Anul XLVIII **2003**

S T U D I A
UNIVERSITATIS BABEŞ-BOLYAI

MATHEMATICA
3

Redacţia: 3400 Cluj-Napoca, str. M. Kogălniceanu nr. 1 • Telefon: 405300

SUMAR – CONTENTS – SOMMAIRE

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

PROFESSOR GHEORGHE MICULA AT HIS 60TH ANNIVERSARY

RADU PRECUP

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Professor Gheorghe Micula was born in Delureni (Bihor County), Romania, on April 23, 1943. He graduated the Faculty of Mathematics and Mechanics of the Babeş-Bolyai University in 1965, with Magna cum Laudae qualification. In 1971 he obtained the Ph. D. degree with the thesis "Contributions to numerical solutions of differential equations by means of splines", elaborated under the guidance of D. V. Ionescu. He was successively Assistant Professor (1965-1971), Lecturer (1971-1990) and Associate Professor (1990-1992) at the Chair of Differential Equations of the Faculty of Mathematics and Computer Science. Since 1992 he is full Professor at the same chair.

He married in 1965 Maria Vasile. They have one daughter, Sanda, who also graduated Faculty of Mathematics and Computer Science of Babeş-Bolyai University.

He was Fellow: A. v. Humboldt (1973-1975) - Universities of Freiburg, Mainz and Berlin; DAAD - Universities of Marburg, Berlin, Darmstadt, Siegen and Würzburg; and Fulbright-University of Kentucky Lexington (USA). Also, between 1974-2003 he was Visiting Professor at several universities, in Germany, Israel, China, New Zealand, Turkey, South Korea and Italy.

Professor Micula obtained the First Prize in Mathematics of Balkan Union of Mathematics (Athens 1973) and "S. Stoilov" Prize of Romanian Academy (1980). He is Doctor Honoris Causa of University of Oradea.

Professor Micula is the president of Cluj section of Romaniam Mathematical Society, vice-president of the Romanian section of GAMM (Germany) and member of the Amer. Math. Soc., European Mech. Soc. and European Math. Soc.

The research interests of Professor Micula go to the numerical solutions of differential, integral and partial differential equations, spline functions, numerical analysis and approximation theory, and are reflected by his 7 books and over than 80 published papers.

The impact of his research on the mathematical community is proved by the quotations of his works in papers by J. Böhmer, J. Butcher, Ju. N. Subbotin, W. Schempp, G. Meinardus, H. Brunner, N. H. Mülthei, G. Hämerlin, Ju. S. Zavjalov, G. Nürnberger, J. W. Schmidt, W. Haussmann, B. D. Bojanov, K. E. Atkinson, J. Györvary, B. Kvasov, etc.

Received by the editors: 20.10.2003.

On behalf of my colleagues in the Chair of Differential Equations, I wish Professor Micula a long life, good health and all the best for many years to come.

LIST OF PUBLICATIONS

I. Monographs, books and textbooks

1. *Spline Functions and Applications* (Romanian). Bucharest Technical Publishing House, 1978, 348 pp. (Book distinguished with "The first Prize of the Romanian Academy of Science", Bucharest 1980).
2. *Teoria funcţiilor spline şi aplicaţii.* Litografia Univ.Cluj-Napoca, 1979, 136 p.
3. *Culegere de Probleme şi Exerciţii de Ecuaţii Diferenţiale şi Integrale*, Litografia Univ. Babeş-Bolyai, Cluj-Napoca, 1980 (joint book with M. Frenkel, P. Pavel, B. Ionescu).
4. *Probleme de ecuaţii diferenţiale şi cu derivate parţiale*, Editura Didactică şi Pedagogică, Bucureţi, 1982 (joint book with I. A. Rus, P. Pavel, B. Ionescu).
5. *Theory and Applications of Spline Functions.* Freie Univerität Berlin, Preprint Nr. A 89/1. Fachbereich Mathematik, Seria A, 1989, Mathematik 249 p.
6. *Differential Equations and Integral Equations by Practical Exercises and Problems* (Romanian). Dacia Publishing House Cluj-Napoca, Romania, 1989, 290 p. (joint book with P. Pavel).
7. *Theory and Applications of Spline Functions, Part. I and Part. II* Preprint Nr. A-91-33, Freie Universität Berlin, Fachbereich Mathematik, Seria A, Mathematik, Berlin (1991), 330 p. (co-editor R. Gorenflo).
8. *Differential and Integral Equations through practical problems and exercises.* Kluwer Academic Publishers, Dordrecht, The Netherlands (1992), 408 p. (joint book with P. Pavel).
9. *Handbook of Splines,* Kluwer Academic Publishers, Dordrecht-Boston-London, 1999, 624 pp. (joint book with Sanda Micula).

II. Articles

1. *O formulă de cuadratură cu cinci noduri cu gradul de exactitate cinci.* Studia Univ. Babeş-Bolyai, Cluj, Series Math. Fasc. 2, (1967), 58-74.
2. *Sur la formule de quadrature de Tricomi.* Bull. Math. Sci. Math. de Roumanie, 12 (1968), 95-105 (joint paper with M. Micula).
3. *Optimal cubature formulas for certain classes of functions.* Anal. Şt. Univ. Al. I. Cuza, Iaşi, 16 (1970), 345-356 (joint paper with G. Coman).
4. *Optimal cubature formulae.* Rendiconti di Mathematica, 4 (1971), Nr. 6, 303-311.
5. *Approximate integration of systems of differential equations by spline functions.* Studia Univ. Babeş-Bolyai, Cluj, Series Math. Fasc. 2 (1971), 27-39.

6. *Fonctions spline d'approximation pour les solutions des systèmes d'équations différentielles.* Anal. St. Univ. Al. I. Cuza, Iaşi 27 (1971), 139-155.
7. *Contributions to the numerical solution of differential equations by spline functions.* (roumanian). Doctoral dissertation, Univ. of Cluj, Romania, 1971.
8. *Spline functions approximating the solution of nonlinear differential equation of n-th order.* ZAMM, 52 (1972), 189-190.
9. *Funcţii spline de grad superior de aproximare a soluţiilor sistemelor de ecuaţii diferenţiale.* Studia Univ. Babeş-Bolyai Cluj, Fasc. 1 (1972), 21-32.
10. *Numerical integration of differential equation* $y^{(n)} = f(x, y)$ *by spline functions.* Rev. Roum. Math. Pures et Appl. (Bucharest), 1 (1972), 1385-1390.
11. *Approximate solution of the differential equation* $y'' = f(x, y)$ *with spline functions.* Math. Comput. 27 (1973), 807-816.
12. *Sur la résolution numérique des équation integrale du type de Volterra de seconde éspèce à l'aide des fonction spline.* Studia Univ. Babeş - Bolyai Cluj, Series Mathematica, Fasc. 2 (1973), 65-68 (joint paper with M. Micula).
13. *Die numerische Lösung nichtlinearer Differentialgleichungen unter Verwendung von Spline-Funktionen.* Lect. Notes in Math. 395, Springer, 1974, 57-83.
14. *Deficient spline approximate solutions to linear differential equations of the second order.* Mathematica (Cluj), 16 (39) (1974), 65-72.
15. *The numerical solution of Volterra integro-differential equations by spline functions.* Rev. Roum. Math. Pures et Appl. (Bucharest), 20 (1975), 349-358.
16. *Über die numerische Lösung nichtlinearer Differentialgleichungen mit Splines von niedriger Ordnung.* Internat. Ser. Numer. Math. 27 (1975), 185-195. Birkhäuser Verlag, Basel-Stuttgart, 1975.
17. *The numerical solution of nonlinear differential equations by deficient spline functions.* ZAMM, 55 (1975), 254-255.
18. *Bemerkungen zur numerischen Lösung von Anfangeswertproblemen mit Hilfe nichtlinearer Spline-Funktionen.* Lect. Notes in Math. 501 (1975), 200-209. Springer, 1975.
19. *On the approximative solution of nonlinear differential equations of the second order by deficient spline functions.* Lucrări Ştiinţifice, Seria A, Inst. Ped. Oradea, (Romania), (1975), 7-11.
20. *Bemerkungen zur numerischen Behandlung von nichtlinearer Volterraschen Integralgleichungen mit Splines.* ZAMM, 56 (1976), 302-304.
21. *Numerische Behandlung der Volterra - Integralgleichungen mit Splines.* Studia Univ. Babeş-Bolyai, Cluj-Napoca, 24 (1979), Fasc. 2, 46-54.

22. *Asupra unor metode de rezolvare aproximativă a ecuaţiilor diferenţiale.* Lucrări Sem. Itinerant de Ec. Funcţ., Aproximare şi Convexitate, Cluj-Napoca, 17-19 mai 1979, 99-105.
23. *Projection method for the numerical solution of Hammerstein equations.* Proc. Seminar of Functional Eqs., Approximation and Convexity, Timişoara (Romania), 7-8 nov. 1980, 137-143.
24. *The "D. V. Ionescu method" of constructing approximation formulas.* Studia Univ. Babeş-Bolyai, Cluj-Napoca, Mathematica, 26 (1981), No 2, 5-13.
25. *On "D. V. Ionescu method" in numerical analysis as a constructing method of spline functions.* Revue Roumaine de Math. Pures et Appl. 24 (1981), No 408, 1131-1141.
26. *Numerische Behandlung von Differentialgleichungen mit modifizierten Argument mit Spline-Funktionen.* Proceedings of Colloq. Approx. and Optimization. Cluj-Napoca (Romania), October 25-27, 1984, 111-128.
27. *Spline-Funktionen und die numerische Verfahren für Differential und Partielledifferentialgleichungen.* Research Seminairies, Faculty of Math. Univ. of Cluj-Napoca, Preprint No. 9, (1985), 27-52.
28. *The spline technique in the theory of differential equations.* Proceedings of the Conference on Differential Equations. Cluj-Napoca, November 21-23, 1985, 29 p.
29. *Numerical solution of some linear elliptic problems by spline functions.* Proc. Itinerant Seminar on Functional Equations, Approximation and Convexity. Iaşi, (Romania), 26 Oct. 1986, 13-17.
30. *Nonpolynomial spline functions for the approximate solution of system of ordinary differential equations.* Research Seminars, Seminar on Numer. and Statist. Calculus, Preprint No 9, (1987), 143-153, Univ. of Cluj-Napoca (joint paper with M. Micula).
31. *A polynomial spline approximation method for solving system of ordinary differential equations.* Studia Univ. Babeş-Bolyai, Mathematica, 4 (1987), 66-60 (joint paper with T. Fawzy and Z. Ramadan).
32. *Approximate solution of the second order differential equations with deviating argument by spline functions.* Mathematica - Revue d'Analyse Numérique et de Théorie de l'Approximation, Cluj-Napoca, Tome 30 (53), No 1, (1988), 37-46 (joint paper with H. Akça).
33. *Numerical solution of differential equations with deviating argument using spline functions.* Studia Univ. Babeş-Bolyai, Mathematica, 23 (1988), No 2, 45-57 (joint paper with H. Akça).
34. *Lösung gewöhnlicher Differentialgleichungen zweiter Ordnung mit nacheilendem Argument durch Spline Funktionen.* Preprint 1181 T. H. Darmstadt, 1988, 18 p.
35. *Über die numerische Lösung gewöhnlicher Differentialgleichungen zweiter Ordnung mit nacheilendem Argument durch Spline-Funktionen.* Rev. Roumaine Math. Pures Appl. Bucharest, 34 (1989), 899-909.

36. *Numerical solution of delay differential equations of higher order by spline functions.* Babeş-Bolyai University Cluj, Research Seminars, Seminar on Differential Equations, Preprint No. 3, (1990), 77-86.
37. *Approximate solutions of the nonlinear n-th order differential equations with deviating argument by polynomial spline functions.* Proceed. Itinerant Seminar an Functional Eqs. Approximation and Convexity, Cluj-Napoca, Oct. (1990), 93-104.
38. *Approximate solution of the nonlinear n-th order differential equations with deviating argument by spline functions.* Buletin Ştiinţific, Seria B, Vol.7, Nr. 1-2, (1991), Univ. Baia Mare, 47-54.
39. *Approximate solution of the second order differential equations with deviating argument by deficient spline functions.* Bul. Şt. Inst. Politehnic Cluj-Napoca, Seria Matematică-Mecanică, Vol.34,(1991) (joint paper with H. Akça and I. Dag).
40. *Direct numerical spline methods for second order Fredholm integro-differential equations.* Studia Univ. Babeş-Bolyai, Cluj-Napoca, 37, 1 (1992), 73-85 (joint paper with M. Micula).
41. *Numerical solutions of system of differential equations with deviating argument by spline functions.* Acta Technica Napocensis, Series Applied Math. and Mech., 35 (1992), 107-116 (joint paper with H. Akça).
42. *Spline approximations for second order neutral delay differential equations.* Studia Univ. Babeş-Bolyai, Cluj-Napoca, Mathematica, 39 (1993),Fasc. 1, 87-97 (joint paper with G. Fairweather).
43. *Direct numerical spline methods for first-order Fredholm integro-differential equations.* Révue d'Analyse Numérique et de Théorie de l'Approximation, Cluj-Napoca, 22 (1993), No. 1, 59-66 (joint paper with G. Fairweather).
44. *Natural spline functions of even degree.* Studia Univ. "Babeş-Bolyai" Cluj-Napoca, Series Math. 38 (1993), Fasc. 2, 31-40 (joint paper with P. Blaga).
45. *Spline approximations for neutral delay differential equations.* Révue d'Analyse Numér et de Theorie de l'Approximation, Cluj-Napoca, 23 (1994), Nr. 2, 117-125 (joint paper with A. Bellen).
46. *Continuous approximate solution to the neutral delay differential equations by a simplified Picard's method.* Studia Univ. "Babeş-Bolyai" Cluj-Napoca, Mathematica, 39 (1994), Nr. 4, 69-78 (joint paper with H. Akça and U. Güray).
47. *On the use of spline functions of even degree for the numerical solution of the delay differential equations.* Calcolo, Vol. 32 (1995), No. 1-2, 83-101 (joint paper with P. Blaga and H. Akça).
48. *Deficient spline approximations for second order neutral delay differential equations.* Studia Babeş-Bolyai, Mathematica, 40 (1995), No. 4, 85-97 (with H. Akça and G. Arslan).

49. *On even degree polynomial spline functions with applications to numerical solution of differential equations with retarded argument.* Preprint, Technische Hochschule Darmstadt, Fachbereich Mathematik, No. 1771, (1995) (joint paper with P. Blaga, M. Micula).
50. *The numerical solution of differential equation with retarded argument by means of natural spline functions of even degree.* Internat. J. Computer Math., UK, 61 (1996), No. 1-2, 1-18 (joint paper with P. Blaga and M. Micula).
51. *The numerical treatment of delay differential equation with constant delay by natural spline functions of even degree.* Libertas Mathematica, 16 (1996), 123-131 (joint paper with P. Blaga and H. Akça).
52. *A polynomial spline approximation method for solving Volterra integro-differential equations.* Studia Univ. Babeş-Bolyai, Mathematica, 41 (1996), No.4, 71-80 (joint paper with A. Ayad).
53. *Polynomial spline approximation method for solving delay differential equations.* In "Proceedings of the Internat. Confer. of Approx. and Optimization (Romania), ICAOR, Cluj-Napoca, July 29- August 1^{st} , 1996, pp. 181-192 (joint paper with P. Blaga, Chung Seiyoung).
54. *Even degree spline technique for numerical solution of delay differential equations.* Preprint No. A-15/96, Freie Univ. Berlin, Fachbereich Math. und Inf., Serie A, Mathematik, 1996 (joint paper with P. Blaga, R. Gorenflo).
55. *Low order splines in solving neutral delay differential equations.* Studia Univ. Babeş-Bolyai, Cluj-Napoca, Mathematica, 41 (1996), No.2, 73-85 (joint paper with J. Kobza, P. Blaga).
56. *Polynomial spline functions of even degree approximating the solution of differential equations.* (I). ZAMM (1996), Suppl. 1, 477-478 (joint paper with P. Blaga).
57. *Remainder in the history of spline functions.* Révue d'Analyse Numérique et de Théorie de l'Approximation, Cluj-Napoca, 26 (1997), No. 1-2, 117-123 (joint paper with M. Micula).
58. *On the spline approximating methods for second order Volterra integro-differential equations.* Studia Univ. "Babeş-Bolyai", Mathematica, Cluj, 42 (1998), No. 1, 101-106 (joint paper with M. Micula and S. Chung).
59. *Polynomial spline functions of even degree approximating the solutions of differential equations.* Analele Univ. din Timişoara Vol. 36, fasc. 2 (1998) 171-190 (joint paper with P. Blaga).
60. *The numerical solution of first order delay Volterra integro-differential equations by spline functions.* Révue d'Analyse Numérique et de Théorie de l'Approximation, Cluj-Napoca, 27 (1998), No. 1, 117-126 (joint paper with A. Ayad).
61. *Even degree spline technique for numerical solution of delay differential equations.* Proceedings of the Annual Meeting of the Romanian Soc. of

Math. Sciences, Bucharest May 29 - June 1, 1997, Ed. Paralela 45, 1998, 29-41 (joint paper with P. Blaga and R. Gorenflo).

62. *On the linear stability of the repeated spline-collocation method for Volterra integro-differential equations.* Bull. Math. Soc. Sci. Math. Roumanie, Tome 41 (89) (1998), No. 4, 237-247 (joint paper with I. Danciu and A. Revnic).
63. *Approximate solutions of the differential equations with delay by nonpolynomial spline functions.* Bull. Math. Soc. Sci. Math. Roumanie, Tome 42 (90), (1999), No.4, 293-300. (joint paper with H. Akça).
64. *Spline Functions and Applications.* References. Preprint 99-15, Universität Stuttgart, Mathematisches Institut A, 1999, 236 pp. (joint paper with W.L.Wendland).
65. *An implicit numerical spline method for systems of ordinary differential equations.* Applied Mathematics and Computation, 111 (2000), 121-132 (joint paper with A. Revnic).
66. *The D. V. Ionescu's contributions to the numerical solution of differential equations to modern numerical analysis and to spline functions theory.* In: "Mathematical Contributions of D. V. Ionescu", Edited by Ioan A. Rus, "Babeş-Bolyai" Univ. Press, 2001, 77 - 94.
67. *The "D. V. Ionescu Method of Function φ" and Spline Functions.* Libertas Mathematica, 21 (2001), 15-26, Volume dedicated to the Centennial Birthday of D. V. Ionescu (joint paper with M. Micula).
68. *Polynomial spline functions of even degree approximating the solutions of (delay) differential equations* II. ZAMM, 81 (2001), Supl. 1, 83 -85 (joint paper with P. Blaga).
69. *Numerical solution of the delay differential equations by nonpolynomial spline functions.* Studia Univ. "Babeş-Bolyai", Informatica, 46 (2001), No. 2, 91-98 (joint paper with V. A. Căuş).
70. *On the numerical approach of Korteweg - de Vries - Burger equations by spline finite element and collocation methods.* Seminar on Fixed Point Theory Cluj-Napoca, 3 (2002), ICNODEA 2001, 261-270 (joint paper with M. Micula).
71. *A new deficient spline functions collocation methods for the second order delay differential equations.* PUMA, Budapest, 13 (2003), 97-109 (joint paper with F. Calio, E. Marchetti, R. Pavani).
72. *A variational approach to spline functions theory.* Rend. Sem. Mat. Univ. Pol. Torino, 61(2003), Nr. 1, 41-59.

