$, then the equation*

$$p^k = 2^\alpha - 1 \tag{21}$$

is not solvable.

Proof. First remark that p must be odd. If $k = 2m + 1$ ($m \geq 1$) is odd, then $p^{2m+1} + 1 = (p + 1)(p^{2m} - \cdots + p + 1)$, where the second term contains an odd number of odd terms, and the signs $+$ or $-$, so it is odd. Thus (21) is impossible. If $k = 2m$ is even, and $p > 5$, then by $p = 6s \pm 1$ (as in the proof of case 2) of Lemma 5) $p^2 = 24M + 1$, so $p^{2m} = \mathcal{M}24 + 1$, $p^{2m} + 1 = 2(\mathcal{M}12 + 1) \neq 2^\alpha$.

For $p = 3, 5$ we must consider separately equation (21) in case $k = 2m$. So $3^{2m} = 9^m = (8 + 1)^m = M8 + 1$, so $M8 + 2 = 2(M4 + 1) \neq 2^\alpha$. Similarly, $5^{2m} = 25^m = (2n + 1)^m = M24 + 1$, i.e. $5^{2m} + 1 = M24 + 2 = 2(M12 + 1) \neq 2^\alpha$. This finishes the proof of Lemma 6'.

The proof of Lemma 6 is similar to that of Lemma 5.

When x is odd, then $p_1^{\alpha_1} - 1 = p^k$ in case a) so $p = 2$, so by (20). This is possible only when $p_1 = 3$, so $k = 3$.

The other cases, when $p_1^{\alpha_1} - 1 = 1$, etc. are impossible, since $p_1 = 2$, contradiction to $x =$ odd. Similarly the case $p_1^{\alpha_1} - 1 = p, \dots, p_r^{\alpha_r} - 1 = p$, $k = 2$, since then $p_1 = \dots = p_r$, impossible.

When x is even, i.e. $p_1 = 2$, in case a) we get $2^{\alpha_1} - 1 = p^k$, and by (21) this is not solvable.

Theorem 7.

$$E_*^*(p^k) = \begin{cases} 2m, & \text{if } \ p = 2, \text{ where } m \text{ is given by Theorem 5,} \\ 2, & \text{if } \ p \geq 5 \text{ is not a Mersenne prime,} \\ 2^k, & \text{if } \ p = 2^k - 1 \text{ is a Mersenne prime.} \end{cases}$$

Proof. This is similar to the proof of Theorem 6 (case $k = 2$), but remarking that for $p = 2$ we must use Theorem 5.

References

[1] Cohen, E., *Arithmetical functions associated with the unitary divisors of an integer*, Math. Z. **74**(1960), 66-80.

[2] Moree, P., Roskam, H., *On an arithmetical function related to Euler's totient and the discriminator*, Fib. Quart. **33**(1995), 332-340.

[3] Niven, I., Zuckerman, H., *An introduction to the theory of numbers*, 4th ed., Wiley, New York, 1980.

[4] Sándor, J., *On certain generalizations of the Smarandache function*, Notes Number Th. Discr. Math. **5**(1999), no. 2, 41-51.

[5] Sándor, J., *On certain generalizations of the Smarandache function*, Smarandache Notion J., **11**(2000), no. 1-3, 202-212.

[6] Sándor, J., *On the Euler minimum and maximum functions*, to appear.

[7] Sándor, J., *The sum-of-divisors and maximum functions*, to appear, C. R. Bulg. Acad. Sci.

[8] Sándor, J., *The Smarandache minimum and maximum functions*, accepted for publication, Smarandache Notions J.

[9] Sándor, J., *The product-of-divisors minimum and maximum functions*, RGMIA Research Report Collection (Australia), **7**(2004), no. 2, art. 18, 11 pp.

[10] Sándor, J., *A note on exponential divisors and related arithmetic functions*, submitted.

[11] Sándor, J., *On the pseudo-Smarandache minimum and maximum functions*, submitted.

Department of Mathematics and Informatics, Babeş-Bolyai University, Str. Kogălniceanu 1, 400084 Cluj-Napoca, Romania

E-mail address: jjsandor@hotmail.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

POLYNOMIAL REPRODUCIBILITY OF A CLASS OF REFINABLE FUNCTIONS: COEFFICIENTS PROPERTIES

E. SANTI AND E. PELLEGRINO

Abstract. In this paper we investigate some properties of the coefficients of the formulas defining the polynomial reproducibility of a class of refinable functions.

1. Introduction

It is well known that constructing refinable approximation operators for real valued functions, it is desiderable to obtain operators approximating smooth functions f with an order of accuracy comparable to the best refinable function approximation.

The key for obtaining operators with such property is to require that they reproduce appropriate classes of polynomials.

Considering the G.P. refinable B-basis on $I = [0, n+1]$, $W_j = \{w_{ji}(x)\}_{i=-n}^{N_j}$, constructed by starting from the class of refinable function defined in [2], it has been proved in [6] that the quasi-interpolatory refinable operators of the form

$$Q_j f(x) = \sum_{i=-n}^{N_j} (\lambda_{ji} f)\, w_{ji}(x) \qquad x \in I = [0, n+1] \tag{1.1}$$

where $N_j = 2^j (n+1) - 1$, and $\{\lambda_{ji}\}_{i=-n}^{N_j}$ is a set of linear functionals, reproduce polynomials $\in \mathbb{P}_l$, the class of polynomials of degree $l-1$, with $1 \le l \le n-1$, if and only if

$$\lambda_{ji} x^{k-1} = \eta_{ji}^{(k)} \quad k = 1, \ldots, l \quad i = -n, \ldots, N_j \tag{1.2}$$

Received by the editors: 12.01.2005.

2000 *Mathematics Subject Classification.* 41A36.

Key words and phrases. Refinable functions, B-splines, optimal bases.

where $\eta_{ji}^{(k)}$ are such that

$$x^{k-1} = \sum_{i=-n}^{N_j} \eta_{ji}^{(k)} w_{ji}(x) \quad x \in I. \tag{1.3}$$

Considering the operators λ_{ji} defined as $\lambda_{ji} := \sum_{k=1}^{l} \alpha_{jik}\lambda_{jik}$ it has been proved in [6], that assuming

$$\alpha_{jik} = \sum_{\nu=0}^{k-1} (-1)^{\nu} \eta_{ji}^{(k-\nu)} symm_{\nu}(\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik-1}) \quad k = 1, \ldots, l \tag{1.4}$$

where $symm_{\nu}(\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik-1})$ is the symmetric function on the distinct points $\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik-1} \in I$, the approximating refinable operator:

$$Q_j f(x) = \sum_{i=-n}^{N_j} \sum_{k=1}^{l} \alpha_{jik} [\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik}] f \; w_{ji}(x) \tag{1.5}$$

where $[\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik}] f$ is the $k-1$ divided difference of f on the points $\tau_{ji1}, \tau_{ji2}, \ldots, \tau_{jik}$ reproduces the polinomials in $\mathbb{P}_l$, $1 \leq l \leq n-1$.

The aim of this paper is to investigate the properties of $\eta_{ji}^{(k)}$ and their localization in the interval I.

We shall make use of such results for proving the convergence of refinable operators (1.1), [7].

The paper is organized as follows. In section 2 we shall give some definitions and the main properties of a wide class of refinable functions Φ_j defined in [2] and we define the B-bases, W_j, constructed starting from Φ_j.

In section 3 we shall consider the evaluation of $\eta_{ji}^{(k)}$ and prove some useful properties.

Section 4 is devoted to prove the relation between the values $\eta_{ji}^{(k)}$ and $\overline{\xi}_{ji}^{(k)}$ that are defined by:

$$x^{k-1} = \sum_{i=-n}^{N_j} \overline{\xi}_{ji}^{(k)} B_{ji}(x), \quad \forall\, x \in I \tag{1.6}$$

where $B_{ji}(x)$ are the normalized B-splines of order $n+1$.

2. Preliminaries

In this section we report for later use some definitions and notations.

It is well known that a scaling function is the solution of a scaling equation

$$\varphi(x) = \sum_{i\in\mathbb{Z}} a_i \varphi(2x - i) \tag{2.1}$$

The class Φ of scaling functions, here considered, consists of the functions $\varphi_h(x)$ solving the scaling equation

$$\varphi_h(x) = \sum_{k=0}^{n+1} a_{kh} \varphi_h(2x - k) \tag{2.2}$$

where h is a real parameter $h \geqslant n \geqslant 2$, and

$$a_{kh} = \frac{1}{2^h}\left[\binom{n+1}{k} + 4\left(2^{h-n} - 1\right)\binom{n-1}{k-1}\right], \quad k = 0, 1, \ldots, n+1. \tag{2.3}$$

Such scaling functions, called G.P. refinable functions, are characterized by the following properties [2]:

(i) supp $\varphi_h = [0, n+1]$;

(ii) $\varphi_h \in C^{n-2}(\mathbb{R})$;

(iii) $\varphi_h(x) > 0 \quad \forall x \in (0, n+1)$;

(iv) φ_h is centrally symmetric, that is $\varphi_h(x) = \varphi_h(n+1-x)$;

(v) its *symbol* $P_n(z) = \sum_{k=0}^{n+1} a_{kh} z^k$ is left-half plane stable;

(vi) $\sum_{k\in\mathbb{Z}} \varphi(x-k) = 1$.