III. Miscellanea

1. *Teoria funcţiilor spline şi aplicaţii. Bibliografie.* Litografia Fac. de Matematică, Univ. Babeş - Bolyai, Cluj-Napoca, 1978, 104 p.
2. *Profesorul emerit D. V. IONESCU la împlinirea vârstei de 80 de ani.* Gazeta Matematică, Bucharest, vol. 86, Nr. 6, 1981, 225-226.

3. *Theorie und Anwendungen von Spline-Funktionen.* Literaturverzeichnis. Preprint Nr. 890, März 1985. Technische Hochschule Darmstadt, B. R. D., Fachbereich Mathematik, 162 pp.
4. *Spline-Funktionen und Anwendungen.* Literaturverzeichnis. Preprint Nr. 1071, Juli 1987. Technische Hochschule Darmstadt, B. R. D. 219 p.
5. *Leonhard Euler. Life and Mathematical Creation.* Plenary Lecture. In: Proceedings of the Annual Meeting of the Romanian Soc. Sc. Math. Roumanie, Cluj-Napoca, May 27-31, 1998. Ed. Digital Data, Cluj-Napoca 1998, 17-24.
6. *O carte în sprijinul înţelegerii matematicii şi informaticii.* In: Proceedings of the Annual Meeting of the Romanian Soc. Sc. Math. Roumanie, Cluj-Napoca, May 27-31, 1998, Ed. Digital Data, Cluj-Napoca 1999, 1-5.
7. *Proceedings of the Annual Meeting of the Romanian Soc. Sc. Math. Roumanie, Cluj-Napoca, May 27-31, 1998.* Ed. Digital Data, Cluj-Napoca 1999, ISBN 973-98789-0-3 (Coeditor P. Mocanu and I. Şerb).

Faculty of Mathematics and Computer Science,
Babeş-Bolyai University, Cluj-Napoca, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume XLVIII, Number 3, September 2003

A COMPARISON OF TWO SYSTEMS DESCRIBING ELECTROMAGNETIC TWO-BODY PROBLEM

V. G. ANGELOV, AND L. GEORGIEV

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. Two systems of equations of motion describing two charged particles in the framework of classical electrodynamics are compared.

1. Introduction

The main purpose of the present paper is to compare two systems equations of motion describing the two-body problem of classical electrodynamics [1], [2] (cf. also [3], [4]). One can see at the end of the paper that even a small difference in the right-hand sides of the equations generates various solutions and completely different physical conclusions for the two-body system.

In 1940 [5] J. L. Synge proposes equations of motion describing the behaviour of two charged particles. His derivations are based on the relativistic form of the pondermotive Lorentz force given by W. Pauli [6] by means of Lienard-Wiechert retarded potentials. J. L. Synge formulates the problem in the Minkowski space, that is, in the framework of the special theory of relativity. Consequently the finite velocity of the propagation of interaction generates delays which are, although implicitly, in the arguments of the unknown velocities of the moving particles in the equations of motion [5]. This does not come as a surprise because the theory of differential equations with retarded argument is formulated about twenty years later (cf. A. D. Myshkis [7]).

In order to overcome this difficulty J. L. Synge [5] builds a sequence of successive approximations such that on every step one has to solve a system of ordinary differential equations. Although there is no a convergence theorem for the successive approximations he proposes some idea for solving of the system. In a recent paper [8], however, we have shown that not only a convergence theorem cannot be proved, but even a sequence of successive approximations could not be constructed in such a way. On the base of the same method [5] J. L. Synge calculates the energy on every step (of successive approximations) and makes a conclusion that the two-body system is not stable (cf. p.139, [5]).

Received by the editors: 10.12.2002.

Later in 1963 R. D. Driver [9] recognizes the system obtained in [5] as a functional differential system with delays depending on the unknown trajectories and obtains a correct formulation of the Synge problem even in 1-dimensional case. Since we have taken the same point of view for the type of the delays we are able to compare the systems for 3-dimensional case considered in [1] and [2]-[4].

Prior to begin the main exposition we want to discuss one more difficulty concerning Synge equations. They are 8 in number, while the unknown functions are 6 in number. The problem mentioned is not considered in [1] and related known papers. In [4] (cf. also [2]) we show that the system of equations of motion is equivalent to the one consisting of 6 equations. More precisely, the 4-th and 8-th equation are a consequence of the rest ones.

In the present paper we recall some formulation from [3] and [2] in order to obtain 3-dimensional case of J. L. Synge equations. Here we succeed to simplify the right-hand sides of the equations in more extent than [2]-[4]. We present the equations of motion from [1] using our denotations which makes the comparison of both systems easier. Thus we see that equations from [1] can be turned into our ones (we pretend they are Synge's equations) if the constant k (from [1]) is chosen to be $k = \frac{1}{c^3}$ (c the speed of light). Then it is not surprise that the right-hand sides of equations from [1] (c^3 times larger than ours) generates unstable solutions. At the same time it is shown in [2] that Kepler problem for two charged particles has a circle solution.

2. J. L. Synge's equations of motion

As in [2]-[5] we denote by $x^{(p)} = (x_1^{(p)}(t), x_2^{(p)}(t), x_3^{(p)}(t), x_4^{(p)}(t) = ict)$ $(p = 1, 2)$ $(i^2 = -1)$ the space-time coordinates of the moving particles, by m_p - their proper masses, by e_p - their charges, c - the speed of light. The coordinates of the velocity vectors are $u^{(p)} = (u_1^{(p)}(t), u_2^{(p)}(t), u_3^{(p)}(t))(p = 1, 2)$. The coordinates of the unit tangent vectors to the world-lines are (cf. [2], [3]):

$$\lambda_\alpha^{(p)} = \frac{\gamma_p u_\alpha^{(p)}(t)}{c} = \frac{u_\alpha^{(p)}(t)}{\Delta_p}(\alpha = 1, 2, 3), \lambda_4^{(p)} = i\gamma_p = \frac{ic}{\Delta_p} \tag{1}$$

where $\gamma_p = (1 - \frac{1}{c^2}\sum_{\alpha=1}^{3}[u_\alpha^{(p)}(t)]^2)^{-\frac{1}{2}}, \Delta_p = (c^2 - \sum_{\alpha=1}^{3}[u_\alpha^{(p)}(t)]^2)^{\frac{1}{2}}$. It follows $\gamma_p = c/\Delta_p$.

By $< .,. >_4$ we denote the scalar product in the Minkowski space, while by $< .,. >$ - the scalar product in 3-dimensional Euclidean subspace. The equations of motion modeling the interaction of two moving charged particles are the following (cf. [5]):

$$m_p \frac{d\lambda_r^{(p)}}{ds_p} = \frac{e_p}{c^2} F_{rn}^{(p)} \lambda_n^{(p)} (r = 1, 2, 3, 4) \tag{2}$$

where the elements of proper time are $ds_p = \frac{c}{\gamma_p}dt = \Delta_p dt (p = 1, 2)$. Recall that in (2) there is a summation in n $(n = 1, 2, 3, 4)$. The elements $F_{rn}^{(p)}$ of the electromagnetic tensors are derived by the retarded Lienard-Wiechert potentials $A_r^{(p)} = -\dfrac{e_p\lambda_r^{(p)}}{\langle\lambda^{(p)},\xi^{(pq)}\rangle_4}(r = 1,2,3,4)$, that is, $F_{rn}^{(p)} = \dfrac{\partial A_n^{(p)}}{\partial x_r^{(p)}} - \dfrac{\partial A_r^{(p)}}{\partial x_n^{(p)}}$. By $\xi^{(pq)}$ we denote the isotropic vectors (cf. [2], [5])

$$\xi^{(pq)} = (x_1^{(p)}(t)-x_1^{(q)}(t-\tau_{pq}(t)), x_2^{(p)}(t)-x_2^{(q)}(t-\tau_{pq}(t)), x_3^{(p)}(t)-x_3^{(q)}(t-\tau_{pq}(t)), ic\tau_{pq}(t)),$$

where $\langle\xi^{(p,q)},\xi^{(p,q)}\rangle_4 = 0$ or

$$\tau_{pq}(t) = \frac{1}{c}\left(\sum_{\beta=1}^{3}[x_\beta^{(p)}(t) - x_\beta^{(q)}(t-\tau_{pq}(t))]^2\right)^{\frac{1}{2}}, ((pq) = (12),(21)). \qquad (3_{pq})$$

Calculating $F_{rn}^{(p)}$ as in [5] we write equations from (2)in the form:

$$\frac{d\lambda_\alpha^{(p)}}{ds_p} = \frac{Q_p}{c^2}\left\{\frac{\xi_\alpha^{(pq)}\langle\lambda^{(p)},\lambda^{(q)}\rangle_4 - \lambda_\alpha^{(q)}\langle\lambda^{(p)},\xi^{(pq)}\rangle_4}{\langle\lambda^{(q)},\xi^{(pq)}\rangle_4^3}\left[1+\left\langle\xi^{(pq)},\frac{d\lambda^{(q)}}{ds_q}\right\rangle_4\right]+\right.$$
$$\left.+\frac{1}{\langle\lambda^{(q)},\xi^{(pq)}\rangle_4^2}\left[\langle\lambda^{(p)},\xi^{(pq)}\rangle_4\frac{d\lambda_\alpha^{(q)}}{ds_q} - \left\langle\lambda^{(p)},\frac{d\lambda^{(q)}}{ds_q}\right\rangle_4\xi_\alpha^{(pq)}\right]\right\} \quad (\alpha = 1,2,3) \qquad (4.\alpha)$$

$$\frac{d\lambda_4^{(p)}}{ds_p} = \frac{Q_p}{c^2}\left\{\frac{\xi_4^{(pq)}\langle\lambda^{(p)},\lambda^{(q)}\rangle_4 - \lambda_4^{(q)}\langle\lambda^{(p)},\xi^{(pq)}\rangle_4}{\langle\lambda^{(q)},\xi^{(pq)}\rangle_4^3}\left[1+\left\langle\xi^{(pq)},\frac{d\lambda^{(q)}}{ds_q}\right\rangle_4\right]+\right.$$
$$\left.+\frac{1}{\langle\lambda^{(q)},\xi^{(pq)}\rangle_4^2}\left[\langle\lambda^{(p)},\xi^{(pq)}\rangle_4\frac{d\lambda_4^{(q)}}{ds_q} - \left\langle\lambda^{(p)},\frac{d\lambda^{(q)}}{ds_q}\right\rangle_4\xi_4^{(pq)}\right]\right\} \qquad (4.4)$$

where $Q_p = e_1e_2/m_p(p = 1,2)$. Further on, we have $u^{(q)} \equiv u^{(q)}(t_{pq})$ $(t_{pq} = t - \tau_{pq})$,

$$\lambda^{(q)} = (\gamma_{pq}u_1^{(q)}/c, \gamma_{pq}u_2^{(q)}/c, \gamma_{pq}u_3^{(q)}/c, i\gamma_{pq}) = (u_1^{(q)}/\Delta_{pq}, u_2^{(q)}/\Delta_{pq}, u_3^{(q)}/\Delta_{pq}, ic/\Delta_{pq})$$

where $\gamma_{pq} = \left(1 - \dfrac{1}{c^2}\displaystyle\sum_{\alpha=1}^{3}[u_\alpha^{(q)}(t-\tau_{pq}(t)]^2\right)^{-\frac{1}{2}}$, $\Delta_{pq} = \left(c^2 - \displaystyle\sum_{\alpha=1}^{3}[u_\alpha^{(q)}(t-\tau_{pq}(t))]^2\right)^{\frac{1}{2}}$

and $\dfrac{d\lambda_\alpha^{(p)}}{ds_p} = \dfrac{d(\frac{\gamma_p}{c}u_\alpha^{(p)})}{\frac{c}{\gamma_p}dt} = \dfrac{d(\frac{u_\alpha^{(p)}}{\Delta_p})}{\Delta_p dt} = \dfrac{1}{\Delta_p^2}\dot{u}_\alpha^{(p)} + \dfrac{u_\alpha^{(p)}}{\Delta_p^4}\langle u^{(p)},\dot{u}^{(p)}\rangle(\alpha = 1,2,3)$

$\dfrac{d\lambda_4^{(p)}}{ds_p} = \dfrac{d(i\gamma_p)}{\frac{c}{\gamma_p}dt} = \dfrac{icd(\frac{1}{\Delta_p})}{\Delta_p dt} = \dfrac{ic}{\Delta_p^4}\langle u^{(p)},\dot{u}^{(p)}\rangle$, where the dot means a differentiation in t.

In order to calculate $\dfrac{d\lambda^{(q)}}{ds_q}$ we need the derivative $\dfrac{dt}{dt_{pq}} \equiv D_{pq}$ which should be calculated from the relation

$$t - t_{pq} = \frac{1}{c}\left(\sum_{\alpha=1}^{3}[x_\alpha^{(p)}(t) - x_\alpha^{(q)}(t_{pq})]^2\right)^{\frac{1}{2}} \quad (t_{pq} < t \text{ by assumption}).$$

So we have $\dfrac{dt}{dt_{pq}}-1=\dfrac{\displaystyle\sum_{\alpha=1}^{3}[x_\alpha^{(p)}(t)-x_\alpha^{(q)}(t_{pq})][u_\alpha^{(p)}(t)\frac{dt}{dt_{pq}}-u_\alpha^{(q)}(t_{pq})]}{c\left(\displaystyle\sum_{\alpha=1}^{3}[x_\alpha^{(p)}(t)-x_\alpha^{(q)}(t_{pq})]^2\right)^{\frac{1}{2}}}$.

Since (3_{pq}) has a unique solution (cf. [3]) we obtain $c^2\tau_{pq}(D_{pq}-1)=\langle\xi^{(pq)},u^{(p)}\rangle D_{pq}-\langle\xi^{(pq)},u^{(q)}\rangle$ and we can solve the above equation with respect to D_{pq}: $D_{pq}=\dfrac{c^2\tau_{pq}-\langle\xi^{(pq)},u^{(q)}\rangle}{c^2\tau_{pq}-\langle\xi^{(pq)},u^{(p)}\rangle}$. We have also $\dfrac{d}{ds_p}=\dfrac{d}{\Delta_p dt}$.

Then $\dfrac{d}{ds_q}=\dfrac{1}{\Delta_{pq}}\dfrac{d}{dt_{pq}}=\dfrac{1}{\Delta_{pq}}\dfrac{d}{dt}\dfrac{dt}{dt_{pq}}=\dfrac{D_{pq}}{\Delta_{pq}}\dfrac{d}{dt}$;

$$\frac{d\lambda_\alpha^{(q)}}{ds_q}=\frac{D_{pq}}{\Delta_{pq}}\frac{d\lambda_\alpha^{(q)}}{dt}=\frac{D_{pq}}{\Delta_{pq}}\frac{d\left(u_\alpha^{(q)}/\Delta_{pq}\right)}{dt}=D_{pq}\left[\dot u_\alpha^{(q)}\frac{1}{\Delta_{pq}^2}+\frac{u_\alpha^{(q)}}{\Delta_{pq}^4}\langle u^{(q)},\dot u^{(q)}\rangle\right]$$

$(\alpha=1,2,3)$;

$\dfrac{d\lambda_4^{(q)}}{ds_q}=\dfrac{icD_{pq}}{\Delta_{pq}^4}\langle u^{(q)},\dot u^{(q)}\rangle$; $\langle\lambda^{(p)},\lambda^{(q)}\rangle_4=\dfrac{\langle u^{(p)},u^{(q)}\rangle-c^2}{\Delta_p\Delta_{pq}}$;

$\langle\lambda^{(p)},\xi^{(pq)}\rangle_4=\dfrac{\langle u^{(p)},\xi^{(pq)}\rangle-c^2\tau_{pq}}{\Delta_p}$; $\langle\lambda^{(q)},\xi^{(pq)}\rangle_4=\dfrac{\langle u^{(q)},\xi^{(pq)}\rangle-c^2\tau_{pq}}{\Delta_{pq}}$;

$\langle\xi^{(pq)},\dfrac{d\lambda^{(q)}}{ds_q}\rangle_4=D_{pq}\left[\dfrac{1}{\Delta_{pq}^2}\langle\xi^{(pq)},\dot u^{(q)}\rangle+\dfrac{\langle\xi^{(pq)},u^{(q)}\rangle-c^2\tau_{pq}}{\Delta_{pq}^4}\langle u^{(q)},\dot u^{(q)}\rangle\right]$;

$\langle\lambda^{(p)},\dfrac{d\lambda^{(q)}}{ds_q}\rangle_4=\dfrac{D_{pq}}{\Delta_p\Delta_{pq}^2}\left[\langle u^{(p)},\dot u^{(q)}\rangle+\dfrac{\langle u^{(p)},u^{(q)}\rangle-c^2}{\Delta_{pq}^2}\langle u^{(q)},\dot u^{(q)}\rangle\right]$.

We note that in the last expressions $\xi^{(pq)}$ is 4-dimensional vector in the left-hand sides, while in the right-hand sides $\xi^{(pq)}$ is 3-dimensional part of the first three coordinates.

Replacing the above expressions in $(4.\alpha)$ and (4.4) and performing some obvious transformations we obtain for $(pq)=(12),(21),\alpha=1,2,3$:

$$\begin{aligned}\frac{1}{\Delta_p}\dot u_\alpha^{(p)}+\frac{u_\alpha^{(p)}}{\Delta_p^3}\langle u^{(p)},\dot u^{(p)}\rangle=\frac{Q_p}{c^2}\Bigg\{&\frac{[c^2-\langle u^{(p)},u^{(q)}\rangle]\xi_\alpha^{(pq)}-[c^2\tau_{pq}-\langle u^{(p)},\xi^{(pq)}\rangle]u_\alpha^{(q)}}{[c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle]^3}\cdot\\ &\cdot\frac{\Delta_{pq}^4+D_{pq}\left[\Delta_{pq}^2\langle\xi^{(pq)},\dot u^{(q)}\rangle+(\langle\xi^{(pq)},u^{(q)}\rangle-c^2\tau_{pq})\langle u^{(q)},\dot u^{(q)}\rangle\right]}{\Delta_{pq}^2}+\\ &+D_{pq}\frac{[\langle u^{(p)},\xi^{(pq)}\rangle-c^2\tau_{pq}][\dot u_\alpha^{(q)}+u_\alpha^{(q)}\langle u^{(q)},\dot u^{(q)}\rangle/\Delta_{pq}^2]}{[c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle]^2}-\\ &-D_{pq}\frac{[\langle u^{(p)},\dot u^{(q)}\rangle+(\langle u^{(p)},u^{(q)}\rangle-c^2)\langle u^{(q)},\dot u^{(q)}\rangle/\Delta_{pq}^2]\xi_\alpha^{(pq)}}{[c^2\tau_{pq}-\langle u^{(q)},\xi_\alpha^{(pq)}\rangle]^2}\Bigg\},\end{aligned}\qquad(5_{p\alpha})$$

$$\frac{1}{\Delta_p^3}\langle u^{(p)},\dot u^{(p)}\rangle=\frac{Q_p}{c^2}\left\{\frac{\langle u^{(p)},\xi^{(pq)}\rangle-\tau_{pq}\langle u^{(p)},u^{(q)}\rangle}{[c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle]^3}\right.\cdot$$

$$\cdot\left[\Delta_{pq}^2 + D_{pq}(\langle\xi^{(pq)},\dot u^{(q)}\rangle + (\langle\xi^{(pq)},u^{(q)}\rangle - c^2\tau_{pq})\langle u^{(q)},\dot u^{(q)}\rangle/\Delta_{pq}^2)\right] + \tag{5_{p4}}$$

$$\left. + D_{pq}\frac{\langle u^{(p)},\xi^{(pq)}\rangle\langle u^{(q)},\dot u^{(q)}\rangle/\Delta_{pq}^2 - \tau_{pq}\langle u^{(p)},\dot u^{(q)}\rangle - \tau_{pq}\langle u^{(p)},u^{(q)}\rangle\langle u^{(q)},\dot u^{(q)}\rangle/\Delta_{pq}^2}{[c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle]^2}\right\}.$$

One can prove (as in [4]) that (5_{p4}) is a consequence of $(5_{p\alpha})$. Indeed, multiplying $(5_{p\alpha})$ by $\overset{(p)}{u_\alpha}$, summing up in α and dividing into c^2 we obtain (5_{p4}). Therefore we can consider a system consisting of the $1^{st}, 2^{nd}, 3^{rd}, 5^{th}, 6^{th}$ and 7^{th} equations. The last equations form a nonlinear functional differential system of neutral type with respect to the unknown velocities (cf. [10]-[13]). The delays τ_{pq} depend on the unknown trajectories by the relations (3_{pq}).

Now we are able to present $(5_{p\alpha})$ in a suitable form in order to make further simplifications (Recall that we shall not consider (5_{p4}) because it is a consequence of $(5_{p\alpha})$):

$$\dot u_\alpha^p + \frac{\langle u^{(p)},\dot u^{(p)}\rangle}{\Delta_p^2}u_\alpha^{(p)} = \frac{Q_p\Delta_p}{c^2(c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle)^2}\left\{\left[\frac{c^2 - \langle u^{(p)},u^{(q)}\rangle}{c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle}\xi_\alpha^{(pq)} - \right.\right.$$

$$\left. - \frac{c^2\tau_{pq} - \langle u^{(p)},\xi^{(pq)}\rangle}{c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle}u_\alpha^{(q)}\right]\left[\Delta_{pq}^2 + D_{pq}\langle\xi^{(pq)},\dot u^{(q)}\rangle + D_{pq}\frac{\langle u^{(q)},\xi^{(pq)}\rangle - c^2\tau_{pq}}{\Delta_{pq}^2}\langle u^{(q)},\dot u^{(q)}\rangle\right] +$$

$$+ D_{pq}(\langle u^{(p)},\xi^{(pq)}\rangle - c^2\tau_{pq})\dot u_\alpha^{(q)} + D_{pq}(\langle u^{(p)},\xi^{(pq)}\rangle - c^2\tau_{pq})\frac{\langle u^{(q)},\dot u^{(q)}\rangle}{\Delta_{pq}^2}u_\alpha^{(q)} - \tag{$6_{p\alpha}$}$$

$$\left. - D_{pq}\langle u^{(p)},\dot u^{(q)}\rangle\xi_\alpha^{(pq)} + D_{pq}\frac{c^2 - \langle u^{(p)},u^{(q)}\rangle}{\Delta_{pq}^2}\langle u^{(q)},\dot u^{(q)}\rangle\xi_\alpha^{(pq)}\right\}.$$

Let us recall that if $(pq) = (12)$ then $u^{(1)} = u^{(1)}(t)$ and $u^{(2)} = u^{(2)}(t - \tau_{12})$, while when $(pq) = (21)$, then $u^{(2)} = u^{(2)}(t)$ and $u^{(1)} = u^{(1)}(t - \tau_{21})$. Further on from $(6_{p\alpha})$ we obtain

$$\dot u_\alpha^p + \frac{\langle u^{(p)},\dot u^{(p)}\rangle}{\Delta_p^2}u_\alpha^{(p)} =$$

$$= \frac{Q_p\Delta_p}{c^2(c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle)^2}\left\{\frac{\left[c^2 - \langle u^{(p)},u^{(q)}\rangle\right]\left[\Delta_{pq}^2 + D_{pq}\langle\xi^{(pq)},\dot u^{(q)}\rangle\right]}{c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle}\xi_\alpha^{(pq)} - \right.$$

$$- D_{pq}\frac{(c^2 - \langle u^{(p)},u^{(q)}\rangle)\langle u^{(q)},\dot u^{(q)}\rangle}{\Delta_{pq}^2}\xi_\alpha^{(pq)} -$$

$$- \frac{c^2\tau_{pq} - \langle u^{(p)},\xi^{(pq)}\rangle}{c^2\tau_{pq} - \langle u^{(q)},\xi^{(pq)}\rangle}\left[\Delta_{pq}^2 + D_{pq}\langle\xi^{(pq)},\dot u^{(q)}\rangle + D_{pq}\frac{\langle u^{(q)},\xi^{(pq)}\rangle - c^2\tau_{pq}}{\Delta_{pq}^2}\langle u^{(q)},\dot u^{(q)}\rangle\right]u_\alpha^{(q)} +$$

$$+ D_{pq}(\langle u^{(p)},\xi^{(pq)}\rangle - c^2\tau_{pq})\dot u_\alpha^{(q)} + D_{pq}\frac{(\langle u^{(p)},\xi^{(pq)}\rangle - c^2\tau_{pq})\langle u^{(q)},\dot u^{(q)}\rangle}{\Delta_{pq}^2}u_\alpha^{(q)} -$$

$$\left. - D_{pq}\langle u^{(p)},\dot u^{(q)}\rangle\xi_\alpha^{(pq)} + D_{pq}\frac{(c^2 - \langle u^{(p)},u^{(q)}\rangle)\langle u^{(q)},\dot u^{(q)}\rangle}{\Delta_{pq}^2}\xi_\alpha^{(pq)}\right\}.$$

Obviously the second summand and the last one cancel each other and after a re-arrangement of the rest summands we have:

$$\dot{u}^p_\alpha + \frac{\langle u^{(p)}, \dot{u}^{(p)}\rangle}{\Delta_p^2} u^{(p)}_\alpha =$$

$$= \frac{Q_p \Delta_p}{c^2(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle)^2} \left\{ \left[\frac{\left(c^2 - \langle u^{(p)}, u^{(q)}\rangle\right)\left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} - D_{pq}\langle u^{(p)}, \dot{u}^{(q)}\rangle \right] \xi^{(pq)}_\alpha + \right.$$

$$+ \left[-\frac{c^2\tau_{pq} - \langle u^{(p)}, \xi^{(pq)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} \left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right) - \right.$$

$$- \frac{c^2\tau_{pq} - \langle u^{(p)}, \xi^{(pq)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} D_{pq} \frac{\langle u^{(q)}, \xi^{(pq)}\rangle - c^2\tau_{pq}}{\Delta^2_{pq}} \langle u^{(q)}, \dot{u}^{(q)}\rangle +$$

$$\left. + D_{pq} \frac{(\langle u^{(p)}, \xi^{(pq)}\rangle - c^2\tau_{pq})\langle u^{(q)}, \dot{u}^{(q)}\rangle}{\Delta^2_{pq}} \right] u^{(q)}_\alpha + D_{pq}(\langle u^{(p)}, \xi^{(pq)}\rangle - c^2\tau_{pq})\dot{u}^{(q)}_\alpha \Bigg\}.$$

It is easy to see that the second and third summands before $u^{(q)}_\alpha$ cancel each other so that we obtain the following simplified form of $(6_{p\alpha})$:

$$\dot{u}^p_\alpha + \frac{\langle u^{(p)}, \dot{u}^{(p)}\rangle}{\Delta_p^2} u^{(p)}_\alpha = \frac{Q_p \Delta_p}{c^2(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle)^2} \cdot$$

$$\cdot \left\{ \left[\frac{\left(c^2 - \langle u^{(p)}, u^{(q)}\rangle\right)\left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} - D_{pq}\langle u^{(p)}, \dot{u}^{(q)}\rangle \right] \xi^{(pq)}_\alpha - \right. \qquad (7_{p\alpha})$$

$$\left. - \frac{\left(c^2\tau_{pq} - \langle u^{(p)}, \xi^{(pq)}\rangle\right)\left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} u^{(q)}_\alpha + D_{pq}\left(\langle u^{(p)}, \xi^{(pq)}\rangle - c^2\tau_{pq}\right)\dot{u}^{(q)}_\alpha \right\}.$$

In the same way we can obtain more suitable (in view of next section) form of the equations (5_{p4}) (although we proved that (5_{p4}) is a consequence of $(5_{p\alpha})$):

$$\frac{1}{\Delta_p^2}\langle u^{(p)}, \dot{u}^{(p)}\rangle = \frac{Q_p\Delta_p}{c^2} \left\{ \frac{\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle}{\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^3} \left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right) - \right.$$

$$- \frac{D_{pq}\langle u^{(q)}, \dot{u}^{(q)}\rangle}{\Delta^2_{pq}} \cdot \frac{\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle}{\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^2} +$$

$$\left. + \frac{D_{pq}\langle u^{(q)}, \dot{u}^{(q)}\rangle}{\Delta^2_{pq}} \cdot \frac{\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle}{\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^2} - \frac{D_{pq}\tau_{pq}\langle u^{(p)}, \dot{u}^{(q)}\rangle}{\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^2} \right\}, \text{ i.e.}$$

$$\frac{1}{\Delta_p^2}\langle u^{(p)}, \dot{u}^{(p)}\rangle = \frac{Q_p\Delta_p}{c^2\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^2} \cdot \qquad (7_{p4})$$

$$\cdot \left[\frac{\left(\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle\right)\left(\Delta^2_{pq} + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} - D_{pq}\tau_{pq}\langle u^{(p)}, \dot{u}^{(q)}\rangle \right]$$

$(p = 1, 2)$.

3. Equations of motion from [1]

Now we consider the equations of motion, considered in [1], p.79-80, using his denotations:

$x_k(s)$ - the position of the particle k in R^3 at the instant s ($k = 1, 2$);

e_k - the charge of the particle k ($k = 1, 2$);

m_k - the mass of the particle k ($k = 1, 2$);

$r_i = r_i(t)$ - the delays, which satisfy the equations:

$(*) \qquad cr_i(t) = |x_i(t) - x_j(t - r_i)|\ (j \neq i)$;

v_i - the normalized velocities, where $x_i' = cv_i$ for $i = 1, 2$;

$u_i = \dfrac{x_i(t) - x_j(t - r_i)}{cr_i}$, $\gamma_i = 1 - v_j(t - r_i).u_i$ $(j \neq i)$, where "." indicates the dot of scalar product in R^3.

REMARK 1: $x_k(s)$ is the restriction of the space-time vector $(x_1^{(k)}(s), x_2^{(k)}(s), x_3^{(k)}(s), x_4^{(k)}(s) = ics)$, while u_i corresponds to the restriction of the isotropic vector $\dfrac{\xi^{(pq)}}{c\tau_{pq}}$ in 3-dimensional Euclidean subspace of the Minkowski space.

The delays $r_i = r_i(t)$ and the equations $(*)$ correspond to $\tau_{pq} = \tau_{pq}(t)$ and to the equations (3_{pq}) respectively. Finally, we use the notation $u^{(k)} = u^{(k)}(s) = (u_1^{(k)}(s), u_2^{(k)}(s), u_3^{(k)}(s))$ for the velocity vectors, where $u_\alpha^{(k)}(s) = \dfrac{dx_\alpha^{(k)}(s)}{ds}$ ($\alpha = 1, 2, 3$), $k = 1, 2$ so the normalized vector $v_k(s)$ would be equaled to $\dfrac{u^{(k)}(s)}{c}$, if we should use the notation for it ($k = 1, 2$).

The equations of motion, given in [1], are the following:

$$v_i' = \frac{e_i(1 - v_i^2)^{1/2}}{m_i c}[E_j + (v_i.E_j)(u_i - v_i) - (v_i.u_i)E_j], \qquad (**)$$

where $v_i^2 = |v_i|^2 (= v_i.v_i)$ and

$$E_j = \frac{kce_j}{r_i^2\gamma_i^3}[u_i - v_j(t - r_i)][1 - v_j^2(t - r_i)] + \frac{kce_j}{r_i\gamma_i^3}u_i \times ([u_i - v_j(t - r_i)] \times v_j'(t - r_i)),$$

where " $\times$ " stands for the cross product in R^3, "$k > 0$ is a constant depending on the units used", and the denotation $v_j'(t - r_i)$ most probably means a derivative with respect to the argument of $v_j(t - r_i)$, (in [1] there is no explanation). Rewrite the right-hand side of $(**)$ in the form:

$$\frac{e_i(1 - v_i^2)^{1/2}}{m_i c}\{[1 - (v_i.u_i)]E_j + (v_i.E_j)(u_i - v_i)\}$$

and calculate the vector cross product from E_j:

$$E_j = \frac{kce_j}{r_i^2\gamma_i^3}[1 - v_j^2(t - r_i)][u_i - v_j(t - r_i)] + \frac{kce_j}{r_i\gamma_i^3}(u_i.v_j'(t - r_i)[u_i - v_j(t - r_i)] - \frac{kce_j}{r_i\gamma_i^2}v_j'(t - r_i).$$

(since $u_i.[u_i - v_j(t - r_i)] = |u_i|^2 - v_j(t - r_i).u_i = 1 - v_j(t - r_i).u_i = \gamma_i$!). Consequently the equations $(**)$ are equivalent to the following ones:

$$v_i' = \frac{e_i(1 - v_i^2)^{1/2}}{m_i c}\left\{[1 - (u_i.v_i)]\left[\frac{kce_j}{r_i^2\gamma_i^3}[1 - v_j^2(t - r_i)][u_i - v_j(t - r_i)] + \right.\right.$$

$$+\frac{kce_j}{r_i\gamma_i^3}(u_i.v_j'(t-r_i))[u_i-v_j(t-r_i)]-\frac{kce_j}{r_i\gamma_i^2}v_j'(t-r_i)\Bigg]+$$

$$+\left[\frac{kce_j}{r_i^2\gamma_i^3}[1-v_j^2(t-r_i)][(v_i.u_i)-(v_i.v_j(t-r_i))]+\right. \tag{8}$$

$$+\frac{kce_j}{r_i\gamma_i^3}(u_i.v_j'(t-r_i))[(v_i.u_i)-(v_i.v_j(t-r_i))]-$$

$$\left.-\frac{kce_j}{r_i\gamma_i^2}(v_i.v_j'(t-r_i))\right](u_i-v_i)\Bigg\}.$$

Then we can arrange the symbols, including the vectors $u_i, v_i, v_j(t-r_i)$ and $v_j'(t-r_i)$ respectively and we obtain the equivalent equations:

$$v_i'=\frac{ke_ie_j(1-v_i^2)^{1/2}}{m_ir_i\gamma_i^2}\left\{u_i\left[\frac{[1-(u_i.v_i)][1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))]}{r_i\gamma_i}+\right.\right.$$

$$\left.+\frac{[1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))][(v_i.u_i)-(v_i.v_j(t-r_i))]}{r_i\gamma_i}-(v_i.v_j'(t-r_i))\right]-$$

$$-v_i\left[\frac{[1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))][(v_i.u_i)-(v_i.v_j(t-r_i))]}{r_i\gamma_i}-(v_i.v_j'(t-r_i))\right]-$$

$$\left.-v_j(t-r_i)\frac{[1-(u_i.v_i)][1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))]}{r_i\gamma_i}-v_j'(t-r_i)[1-(u_i.v_i)]\right\}$$

and finally one has

$$v_i'=\frac{ke_ie_j(1-v_i^2)^{1/2}}{m_ir_i\gamma_i^2}\left\{u_i\left[\frac{[1-(v_i.v_j(t-r_i))][1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))]}{r_i\gamma_i}-\right.\right.$$

$$\left.-(v_i.v_j'(t-r_i))\right]- \tag{9}$$

$$-v_i\left[\frac{[1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))][(v_i.u_i)-(v_i.v_j(t-r_i))]}{r_i\gamma_i}-(v_i.v_j'(t-r_i))\right]-$$

$$\left.-v_j(t-r_i)\frac{[1-(u_i.v_i)][1-v_j^2(t-r_i)+r_i(u_i.v_j'(t-r_i))]}{r_i\gamma_i}-v_j'(t-r_i)[1-(u_i.v_i)]\right\}.$$

To compare both systems we present the equations (9), using the denotations from our previous section II.