For the above properties, for any admissible h, the system of linearly independent functions

$$\Phi_{0,h} = \{\varphi_h(x-k), \quad k \in \mathbb{Z}\}, \quad h \geq n \geq 2, \tag{2.4}$$

provides a normalized totally positive (NTP) basis in $\mathbb{R}$. Moreover, φ_h generates a multiresolution analysis (MRA) in $L^2(\mathbb{R})$, whose approximating space V_j are defined by

$$V_j = clos_{L^2}\left\{2^{j/2}\varphi_h\left(2^j x - k\right),\ k \in \mathbb{Z}\right\}, \quad j \in \mathbb{Z}^+.$$

Considering a bounded interval $J=[a,b]$, the system

$$\Phi_{j,h}=\left\{\varphi_{jhk}(x)=2^{j/2}\varphi_h\left(2^jx-k\right),\ 2^ja-n\le k\le 2^jb-1,\ x\in[a,b]\right\} \tag{2.5}$$

with $j\ge j_0$, where j_0 is the first integer such that $2^{j_0}(b-a)\ge n+1$, constitutes a NTP basis. For the sake of short notation we shall eliminate the explicit dependence on h of $\Phi_{j,h}$, we write φ_{jk} for φ_{jhk} and we shall set

$$\underline{N}_j=2^ja-n \quad \overline{N}_j=2^jb-1;$$

therefore, Φ_j is a NTP basis for the space $\widetilde{V}_j$ generated on J.

In [3], it has been proved that there exist $\overline{N}_j-\underline{N}_j+1$ real numbers $\xi_{jk}^{(l)}$ such that,

$$x^{l-1}=\sum_{k=\underline{N}_j}^{\overline{N}_j}\xi_{jk}^{(l)}\varphi_{jk}(x) \quad x\in J, \quad l=1,\ldots,n-1. \tag{2.6}$$

We remark that for the property (vi) of refinable functions φ_{jk}, there results $\xi_{jk}^{(1)}=1\ \forall\ k=\underline{N}_j,\ldots,\overline{N}_j$.

We recall that a TP system of linearly independent functions $W_j=(w_{j0}\ldots w_{jm})$ defined on the bounded interval J is said to be a B-basis (or optimal basis) if each TP basis $U_j=(u_{j0}\ldots u_{jm})$ of the space generated by W_j satisfies the relation

$$U_j=W_jA_j \tag{2.7}$$

where A_j is a non singular TP and stochastic matrix.

In [1] an algorithm for the construction of W_j starting from any U_j is given. This algorithm can be applied, in particular, when the TP system U_j under consideration is constituted by suitable integer shifts of TP refinable functions as considered in (2.5).

Therefore, let $W_j=\{w_{ji}(x)\}_{i=\underline{N}_j}^{\overline{N}_j}$ be the B-basis associated to Φ_j, W_j generates a MRA on J [3]. Moreover, this MRA reproduces polynomials up to the order $d=n-1$, that is, there exist $\overline{N}_j-\underline{N}_j+1$ real numbers $\eta_{jk}^{(l)}$ such that

$$x^{l-1}=\sum_{k=\underline{N}_j}^{\overline{N}_j}\eta_{jk}^{(l)}w_{jk}(x) \quad x\in J, \quad l=1,\ldots,n-1, \tag{2.8}$$

with

$$\underline{\eta}_j^{(l)} = A_j \underline{\xi}_j^{(l)}, \tag{2.9}$$

and A_j is the matrix in (2.7).

By exploiting some properties of the system of functions W_j, we shall establish some properties of the values $\eta_{jk}^{(l)}$ that can be very profitable in proving convergence properties of approximating (in particular quasi-interpolatory) refinable operators.

We assume $J = I = [0, n+1]$ an assumption by no means restrictive, and in such a case there results $\underline{N}_j = -n$, $\overline{N}_j = 2^j(n+1) - 1 = N_j$ and

$$\text{supp } w_{jk} = \left[\max\left(0, \frac{k}{2^j}\right), \min\left(\frac{k+n+1}{2^j}, n+1\right)\right] \tag{2.10}$$

In Fig.1 we show an example of B-basis constructed on $I = [0, 6]$ starting from Φ_0, NTP basis with support $[0, 6]$ and $h = 10$.

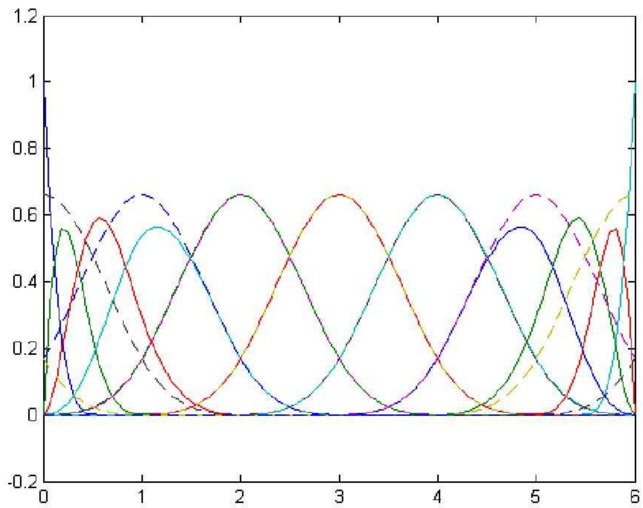


FIGURE 1. The normalized B-bases (solid line) in the interval $[0, 6]$. The dashed line represents the starting NTP basis Φ_0 with $n = 5$ and $h = 10$

TABLE 1

	$L=1$	$L=2$	$L=3$
x_i	err	err	err
0.000	5.4 (-17)	7.1 (-16)	5.4 (-15)
1.000	1.1 (-16)	4.4 (-16)	3.7 (-15)
2.000	1.1 (-16)	0.0 (00)	3.3 (-16)
3.000	0.0 (00)	0.0 (00)	1.3 (-16)
4.000	0.0 (00)	1.1 (-16)	2.2 (-16)
5.000	1.8 (-16)	1.4 (-16)	1.1 (-16)
6.000	1.5 (-16)	2.0 (-16)	0.0 (00)

The table 1 shows the relative error, valued in knots x_i, given by the difference between first and second member of (2.8), for several values of L, confirming the polynomial reproducibility.

3. Evaluation and localization of $\eta_{jk}^{(l)}$

In [3] has been proved that for the values $\xi_{0k}^{(l)}$, at level $j=0$, the following relations

$$\xi_{0k}^{(l)} = \sum_{r=0}^{l-1} \binom{l-1}{r} k^{l-1-r} C_r \tag{3.1}$$

$$C_0 = 1, \quad C_1 = \mu_1, \quad C_r = \sum_{s=1}^{r} (-1)^{s-1} \binom{r}{s} \mu_s C_{r-s} \tag{3.2}$$

hold, where $\mu_i := \int_{\mathbb{R}} x^i \varphi_0(x)\, dx$ denotes the i-th moment of $\varphi_0(x) = \varphi(x)$.

By rearranging (3.1), we estabilish a formula that allows the evaluation in I of $\xi_{jk}^{(l)}$ for any $j \geqslant 0$, only using the values of $\xi_{jk}^{(2)} = 2^{-j}\left(k + \frac{n+1}{2}\right)$, $k = -n, \ldots, N_j$ and the coefficients C_i $i = 0, \ldots, l-1$ in (3.2). This procedure permits to avoid, for each j, the evaluation of unnecessary values of $\xi_{jk}^{(l)}$, that should be used for the level $j+1$.