We have for $i\equiv p, j\equiv q, r_i\equiv\tau_{pq}$, and in view of Remark 1:

$$v_i=v_i(t)\equiv\frac{u^{(p)}}{c};\quad (1-v_i^2)^{1/2}=\left(1-\left\langle\frac{u^{(p)}}{c},\frac{u^{(p)}}{c}\right\rangle\right)^{1/2}=\frac{\Delta_p}{c};$$

$$v_i'=\frac{d(u^{(p)}/c)}{dt}=\dot{u}^{(p)}/c\quad (u^{(p)}=u^{(p)}(t));$$

$$u_i \equiv \frac{x^{(p)}(t) - x^{(q)}(t-\tau_{pq})}{c\tau_{pq}} = \frac{1}{c\tau_{pq}}\left(\xi_1^{(pq)}, \xi_2^{(pq)}, \xi_3^{(pq)}\right);$$

$$v_j(t-r_i) \equiv u^{(q)}/c; \quad v_j'(t-r_i) = \frac{d(u^{(q)}/c)}{dt_{pq}} = \frac{1}{c}\cdot\frac{du^{(q)}}{dt}\cdot\frac{dt}{dt_{pq}} = \frac{D_{pq}\dot{u}^{(q)}}{c}$$

$$(u^{(q)} = u^{(q)}(t-\tau_{pq}) = u^{(q)}(t_{pq}));$$

$$\gamma_i \equiv 1 - \left\langle \frac{u^{(q)}}{c}, \frac{\xi^{(pq)}}{c\tau_{pq}} \right\rangle = \frac{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle}{c^2\tau_{pq}};$$

$$1 - (v_i.v_j(t-r_i)) = 1 - \left\langle \frac{u^{(p)}}{c}, \frac{u^{(q)}}{c} \right\rangle = \frac{c^2 - \langle u^{(p)}, u^{(q)}\rangle}{c^2};$$

$$\frac{1 - v_j^2(t-r_i) + r_i(u_i.v_j'(t-r_i))}{r_i\gamma_i} \equiv \frac{1 - \left\langle \frac{u^{(q)}}{c}, \frac{u^{(q)}}{c} \right\rangle + \tau_{pq}\left\langle \frac{\xi^{(pq)}}{c\tau_{pq}}, \frac{D_{pq}\dot{u}^{(q)}}{c} \right\rangle}{\tau_{pq}\left(1 - \left\langle \frac{u^{(q)}}{c}, \frac{\xi^{(pq)}}{c\tau_{pq}} \right\rangle\right)} =$$

$$= \frac{\Delta_{pq}^2 + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle};$$

$$1 - (u_i.v_i) \equiv 1 - \left\langle \frac{\xi^{(pq)}}{c\tau_{pq}}, \frac{u^{(p)}}{c} \right\rangle = \frac{c^2\tau_{pq} - \langle u^{(p)}, \xi^{(pq)}\rangle}{c^2\tau_{pq}};$$

$$(v_i.u_i) - (v_i.v_j(t-r_i)) \equiv \frac{\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle}{c^2\tau_{pq}}$$

and replacing in (9) we obtain the following 6 scalar equations:

$$\frac{1}{c}\dot{u}_\alpha^p = \frac{ke_pe_q\Delta_p(c^2\tau_{pq})^2}{cm_p\tau_{pq}(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle)^2}\cdot$$

$$\cdot\left\{ \frac{\xi_\alpha^{(pq)}}{c\tau_{pq}}\left[\frac{c^2 - \langle u^{(p)}, u^{(q)}\rangle}{c^2}\cdot\frac{\Delta_{pq}^2 + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} - \frac{D_{pq}}{c^2}\langle u^{(p)}, \dot{u}^{(q)}\rangle\right] - \right.$$

$$-\frac{u_\alpha^{(p)}}{c}\left[\frac{\Delta_{pq}^2 + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle}\cdot\frac{\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle}{c^2\tau_{pq}} - \frac{D_{pq}}{c^2}\langle u^{(p)}, \dot{u}^{(q)}\rangle\right] -$$

$$-\frac{u_\alpha^{(q)}}{c}\cdot\frac{c^2\tau_{pq} - \langle u^{(p)}, \xi^{(pq)}\rangle}{c^2\tau_{pq}}\cdot\frac{\Delta_{pq}^2 + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} +$$

$$\left. + \frac{D_{pq}\dot{u}_\alpha^{(q)}}{c}\cdot\frac{\langle u^{(p)}, \xi^{(pq)}\rangle - c^2\tau_{pq}}{c^2\tau_{pq}}\right\} \qquad (\alpha = 1,2,3;\ (pq) = (12),\ (21)),$$

The above system is obviously equivalent to the following one (with $Q_p = \dfrac{e_pe_q}{m_p}$):

$$\dot{u}_\alpha^p + \frac{kcQ_p\Delta_p}{\left(c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle\right)^2}\cdot\Big[\frac{\left(\langle u^{(p)}, \xi^{(pq)}\rangle - \tau_{pq}\langle u^{(p)}, u^{(q)}\rangle\right)\left(\Delta_{pq}^2 + D_{pq}\langle \xi^{(pq)}, \dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq} - \langle u^{(q)}, \xi^{(pq)}\rangle} -$$

$$-D_{pq}\tau_{pq}\langle u^{(p)},\dot{u}^{(q)}\rangle\Big]u_\alpha^{(p)}=$$

$$=\frac{kcQ_p\Delta_p}{(c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle)^2}\cdot\Bigg\{\Bigg[\frac{\left(c^2-\langle u^{(p)},u^{(q)}\rangle\right)\left(\Delta_{pq}^2+D_{pq}\langle\xi^{(pq)},\dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle}-D_{pq}\langle u^{(p)},\dot{u}^{(q)}\rangle\Bigg]\xi_\alpha^{(pq)}-$$

$$\frac{\left(c^2\tau_{pq}-\langle u^{(p)},\xi^{(pq)}\rangle\right)\left(\Delta_{pq}^2+D_{pq}\langle\xi^{(pq)},\dot{u}^{(q)}\rangle\right)}{c^2\tau_{pq}-\langle u^{(q)},\xi^{(pq)}\rangle}u_\alpha^{(q)}+ \qquad (10_{p\alpha})$$

$$+D_{pq}\left(\langle u^{(p)},\xi^{(pq)}\rangle-c^2\tau_{pq}\right)\dot{u}_\alpha^{(q)}\Bigg\}\qquad(\alpha=1,2,3;\ (pq)=(12),\ (21)).$$

Conclusion remarks

Our goal is to point out the difference between the system of equations of motion (9) (or equivalently $(10_{p\alpha})$) from [1] and Synge's equations $(7_{p\alpha})$, (7_{p4}).

1) equations (9) are obtained under assumption that (7_{p4}) should be identities which is not discussed in [1]. On the other hand in [4] we have already proved that (7_{p4}) is a consequence of $(7_{p\alpha})$, $\alpha=1,2,3$. It is not obvious that (7_{p4}) is an identity.

2) the right-hand sides of $(10_{p\alpha})$ and Synge's equations $(7_{p\alpha})$ differ each other by the multiplier c^3 which is a consequence from $kc=\dfrac{1}{c^2}$. But this means that the right-hand sides of $(10_{p\alpha})$ are c^3-times larger than the right-hand sides of Synge equations $(7_{p\alpha})$, that is, they have another dimension. Therefore it is not surprised that they possess only unstable solutions, while $(7_{p\alpha})$ have a circle solution [2].

References

[1] Driver, R. D., *A neutral system with state-dependent delay*, J. Diff. Equations, v.54, N1 (1984), 73-86.

[2] Angelov, V. G., *Plane orbits for Synge's electromagnetic two-body problem (I).* Seminar on Fixed Point Theory, Cluj-Napoca, v.3 (2002), 3-12.

[3] Angelov, V. G., *On the Synge equations in three-dimensional two-body problem of classical electrodynamics*, J. Math. Anal. Appl. v.151, N2 (1990), 488-511.

[4] Angelov, V. G., *Escape trajectories of J. L. Synge equations*, J. Non Anal. RWA, ser. B, v.1 (2000), 189-204.

[5] Synge, J. L., *On the electromagnetic two-body problem*, Proc. Roy. Soc. (London), ser. A, 177 (1940), 118-139.

[6] Pauli, W., *Relativitatstheorie*, Encycl. der Math. Wissenschaften. B.5, H.S. Art. 19, 1921.

[7] Myshkis, A. D., *On some probrlem of the theory of differential equations with deviating argument*, Uspekhi Mat. Nauk, v.4, N5 (1949), 99-141 (in Russian).

[8] Angelov, V. G., *On the method of sucsesssive approximations for J. L. Synge electromagnetic two-body problem* (to appear).

[9] Driver, R. D., *A two-body problem of classical electrodynamics: the one-dimensional case*, Ann. Physics, v.21, N1 (1963), 122-142.

[10] El'sgoltz, L. E., Norkin, S. B., *Introduction in the theory of differential equations with deviating argument*, Moscow, Nauka, 1971 (in Russian).

[11] Myshkis, A. D., *Linear differential equations with retarded argument*, Moscow, Nauka, 1972 (in Russian).

[12] Kamenskii, G. A., Skubachevskii, A. L., *Linear boundary value problems for differential-difference equations*, Moscow, 1992 (in Russian).

[13] Hale, J.K., Verduyn-Lunel, S.M., *Functional Differential Equations*, Springer Verlag, 1993.

University of Mining and Geology "St. I. Rilski",
Department of Mathematics, 1700 Sofia, Bulgaria
E-mail address: angelov@staff.mgu.bg

GENERAL EXISTENCE RESULTS FOR THE ZEROS OF A COMPACT NONLINEAR OPERATOR DEFINED IN A FUNCTIONAL SPACE

CEZAR AVRAMESCU

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. Let X be a Banach space whose the elements are functions defined on a non-empty set Ω with values in a prehilbertian space H. Let $B := \{x \in X,\ \|x\| \leq 1\}$, $S := \{x \in X,\ \|x\| = 1\}$ and let $f : \overline{B} \to X$ be a compact operator. one shows that if f fulfills on S certain conditions, then the equation $(*)$ $f(x) = 0$ admits solutions. The particular case when Ω is a topological compact space and $X = C(\Omega, \mathbb{R}^n)$ is also considered.

1. Many existence problems in analysis are reduced to an equation of type

$$f(x) = 0, \tag{1}$$

where f is an operator defined between two adequate functional spaces. Generally, the problem (1) is reduced many times to a fixed point problem for the mapping $x \to x + f(x)$, but not always this reducing is adequate.

Through the results concerning directly the equation (1) we mention the one of Miranda [3], which considers the particular case when f maps in a finite dimensional space. The case considered in what follows is much more general.

2. Let Ω be a non-empty arbitrary set, H be a real prehilbertian space and X be a subset of H^Ω; suppose that X is a Banach space endowed with the norm $\|\cdot\|$.

Denote by $\langle | \rangle$ the scalar product of H and define a mapping from $X \times X$ to $\mathbb{R}^\Omega$,

$$(x, y) \to [x \mid y](t) := \langle x(t) \mid y(t) \rangle, \text{ for all } t \in \Omega. \tag{2}$$

Let us consider

$$[x \mid y] > 0 \tag{3}$$

if

$$[x \mid y](t) > 0, \text{ for all } t \in \Omega, \tag{4}$$

for the inequality "<" the convention being the same.

Denote

$$\overline{B} := \{x \in X,\ \|x\| \leq 1\},\ S := \{x \in X,\ \|x\| = 1\}.$$

Received by the editors: 18.06.2002.
2000 *Mathematics Subject Classification.* 47H10, 47H99.
Key words and phrases. Fixed point theorems, Zeros of nonlinear operator.

Let $f:\overline{B}\to X$ be a given operator; one can proof the following result.

Theorem 1. *Suppose that:*
i) f is a compact operator;
ii) $[x \mid f(x)] < 0$, for all $x\in S$;
iii) $0\notin \overline{f(\overline{B})}\backslash f(\overline{B})$.
Then the equation (1) *admits solutions.*

Proof. By means of contradiction suppose that

$$f(x)\neq 0,\ x\in\overline{B}. \tag{5}$$

Then the operator

$$F(x):=\frac{1}{\|f(x)\|}\cdot f(x) \tag{6}$$

is defined on $\overline{B}$ and is continuous; in addition,

$$F(\overline{B})\subset S. \tag{7}$$

We shall proof that $F(\overline{B})$ is a relatively compact set.

If $y_n\in F(\overline{B})$, then $y_n=F(x_n)$, $x_n\in\overline{B}$, i.e.

$$y_n=\frac{1}{\|f(x_n)\|}\cdot f(x_n).$$

Since $f(\overline{B})$ is relatively compact, it results that $(f(x_n))_n$ contains a convergent subsequence; one can admit that

$$\lim_{n\to\infty} f(x_n)=z\in\overline{f(\overline{B})}. \tag{8}$$

It remains to show that $z\neq 0$ to conclude that $(y_n)_n$ given by (7) is convergent. But

$$0\in\overline{f(\overline{B})} \tag{9}$$

implies by (5)

$$0\in\overline{f(\overline{B})}\backslash f(\overline{B}),$$

which is not true.

Hence, F fulfills the hypotheses of Schauder's fixed point theorem and so it will admit a fixed point which, by (6) will belong to S.

By

$$x=F(x),\ x\in S$$

it results

$$x\cdot\|f(x)\|=f(x),\ x\in S, \tag{10}$$

therefore

$$[x \mid x]^2\cdot\|f(x)\|=[x \mid f(x)]. \tag{11}$$

But

$$[x \mid x]^2\geq 0,$$

which contradicts hypothesis ii).

The theorem is proved. □

Remark 1. *Hypothesis iii) can be replaced with a formulation of "aprioric estimate" type, i.e.*

ii) if $0 \in \overline{f\left(\overline{B}\right)}$, *then* $0 \in f\left(\overline{B}\right)$.

Remark 2. *The importance of the result is the fact that one doesn't suppose any link between the topologies of* H *and* X *and the special properties for the applications* $x : \Omega \to X$, *too.*

Remark 3. *Hypothesis iii) is useless if* f *is a closed operator or if* $\dim X < \infty$.

3. In this section we consider the case

$$X = C\left(\Omega, \mathbb{R}^n\right) := \left\{x : \Omega \to \mathbb{R}^n,\ x \text{ continuous}\right\}.$$

Suppose that Ω is a compact topological space and consider in X the norm

$$\|x\| := \sup_{t\in\Omega} |x(t)|,$$

where the norm in $\mathbb{R}^n$ is given by

$$|x| = \max_{1\le i\le n} \left\{|x_i|\right\},\ x = (x_i)_{i\in\overline{1,n}} \in \mathbb{R}^n.$$

Obviously, the result contained in Theorem 1 yields, but in this case one can replace hypothesis ii) with another weaker one. To this aim, set

$$\begin{aligned} S_i^+ &: = \left\{x \in \overline{B},\ x(t) = (x_j(t))_{j\in\overline{1,n}},\ x_j(t) \equiv 1\right\} \\ S_i^- &: = \left\{x \in \overline{B},\ x(t) = (x_j(t))_{j\in\overline{1,n}},\ x_j(t) \equiv -1\right\}. \end{aligned}$$

Clearly,

$$\cup_{i=1}^n \left(S_i^+ \cup S_i^-\right) \subset S.$$

Theorem 2. *Suppose that:*

i) $f = (f_i)_{i\in\overline{1,n}} : \overline{B} \to X$ *is a compact operator;*

ii) $\begin{cases} (f_i(x))(t) \le 0,\ x \in S_i^+,\ t \in \Omega,\ i \in \overline{1,n} \\ (f_i(x))(t) \ge 0,\ x \in S_i^-,\ t \in \Omega,\ i \in \overline{1,n} \end{cases}$;

iii) $0 \notin \overline{f\left(\overline{B}\right)} \backslash f\left(\overline{B}\right)$.

Then the equation (1) *admits solutions.*

Proof. As in Theorem 1, if (5) holds, then by using again the operator F, one can deduce similarly the relation (10).

One gets

$$(x \in S) \Longleftrightarrow \left(\sup_{t\in\Omega} |x(t)| = 1\right) \Longleftrightarrow \left((\exists)\ t_0 \in \Omega,\ \ |x(t_0)| = 1\right).$$

Hence, since $x \in S$,

$$(\exists)\ t_0 \in \Omega,\ \ (\exists)\ i \in \overline{1,n},\ |x_i(t_0)| = 1. \tag{12}$$

Suppose firstly that $x_i(t_0) = 1$; by (5) it follows

$$x_i(t_0) \cdot \|f(x)\| = (f_i(x))(t_0),$$

therefore

$$f_i(x)(t_0) > 0.$$

By starting from the fixed point $x = (x_i)_{i \in \overline{1,n}}$ we build $\widetilde{x} : \Omega \to \mathbb{R}^n$ by setting

$$\widetilde{x}(t) = (x_1(t), ..., x_{i-1}(t), 1, x_{i+1}(t), ..., x_n(t)).$$

Obviously,

$$\widetilde{x} \in S_i^+$$

and so, by hypotheses,

$$(f_i(\widetilde{x}))(t) \leq 0, \ t \in \Omega.$$

Since

$$\widetilde{x}(t_0) = x(t_0)$$

and (14) one obtains

$$(f_i(x))(t_0) \leq 0,$$

which contradicts (13). □

References

[1] Avramescu, C., *Some remarks about Miranda's theorem,* An. Univ. Craiova, Vol. XXVII (2000), p. 9-13.
[2] Avramescu, C., *A generalization of Miranda's theorem,* SFPT Cluj (2002), p. 121-129.
[3] Miranda, C., *Un'osservazione su un teorema di Brouwer,* Bul. U.M.I. 2, Vol. 3 (1940-1941), p. 82-87.

Department of Mathematics, University of Craiova,
13 A.I. Cuza Street, 1100 Craiova, Romania
E-mail address: cezaravramescu@hotmail.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

SOME APPLICATIONS OF AN ASYMPTOTICAL FIXED POINT THEOREMS FOR INTEGRAL EQUATIONS WITH DEVIATING ARGUMENT

T. BARANYAI

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. In this paper we will present an application of an asymptotical fixed point theorem for integral equation with deviating argument.

The following result is well known : ([1], [2])

Theorem 1. *Let the following integral equation with deviating argument:*

$$x(t) = h(t) + \int_a^t f(s, x(g(s)))ds,\ t \in [a, b]. \tag{1}$$

We suppose that:

(a) $h \in C([a, b], [a, b]),\ h(a) = 0$
(b) $g : [a, b] \longrightarrow [a, b],\ a \leq g(t) \leq t \leq b$
(c) $f \in C([a, b] \times \mathbb{R})$
$\exists L_f > 0,\ |f(t, u) - f(t, v)| \leq L_f\,|u - v|$ *for all* $t \in [a, b],\ u, v \in \mathbb{R}$

Then the equation (1) has an unique solution in $C[a, b]$.

In proving of this theorem are apply the contraction principle for the following operator:

$$A : C[a, b] \longrightarrow C[a, b]\,,$$

$$A(x)(t) := h(t) + \int_a^t f(s, x(g(s)))ds,\ t \in [a, b].$$

In the following we prove the existence and the unicity of the solution of the integral equation (1) without using condition (b) for the operator g. In the proof of theorem 1 are use the Bielicki norm, but in the following theorem we use the Cebîşev norm and an asymptotic fixed point principle.

Received by the editors: 29.09.2003.
2000 *Mathematics Subject Classification.* 45D05, 47H10.
Key words and phrases. Banach space, asymptotical fixed point theorems, integral equations with deviating argument.

Let X be a Banach space. We consider the following integral equation:

$$x(t) = h(t) + \int_a^t f(s, x(g(s)))ds,\ t \in [a,b] \tag{2}$$

Theorem 2. *We suppose that:*

(a) $g \in C([a,b],[a,b])$
(b) $h \in C([a,b],[a,b]),\ h(a) = 0$
(c) $f \in C([a,b] \times X, X)$
$\exists L_f > 0,\ \|f(t,u) - f(t,v)\|_X \leq L_f \|u - v\|_X$ *for all* $t \in [a,b],\ u,\ v \in X$.

Then the equation (2) has an unique solution in $C([a,b],X)$.

Proof. We consider the operator

$$A : C([a,b],X) \longrightarrow C([a,b],X)$$

$$A(x)(t) := h(t) + \int_a^t f(s, x(g(s)))ds,$$

Then the iterates of A are:

$$A^2(x)(t) = h(t) + \int_a^t f(s, A(x)(g(s)))ds,$$

$$\ldots$$

$$A^{n+1}(x)(t) = h(t) + \int_a^t f(s, A^n(x)(g(s)))ds$$

We have the following estimations ([3]):

$$|A(x)(t) - A(y)(t)| \leq L_f \int_a^t |x(g(s)) - y(g(s))|\, ds \leq$$

$$\leq L_f \|x - y\|_C \frac{t-a}{1!}, \quad \forall t \in [a,b] \quad (\|\ \|_C \text{ is the Cebîşev norm})$$

$$\left|A^2(x)(t) - A^2(y)(t)\right| \leq L_f \int_a^t |x(A(s)) - y(A(s))|\, ds \leq$$

$$\leq L_f \|x - y\|_C \int_a^t \frac{s-a}{1!} ds \leq L_f^2 \|x - y\|_C \frac{(t-a)^2}{2!}, \quad \forall t \in [a,b]$$

$$\ldots$$

$$\left|A^k(x)(t) - A^k(y)(t)\right| \leq L_f^k \|x - y\|_C \frac{(t-a)^k}{k!}, \quad \forall t \in [a,b], \forall k \in \mathbb{N}$$

$$\left\|A^k(x) - A^k(y)\right\| \leq \frac{[L_f(b-a)]^k}{k!} \|x - y\|_C, \quad \forall k \in \mathbb{N}.$$

So there exists a natural number k such that that:

A^k is contraction with the contraction constant $\alpha = \dfrac{[L_f(b-a)]^k}{k!} < 1$.

Now we apply an asymptotical variant of contraction principle ([2]) and we have that, the integral equation (2) has an unique solution. Q.E.D.

Remarks.

1. When we take $X = \mathbb{R}^m$ we have a result for the following system of integral equations:

$$\begin{aligned} x_1(t) &= h_1(t) + \int_a^t f_1(s, x_1(g(s)), \ldots, x_m(g(s))))ds \\ x_2(t) &= h_2(t) + \int_a^t f_2(s, x_1(g(s)), \ldots, x_m(g(s))))ds \qquad t \in [a,b] \\ &\ldots \\ x_m(t) &= h_m(t) + \int_a^t f_m(s, x_1(g(s)), \ldots, x_m(g(s))))ds \end{aligned}$$

2. When $X = l^2(\mathbb{R})$ we have a result for the following infinit sistem of integral equations:

$$x_i(t) = h_i(t) + \int_a^t f_i(s, x_1(g(s)), \ldots, x_n(g(s)), \ldots)ds, \quad t \in [a,b],\ i \in \mathbb{N}^*.$$

References

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

[2] I. A. Rus, *Generalized contractions and applications*, Cluj University Press, Cluj-Napoca, 2001.

[3] I. A. Rus, *Aplicaţii cu iterate φ-contracţii*, Studia Univ. Babeş-Bolyai, Mathematica, **25**, (1980), fasc.4, 47-50.

Department of Mathematics, Babeş-Bolyai University,
Teacher Training College, Satu Mare
E-mail address: baratun@yahoo.com

SCHURER-STANCU TYPE OPERATORS

DAN BĂRBOSU

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. Considering two non-negative parameters α, β which satisfy $0 \leq \alpha \leq \beta$ and a given non-negative integer p, the Stancu-Schurer type operators $\widetilde{S}_{m,p}^{(\alpha,\beta)} : C(0, 1+p]) \to C([0,1])$

$$\left(\widetilde{S}_{m,p}^{(\alpha,\beta)} f\right)(x) = \sum_{k=0}^{m+p} \widetilde{p}_{m,k}(x) f\left(\frac{k+\alpha}{m+\beta}\right)$$

are introduced and some approximation properties of these operators are studied.

1. **Preliminaries**

Let $p \geq 0$ be a given integer. In 1962, F. Schurer (see ([7])), introduced and studied the linear positive operator $\widetilde{B}_{m,p} : C([0, 1+p]) \to C([0, 1])$, defined for any $f \in C([0, 1+p])$ and any $m \in \mathbb{N}$ by

$$\left(\widetilde{B}_{m,p} f\right)(x) = \sum_{k=0}^{m+p} \widetilde{p}_{mk}(x) f(k/m) \tag{1.1}$$

where $\widetilde{p}_{mk}(x) = \binom{m+p}{k} x^k (1-x)^{m+p-k}$ are the fundamental Schurer polynomials.

Considering the given real parameters α, β which satisfy $0 \leq \alpha \leq \beta$, in 1968, D.D. Stancu (see ([9])), constructed the linear positive operators $P_m^{(\alpha,\beta)} : C([0,1]) \to C([0,1])$ defined for any $f \in C([0,1])$ and any $m \in \mathbb{N}$ by

$$\left(P_m^{(\alpha,\beta)} f\right)(x) = \sum_{k=0} p_{mk}(x) f\left(\frac{k+\alpha}{m+\beta}\right) \tag{1.2}$$

where $p_{mk}(x) = \binom{m}{k} x^k (1-x)^{m-k}$ are the fundamental Bernstein polynomials.

Note that for $p = 0$, the operator (1.1) reduces to the classical Bernstein operator and for $\alpha = \beta = 0$, the operator (1.2) reduces also to the classical Bernstein operator. Follows that the above operators generalize the classical Bernstein operator.

Received by the editors: 29.04.2003.

2000 *Mathematics Subject Classification.* 41A36, 41A25, 41A63.

Key words and phrases. linear positive operators, Bohman-Korovkin theorem, first order modulus of smoothness, Shisha-Mond theorem.

Let $\widetilde{S}_{m,p}^{(\alpha,\beta)} : C([0,1+p] \to C([0,1])$ be defined for any $f \in C([0,1+p])$ and any $m \in \mathbb{N}$, by:

$$\left(\widetilde{S}_{m,p}^{(\alpha,\beta)} f\right)(x) = \sum\nolimits_{k=0}^{m+p} \widetilde{p}_{m,k}(x) f\left(\frac{k+\alpha}{m+\beta}\right) \tag{1.3}$$

For $\alpha = \beta = 0$ the operator (1.3) reduces to the Schurer operator (1.1) and for $p = 0$, (1.3) reduces to the Stancu operator (1.2).

In what follows the operator defined by (1.3) will be called Schurer-Stancu type operator.

The focus of the paper is to investigate approximation properties of operator (1.3).

2. Main results

Lemma 2.1. *The Shurer-Stancu operators, defined by* (1.3)*, are linear and positive.*

Proof. The assertions follows from definition (1.3). □

Like usually, let us to denote by $e_k(s) = s^k, k \in \mathbb{N}$ the test functions.

Lemma 2.2. *For any $x \in [0, 1+p]$ and any $m \in \mathbb{N}$ the Schurer-Stancu operators* (1.3) *verify*

$$\left(\widetilde{S}_{m,p}^{(\alpha,\beta)} e_0\right)(x) := \widetilde{S}_{m,p}^{(\alpha,\beta)}(1;x) = 1 \tag{2.1}$$

$$\left(\widetilde{S}_{m,p}^{(\alpha,\beta)} e_1\right)(x) := \widetilde{S}_{m,p}^{(\alpha,\beta)}(s;x) = \frac{m+p}{m+\beta}x + \frac{\alpha}{m+\beta} \tag{2.2}$$

$$\begin{aligned}\left(\widetilde{S}_{m,p}^{(\alpha,\beta)} e_2\right)(x) &= \widetilde{S}_{m,p}^{(\alpha,\beta)}(s^2;x) = \\ &= \frac{1}{(m+\beta)^2}\left\{(m+p)^2x^2 + (m+p)x(1-x) + \right. \\ &+ \left. 2\frac{\alpha m(m+p)}{m+\beta}x + \frac{\alpha^2(3m+\beta)}{m+\beta}\right\}\end{aligned} \tag{2.3}$$

Proof. Using the definition (1.3), we get

$$\widetilde{S}_{m,p}^{(\alpha,\beta)}(1;x) = \sum\nolimits_{k=0}^{m+p} \widetilde{p}_{m,k}(x) = \widetilde{B}_{m,k}(x) = \widetilde{B}_{m,p}(1;x) = 1,$$

where we used a well known property of $\widetilde{B}_{m,p}$ (see([7])).

Next

$$\begin{aligned}\widetilde{S}_{m,p}^{(\alpha,\beta)}(s;x) &= \sum\nolimits_{k=0}^{m+p} \widetilde{p}_{m,k}(x)\frac{k+\alpha}{m+\beta} = \\ &= \frac{m}{m+\beta}\sum\nolimits_{k=0}^{m+\beta} \widetilde{p}_{m,k}(x) \cdot \frac{k}{m} + \frac{\alpha}{m+\beta}\sum\nolimits_{k=0}^{m+p} \widetilde{p}_{m,k}(x) = \\ &= \frac{m}{m+\beta}\widetilde{B}_{m,p}(s;x) + \frac{\alpha}{m+\beta}\widetilde{B}_{m,p}(1;x)\end{aligned}$$

But (see ([7])):

$$\widetilde{B}_{m,p}(s;x) = \left(1+\frac{p}{m}\right)x, \widetilde{B}_{m,p}(1;x) = 1$$

We can then conclude that

$$\widetilde{S}_{m,p}^{(\alpha,\beta)}(s;x)=\frac{m+\beta}{m+\beta}x+\frac{\alpha}{m+\beta},$$

i.e. (2.2) holds.
In a same way, we obtain

$$\widetilde{S}_{m,p}^{(\alpha,\beta)}(s^2;x)=\sum\nolimits_{k=0}^{m+p}\widetilde{p}_{m,k}(x)\cdot\left(\frac{k+\alpha}{m+\beta}\right)^2=$$

$$=\frac{1}{(m+\beta)^2}\left\{m^2\sum\nolimits_{k=0}^{m+p}\widetilde{p}_{m,k}(x)\cdot\left(\frac{k}{m}\right)^2+\right.$$

$$\left.+2\alpha m\sum\nolimits_{k=0}^{m+p}\widetilde{p}_{m,k}(x)\frac{k}{m}+\alpha^2\sum\nolimits_{k=0}^{m+p}\widetilde{p}_{m,k}(x)\right\}=$$

$$=\frac{1}{(m+\beta)^2}\left\{m^2\widetilde{B}_{m,p}(s^2;x)+2\alpha m\widetilde{B}_{m,p}(s;x)+\alpha^2\widetilde{B}_{m,p}(1;x)\right\}$$

But (see ([7]))

$$\widetilde{B}_{m,p}(s^2;x)=\frac{m+p}{m^2}\left\{(m+p)x^2+x(1-x)\right\}$$

Taking into account of the above equalities, we get

$$\begin{aligned}\widetilde{S}_{m,p}^{(\alpha,\beta)}(s^2;x) &= \frac{1}{(m+\beta)^2}\left\{(m+p)^2x^2+(m+p)x(1-x)+\right.\\ &+ \left.2\alpha m\cdot\frac{m+p}{m+\beta}x+2\alpha^2\cdot\frac{m}{m+\beta}+\alpha^2\right\}=\\ &= \frac{1}{(m+\beta)^2}\left\{(m+p)^2x^2+(m+p)x(1-x)+\right.\\ &+ \left.2\frac{\alpha m(m+p)}{m+\beta}x+\frac{\alpha^2(3m+\beta)}{m+\beta}\right\}\end{aligned}$$

i.e. (2.3) holds and the proof ends. □

Lemma 2.3. *The operators* (1.3) *verify*

$$\widetilde{S}_{m,p}^{(\alpha,\beta)}((e_1-x)^2;x)=\frac{(p-\beta)^2}{(m+\beta)^2}x^2+\frac{m+p}{(m+\beta)^2}x(1-x)+$$

$$+\frac{2\alpha(mp-2m\beta-\beta^2)}{(m+\beta)^3}x+\frac{\alpha^2(3m+\beta)}{(m+\beta)^3}\tag{2.4}$$

Proof. The linearity of $\widetilde{S}_{m,p}^{(\alpha,\beta)}$ (see Lemma 2.1) leads us to

$$\begin{aligned}\widetilde{S}_{m,p}^{(\alpha,\beta)}((e_1-x)^2;x) &= \widetilde{S}_{m,p}^{(\alpha,\beta)}(s^2;x)-2x\widetilde{S}_{m,p}^{(\alpha,\beta)}(s;x)+\\ &+ x^2\widetilde{S}_{m,p}^{(\alpha,\beta)}(1;x)\end{aligned}$$

Applying next Lemma 2.2, we get (2.4). □

We are now ready to establish an important convergence property of the sequence $\left\{\widetilde{S}_{m,p}^{(\alpha,\beta)}f\right\}_{m\in\mathbb{N}}$ contained in

Theorem 2.1. *The sequence* $\left\{\widetilde{S}_{m,p}^{(\alpha,\beta)}f\right\}_{m\in\mathbb{N}}$ *converges to* f*, uniformly on* $[0,1]$*, for any* $f\in C([0,1+p])$.

Proof. Because

$$\lim_{m\to\infty}\left\{\frac{(p-\beta)^2}{(m+\beta)^2}x^2\frac{m+p}{(m+\beta)^2}x(1-x)+\frac{2\alpha(mp-2m\beta-\beta^2)}{(m+\beta)^3}x+\frac{\alpha^2(3m+\beta)}{(m+\beta)^3}\right\}=0$$

uniformly on $[0,1]$, we can apply the well known Bohman-Korovkin Theorem and we arrive to the desired result. □

For evaluating the rate of convergence, we will use the first order modulus of smoothness (see ([1])). Let us to recall the definition of this modulus.

Definition 2.1. Let $f:[a,b]\to\mathbb{R}$ be a real valued function, bounded on $[a,b]$. The first order modulus of smoothness is the function $\omega_1:[0,b-a]\to[0,+\infty)$, defined for any $\delta\in[0,b-a]$ by

$$\omega_1(f;\delta)=\sup\{|f(x)-f(x')|:x,x'\in[0,b-a],\ |x-x'|\le\delta\} \tag{2.5}$$

It is well known the following result, due to O. Shisha and B. Mond (see([8])).

Theorem 2.2. *Let* $(L_m)_{m\in\mathbb{N}}$*,* $L_m:C([a,b])\to B([a,b])$ *be a sequence of linear positive operators, reproducing the constant functions. For any* $f\in C([a,b])$*, any* $x\in[a,b]$ *and any* $\delta\in[0,b-a]$*, the following*

$$|(L_mf)(x)-f(x)|\le\left\{1+\delta^{-1}\sqrt{L_m((e_1-x)^2;x)}\right\}\omega_1(\delta) \tag{2.6}$$

holds.

Theorem 2.3. *For any* $f\in C([0,1+p])$ *and any* $x\in[0,1]$ *the Schurer-Stancu operators* (1.3) *verify*

$$\left|\left(\widetilde{S}_{m_1}^{(\alpha,\beta)}f\right)(x)-f(x)\right|\le 2\omega_1\left(\sqrt{\delta_{m,p,\alpha,\beta,x}}\right) \tag{2.7}$$

where:

$$\begin{aligned}\delta_{m,p,\alpha,\beta,x} &= \frac{(p-\beta)^2}{(m+\beta)^2}+\frac{m+p}{(m+\beta)^2}x(1-x)+\\ &+ \frac{2\alpha(mp-2m\beta-\beta^2)}{(m+\beta)^3}\,x+\frac{\alpha^2(3m+\beta)}{(m+\beta)^2}\end{aligned} \tag{2.8}$$

$$\beta\in\left[0,\sqrt{m^2+mp}\,\right] \tag{2.9}$$

Proof. Applying Theorem 2.2 and Lemma 2.3, follows

$$\left|\left(S_m^{(\alpha,\beta)}f\right)(x)-f(x)\right|\le\left(1+\delta^{-1}\cdot\sqrt{\delta_{m,p,\alpha,\beta,x}}\right)\omega_1(\delta)$$

for any $\delta>0$. Choosing $\delta=\sqrt{\delta_{m,p,\alpha,\beta,x}}$ in the above inequality we arrive to (2.8) and the proof ends. □

Remark 2.1. In Theorem 2.3 is expressed the order of local approximation of f by $\widetilde{S}_m^{(\alpha,\beta)}f$. For obtaining the order of global approximation, we must take in (2.8) the maximum of $\delta_{m,p,\alpha,\beta,x}$ when $x \in [0,1]$. Clearly, this maximum depends of the relations between α, β, p.

References

[1] Agratini, O., *Aproximare prin operatori liniari*, Presa Universitară Clujeană, Cluj-Napoca, 2000 (Romanian).

[2] Bărbosu, D., *A Voronovskaja type theorem for the operator of D.D. Stancu*, BULLETINS for APPLIED & COMPUTER MATHEMATICS, BAM 1998 - C/2000, T.U. Budapest (2002), 175-182.

[3] Bărbosu, D., Bărbosu, M., *Properties of the fundamental polynomials of Bernstein-Schurer* (to appear in Procced. of "icam3", International Conference on Applied Mathematics, 3-th Edition, Baia Mare - Borşa, october 10-13, 2002).