Proposition 1. *For $j \geq 0$, there results*

$$\xi_{jk}^{(l)} = \left(\xi_{jk}^{(2)}\right)^{l-1} + 2^{-j(l-1)} \sum_{s=2}^{l-1} \binom{l-1}{s} k^{l-1-s} \left(C_s - C_1^s\right). \tag{3.3}$$

Proof. From (3.1), it is straightforward to verify that for $j = 0$

$$\xi_{0k}^{(l)} = k\xi_{0k}^{(l-1)} + k\sum_{r=1}^{l-2}\binom{l-2}{r-1}k^{l-2-r}C_r + C_{l-1}, \quad k = -n, \ldots, N_0 \tag{3.4}$$

and in particular, being $\xi_{0k}^{(2)} = k + C_1$, there results

$$\begin{aligned} \xi_{0k}^{(3)} &= k\xi_{0k}^{(2)} + kC_1 + C_2 = \left(k^2 + 2kC_1 + C_1^2\right) + C_2 - C_1^2 \\ &= \left(\xi_{0k}^{(2)}\right)^2 + \left(C_2 - C_1^2\right), \quad k = -n, \ldots, N_0. \end{aligned} \tag{3.5}$$

By applying iteratively (3.4) and rearranging the terms, we obtain

$$\xi_{0k}^{(l)} = \left(\xi_{0k}^{(2)}\right)^{l-1} + \sum_{s=2}^{l-1}\binom{l-1}{s}k^{l-1-s}\left(C_s - C_1^s\right), \quad k = -n, \ldots, N_0. \tag{3.6}$$

Now considering that

$$\mu_{ji} = 2^j \int_{I\!R} \left(2^j x\right)^i \varphi\left(2^j x\right) dx = \int_{I\!R} t^i \varphi(t)\, dt = \mu_{0i} \tag{3.7}$$

and

$$\sum_{k\in Z\!\!Z} \left(2^j x - k\right)^i \varphi_{jk}(x) = \mu_{ji}, \tag{3.8}$$

we obtain, for $l = 1, 2, \ldots, n-1$

$$x^{l-1} = \sum_{k\in Z\!\!Z} 2^{-j(l-1)}\xi_{0k}^{(l)}\varphi_{jk}(x) \tag{3.9}$$

and we get (3.3) by assuming $\xi_{jk}^{(l)} = 2^{-j(l-1)}\xi_{0k}^{(l)}$. □

Therefore the procedure for evaluating for any fixed integer $l,\ 1 \le l \le n-1$, $\xi_{jk}^{(l)},\ k = -n, \ldots, N_j$, consists in the following steps:

- evaluation of $C_i,\ i = 0, 1, \ldots, l-1$ using (3.2);
- evaluation of $\xi_{jk}^{(2)} = 2^{-j}\left(k + \frac{n+1}{2}\right) \quad k = -n, \ldots, N_j$;
- evaluation of $\xi_{jk}^{(l)}$ by means of (3.3).

Once determined $\underline{\xi}_j^{(l)} = \left[\xi_{j,-n}^{(l)} \cdots \xi_{j,N_j}^{(l)}\right]$ we determine by means of (2.9) the vector $\underline{\eta}_j^{(l)} = A_j \underline{\xi}_j^{(l)}$.

We are interested now, in investigate the properties of vector $\underline{\eta}_j^{(l)}$ and the localization of its components, useful for proving the convergence of operators (1.5) as $j \to \infty$.

Starting again from (3.1) and considering that the values of $\xi_{0,k}^{(l)}$, $k = -n, \ldots, N_0$ are obtained evaluating a polynomial of $l-1$ degree with coefficients $C_0, C_1, \ldots, C_{l-1}$ at the point k, and the first derivative of such polynomial evaluated at k coincides with $(l-1)\ \xi_{0,k}^{(l-1)}$ we can write

$$D\underline{\xi}_0^{(l)} = (l-1)\,\underline{\xi}_0^{(l-1)}. \tag{3.10}$$

Therefore we can prove the following

Proposition 2. *All the nodes $\eta_{0,k}^{(l)}$, $k = -n, \ldots, n$, $l = 1, \ldots, n-1$ are non decreasing and they satisfy, for $l > 1$:*

$$0 = \eta_{0,-n}^{(l)} \leq \eta_{0,-n+1}^{(l)} \leq \ldots \leq \eta_{0,-n+l}^{(l)} \leq \ldots \leq \eta_{0,n}^{(l)} = (n+1)^{l-1}. \tag{3.11}$$

Proof. We know that for $l = 1$ there results

$$\eta_{0,k}^{(1)} = 1 \quad k = -n, \ldots, N_0 = n. \tag{3.12}$$

Using (2.8) with $j = 0$, and taking into account that $w_{0k}(0) = \delta_{k,-n}$ and $w_{0k}(n+1) = \delta_{k,n}$ one has, for $1 < l \leq n-1$,

$$0 = \sum_{k=-n}^{n} \eta_{0,k}^{(l)} w_{0k}(0) = \eta_{0,-n}^{(l)}, \quad (n+1)^{l-1} = \sum_{k=-n}^{n} \eta_{0,k}^{(l)} w_{0k}(n+1) = \eta_{0,n}^{(l)}. \tag{3.13}$$

Finally by (2.9) and (3.10)

$$D\underline{\eta}_0^{(l)} = A_0 D\underline{\xi}_0^{(l)} = (l-1)\,A_0\underline{\xi}_0^{(l-1)} = (l-1)\,\underline{\eta}_0^{(l-1)}. \tag{3.14}$$

In [4] has been proved that

$$0 = \eta_{0,-n}^{(2)} \leq \eta_{0,-n+1}^{(2)} \leq \ldots \leq \eta_{0,n}^{(2)} = n+1, \tag{3.15}$$

that can be also deduced considering (3.12) - (3.14).

Therefore, for $l = 3$ the sequence $\eta_{0,i}^{(3)}$ $i = -n, \ldots, n$, using (3.14), is non decreasing and bounded by 0 and $(n+1)^2$. By induction we prove (3.11) for each $1 < l \leq n-1$. □

Remark 1. *We remark that being* $\eta^{(2)}_{0,-n}=0$, *when* $l=3$ *one has for (3.13)* $\eta^{(3)}_{0,-n}=0$ *and using (3.14)* $\eta^{(3)}_{0,-n+1}=0$. *Then for* $l\geq 2$, *there results*

$$\eta^{(l)}_{0,-n}=\eta^{(l)}_{0,-n+1}=\ldots=\eta^{(l)}_{0,-n+l-2}=0.$$

The following proposition gives the precise localization of $\eta^{(l)}_{0,i}$ for any $i=-n,\ldots,n$ and $l\geq 2$.

Proposition 3. *There results:*

$$\begin{cases}\eta^{(l)}_{0,i}\in\left[0,(i+n+1)^{l-1}\right] & \text{for } i=-n,\ldots,0\\ \\ \eta^{(l)}_{0,i}\in\left[i^{l-1},(n+1)^{l-1}\right] & \text{for } i=1,\ldots,n\end{cases}\quad. \tag{3.16}$$

We recall that $[0,i+n+1]$ *are the supports of* w_{0i} *for* $i=-n,\ldots,0$ *and* $[i,n+1]$ *are that ones of* w_{0i} *for* $i=1,\ldots,n$.

Proof. Let us demonstrate (3.16) just for $i=-n+1,\ldots,-1$ and for $i=1,\ldots,n-1$, since for $i=-n,n$ and $i=0$ it is obvious.

Recalling that W_0 is TP and normalized, for $s=-n+1,\ldots,-1$ using (2.8) and (3.11) with $j=0$, we obtain

$$x^{l-1}=\sum_{i=-n}^{n}\eta^{(l)}_{0i}w_{0i}(x)\geq\sum_{i=s}^{n}\eta^{(l)}_{0s}w_{0i}(x)=\eta^{(l)}_{0s}\sum_{i=s}^{n}w_{0i}(x).$$

Since $1=\sum\limits_{i=-n}^{n}w_{0i}(x)=\sum\limits_{i=-n}^{s-1}w_{0i}(x)+\sum\limits_{i=s}^{n}w_{0i}(x)$, one has

$$\sum_{i=s}^{n}w_{0i}(x)=1-\sum_{i=-n}^{s-1}w_{0i}(x).$$

If it were $\eta^{(l)}_{0s}>(s+n+1)^{l-1}$, for $x\in[n+s-1,n+s]$ we had

$$x^{l-1}>(s+n+1)^{l-1}\left(1-w_{0s-1}(x)\right) \tag{3.17}$$

since $w_{0k}(x)\equiv 0$, $k=-n,\ldots,s-2$ for $x>n+s-1$.

But $w_{0s-1}(n+s)=0$, thus in $[n+s-1,n+s]$ there are points x for which $(1-w_{0s-1}(x))>\frac{(s+n)^{l-1}}{(s+n+1)^{l-1}}$, and for the same points, (3.17) would give $x^{l-1}>(s+n)^{l-1}$, a contradiction.

Consider now $s = 1, \ldots, n-1$; for each fixed s, supp $w_{0s} = [s, n+1]$. We had to prove that $\eta_{0s}^{(l)} > s^{l-1}$. For $x \in [s, s+1]$, there results

$$x^{l-1} = \sum_{i=-n}^{n} \eta_{0i}^{(l)} w_{0i}(x) = \sum_{i=-n+s}^{s} \eta_{0i}^{(l)} w_{0i}(x) \leq \eta_{0s}^{(l)} \sum_{i=-n+s}^{s} w_{0i}(x) \leq \eta_{0s}^{(l)}.$$

If we were to take $\eta_{0s}^{(l)} < s^{l-1}$, we would have $x^{l-1} < s^{l-1}$ a contradiction. □

For extending the results to the level j we recall relation (2.10) and, by denoting $y_{j,i}$, $y_{j,i+n+1}$ the bounds of supp w_{ji}, we can write:

$$y_{j,i}^{l-1} \leq \eta_{ji}^{(l)} \leq y_{j,i+n+1}^{l-1} \tag{3.18}$$

Remark 2. *In the supports $[y_{j,i}, y_{j,i+n+1}]$ there are, for any i, $n+2$ uniformly spaced (norm $\triangle_j = 2^{-j}$) points, also partially coinciding with 0 or $n+1$.*

4. Relation between $\eta_{jk}^{(l)}$ and $\xi_{jk}^{(l)}$

In the following Proposition 4 we determine a relation between the values $\eta_{jk}^{(l)}$ and $\overline{\xi}_{jk}^{(l)}$ that are such that:

$$x^{l-1} = \sum_{i=-n}^{N_j} \overline{\xi}_{ji}^{(l)} B_{ji}(x) \quad x \in I, \ 1 \leq l \leq n,$$

where $B_{ji}(x)$ are the normalized B-splines of order $n+1$ and uniformly spaced knots, that, as we know, have the same supports of $w_{ji}(x)$.

Proposition 4. *For any integer l, $1 \leq l \leq n-1$, and $j \geq 0$, there results*

$$\eta_{ji}^{(l)} = C_j(i,l)\, \overline{\xi}_{ji}^{(l)} \tag{4.1}$$

where $0 \leq C_j(i,l) \leq [2(n+1)]^{l-1}$.

Proof. Let $suppB_{ji} = [y_{j,i}, y_{j,i+n+1}]$ we know [8] that

$$\overline{\xi}_{ji}^{(l)} = \frac{symm_{l-1}(y_{j,i+1}, \ldots, y_{j,i+n})}{\binom{n}{l-1}} \quad i = -n, \ldots, N_j, \tag{4.2}$$

where the symmetric function is defined by:

$$symm_k\left(t_1, t_2, \ldots, t_p\right) = \sum_{1 \le i_1 \le i_2 \le \ldots \le i_k \le p} t_{i_1} t_{i_2} \ldots t_{i_k} \tag{4.3}$$

and the sum is over $\binom{p}{k}$ terms.

For $l = 1$ there results: $\overline{\xi}_{ji}^{(1)} = \eta_{ji}^{(1)} = 1$, $\forall i = -n, \ldots, N_j$.

Therefore we consider $2 \le l \le n-1$ and recall that for each l, $\overline{\xi}_{j,-n}^{(l)} = \overline{\xi}_{j,-n+1}^{(l)} = \ldots = \overline{\xi}_{j,-n+l-2}^{(l)} = 0$ and $\eta_{j,-n}^{(l)} = \eta_{j,-n+1}^{(l)} = \ldots = \eta_{j,-n+l-2}^{(l)} = 0$, and $\overline{\xi}_{jN_j}^{(l)} = \eta_{jN_j}^{(l)} = (n+1)^{l-1}$.

Then, we can consider the cases: a) $i = -n+l-1, \ldots, -1$; b) $i = 0, \ldots, N_j - n$; c) $i = N_j - n + 1, \ldots, N_j - 1$.

a) Since $y_{j,-n+1} = \ldots = y_{j,0} = 0$, $symm_{l-1}\left(y_{j,i+1}, \ldots, y_{j,i+n}\right)$ reduces to the sum of $\binom{i+n}{l-1}$ terms, and

$$\binom{i+n}{l-1} \frac{y_{j,1}^{l-1}}{\binom{n}{l-1}} \le \overline{\xi}_{ji}^{(l)} \le \binom{i+n}{l-1} \frac{y_{j,i+n}^{l-1}}{\binom{n}{l-1}}.$$

Therefore

$$0 = \frac{\binom{n}{l-1}}{\binom{i+n}{l-1}} \left(\frac{y_{j,i}}{y_{j,i+n}}\right)^{l-1} \le \frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} \le \left(\frac{y_{j,i+n+1}}{y_{j,1}}\right)^{l-1} \frac{\binom{n}{l-1}}{\binom{i+n}{l-1}} = \tag{4.4}$$

$$= (n+1)^{l-1} \frac{n(n-1)\ldots(n-l+2)}{(i+n)(i+n-1)\ldots(i+n-l+2)} =$$

$$= (n+1)^{l-1} \prod_{h=0}^{l-2} \left(1 - \frac{i}{i+n-h}\right) \le \left[2(n+1)\right]^{l-1}.$$

b) In such case

$$\left(\frac{y_{j,i}}{y_{j,i+n}}\right)^{l-1} \le \frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} \le \left(\frac{y_{j,i+n+1}}{y_{j,i+1}}\right)^{l-1}$$

and then

$$0 \le \left(1 - \frac{n}{i+n}\right)^{l-1} \le \frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} \le (n+1)^{l-1} < \left[2(n+1)\right]^{l-1}. \tag{4.5}$$

c) Finally when $i = N_j - n + 1, \ldots, N_j - 1$, taking into account that $y_{j,N_j+1} = \ldots = y_{j,N_j+n+1} = n + 1$, there results:

$$\left(\frac{N_j - n + 1}{2^j (n+1)}\right)^{l-1} \leq \frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} \leq \left(\frac{2^j (n+1)}{N_j - n + 2}\right)^{l-1};$$

therefore recalling the definition of N_j,

$$0 \leq \frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} < [2(n+1)]^{l-1}. \tag{4.6}$$

By denoting $\frac{\eta_{ji}^{(l)}}{\overline{\xi}_{ji}^{(l)}} = C_j(i,l)$, from (4.4), (4.5), (4.6) we get the thesis. □

Now we show, in Fig. 2, the behaviour of the operator $Q_j f(x)$, given by (1.5), for some values j, for the functions $f(x) = sin(2\pi x)$, and $f(x) = x^4 + |x| * x$. The operator $Q_j f(x)$ has been constructed by starting from $\{\Phi_0\}$ with $n = 5$ and $h = 6$.

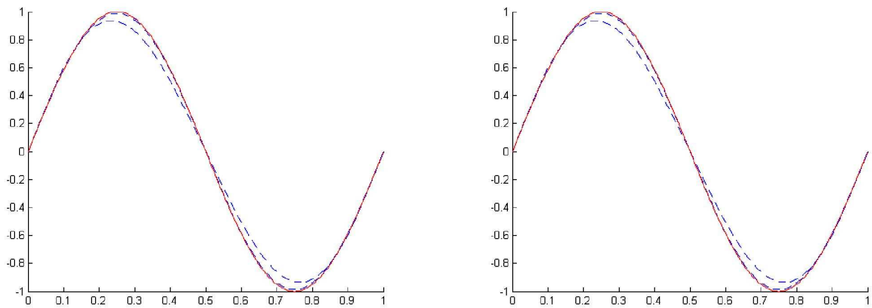


FIGURE 2. The operator $Q_j f(x)$ (dashed line), for some j, for the functions $f(x) = sin(2\pi x)$ (left), and $f(x) = x^4 + |x| * x$ (right).

In the following table 2 we report the infinite norm of the error $|Q_j f - f|$ in $[0,1]$ where $f(x) = sin(2\pi x)$ and $Q_j f$ is constructed by using refinable functions or B-splines of order n+1 in $errQf$ and $errQf\_b$ respectively.

The table 3 is relative to the function $f(x) = x^4 + |x| * x$ $x \in [-1,1]$.

TABLE 2

$f(x) = sin(2\pi x)$		
j	$errQf$	$errQf\_b$
3	9.58 (-2)	9.90 (-2)
4	1.25 (-2)	1.34 (-2)
5	8.32 (-4)	8.97 (-4)
6	5.28 (-5)	5.70 (-5)
7	3.32 (-6)	3.58 (-6)
8	2.07 (-7)	2.24 (-7)
9	1.30 (-8)	1.40 (-8)

TABLE 3

$f(x) = x^4 + \|x\| * x$		
j	$errQf$	$errQf\_b$
3	1.17(-2)	1.22(-2)
4	2.31(-3)	2.36(-3)
5	5.37(-4)	5.48(-4)
6	1.32(-4)	1.34(-4)
7	3.28(-5)	3.34(-5)
8	8.20(-6)	8.35(-6)
9	2.05(-6)	2.09(-6)

References

[1] Carnicer, J. M., Peña, J. M., *Total positivity and optimal bases*, in Total positivity and its Applications (M. Gasca and C.A. Micchelli eds.), Kluwer Academic Publishers, 1996, pp. 133-155.