[4] Bărbosu, D., *The Voronovskaja theorem for Bernstein-Schurer operators*, (to appear in Procced. of "icam3", International Conference on Applied Mathematics, 3-th Edition, Baia Mare - Borşa, october 10-13, 2002).

[5] Bohman, H., *On approximation of continuous and analitic functions*, Ark. Mat., 2(1952), 43-56.

[6] Korovkin, P.P., *On convergence of linear positive operators in the space of continuous functions*, (Russian), Dokl. Akad. Nauk SSSR (N.S.), 90(1953), 961-964.

[7] Schurer, F., *Linear positive operators in approximation theory*, Math. Inst. Techn. Univ. Delft:Report, 1962.

[8] Shisha, O., Mond, B., *The degree of convergence of linear positive operators*, Proc. Nat. Acad. Sci. U.S.A, 60(1968), 1196-1200.

[9] Stancu D.D., *Approximation of functions by a new class of linear polynomials operators*, Rev. Roum. Math. Pures et Appl., 13(1968), No.8, 1173-1194.

[10] Stancu, D.D., Coman, Gh., Agratini, O., *Analiză numerică şi teoria aproximării*, Presa Universitară Clujeană, Cluj-Napoca, 2001 (Romanian).

North University of Baia Mare, Faculty of Sciences,
Department of Mathematics and Computer Science, Victoriei 76,
4800 Baia Mare, Romania
E-mail address: dbarbosu@ubm.ro, danbarbosu@yahoo.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

WEIGHTED UNIFORM SAMPLING METHOD BASED ON SPLINE FUNCTIONS

PETRU P. BLAGA

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. The weighted uniform sampling method to reduce of variance is investigated by using the multivariate Schoenberg spline operator on the unit hypercube. The new estimators obtained for the random numerical integration are numerically compared with the crude Monte Carlo estimators.

1. Introduction

It is known that definite integrals can be estimated by probabilistic considerations, and these are rather when multiple integrals are concerned. The integral is interpreted as the mean value of certain random variable, which is an unknown parameter. To estimate this parameter, i.e. the definite integral, one regards the sample mean of the sampling from a suitable random variable. This sample mean is an unbiased estimator for the definite integral and is referred as *the crude Monte Carlo estimator.*

Generally, this method is not fast–converging ratio to the volume of sampling, and efficiency depends on the variance of the estimator, which is expressed by the variance of the integrand. Consequently, for improving the efficiency of Monte Carlo method, it must reduce as much as possible the variance of the integrated function. There is a lot of procedures for reducing of the variance in the Monte Carlo method. In the following we approach the reducing of variance by the so–called *weighted uniform sampling method*, using the multivariate Schoenberg spline operator on the unit hypercube.

Numerical experiments are considered comparatively with the crude Monte Carlo estimates.

2. Multivariate B-spline functions

Let $\boldsymbol{D}_n = [0,1]^n$ be the n–dimensional unit hypercube. We consider the fixed vectors $\boldsymbol{m} = (m_1, \ldots, m_n)$ and $\boldsymbol{k} = (k_1, \ldots, k_n)$, whose components are integer positive values, namely $m_i > 0$ and $k_i > 1$, $i = \overline{1, n}$.

Received by the editors: 09.04.2003.

An extended rectangular partition $\boldsymbol{\Delta}$ of the domain $\boldsymbol{D}_n$ is defined by the following one–dimensional extended partitions:

$$\Delta_i \colon t_1^{(i)} = \cdots = t_{k_i}^{(i)} = 0 < t_{k_i+1}^{(i)} \leqslant \cdots \leqslant t_{k_i+m_i-1}^{(i)} < 1 = t_{k_i+m_i}^{(i)} = \cdots = t_{2k_i+m_i-1}^{(i)},$$

for all $i = \overline{1,n}$, where $t_j^{(i)} < t_{k_i+j}^{(i)}$, $j = \overline{1, k_i + m_i - 1}$.

If one denotes the multi–index set

$$\boldsymbol{J} = \left\{ \boldsymbol{j} = (j_1, \ldots, j_n) \;\middle|\; j_i = \overline{1, m_i + 2k_i - 1},\, i = \overline{1,n} \right\},$$

the partition $\boldsymbol{\Delta}$ is given by the cartesian product

$$\boldsymbol{\Delta} = \Delta_1 \times \cdots \times \Delta_n = \left\{ \boldsymbol{t_j} = \left(t_{j_1}^{(1)}, \ldots, t_{j_n}^{(n)} \right) \;\middle|\; \boldsymbol{j} \in \boldsymbol{J} \right\}.$$

The points of $\boldsymbol{\Delta}$ are called *knots* of the partition.

Using the knots of the partition $\boldsymbol{\Delta}$, one defines *the* (n*–variate*) *B–spline functions*

$$\boldsymbol{M_{i,k}}(\boldsymbol{x}) = M_{i_1,k_1}^{(1)}(x_1) \cdots M_{i_n,k_n}^{(n)}(x_n), \quad \boldsymbol{x} = (x_1, \ldots, x_n) \in \boldsymbol{D}_n, \tag{1}$$

for every multi–index $\boldsymbol{i} = (i_1, \ldots, i_n) \in \boldsymbol{I}$, where

$$\boldsymbol{I} = \left\{ \boldsymbol{i} = (i_1, \ldots, i_n) \;\middle|\; i_j = \overline{1, m_j + k_j - 1},\, j = \overline{1,n} \right\}.$$

The factors from the right side of the formula (1) are *the* (*one–variate*) *B–spline functions*, i.e.

$$M_{i_j,k_j}^{(j)}(x_j) = \left[t_{i_j}^{(j)}, \ldots, t_{i_j+k_j}^{(j)}; k_j \left(t - x_j\right)_+^{k_j-1} \right], \; i_j = \overline{1, m_j + k_j - 1}, \; j = \overline{1,n},$$

where $[z_0, z_1, \ldots, z_r; f(t)]$ denotes the r-th divided difference relative to the knots z_0, $z_1, \ldots, z_r$ of the function $f(t)$.

The normalized (n*–variate*) *B–spline* functions are defined by

$$\boldsymbol{N_{i,k}}(\boldsymbol{x}) = N_{i_1,k_1}^{(1)}(x_1) \cdots N_{i_n,k_n}^{(n)}(x_n), \quad \boldsymbol{x} = (x_1, \ldots, x_n) \in \boldsymbol{D}_n, \tag{2}$$

for every $\boldsymbol{i} = (i_1, \ldots, i_n) \in \boldsymbol{I}$, where

$$N_{i_j,k_j}^{(j)}(x_j) = \frac{t_{i_j+k_j}^{(j)} - t_{i_j}^{(j)}}{k_j} M_{i_j,k_j}^{(j)}(x_j)$$

are *the* (*one–variate*) *normalized B–spline functions.* We recall the following properties of B–spline functions:

(i) $\boldsymbol{N_{i,k}}(\boldsymbol{x}) \geqslant 0$,

(ii) $\boldsymbol{N_{i,k}}(\boldsymbol{x}) = A_{\boldsymbol{i,k}} \boldsymbol{M_{i,k}}(\boldsymbol{x}), \quad A_{\boldsymbol{i,k}} = \displaystyle\prod_{j=1}^{n} \frac{t_{i_j+k_j}^{(j)} - t_{i_j}^{(j)}}{k_j}$,

(iii) $\displaystyle\sum_{\boldsymbol{i} \in \boldsymbol{I}} \boldsymbol{N_{i,k}}(\boldsymbol{x}) = 1$,

(iv) $\displaystyle\int_{\boldsymbol{D}_n} \boldsymbol{M_{i,k}}(\boldsymbol{x})\, \boldsymbol{dx} = 1$.

3. Schoenberg spline operator

Using the knots of partition $\boldsymbol{\Delta}$, one defines *the nodes*

$$\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}} = \left(\xi^{(1)}_{i_1,k_1}, \ldots, \xi^{(n)}_{i_n,k_n}\right), \quad \boldsymbol{i} = (i_1, \ldots, i_n) \in \boldsymbol{I},$$

where

$$\xi^{(j)}_{i_j,k_j} = \frac{t^{(j)}_{i_j+1} + \cdots + t^{(j)}_{i_j+k_j-1}}{k_j - 1}, \quad i_j = \overline{1, m_j + k_j - 1},\ j = \overline{1,n}.$$

We remark that $0 = \xi^{(j)}_{1,k_j} < \xi^{(j)}_{2,k_j} < \cdots < \xi^{(j)}_{m_j+k_j-1,k_j} = 1, \quad j = \overline{1,n}$, and consequently the nodes $\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}$, $\boldsymbol{i} \in \boldsymbol{I}$, belong to $\boldsymbol{D}_n$.

The (n*–variate*) *Schoenberg spline* operator relative to a real function f defined on $\boldsymbol{D}_n$ is given by

$$S_{\boldsymbol{\Delta}}(f)(\boldsymbol{x}) = (S_{\boldsymbol{\Delta}} f)(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \boldsymbol{I}} \boldsymbol{N}_{\boldsymbol{i},\boldsymbol{k}}(\boldsymbol{x})\, f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right), \quad \boldsymbol{x} \in \boldsymbol{D}_n. \tag{3}$$

Some important properties are recalled here:

(i) $S_{\boldsymbol{\Delta}}(f)(\boldsymbol{x})$ is a polynomial spline of degree k_i in the i-th variable,

(ii) $S_{\boldsymbol{\Delta}}(f)$ defines a positive linear operator,

(iii) If $m_i = 1$, then $S_{\boldsymbol{\Delta}}(f)(\boldsymbol{x})$ is a polynomial of degree $k_i - 1$ in the i-th variable, and consequently if $m_i = 1$, for all $i = \overline{1,n}$, the $S_{\boldsymbol{\Delta}}(f)(\boldsymbol{x})$ is the multivariate Bernstein polynomial,

(iv) $S_{\boldsymbol{\Delta}}(f) = f$, for all $f(\boldsymbol{x}) = x_1^{s_1} \ldots x_n^{s_n}, \quad s_i = 0, 1,\ i = \overline{1,n}$,

(v) If $f \in C(\boldsymbol{D}_n)$, then $S_{\boldsymbol{\Delta}}(f)$ converges uniformly to the function f as $\frac{\|\Delta_1\|}{k_1} + \cdots + \frac{\|\Delta_n\|}{k_n} \to 0$, where $\|\Delta_i\|$ denotes the norm of the partition Δ_i.

Taking into account that

$$N^{(j)}_{i_j,k_j}(0) = \delta_{1,i_j}, \quad N^{(j)}_{i_j,k_j}(1) = \delta_{m_j+k_j-1,i_j}, \quad i_j = \overline{1, m_j + k_j - 1},\ j = \overline{1,n},$$

where $\delta_{r,s}$ denotes the Kronecker symbol, we have $S_{\boldsymbol{\Delta}}(f)(\boldsymbol{e}) = f(\boldsymbol{e})$, for all the vertices $\boldsymbol{e}$ of the hypercube $\boldsymbol{D}_n$.

4. Crude Monte Carlo method

Let $\boldsymbol{X}$ be an n–dimensional random variable having the probability density function $\rho\colon \mathbb{R}^n \to \mathbb{R}$. In the random numerical integration the multidimensional integral

$$I[\rho; f] = \int_{\mathbb{R}^n} \rho(\boldsymbol{x})\, f(\boldsymbol{x})\, \boldsymbol{dx} \tag{4}$$

is interpreted as the mean value of the random variable $f(\boldsymbol{X})$, where $f\colon \mathbb{R}^n \to \mathbb{R}$ usually belongs to $L^2_\rho(\mathbb{R}^n)$, in other words $\int_{\mathbb{R}^n} \rho(\boldsymbol{x})\, f^2(\boldsymbol{x})\, \boldsymbol{dx}$ exists, and therefore the mean value $I[\rho; f]$ exists.

Using a basic statistical technique, the mean value given by (4) can be estimated by taking N independent samples (random numbers) $\boldsymbol{x}_i$, $i = \overline{1,N}$, with the probability density function ρ. These random numbers are regarded as values of the independent identically distributed random variables $\boldsymbol{X}_i$, $i = \overline{1,N}$, i.e. sample variables with the common probability function ρ.

We use the same notation $\bar{I}_N[\rho;f]$ for the sample mean of random variables $f(\boldsymbol{X}_i)$, $i=\overline{1,N}$, and respectively for its value, i.e.

$$\bar{I}_N[\rho;f]=\frac{1}{N}\sum_{i=1}^{N}f(\boldsymbol{X}_i),$$

$$\bar{I}_N[\rho;f]=\frac{1}{N}\sum_{i=1}^{N}f(\boldsymbol{x}_i).$$

The estimator $\bar{I}_N[\rho;f]$ satisfies the following properties:

$$E\left(\bar{I}_N[\rho;f]\right)=I[\rho;f],\quad \text{(unbiased estimator of } I[\rho;f]\text{)},$$

$$Var\left(\bar{I}_N[\rho;f]\right)\to 0,\quad N\to\infty,$$

$$\bar{I}_N[\rho;f]\to I[\rho;f],\quad N\to\infty,\quad \text{(with probability 1)}.$$

Taking into account these results, the crude Monte Carlo integration formula is defined by

$$I[\rho;f]=\int_{\mathbb{R}^n}\rho(\boldsymbol{x})f(\boldsymbol{x})\,\boldsymbol{dx}\approxeq\frac{1}{N}\sum_{i=1}^{N}f(\boldsymbol{x}_i).\tag{5}$$

It must remark that in (4) the domain of integration is only apparently the whole n–dimensional Euclidean space. Thus, it is possible that the density $\rho(\boldsymbol{x})=0$, $\boldsymbol{x}\notin\boldsymbol{D}$, where $\boldsymbol{D}$ is a region of the n–dimensional Euclidean space $\mathbb{R}^n$, therefore the integral (4) becomes

$$I[\rho;f]=\int_{\boldsymbol{D}}\rho(\boldsymbol{x})f(\boldsymbol{x})\,\boldsymbol{dx},$$

and the crude Monte Carlo method must be interpreted in an appropriate manner.

5. Weighted uniform sampling method

This method was given in [8], reconsidered in [11], and recently in [6] it was compared with other methods for reducing of the variance.

Let us consider the integral

$$I[f]=\int_{\boldsymbol{D}}f(\boldsymbol{x})\,\boldsymbol{dx}=V\int_{\boldsymbol{D}}\frac{1}{V}f(\boldsymbol{x})\,\boldsymbol{dx},$$

where $\boldsymbol{D}\subset\mathbb{R}^n$ and $V=\text{Volume}(\boldsymbol{D})$.

The crude Monte Carlo estimator for $I[f]$ is

$$\bar{I}_N^c[f]=\frac{V}{N}\sum_{i=1}^{N}f(\boldsymbol{X}_i),$$

with the sampling variables $\boldsymbol{X}_i$, $i=\overline{1,N}$, independent uniformly in the region $\boldsymbol{D}$, i.e. these have the common density probability function

$$\rho(\boldsymbol{x})=\begin{cases}\dfrac{1}{V}, & \text{if } \boldsymbol{x}\in\boldsymbol{D},\\[2ex] 0, & \text{if } \boldsymbol{x}\notin\boldsymbol{D}.\end{cases}$$

The method of weighted uniform sampling consists in the considering a function $g\colon \boldsymbol{D} \to \mathbb{R}$ such that

$$\int_{\boldsymbol{D}} g(\boldsymbol{x})\,\boldsymbol{dx} = 1,$$

and the corresponding sampling function

$$\bar{I}_N^w[g;f] = \left(\sum_{i=1}^{N} f(\boldsymbol{X}_i)\right) \Big/ \left(\sum_{i=1}^{N} g(\boldsymbol{X}_i)\right),$$

where $\boldsymbol{X}_i$, $i = \overline{1,N}$, are the same above sampling variables.

If one denotes by Θ_N and $\tilde{\Theta}_N$, the sample means of $f(\boldsymbol{X}_i)$ and respectively of $g(\boldsymbol{X}_i)$, $i = \overline{1,N}$, we have

$$\bar{I}_N^w[g;f] = \frac{\Theta_N}{\tilde{\Theta}_N} \cdot$$

Taking into account that the sample means

$$\Theta_N = \frac{1}{N}\sum_{i=1}^{N} f(\boldsymbol{X}_i) = \bar{I}_N^c[f]/V,$$

$$\tilde{\Theta}_N = \frac{1}{N}\sum_{i=1}^{N} g(\boldsymbol{X}_i) = \bar{I}_N^c[g]/V$$

are unbiased estimators for $I[f]$ and $I[g] = 1$ respectively, it results that

$$\frac{E(\Theta_N)}{E\left(\tilde{\Theta}_N\right)} = I[f] \cdot$$

However $\bar{I}_N^w[g;f]$ is a biased estimator for $I[f]$, satisfying only asymptotical relation

$$E\left(\bar{I}_N^w[g;f]\right) \approxeq I[f].$$

For the variance of the estimator $\bar{I}_N^w[g;f]$ we have [6]:

$$Var(\bar{I}_N^w[g;f]) = \frac{V^2}{N} Var[f(\boldsymbol{X}) - I[f]\,g(\boldsymbol{X})] + O\left(\frac{1}{N^2}\right),$$

that is

$$Var(\bar{I}_N^w[g;f]) \approxeq \frac{V^2}{N} Var[f(\boldsymbol{X}) - I[f]\,g(\boldsymbol{X})].$$

On the other hand we have

$$Var(\bar{I}_N^c[f]) = \frac{V^2}{N} Var[f(\boldsymbol{X})] \cdot$$

In this manner, the comparison of the variances of the two estimators $\bar{I}_N^c[f]$ and $\bar{I}_N^w[g;f]$ is reduced to compare the variances $Var[f(\boldsymbol{X})]$ and $Var[f(\boldsymbol{X}) - I[f]\,g(\boldsymbol{X})]$.

Because

$$Var[f(\boldsymbol{X})] - Var[f(\boldsymbol{X}) - I[f]\,g(\boldsymbol{X})] = I[f]\left[2Cov(f(\boldsymbol{X}), g(\boldsymbol{X})) - I[f]\,\sigma^2[g]\right],$$

the covariance $Cov\left(f\left(\boldsymbol{X}\right), g\left(\boldsymbol{X}\right)\right)$ controls the magnitude of difference of the two variances, it must that g has the same monotonicity as f.

In the following we consider the function g from the weighted uniform sampling method given by the multivariate spline function corresponding to the integrand f defined by Schoenberg spline operator.