[2] Gori, L., Pitolli, F., *Multiresolution analysis based on certain compactly supported functions*, in "Proceedings of ICAOR", Cluj-Napoca, Romania, 1996.

[3] Gori, L., Pitolli, F., *Refinable functions and positive operators*, in "Proceedings of MASCOT 01" Rome, 2001.

[4] Gori, L., Pitolli, F., *On some applications of a class of totally positive bases*, in "Proceedings of International Conference on Wavelet Analysis and its Applications", Guangzhou, China, 1999.

[5] Gori, L., Pitolli, F., Santi, E., *Positive operators based on scaling functions*, in "Proceedings of International Conference OFEA", St. Petersburg, Russia, 2001.

[6] Gori, L., Santi, E., *Refinable quasi-interpolatory operators*, in "Proceedings of International Conference on Constructive Theory of Functions", Varna,2002 (B. Bojanov, Ed.), Darba, Sofia, 2003, pp. 288-294.

[7] Gori, L., Santi, E., *On the convergence of refinable quasi-interpolatory operators*, submitted.

[8] Schumaker, L. L., *Spline Functions: Basic Theory* (Jon Wiley and Sons, New York, 1981).

Dip. Ingegneria Meccanica Energetica e Gestionale
Università dell'Aquila, P.zza E. Pontieri, 2
Località Monteluco di Roio, 67040 L'Aquila, Italy
E-mail address: esanti@ing.univaq.it

E-mail address: enza.pellegrino@ing.univaq.it

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **L**, Number 2, June 2005

THE FINITE-VOLUME PARTICLE METHOD ON A MOVING DOMAIN

J. STRUCKMEIER AND DELIA TELEAGA

Abstract. In the present work we apply the Finite-Volume particle method (FVPM) to a test problem posed on a moving geometry. The FVPM is a relatively new meshless method for discretizing conservation laws, which combines the generic features of a Finite-Volume scheme and a particle method. After a brief derivation of the method, a special Ansatz for the movement of the particles is proposed. Finally we present numerical results obtained for the test problem using FVPM.

1. Introduction

The Finite-Volume Particle Method (FVPM) is a relatively new meshless method for solving hyperbolic systems of conservation laws. The motivation for developing a new scheme was to unify advantages of particles methods and Finite-Volume methods (FVM) in one scheme. The FVPM combines the generic features of a Finite-Volume scheme and a particle method, namely the concept of a numerical flux function and the flow description using moving particles. This method was studied in detail in [1 - 9].

Here we shortly present the application of the FVPM to a test problem posed on a moving domain, a problem which was discussed in detail in [8].

Received by the editors: 10.10.2004.

2000 *Mathematics Subject Classification.* 65M99, 76M12, 76M25, 76M28, 76N99.

Key words and phrases. Finite-Volume methods, non-Lagrangian particles, moving domain.

2. Derivation of the method

We consider conservation laws written in the form

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{\mathcal{F}}(\boldsymbol{u}) = \boldsymbol{0}, \quad \forall \boldsymbol{x} \in \Omega(t) \subset \mathbb{R}^2,\, t > 0 \tag{1}$$

with initial conditions $\boldsymbol{u}(\boldsymbol{x},0) = \boldsymbol{u}_0(\boldsymbol{x})$, $\forall \boldsymbol{x} \in \Omega(t)$, and suitable boundary conditions, where $\Omega(t)$ is a moving, bounded domain in $\mathbb{R}^2$, $\boldsymbol{u}(\boldsymbol{x},t) \in \mathbb{R}^m$, $m > 0$ denotes the vector of conservative quantities, and $\boldsymbol{\mathcal{F}}(\boldsymbol{u}(\boldsymbol{x},t))$ denotes the flux function of the conservation law.

A natural approach to discretize conservation laws is to evaluate the weak formulation of (1) with a discrete set of test functions ψ_i, $i = 1, ..., N$. In classical FVM, the test functions are taken as the characteristic functions $I_{\Omega_i}(x)$ of the control volumes Ω_i on a spatial grid. The discrete quantities are obtained from cell averages. Note that characteristic functions form a partition of unity, i.e. $\sum_{i=1}^{N} I_{\Omega_i}(\boldsymbol{x}) = 1$, $\forall \boldsymbol{x} \in \Omega$.

A similar approach is used in the following, but we introduce a different set of test functions. Since we want to derive a mesh-free method, we should not make use of a mesh. Therefore the conservative variables are approximated at each time step by a finite set of particles located in the spatial domain $\Omega(t)$. From this point of view, the FVPM is a particle method with particle positions $\boldsymbol{x}_i(t)$, which may be irregularly spaced and moving. To each position $\boldsymbol{x}_i(t)$ we associate a function $\psi_i(\boldsymbol{x},t)$ - the particle. As in the Finite-Volume approach, let $\{\psi_i : i = 1, ..., N\}$ be a partition of unity, but the supports of the functions should overlap. More exactly, we assume that the particles are smooth functions localized around the particle positions $\boldsymbol{x}_i(t)$ and satisfy

$$\sum_{i=1}^{N} \psi_i(\boldsymbol{x},t) = 1,\, \forall \boldsymbol{x} \in \Omega(t),\, t \geq 0. \tag{2}$$

We construct this partition of unity in the following way:

Taking a Lipschitz continuous function $W : \mathbb{R} \to \mathbb{R}_+$ with compact support (otherwise one has to consider long-range interactions between particles), we define

$$\psi_i(\boldsymbol{x},t) = \frac{W_i(\boldsymbol{x},t)}{\sigma(\boldsymbol{x},t)}, \tag{3}$$

where $\sigma(\boldsymbol{x},t) = \sum_{i=1}^{N} W_i(\boldsymbol{x},t)$, $W_i(\boldsymbol{x},t) = W(\boldsymbol{x}-\boldsymbol{x}_i(t))$, $i = 1,...,N$. Such a partition of unity used in FVPM is shown in Figure 1.

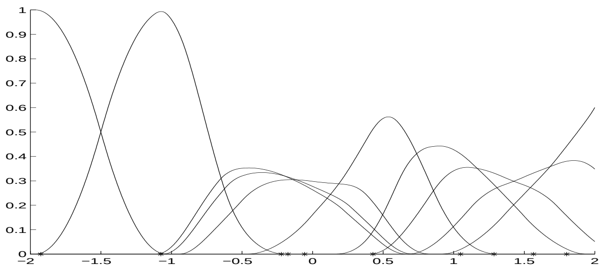


FIGURE 1. A partition of unity used in FVPM

In the FVPM, the particles generically move through the domain, following a prescribed velocity field $\boldsymbol{a}(\boldsymbol{x},t) \in C^0(C^1(\mathbb{R}^2),\mathbb{R}_+)$, i.e. we have $\dot{\boldsymbol{x}}_i = \boldsymbol{a}(\boldsymbol{x}_i,t)$. For $\boldsymbol{a} = \boldsymbol{0}$, one obtains fixed particles, and for $\boldsymbol{a}$ being, for example, the fluid velocity in the case of Euler's equations, one obtains a Lagrangian scheme.

To each particle, one associates a volume $V_i(t)$ and a discrete quantity $\boldsymbol{u}_i(t)$ which is the integral mean value with respect to the test function

$$\boldsymbol{u}_i(t) = \frac{1}{V_i(t)} \int_\Omega \boldsymbol{u}(\boldsymbol{x},t)\psi_i(\boldsymbol{x},t)d\boldsymbol{x}, \quad \text{where} \quad V_i(t) = \int_\Omega \psi_i(\boldsymbol{x},t)d\boldsymbol{x}. \tag{4}$$

Testing the conservation law (1) against the new set of test functions $\psi_i(\boldsymbol{x},t)$ and using the quantities defined above, one ends up with a system of ordinary differential equations (see [8] for details)

$$\frac{d}{dt}(V_i\boldsymbol{u}_i) = -\sum_{j=1}^{N} |\boldsymbol{\beta}_{ij}|\boldsymbol{g}_{ij} - \int_{\partial\Omega} \psi_i(\boldsymbol{\mathcal{F}}(\boldsymbol{u}) - \boldsymbol{u}\cdot\boldsymbol{b})\cdot\boldsymbol{n}d\sigma, \tag{5}$$

with the initial condition

$$\boldsymbol{u}_i(0) = \frac{1}{V_i(0)} \int_\Omega \boldsymbol{u}_0(\boldsymbol{x})\psi_i(\boldsymbol{x},0)d\boldsymbol{x}. \tag{6}$$

Note that the boundary term appears only for particles i which are near the boundary, i.e. $supp\,\psi_i \cap \partial\Omega \neq 0$, and consists of a term containing the flux of the given conservation law, as well of a contribution due to the moving boundary with the velocity $\boldsymbol{b}$.

The coefficients $\boldsymbol{\beta}_{ij}$ and $\boldsymbol{g}_{ij}$ are defined as

$$\boldsymbol{\beta}_{ij}(t) = \boldsymbol{\gamma}_{ij}(t) - \boldsymbol{\gamma}_{ji}(t), \quad \boldsymbol{\gamma}_{ij}(t) = \int_{\Omega(t)} \frac{\psi_i}{\sigma} \nabla W_j d\boldsymbol{x} \tag{7}$$

$$\boldsymbol{g}_{ij}(t) = \boldsymbol{g}(t, \boldsymbol{x}_i, \boldsymbol{u}_i, \boldsymbol{x}_j, \boldsymbol{u}_j, \boldsymbol{n}_{ij}), \quad \boldsymbol{n}_{ij} = \frac{\boldsymbol{\beta}_{ij}}{|\boldsymbol{\beta}_{ij}|} \tag{8}$$

where $\boldsymbol{g}$ is a numerical flux function consistent with the modified flux $\boldsymbol{\mathcal{G}}(t, \dot{\boldsymbol{x}}, \boldsymbol{u})$:

$$\boldsymbol{\mathcal{G}}(t, \dot{\boldsymbol{x}}, \boldsymbol{u}) = \boldsymbol{\mathcal{F}}(\boldsymbol{u}) - \boldsymbol{u} \cdot \dot{\boldsymbol{x}}, \tag{9}$$

where $\dot{\boldsymbol{x}}$ is the particle movement given by $\dot{\boldsymbol{x}}(t) = \boldsymbol{a}(\boldsymbol{x}, t)$. The numerical flux function $\boldsymbol{g}$ can be any numerical flux function used in FVM, but it has to be consistent with the modified flux function $\boldsymbol{\mathcal{G}}$, not with $\boldsymbol{\mathcal{F}}$.

Using an explicit Euler discretization of the time derivative, one obtains

$$V_i^{n+1} \boldsymbol{u}_i^{n+1} = V_i^n \boldsymbol{u}_i^n - \Delta t \sum_{j \in N(i)} |\boldsymbol{\beta}_{ij}^n| \boldsymbol{g}_{ij}^n - \boldsymbol{\mathcal{B}}_i, \tag{10}$$

with $\boldsymbol{u}_i^0 = \frac{1}{V_i^0} \int_\Omega \boldsymbol{u}_0(\boldsymbol{x}) \psi_i(\boldsymbol{x}, 0) d\boldsymbol{x}$, where $\boldsymbol{\mathcal{B}}_i$ is a discretization of the boundary term explained in [8].

A natural reconstruction of a function from the discrete values is given by

$$\tilde{\boldsymbol{u}}(\boldsymbol{x}, t) = \sum_{i=1}^{N} \boldsymbol{u}_i^n \psi_i(\boldsymbol{x}, t) I_{[t_n, t_{n+1})}(t), \quad \boldsymbol{x} \in \Omega, \ t \in [0, T]. \tag{11}$$

3. Special Ansatz for the particles movement

We concentrate on simulating a flow around an oscillating circle in a spatial two-dimensional geometry. The computational domain is given by $\Omega(t) = [0,1] \times [0,1] \setminus B_R(t)$, where $B_R(t) = \{(x,y) \in \mathbb{R}^2 : \|x - x_c(t), y - y_c(t)\| \leq R\}$ is the circle of center $(x_c(t), y_c(t))$ and radius R. Let us denote the domain's boundary by $\partial\Omega(t) := \Gamma_0 \cup \Gamma_R(t)$, where Γ_0 is the exterior boundary and $\Gamma_R(t)$ is the boundary of the moving circle.

We consider a simple, rigid movement of the circle, although one may consider another types of movements. In our example the circle is oscillating up and down, for example with respect to the following equations:

$$\dot{x}_c(t) = 0, \quad x_c(0) = x_c^0 \tag{12}$$

$$\dot{y}_c(t) = A\omega\cos(\omega t), \quad y_c(0) = y_c^0, \tag{13}$$

where A is *the amplitude of the motion* and ω is *the frequency.*

For the fluid-structure interaction problem which is considered here the effects due to viscosity can be neglected. Hence, the fluid is modeled by Euler's equations for compressible inviscid flow.

In formula (5) there are incorporated two movements: $\boldsymbol{a}$, the movement of the particles (through the numerical flux function $\boldsymbol{g}$), and $\boldsymbol{b}$, the movement of the boundary. Now we have to answer the question: being given the velocity field $\boldsymbol{b}$, how should the particles move?

One may observe that it is not suitable to move the particles with the flow velocity if a smoothly varying particle distribution is desired. Therefore we consider that the movement of the particles $\boldsymbol{a}$ is given by the solution of a Laplace equation with corresponding boundary conditions, namely *zero* velocity at the exterior boundaries and velocity of the circle at the interior boundary:

$$\begin{cases} \Delta\boldsymbol{a}(\boldsymbol{x},t) = \boldsymbol{0}, & \Omega(t) \\ \boldsymbol{a}(\boldsymbol{x},t) = \boldsymbol{0}, & \Gamma_0(t) \\ \boldsymbol{a}(\boldsymbol{x},t) = (\dot{x}_c(t), \dot{y}_c(t)), & \Gamma_R(t) \end{cases} \tag{14}$$

In this way the particles follow the domain geometry. In this example, since the movement of the boundary is restricted to a rigid body movement of an isolated object, the whole distribution of particles may be moved with the boundary. In this way the particles remain rigid, i.e. there is no relative motion between the particles. The advantage of this rigid movement is clear, we do not have to recompute every time the coefficients $\boldsymbol{\beta}_{ij}$ for example. However, the rigid movement approach is less general than the one proposed here.

In [8] we also investigated under which conditions on the motion of the circle and the smoothing length of the particles no 'holes' are developed in the domain. By a 'hole' we understand a space which is not covered by the support of any particle.

4. Numerical results

Here we present numerical results concerning the test problem defined in the previous section.

If the circle moves periodically up and down, like specified in (12), (13), there exists a periodic solution, i.e. after a few oscillations up and down the flow becomes periodic, with the same period as the circle's movement. To see this, we compute the difference between the solution every time when the circle attains its initial position, moving upwards, i.e. exactly after a complete period:

$$e_k = \sum_{i \in N} \left| \rho_i^k V_i^k - \rho_i^{k+1} V_i^{k+1} \right|, \quad k = 0, 1, \ldots, k_{max},$$

where $k_{max} = [T/P]$, $P = 2\pi/\omega$ is the period of the movement, T is the final time, t_0 is the time when the circle starts to move, $\rho_i^k = \rho_i(t_0 + kP)$ and $V_i^k = V_i(t_0 + kP)$. For this computation we choose $N = 50 \times 50$ uniform distributed particles, $t_0 = 0$, $\omega = 10\pi$, $A = 0.1$, $P = 2\pi/\omega = 0.2$, and $T = 4.05$. Hence, $k_{max} = 20$. As can be seen in Figure 2, after around 10 complete oscillations, the differences e_k are so small that the flow can be considered to be periodic.

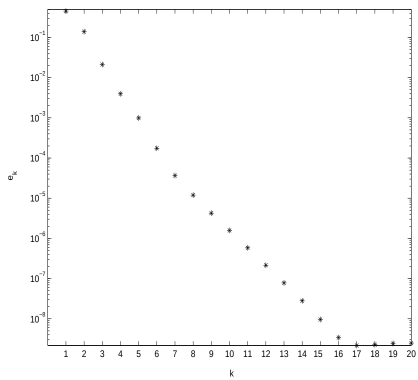


FIGURE 2. Differences e_k versus k

Now we choose $N = 100 \times 100$ quasi-random distributed and moving particles. The movement of the circle is as before, i.e. $A = 0.1$ and $\omega = 10\pi$. The solution at time $T = 0.55$ is presented in Figure 3 and 4. In Figure 3(left) one may see the irregular particle positions together with their corresponding density. The solution reconstructed on a uniform grid is shown in Figure 3(right) (isolines of the density) and Figure 4 (isolines of the velocity components).

These results show that the method works also in the case of a time-dependent domain and using irregular distributed and moving particles.