If the integration region is the unit hypercube $\boldsymbol{D}_n$, the crude Monte Carlo estimator is

$$\bar{I}_N^c\left[f\right] = \frac{1}{N}\sum_{i=1}^{N} f\left(\boldsymbol{X}_i\right),$$

where the sampling variables $\boldsymbol{X}_i$, $i = \overline{1, N}$, are independent uniformly distributed on $\boldsymbol{D}_n$.

The function g from presented weighted uniform sampling method is the following

$$g\left(\boldsymbol{x}\right) = K \cdot \left(S_{\boldsymbol{\Delta}} f\right)\left(\boldsymbol{x}\right),$$

where $\left(S_{\boldsymbol{\Delta}} f\right)\left(\boldsymbol{x}\right)$ is given by (3) and the constant K is such that

$$\int_{\boldsymbol{D}_n} g\left(\boldsymbol{x}\right) \boldsymbol{dx} = 1.$$

From this condition we have that

$$g\left(\boldsymbol{x}\right) = \frac{\left(S_{\boldsymbol{\Delta}} f\right)\left(\boldsymbol{x}\right)}{\sum_{\boldsymbol{i}\in\boldsymbol{I}} A_{\boldsymbol{i},\boldsymbol{k}} f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right)} = \frac{\sum_{\boldsymbol{i}\in\boldsymbol{I}} \boldsymbol{N}_{\boldsymbol{i},\boldsymbol{k}}\left(\boldsymbol{x}\right) f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right)}{\sum_{\boldsymbol{i}\in\boldsymbol{I}} A_{\boldsymbol{i},\boldsymbol{k}} f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right)}.$$

Finally, the random numerical integration formula is given by

$$\bar{I}_N^w\left[g; f\right] = \frac{\left[\sum_{\boldsymbol{i}\in\boldsymbol{I}} A_{\boldsymbol{i},\boldsymbol{k}} f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right)\right]\left[\sum_{i=1}^{N} f\left(\boldsymbol{x}_i\right)\right]}{\sum_{i=1}^{N}\left[\sum_{\boldsymbol{i}\in\boldsymbol{I}} \boldsymbol{N}_{\boldsymbol{i},\boldsymbol{k}}\left(\boldsymbol{x}_i\right) f\left(\boldsymbol{\xi}_{\boldsymbol{i},\boldsymbol{k}}\right)\right]}. \tag{6}$$

The random points $\boldsymbol{x}_i$, $i = \overline{1, N}$, are independent uniformly distributed in the hypercube $\boldsymbol{D}_n$.

We must remark that the spline functions give a more flexible method than Bernstein polynomials, which have been used in [3], for the same uniform sampling method. This is because the nodes in the Schoenberg spline operator are not necessarily equidistant, like in the Bernstein operator. Consequently, if some smoothness informations for the integrand f are known, we can require more nodes in the domain of integration where the function f has a bed smoothness.

6. Numerical experiments

Numerical examples are considered in the unidimensional ($n = 1$) and bidimensional ($n = 2$) cases for the estimator (6) with the integrand f given by $f\left(x\right) = \frac{1}{1+x}$ and $f\left(x, y\right) = \frac{1}{1+x+y}$ respectively. The interior knots for the variable x are $0.1, 0.3, 0.5, 0.7, 0.9$, and $0.2, 0.5, 0.8$ for the variable y. The numerical results comprised in the following two tables compare the estimates obtained by the weighted uniform sampling technique and the crude Monte Carlo method.

Each table contains: the sampling volume N, the orders k (or k_1, k_2) of the spline functions, the estimates given by the two methods $\bar{I}_N^c[f]$ and $\bar{I}_N^w[g;f]$, the error estimates Err^c and Err^w respectively, and the ratio Err^c/Err^w. We also remark that the estimations and the error estimates from each row of tables represent the mean values in one hundred of samplings.

$$I[f] = \log 2 = 0.69314718...$$

N	k	$\bar{I}_N^c[f]$	$\bar{I}_N^w[g;f]$	Err^c	Err^w	Err^c/Err^w
50	2	0.6911300	0.6931458	2.017e-03	1.394e-06	1447.4
100	2	0.6910238	0.6931734	2.123e-03	2.621e-05	81.0
300	2	0.6921497	0.6931548	9.975e-04	7.649e-06	130.4
500	2	0.6922726	0.6931520	8.745e-04	4.836e-06	180.9
50	3	0.6911300	0.6931688	2.017e-03	2.161e-05	93.3
100	3	0.6910238	0.6931799	2.123e-03	3.274e-05	64.9
300	3	0.6921497	0.6931633	9.975e-04	1.609e-05	62.0
500	3	0.6922726	0.6931567	8.745e-04	9.477e-06	92.3
50	5	0.6911300	0.6931639	2.017e-03	1.674e-05	120.5
100	5	0.6910238	0.6931832	2.123e-03	3.598e-05	59.0
300	5	0.6921497	0.6931694	9.975e-04	2.219e-05	44.9
500	5	0.6922726	0.6931605	8.745e-04	1.336e-05	65.5

$$I[f] = \log \frac{27}{16} = 0.5232481...$$

N	k_1	k_2	$\bar{I}_N^c[f]$	$\bar{I}_N^w[g;f]$	Err^c	Err^w	Err^c/Err^w
50	2	2	0.5208017	0.5233035	2.446e-03	5.539e-05	44.2
100	2	2	0.5216731	0.5232890	1.575e-03	4.082e-05	38.6
300	2	2	0.5225268	0.5232643	7.214e-04	1.615e-05	44.7
50	2	3	0.5208017	0.5233042	2.446e-03	5.610e-05	43.6
100	2	3	0.5216731	0.5232827	1.575e-03	3.452e-05	45.6
300	2	3	0.5225268	0.5232604	7.214e-04	1.227e-05	58.8
50	2	4	0.5208017	0.5233092	2.446e-03	6.108e-05	40.1
100	2	4	0.5216731	0.5232843	1.575e-03	3.618e-05	43.5
300	2	4	0.5225268	0.5232619	7.214e-04	1.380e-05	52.3
50	3	4	0.5208017	0.5233095	2.446e-03	6.133e-05	39.9
100	3	4	0.5216731	0.5232821	1.575e-03	3.393e-05	46.4
300	3	4	0.5225268	0.5232607	7.214e-04	1.260e-05	57.3
50	4	4	0.5208017	0.5233096	2.446e-03	6.146e-05	39.8
100	4	4	0.5216731	0.5232812	1.575e-03	3.310e-05	47.6
300	4	4	0.5225268	0.5232622	7.214e-04	1.403e-05	51.4
50	5	5	0.5208017	0.5233079	2.446e-03	5.980e-05	40.9
100	5	5	0.5216731	0.5232791	1.575e-03	3.097e-05	50.8
300	5	5	0.5225268	0.5232640	7.214e-04	1.582e-05	45.6

References

[1] Blaga, P. P., *Monte Carlo integration on simplex*, Studia Univ. "Babeş–Bolyai", Mathematica, Vol. **XXXIII**, no. 3 (1988), 19–26.
[2] Blaga, P. P., *Reducing of variance by spline functions in Monte Carlo integration*, Studia Univ. "Babeş–Bolyai", Mathematica, Vol. **XXXIV**, no. 4 (1989), 69–78.
[3] Blaga, P. P., *Bernstein operators to reduce of variance in random numerical integration*, *in* Proceedings of the International Symposium on Numerical Analysis and Approximation Theory, Cluj–Napoca, May 9–11, 2002, Cluj University Press, pp. 76–91.
[4] de Boor, C., *A practical guide to splines*, Springer, Berlin–Heidelberg–New York, 1978.
[5] Coman, Gh., Frenţiu, M., *Bivariate spline approximation*, Studia Univ. "Babeş–Bolyai", Mathematica, Vol. **XIX**, no. 1 (1974), 59–64.
[6] Gorbacheva, N. B., Sobol´, I. M., Trikuzov, A. I., *Variance–reducing multipliers in the computation of integrals by the Monte Carlo method* (Russian), Zh. Vychisl. Mat. Mat. Fiz, Vol. **41** (2001), 1310–1314.
[7] Hammersley, J. M., Handscomb, D. C., *Monte Carlo Methods*, Methuen–John Wiley and Sons, London–New York, 1964.
[8] Handscomb, D. C., *Remarks on a Monte Carlo integration method*, Numer. Math., Vol. **6** (1964), 261–268.
[9] Marsden, M. J., Schoenberg, I. J., *On variation diminishing spline approximation*, Mathematica (Cluj), Vol. **8(31)** (1966), 61–82.
[10] Micula, Gh., *Funcţii spline şi aplicaţii*, Editura Tehnică, Bucureşti, 1978.
[11] Powell, M. J. D., Swann, J., *Weighted uniform sampling–a Monte Carlo technique for reducing variance*, J. Inst. Math. Appl., Vol. **2** (1966), 228–236.
[12] Schoenberg, I. J., *On spline functions*, in "Inequalities" (O. Shisha ed.), Symposium at Wright–Paterson Air Force Base, Academic Press, New York, 1967, pp. 255–291.
[13] Stancu, D. D., Coman, Gh., Blaga, P., *Numerical Analysis and Approximation Theory*, Vol. II, Cluj University Press, 2002.

"Babeş–Bolyai" University, Cluj–Napoca

ABOUT SOME VOLTERRA PROBLEMS SOLVED BY A PARTICULAR SPLINE COLLOCATION

F. CALIO, E. MARCHETTI, R. PAVANI, AND GH. MICULA

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. In this paper we propose a deficient spline collocation method for a special Volterra integral equation problem. The existence and uniqueness of the approximating spline are investigated. Some numerical examples illustrate the efficiency of the proposed numerical method.

1. **Introduction**

The theory and applications of the Volterra integral equations of the form

$$y(x) = \int_0^x K(x,t,y(t))dt + g(x), \quad x \in [0,T]$$

is an important subject within applied mathematics. Volterra integral equations are used as mathematical models for many and varied physical phenomena and processes but they occur as reformulations of other mathematical problems.

In the present work we consider the following Volterra equation with constant delay $\tau > 0$:

$$y(x) \ = \int_0^x K_1(x,t,y(t))dt + \int_0^{x-\tau} K_2(x,t,y(t))dt + g(x), x \in J = [0,T] \qquad (1)$$

with $y(x) = \varphi(x)$, $x \in [\tau, 0)$.

Equation (1) is worth studying as it is frequently encountered in physical and biological modeling processes (e.g. [5]).

We assume that the given functions

$$\varphi : [-\tau, 0] \to \mathbb{R}, g : J \to \mathbb{R}, K_1 : \Omega \times \mathbb{R} \to \mathbb{R}(\Omega := [(x,t) : 0 \leq t \leq x \leq T]),$$
$$K_2 : \Omega_\tau \times \mathbb{R} \to \mathbb{R}(\Omega_\tau := J \times [-\tau, T-\tau])$$

are at least continuous on their domains such that (1) possesses a unique solution $y \in C(J)$, and $\varphi \in C^{m-2}[-\tau, 0]$, $g \in C^{m-2}[0,T]$.

In the following, let us assume in (1) that :

1. K_1 satisfies the following Lipschitz condition :

$$\| K_1(x,t,y_1) - K_1(x,t,y_2) \| \leq L_1 \| y_1 - y_2 \|$$

Received by the editors: 01.07.2003.
2000 *Mathematics Subject Classification.* 65R20, 65D07.

$$\forall(x,t,y_1),(x,t,y_2) \in J \times J \times \mathbb{R}, t \le x$$

with $L_1 \ge 0$ a constant independent of x and t.

2. K_2 satisfies the following Lipschitz condition :

$$\| K_2(x,t,z_1) - K_2(x,t,z_2) \| \le L_2 \| z_1 - z_2 \|$$

$$\forall(x,t,z_1),(x,t,z_2) \in J \times J \times \mathbb{R}, t \le x$$

with $L_2 \ge 0$ also a constant independent of x and t.

Recently, various aspects of numerical methods for (1) have been studied from the point of view of polynomial collocation methods (e.g. [1], [2], [8]). In this context here we propose to approximate the solution of (1) by means of functions pertaining to the class of splines $s: J \to \mathbb{R}$,($s \in \mathcal{S}_m$, $s \in C^{m-2}$) of degree $m \ge 2$ and deficiency $d \ge 2$.

We already used an analogous deficient spline collocation method in the case of delay differential equations [3], [4]. As it revealed simple and efficient, here we propose to extend it to Volterra integral equations with delay argument to provide an alternative to the discrete collocation method proposed in [1]. Indeed our method presents some advantages:

- it does not require any additional initial value
- it provides a global approximation of the solution
- in case of need, the lenght of each collocation step can be modified, and similarly the degree m of the used spline functions and deficiency d
- the proposed numerical method reveals extremely easy to be implemented in the linear case.

We emphasize that this method is peculiar for solutions belonging to low regularity class. Indeed we use $m = 2$ or $m = 3$ only; so that the used splines are $s \in C^0$ or at most $s \in C^1$; there is numerical evidence that it suffices in order to approximate solutions belonging to class C^0 or C^1.

In Section 2 we present the numerical method to approximate the solution by collocation of deficient spline functions; Section 3 is devoted to theoretical results referring to existence and unicity of the numerical solution and there we recall also some results about convergence and numerical stability. In the last Section we report some examples relating to integral equations with solutions characterized by low regularity.

2. **Construction of approximating spline solution**

In this section we describe the numerical model used to approximate the solution of (1).

Firstly we shall construct a polynomial spline function of degree $m > 1$, which we denote by s. On the interval $J := [0,T]$ the spline function s is defined in $[t_k, t_{k+1}]$ where $t_k := t_0 + kh$, $k = 0, 1, \cdots, N$; $t_0 := 0$, $t_N = T$, $h := \frac{T}{N}$ as :

$$s_k(t) := \sum_{j=0}^{m-2} \frac{s_{k-1}^{(j)}(t_k)}{j!}(t-t_k)^j + \frac{a_k}{(m-1)!}(t-t_k)^{m-1} + \frac{b_k}{m!}(t-t_k)^m$$

We choose to determine coefficients a_k, b_k by the following system of collocation conditions:

$$s_k(t_k+\frac{h}{2})=\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}K_1(t_k+\frac{h}{2},t,s_j(t))dt+\int_{t_k}^{t_k+\frac{h}{2}}K_1(t_k+\frac{h}{2},t,s_k(t))dt+$$
$$+\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}-\tau}K_2(t_k+\frac{h}{2},t,s_{j-1}(t))dt+\int_{t_k}^{t_k+\frac{h}{2}-\tau}K_2(t_k+\frac{h}{2},t,s_{k-1}(t))dt+g(t_k+\frac{h}{2})$$
$$s_k(t_{k+1})=\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}K_1(t_{k+1},t,s_j(t))dt+\int_{t_k}^{t_{k+1}}K_1(t_{k+1},t,s_k(t))dt+\qquad(2)$$
$$+\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}-\tau}K_2(t_{k+1},t,s_{j-1}(t))dt+\int_{t_k}^{t_{k+1}-\tau}K_2(t_{k+1},t,s_{k-1}(t))dt+g(t_{k+1})$$

provided that

$$s_{-1}(0)=y(0)=\varphi(0),s'_{-1}(0)=y'(0)=\varphi'(0),\cdots,s_{-1}^{(m-2)}(0)=y^{(m-2)}(0)=\varphi^{(m-2)}(0)$$

Our model is thus reduced to compute the solution of the non-linear system (2), through which the spline is globally determined on the interval J.

3. The theoretical results

It remains to prove that for h sufficiently small, the parameters a_k, b_k, $0\le k\le N-1$ can be uniquely determined from (2).

Theorem. *Let as consider the Volterra equation* (1)*. If K_1 and K_2 satisfy the hypotheses* 1 *and* 2*, and if h is small enough, then there exists a unique spline solution s of* (1) *given by the above construction.*

Proof. If we set

$$A_k(t)=\sum_{j=0}^{m-2}\frac{s_{k-1}^{(j)}(t_k)}{j!}(t-t_k)^j$$

then (2) becomes:

$$a_k = \left(\frac{2}{h}\right)^{m-1}(m-1)![-A_k(t_k+\frac{h}{2}) - \frac{b_k}{m!}\left(\frac{h}{2}\right)^m +$$
$$+\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}} K_1(t_k+\frac{h}{2}, t, A_j(t) + \frac{a_j}{(m-1)!}(t-t_j)^{m-1} + \frac{b_j}{m!}(t-t_j)^m)dt +$$
$$+\int_{t_k}^{t_k+\frac{h}{2}} K_1(t_k+\frac{h}{2}, t, A_k(t) + \frac{a_k}{(m-1)!}(t-t_k)^{m-1} + \frac{b_k}{m!}(t-t_k)^m)dt +$$
$$+\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}-\tau} K_2(t_k+\frac{h}{2}, t, A_{j-1}(t) + \frac{a_{j-1}}{(m-1)!}(t-t_{j-1})^{m-1} + \frac{b_{j-1}}{m!}(t-t_{j-1})^m)dt +$$
$$+\int_{t_k}^{t_k+\frac{h}{2}-\tau} K_2(t_k+\frac{h}{2}, t, A_{k-1}(t) + \frac{a_{k-1}}{(m-1)!}(t-t_{k-1})^{m-1} + \frac{b_{k-1}}{m!}(t-t_{k-1})^m)dt + g(t_k+\frac{h}{2})]$$
$$b_k = \frac{m!}{h^m}[-A_k(t_{k+1}) - a_k\frac{h^{m-1}}{(m-1)!} +$$
$$+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} K_1(t_{k+1}, t, A_j(t) + \frac{a_j}{(m-1)!}(t-t_j)^{m-1} + \frac{b_j}{m!}(t-t_j)^m)dt +$$
$$+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}-\tau} K_2(t_{k+1}, t, A_{j-1}(t) + \frac{a_{j-1}}{(m-1)!}(t-t_{j-1})^{m-1} + \frac{b_{j-1}}{m!}(t-t_{j-1})^m)dt + g(t_{k+1})]$$

thus we can deduce
$a_k = F_1(a_k, b_k)$
$b_k = F_2(a_k, b_k)$
where F_1 and F_2 are the right hand side of the above equations.