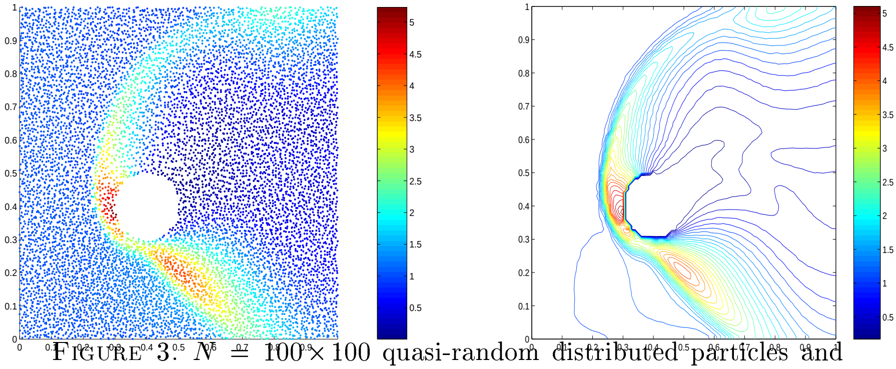


FIGURE 3. $N = 100\times 100$ quasi-random distributed particles and their corresponding density (left) and isolines of the density (right)

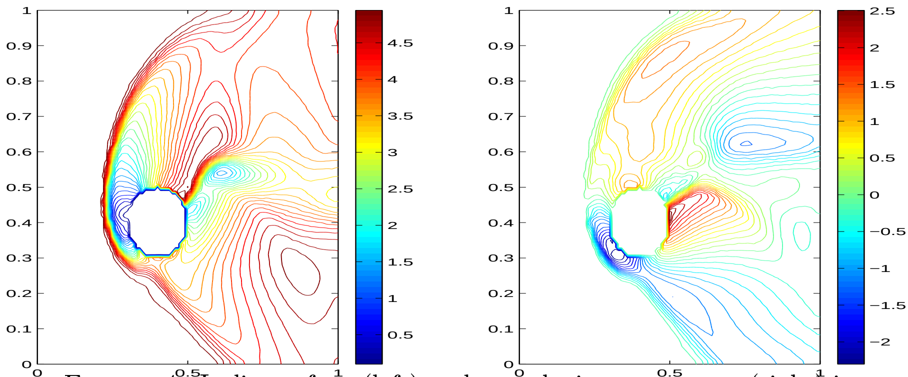


FIGURE 4. Isolines of u - (left) and v - velocity component (right) in the same case as in Figure 3

5. Conclusions

We presented here an application of the FVPM to a spatial two-dimensional problem posed on a moving domain, where the meshless character of the method is fully exploited. The particles are irregularly distributed in the domain and they are moving in a non-Lagrangian way such that they smoothly follow the time-dependent computational domain.

Numerical results indicate that the method is well-suited for such problems. Also the discretization of the boundary conditions works very satisfactory.

Thus, a first step to applying the FVPM to real fluid-structure interaction problems, which in general limit the use of grid-based methods, is done.

References

[1] Hietel, D., Steiner, K., Struckmeier, J., *A Finite-Volume Particle Method for Compressible Flows*, Math. Models Methods Appl., 10(9), 1363-1382, 2000.

[2] Junk, M., *Do finite volume methods need a mesh?*, Intern. Workshop on meshfree methods for PDE, Bonn, 2001.

[3] Junk, M., Struckmeier, J., *Consistency analysis of mesh-free methods for conservation laws*, GAMM-Mitteilungen 2, 96-126, 2001.

[4] Keck, R., *The Finite Volume Particle Method: A Meshless Projection Method for Incompressible Flow*, PhD Thesis, Univ. Kaiserslautern, Shaker Verlag, 2002.

[5] Lamichhane, B. H., *The Applications of Finite-Volume Particle Method for Moving Boundary*, Master Thesis, Univ. Kaiserslautern, 2001.

[6] Schick, C., *Adaptivity for Particle Methods in Fluid Dynamics*, Diploma Thesis, Univ. Kaiserslautern, 2000.

[7] Teleaga, D., *Numerical Studies of a Finite-Volume Particle Method for Conservation Laws*, Master Thesis, Univ. Kaiserslautern, 2000.

[8] Teleaga, D., *A Finite-Volume Particle Method for Conservation Laws*, PhD Thesis, Univ. Hamburg, 2004.

[9] Yang, Z., *Efficient Calculation of Geometric Parameters in Finite Volume Particle Method*, Master Thesis, Univ. Kaiserslautern, 2001.

University of Hamburg, Department of Mathematics, Germany,
TU Darmstadt, Department of Mathematics, Germany

BOOK REVIEWS

Titu Andreescu and Bogdan Enescu, ***Mathematical Olympiad Treasures***, Birkhäuser Verlag, Boston-Basel-Berlin, 2004, 234 pp., ISBN 0-8176-4305-2.

This excellent book deals with some important topics of elementary mathematics necessarily in the process of training students for various contests and olympiads. One of the main intention of the authors is to build a bridge between ordinary high school exercises and more sophisticated, intricate and abstract concepts and problems in undergraduate mathematics. Mathematical Olympiad Treasures reflects the depth of experience of two seasoned professors and coaches from the USA and Romanian Olympiad teams. The book is organized into three chapters each of them containing eight sections. Each sections contains some suggestive completely solved problems and some proposed problems which are solved in the second part of the book. In what follows we will briefly present each of them.

Chapter 1 is entitled *Algebra* and contains some basic notions and results concerning the following topics: An algebraic identity, Cauchy-Schwartz revisited, Easy ways through absolute values, Parameters, Take the conjugate, Inequalities with convex functions, Induction at work, Roots and coefficients.

Chapter 2, *Geometry and Trigonometry*, contains eight sections dealing with Geometric inequalities, An interesting locus, Cyclic quads, Equiangular polygons, More on equilateral triangles, The carpets theorem, Quadrilaterals with an inscribed circle, Dr. Trig learns complex numbers.

In Chapter 3, *Number Theory and Combinatorics*, the authors present some fundamental ideas and interesting problems concerning Arrays of numbers, Functions defined on sets of points, Count twice, Sequences of integers, Equations with infinitely many solutions, Equations with no solutions, Powers of 2, Progressions.

The book ends with a useful Glossary, an Index of notation and an Index containing the sources of the problems. The book is written in a very clear and rigorous manner and it is recommended for students, graduated students and their teachers and for anyone interested in mathematical contests and olympiads.

Dorin Andrica

Titu Andreescu and Zuming Feng, ***A Path to Combinatorics for Undergraduates. Counting Strategies***, Birkhäuser Verlag, Boston-Basel-Berlin, 2004, 228 pp., ISBN 0-8176-4288-9.

This is a unique approach to combinatorics centered around challenging examples, fully-worked solutions, and hundreds of problems many from mathematical contests and olympiads. This excellent book deals with some important topics of combinatorics which are very useful in the process of training students for various contests and olympiads and it reflects the depth of experience of two seasoned professors and coaches from the USA Olympiad team. The book is organized into nine chapters containing the basic notions and results in the following topics : Addition or multiplication, Combinations, Properties of binomial coefficients, Bijections, Recursions, Inclusion and exclusion, Calculating in two ways: Fubinis principle, generating functions, review exercises. The book ends with a useful Glossary, an Index of notions and a suggestive list of references. The book is written in a very clear and rigorous manner and it is recommended for students, graduated students and their teachers and for anyone interested in challenging mathematics. It can be used as a solid stepping stone for other advanced mathematical readings.

Dorin Andrica

Titu Andreescu and Zuming Feng, ***103 Trigonometry Problems From the Training of the USA IMO Team***, Birkhäuser Verlag, Boston-Basel-Berlin, 2005, 214 pp., ISBN 0-8176-4334-6.

This excellent book contains 103 highly selected problems used in the training and testing of the USA IMO (International Mathematical Olympiad) team. From the authors preface it follows that "It is not a collection of very difficult, impenetrable questions. Instead, the book gradually builds students trigonometric skills and techniques". The book contains five chapters. The first chapter provides a comprehensive introduction to trigonometric functions, their relations and functional properties, and their applications into plane and solid geometry. Chapters two and three contain 52 introductory and 51 advanced proposed problems, respectively. In Chapters four and five the authors present the solutions and some comments to the proposed problems. The book also contains a Glossary and a rich list of references.

This work aims to broaden students view of mathematics and better prepare them for possible participation in various mathematical contests and olympiads. The book further stimulates interest for the future study of mathematics.

Dorin Andrica

Vasile Berinde, ***Exploring, Investigating and Discovering in Mathematics***, Birkhäuser Verlag, Basel-Boston-Berlin, 2004, 246 +xix pp., ISBN 3-7643-7019-x.

This book represent the English version of the Romanian edition (V.Berinde, Explorare, investigare si descoperire in matematică, Editura Efemeride, Baia Mare, 2001). The author writes in his Preface to the book: "The book is addressed mainly to students, young mathematicians, and teachers, involved or/and actively working in mathematics competitions and training gifted people. It collects many valuable techniques for solving various classes of difficult problems and, simultaneously, offers a comprehensive introduction to creating new problems. The book should also be of interest to anybody who is in any way connected to mathematics or interested in the creative process and in mathematics as a art". Indeed the author has been greatly successful. The reader can find here ideas and problems which combine a number of classical topics from various fields of mathematics.

The book is organized into 24 chapters, most of them independent, involving the following topics: Chase problems, Sequences of integers simultaneously prime, A geometric construction using ruler and compass, Solving a class of nonlinear systems, A class of homogeneous inequalities, The first decimal of some irrational numbers, Polynomial approximation of continuous functions, On an interesting divisibility problem, Determinants with alternate entries, Solving some cyclic systems, On a property of recurrent affine sequences, Binomial characterizations of arithmetic progressions, Using duality in studying homographic recurrences, Exponential equations having exactly two solutions, A class of functional equations, An extension of Leibniz-Newton formula, A measurement problem, A class of discontinuous functions admitting primitives, On two classes of inequalities, Another problem of geometric construction, How can we discover new problems by means of the computer, On the convergence of some sequences of real numbers, An applications of the integral mean, Difference and differential equations.

Each chapter ends with a suggestive and useful bibliography concerning the topic. Some basic and general principles regarding creativity in solving problems are discussed in an Addendum at the end of the book. The book is strongly recommended to all students and teachers but also to everyone who has a special love for mathematical problems that are stated and also solved in a simple and in an ingenious way.

Dorin Andrica

Advanced Courses in Mathematical Analysis, I. A. Aizpuru-Tomas and F. Leon-Saavedra (Editors), World Scientific Publishers, London-Singapore 2004, vii+155 pp., ISBN 981-256-060-2.

The volume contains the written versions of the lectures delivered at the First International Course of Mathematical Analysis in Andalucia, organized by the University of Cadiz from 23 to 27 September, 2005. The aim of the course was to bring together different research groups working in mathematical analysis and to provide the young researchers of these groups with access to the most advanced lines of research. A second course took place in September 2004 in Granada.

There are included five survey papers: 1. Y. Benyamini, *Introduction to uniform classification of Banach spaces* ; 2. M. Gonzáles, *An introduction to local duality for Banach spaces* ; 3. V. Müller, *Orbits of operators* ; 4. E. Matoušková, S. Reich and A. J. Zaslavski, *Genericity in nonexpansive mapping theory* ; 5. A. R. Palacios, *Absolute-valued algebras, and absolute-valuable Banach spaces.*

The first paper is the only updated survey on the classification of Banach spaces under uniformly continuous mappings. Its aim is to introduce the reader to this area and to present some results and open questions, a complete presentation of these problems and of other related ones being given in the recent treatise of Y. Beniaminy and J. Lindenstrauss, *Nonlinear Geometric Functional Analysis,* I., AMS, 2000.

The local duality for Banach spaces is a tool recently developed by the author of the second paper and some co-workers, which turned to be very useful in the study Banach spaces, mainly in the case when the dual of a Banach space is too large.

V. Müller emphasizes in the third paper the relevance of the orbit method and of Scott Brown's technique in the study of invariant subspaces. A more comprehensive treatment is given in his recent book ????.

It is known that nonexpansive mappings could note have fixed points, but, as it was shown by F. S. De Blasi and J. Myjak in 1976, most of them (in the sense of Baire category) do have. The fourth paper surveys various category and porosity results concerning the well-posedness of the fixed point problem for nonexpansive mappings, most of them being obtained recently by the authors.

An absolute-valued algebra is a normed algebra A such that $\|xy\| = \|x\|\|y\|$, for all $x, y \in A$. As it is well-known, if A is associative and commutative then it agrees with $\mathbb{R}$ or $\mathbb{C}$, and with the quaternion field $\mathbb{H}$ if A is only associative. Therefore, the interesting case is that of non-associative absolute-valued algebras, presented in the last paper of the book. The results are presented from historical perspective to the frontier of current research in the field.

The book contains surveys of some topics of interest in the current research in functional analysis, written by leading experts in the area. It can be used as an introductory material for young researchers, as a guide to more advanced books or research papers.

S. Cobzaş

Michael Růžička, ***Nicthlineare Funktionalanalysis – Einr Einführung***, Springer Verlag, Berlin-Heidelberg-New York 2004, xii+208 pp., ISBN 3-540-20066-5.

The book is based on a one-semester course (the 6th semester) taught by the author at the Universities of Bonn (1999) and Freiburg (2002 and 2003). Its aim is to provide the reader with the basic results and techniques in the field, which can form a basis for the reading of more advanced books, as, e.g, the monumental four volume treatise of E. Zeidler, *Nonlinear Functional Analysis and Applications*, Springer Verlag, 1985-1990.

The topics covered by the present volume are best illustrated by the headings of the chapters: 1. *Fixed point theorems*; 2. *Integration and differentiation in Banach spaces*; 3. *The theory of monotone operators*; 4. *The topological degree theory.*

The first chapter of the book contains the basic fixed point theorems: Banach's contraction principle, Brouwer and Schauder fixed point theorems. The proof of Brouwer fixed point theorem uses some techniques from variational calculus. Applications are given to existence results for ordinary differential equations (Picard's iterative method).

The second chapter is concerned with Bochner integration, L^p-spaces of Banach valued functions and differential calculus in Banach spaces (Gâteaux and Fréchet derivatives).

The third chapter contains an introduction to monotone operators and includes results of Browder, Minty and Brezis. Maximal monotone operators, subdifferentials of convex functions, and the duality mapping are also included. Applications are given to quasilinear parabolic and elliptic partial differential equations.

The last chapter of the book is devoted to a presentation of Brouwer and Leray-Schauder topological degree theories wit applications to Brouwer fixed point theorem and to quasilinear elliptic equations.

The prerequisites from topology, measure theory and linear functional analysis, needed for the reading of the book are included in an Appendix (37 pp.). There are no problems and exercises in the book.

The book is clearly written, with complete and carefully written proofs and illuminating examples. It can serve as a base text for introductory courses in nonlinear functional analysis.

S. Cobzaş

Kehe Zhu, ***Spaces of Holomorphic Functions in the Unit Ball***, Graduate Texts in Mathematics, Vol. 226, Springer, Berlin-Heidelberg-New York, 2005, x+271 pp., ISBN 0-387-22036-4.

The book is concerned with the basic properties of the most well-known and widely used spaces of holomorphic functions in the open unit ball $\mathbb{B}_n$ of $\mathbb{C}^n$. The restriction to the unit ball of $\mathbb{C}^n$ allows the author to present direct proofs of most of the results by straightforward formulas. The central idea of these proofs is to use integral representations for holomorphic functions and elementary properties of the Bergman kernel, the Bergman metric, and the automorphism group. In this way, although few of the results are new, most of the proofs are new and simpler than the existing ones. On the other hand, this reduces the prerequisites to a minimum: only familiarity with single variable complex analysis, no prior knowledge of several complex variables theory being required.

The first chapter of the book, Ch.1, *Preliminaries*, has an introductory character and contains some results on holomorphic functions, the authomorphism group of $\mathbb{B}_n$, Lebesgue spaces, Bergman metric, subharmonic functions, complex interpolation.

Each of the remaining chapters is devoted to a class of spaces of holomorphic functions in the unit ball: Ch. 2. *Bergman spaces*; Ch. 3. *The Bloch space*; Ch. 4. *Hardy spaces*; Ch. 5. *Functions of bounded mean oscillation* (the study of BMOA spaces); Ch. 6. *Besov spaces*; Ch. 7. *Lipschitz spaces.* For each class of spaces, the author discusses integral representations, characterizations in terms of various derivatives (radial derivatives, holomorphic gradients, fractional derivatives), atomic decompositions, complex interpolation and duality.

All these spaces are intimately related as it is emphasized in the book: the Bloch space $\mathcal{B}$ can be thought as a limit case of Bergman space A^p_α as $p \to \infty$. In particular, $\mathcal{B}$ can be naturally identified with the dual of the Hardy space H^1. The Besov space B_p is the image of the Bergman space A^p_α under a suitable fractional integral operator, and B_∞ is just the Bloch space. In their turn Lipschitz spaces Λ_α are images of Bloch spaces under some fractional integral operator. In fact, as the author points out in the Preface to the book, all these spaces are special cases of a more general family of holomorphic Sobolev spaces, but the direct treatment of these particular cases is far more interesting and appealing than a cumbersome presentation of an exhaustive class of functions containing all of them.

Each chapter ends with a set of exercises of varying difficulty. For the difficult results, completing the main text, exact references are given. The Notes sections contain bibliographical mentions and references to related results.

Of course that the title of the book automatically directs us to the classic book of W. Rudin, *Function Theory in the Unit Ball of* $\mathbb{C}^n$ (Springer 1980). Although there are some inevitable overlaps, they are not substantial and the two books complement each other.

Professor Zhu is an authoritative personality in the area and the author of the books: *Operator Theory in Function Spaces* (M. Dekker 1990), *Theory of Bergman Spaces*, with H. Hedenmalm and B. Korenblum (Springer 2000), and *An Introduction to Operator Algebras* (CRC Press 1993).

The book is well written and can be used as a textbook for advanced graduate courses in complex analysis and spaces of holomorphic functions.

Mirela Kohr