Now we define the application $F : \mathbb{R}^2 \to \mathbb{R}^2$ as $(a_k, b_k) \to F(a_k, b_k) := (F_1(a_k, b_k), F_2(a_k, b_k))$ and

$$d(F(a'_k, b'_k), F(a''_k, b''_k)) := |F_1(a'_k, b'_k) - F_1(a''_k, b''_k)| + |F_2(a'_k, b'_k) - F_2(a''_k, b''_k)|$$

At first, for $m > 1$, we have

$$|F_1(a'_k, b'_k) - F_1(a''_k, b''_k)| \le \frac{hL_1}{m}(2|a'_k - a''_k| + \frac{3h}{2(m+1)}|b'_k - b''_k|)$$

and similarly

$$|F_2(a'_k, b'_k) - F_2(a''_k, b''_k)| \le L_1(3|a'_k - a''_k| + \frac{5h}{2(m+1)}|b'_k - b''_k|)$$

and taking account that

$$|F_2(a'_k, b'_k) - F_2(a''_k, b''_k)| = |F_2(F_1(a'_k, b'_k), b'_k) - F_2(F_1(a''_k, b''_k), b''_k)|$$

from the previous relations at last it follows that

$$\begin{aligned} d(F(a'_k, b'_k), F(a''_k, b''_k)) &\leq hL_1(\tfrac{2+6L_1}{m}|a'_k - a''_k| + \\ + \tfrac{3h+9L_1h+5m}{2m(m+1)}|b'_k - b''_k|) &\leq MhL_1 d((a'_k, b'_k), (a''_k, b''_k)) \end{aligned}$$

where $M = \max\{(1+3L_1), \frac{1}{6}(\frac{3h}{2} + \frac{9h}{2}L_1 + 5)\}$. The upper bound was obtained using $m = 2$.

Therefore, for $MhL_1 < 1$, that is $h < \frac{1}{ML_1}$, F is a contraction and system (2) has a unique solution, which can be found by iterative method.

It is worth noting that these conditions can be greatly simplified for the linear case when in (1) we have $K_1(x, t, y(t)) = k_1(x, t)y(t)$ and $K_2(x, t, y(t)) = k_2(x, t)y(t)$; this case can be treated in a very simple and efficient way.

About the convergence and the numerical stability, we recall results presented in [6], where the case of integral equations without delay arguments is studied. The comprehensive investigation of the convergence will be approached elsewhere.

4. Numerical examples

In the following we present some numerical results to enlighten the features of the proposed numerical method. We emphasize that we will show examples just for the linear case and with exact solution belonging to a low regularity class, because our method is dedicated just to these cases, even though it works also for general cases.

Our computer programs are written in MATLAB5.3, which has a machine precision $\varepsilon \simeq 10^{-16}$.

Example 1.

Consider the following integral equation with delay arguments:

$$\begin{aligned} y(x) &= g(x) + \int_0^x y(s)ds - \int_0^{x-\tau} y(s)ds \\ \tau &= 1, \quad y(x) = 0 \text{ for } x \in [-1, 0] \\ g(x) &= \begin{cases} x - \frac{x^2}{2} & \text{for} \quad x \in [0, 1/2] \\ \frac{x^2}{2} - 2x + \frac{5}{4} & \text{for} \quad x \in [1/2, 1] \end{cases} \end{aligned}$$

The exact solution is:

$$y(x) = \begin{cases} x & \text{for} \quad x \in [0, 1/2] \\ 1 - x & \text{for} \quad x \in [1/2, 1] \end{cases}$$

where $y \in C^0[0, 1]$.

Using $m = 2$ and $d = 2$, we built spline $s \in C^0$. With integration step $h = 0.5$, we obtain numerical results with an error of order 10^{-15}, which means that in practice our results are exact within the machine precision.

Figure 1 refers just to the case $h = 0.5$; there solid line shows the exact solution in $[0,1]$ together with the history in $[-1,0]$; squares show the integration points and circles show intermediate points of the numerical solution computed by means of the analytical expression of spline relating to each integration interval.

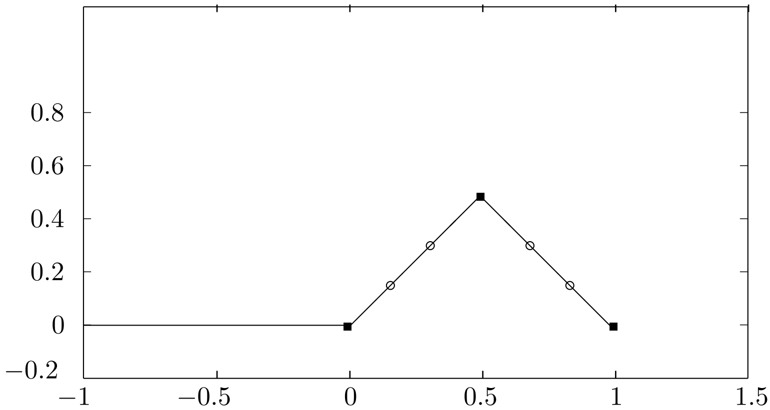


Fig. 1

It is evident that the numerical solution coincides with the exact solution.

Example 2.

Consider the following integral equation with delay arguments:

$$\begin{aligned} y(x) &= g(x) + \int_0^x y(s)ds - \int_0^{x-\tau} y(s)ds \\ \tau &= 1, \quad y(x) = 0 \text{ for } x \in [-1,0] \\ g(x) &= \begin{cases} 100x - 50x^2 & \text{for} \quad x \in [0,1/2] \\ -400(x-1)^3 + 100(x-1)^4 - \frac{75}{4} & \text{for} \quad x \in [1/2,1] \end{cases} \end{aligned}$$

The exact solution is:

$$y(x) = \begin{cases} 100x & \text{for} \quad x \in [0,1/2] \\ \\ -400(x-1)^3 & \text{for} \quad x \in [1/2,1] \end{cases}$$

Using $m = 2$ and $d = 2$, we built splines $s \in C^0$. Even in this case the solution y to be approximated belongs to class C^0, but it is the linear in the first integration subinterval only. Therefore we used a large integration step $h_1 = 0.5$ in $[0,1/2]$ and a shorter step h_2 in $[1/2,1]$.

Figure 2 refers just to the case $h_1 = 0.5$ and $h_2 = 0.125$; there solid line shows the exact solution in $[0, 1]$ together with the history in $[-1, 0]$; rectangles show the integration points and circles show intermadiate points of the numerical solution computed by means of the analytical expression of spline relating just to the first three integration intervals (for graphical convenience).

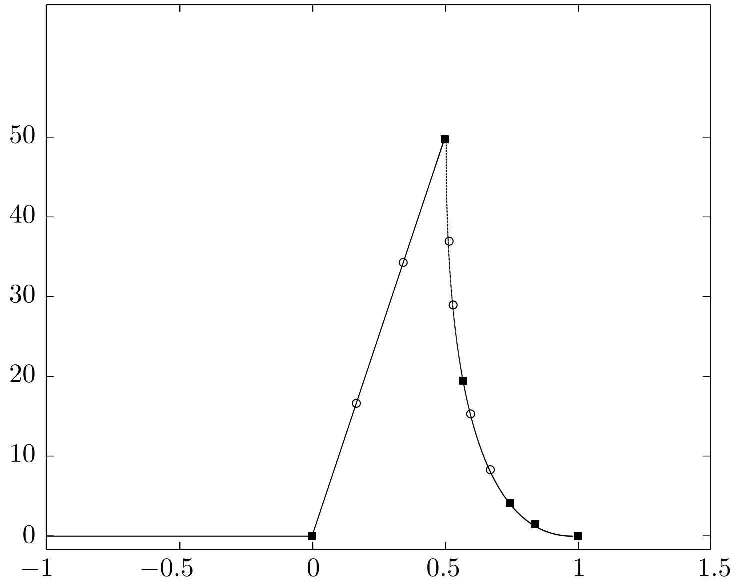


Fig. 2

It is evident that even in this case results are very satisfactory.

In more details, the numerical solution in $x = 1$ is computed with an error equal to $1.0E{-}2$ when $h_2 = 0.25$ and with an error equal to $6.6E{-}4$ when $h_2 = 0.125$.

References

[1] Brunner H., Hu Q., Lin. Q., *Geometric meshes in collocation method for Volterra integral equations with proportional delays*, IMA J. Numer. Anal., 21 (2001), 783-798.

[2] Brunner H., Yatsenko Y., *Spline collocation methods for nonlinear Volterra integral equations with unknown delay*, J. Comput. Appl., 71 (1996), 67-81.

[3] Calio F., Marchetti E., Pavani R., *Peculiar spline collocation method to solve rough and stiff delay equations*, (2003), submited.

[4] Calio F., Marchetti E., Micula G., Pavani R., *A new deficient spline functions collocation method for the second order delay diffrential equations*, Pure Math. Appl., No.1-2, 13 (2003), 97 - 109.

[5] Cooke K. L., *An epidemic equation with immigration*, Math. Biosci., 29 (1976), 135-158.
[6] Micula G., *Numerische Behandlung der Volterra - Integralgleichungen mit Splines*, Studia Univ. Babeş - Bolyai, Mathematica, XXIV, Nr.2, (1979), 46 - 54.
[7] Micula G., Micula S., *Handbook of Splines*, Kluwer Acad. Publishers, Dordrecht-Boston London, 1999 (625 pp.).
[8] Hu Q. Y., *Multilevel correction for discrete collocation solutions of Volterra integral equations with delay arguments*, Appl. Numer. Math., 31 (1999), 159-171.

Politecnico di Milano, Dipartimento di Matematica

Politecnico di Milano, Dipartimento di Matematica

Politecnico di Milano, Dipartimento di Matematica
E-mail address: raffaella.pavani@mate.polimi.it

Faculty of Mathematics, Babeş-Bolyai University,
Cluj-Napoca, Romania
E-mail address: ghmicula@math.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

MINIMUM VALUE OF A MATRIX NORM WITH APPLICATIONS TO MAXIMUM PRINCIPLES FOR SECOND ORDER PARABOLIC SYSTEMS

CRISTIAN CHIFU-OROS

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. The purpose of this paper is to use an estimation of minimum value of a matrix norm to improve some maximum principles given by I.A. Rus in 1968.

1. Introduction

Let M be the linear space of vectorial functions $u = (u_1(x,t), ..., u_n(x,t))$ which belongs to $C(\Omega)$ and are twice continuous differentiable in x and continuous differentiable in t. $\Omega \subseteq \mathbb{R}^2$ is a bounded domain. In M we consider the following system:

$$Lu := p^2 I_n \frac{\partial^2 u}{\partial x^2} + B\frac{\partial u}{\partial x} + Cu - I_n \frac{\partial u}{\partial t} = 0 \tag{1}$$

where $p \in \mathbb{R}^*, B = (b_{ij}(x,t)), C = (c_{ij}(x,t))$ are squared matrixes defined on Ω .

Let $P_o(x_o,t_o) \in \Omega$. We will denote by $S(P_o)$ the set of points Q for which there exist an arch on which the ordinate t is non-decreasing beginning with the point Q.

There are some maximum principles for the solution of system (1) (see for example [2] and [3]).

Let $u = u(x,t)$ be a solution of the system (1). In [3] the following principle is given:

Theorem 1. *Suppose that for each $(x,t) \in \Omega$, there exist $\widetilde{\beta}(x,t) \in \mathbb{R}$ such that:*

$$\xi \begin{pmatrix} -p^2 I_n & 0 \\ B(x,t) - \widetilde{\beta}(x,t) I_n & C(x,t) \end{pmatrix} \xi^* < 0, \forall \xi \in \mathbb{R}^{2n}, \xi \neq 0 \tag{2}$$

If $R(x,t) := \left(\sum\limits_{i=1}^{n} u_i^2\right)^{1/2}$ attains his maximum in $P_o \in \Omega$, then $R(Q) = R(P_o)$, for each $Q \in S(P_o)$.

Received by the editors: 01.07.2003.
2000 *Mathematics Subject Classification.* 35B50, 35C, 15A60.
Key words and phrases. Second order parabolic systems and elliptic-parabolic systems, maximum principles, estimation of the minimum value of a matrix norm.

Remark 1. *If, for each* $(x,t) \in \Omega$*, there exist* $\widetilde{\beta}(x,t) \in \mathbb{R}$ *and* $\varepsilon(x,t) > 0$ *such that:*

(i) $\xi C(x,t)\xi^* < -\left(\frac{\varepsilon(x,t)}{p}\right)^2 \|\xi\|^2, \forall \xi \in \mathbb{R}^n, \xi \neq 0;$

(ii) $\left\|B(x,t) - \widetilde{\beta}(x,t)I_n\right\|_2 \leq 2\varepsilon(x,t),$

where $\|\cdot\|_2$ *is the spectral norm, then (2) holds.*

The aim of this paper is to give some conditions which imply (ii).

Let $A \in M_n(\mathbb{R}), J$ the Jordan normal form of A. We know that there exist a nonsingular matrix T such that $A = TJT^{-1}$.

We shall denote:

$$\begin{aligned}
\widetilde{\alpha} &= \begin{cases} \frac{1}{n}\sum\limits_{k=1}^{s} n_k\lambda_k, \lambda_k \in \mathbb{R} \\ \frac{1}{n}\sum\limits_{k=1}^{s} n_k \; Re\lambda_k, \lambda_k \in \mathbb{C}\backslash\mathbb{R} \end{cases} \\
\gamma_F &= \|T\|_F \cdot \left\|T^{-1}\right\|_F \\
m_F &= \|J - \widetilde{\alpha}I_n\|_F
\end{aligned}$$

where λ_k are the eigenvalues of A, n_k is the number of λ_k which appears in Jordan blocks (generated by λ_k) and $\|\cdot\|_F$ is the euclidean norm of a matrix (see [1]).

We shall use the following result given in [1]:

Theorem 2. *Let* $\varphi_{\|\cdot\|} : \mathbb{R} \to \mathbb{R}, \varphi_{\|\cdot\|}(\alpha) = \|A - \alpha I_n\|$, $\|\cdot\|$ *being one of the following norms:* $\|\cdot\|_F$, $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$. *In these conditions:*

$$\varphi_{\|\cdot\|}(\widetilde{\alpha}) \leq \sqrt{n}\gamma_F m_F$$

In section 2 of this paper we shall give the main result in case of system 1 and in section 3, using the same instrument, we shall try to improve a maximum principle in case of elliptic-parabolic systems.

2. **Main result in parabolic case**

Using Theorem 2 and choosing $\varepsilon(x,t) = \frac{1}{2}\sqrt{n}\gamma_F m_F$, Theorem 1 becomes:

Theorem 3. *Suppose that* $\xi C(x,t)\xi^* < -\frac{1}{4p^2}n\gamma_F^2 m_F^2 \|\xi\|^2, \forall \xi \in \mathbb{R}^n, \xi \neq 0,$ $\forall (x,t) \in \Omega$*. If* $R(x,t) = \left(\sum\limits_{i=1}^{n} u_i^2\right)^{1/2}$ *attains his maximum in* $P_o \in \Omega$*, then* $R(Q) = R(P_o)$*, for each* $Q \in S(P_o)$*.*

Example 1. *Let us consider the system (1) with* $B = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_1 \end{pmatrix}$ *and without restraining the generality we shall suppose that* $a_2, a_3 > 0$*. In this case we shall have:* $\widetilde{\beta} = a_1$, $\varepsilon = a_2 + a_3$ *and:*

$$\|B - a_1 I_2\|_2 \leq \|B - a_1 I_2\|_F = \sqrt{a_2^2 + a_3^2} < 2(a_2 + a_3) = \sqrt{2}\gamma_F m_F = 2\varepsilon$$

$$\xi C(x,t)\xi^* < -\frac{1}{p^2}(a_2 + a_3)^2 \|\xi\|^2 \tag{3}$$

So if (3) holds than we have:

$$\xi \begin{pmatrix} -p^2 I_2 & 0 \\ B - \widetilde{\beta} I_2 & C \end{pmatrix} \xi^* < \frac{1}{4p^2} \left[a_2^2 + a_3^2 - 4\left(a_2 + a_3\right)^2 \right] \|\xi'\|^2 < 0$$

where $\xi = (\xi_1, \xi_2, \xi_3, \xi_4) \in \mathbb{R}^4, \xi \neq 0, \xi' = (\xi_3, \xi_4) \in \mathbb{R}^2, \xi' \neq 0$.

3. Elliptic-parabolic case

Let us consider now the following system:

$$Lu := \frac{\partial^2 u}{\partial x^2} + y^p \frac{\partial^2 u}{\partial y^2} + A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} + Cu = 0 \tag{4}$$

where L is defined in $M = C^{2,n}(\Omega) \cap C^{0,n}\left(\overline{\Omega}\right)$, $A = (a_{ij}(x, y)), B = (b_{ij}(x, y))$, $C = (c_{ij}(x, y))$ are squared matrixes defined on $\overline{\Omega}$, $p \in \mathbb{R}_+$.

Ω is a domain included in the half-plan $y > 0$ and which has a part of frontier laying on $y = 0$, between the points $P(0, 0)$ and $Q(1, 0)$. The operator L is elliptic in Ω and parabolic on $\widehat{PQ}$.

Let $u \in C^2(\Omega, \mathbb{R}^n) \cap C\left(\overline{\Omega}, \mathbb{R}^n\right), u = u(x, y)$, be a solution of (3) and

$$R(x, y) := \left(\sum_{i=1}^{n} u_i^2 \right)^{1/2}.$$

Theorem 4. *([3])If:*

1. *for each $(x, y) \in \Omega$, there exist $\widetilde{\alpha}(x, y), \widetilde{\beta}(x, y) \in \mathbb{R}$ such that*

$$\xi \begin{pmatrix} -I_n & 0 & 0 \\ 0 & -y^p I_n & 0 \\ A(x, y) - \widetilde{\alpha}(x, y) I_n & B(x, y) - \widetilde{\beta}(x, y) I_n & C(x, y) \end{pmatrix} \xi^* < 0, \tag{5}$$

   for all $\xi \in \mathbb{R}^{3n}, \xi \neq 0$;
2. *B is symmetric such that if $\lambda_1(x, y)$ is the first eigenvalue, then $\lambda_1(x, 0) > 0$;*
3. *u is a regular solution of (3) and $R > 0$ in Ω;*
4. *$\lim\limits_{y \to 0} \frac{\partial R(x,y)}{\partial y}$ exist and is bounded,*

then $R = R(x, y)$ cannot attain his maximum value on $\widehat{PQ}$(open).

Remark 2. *If, for each $(x, y) \in \Omega$, there exist $\widetilde{\alpha}(x, y), \widetilde{\beta}(x, y) \in \mathbb{R}$ and $\varepsilon_1(x, y), \varepsilon_2(x, y) > 0$ such that:*

$$(i) \xi C(x, y) \xi^* < -\left(\varepsilon_1^2(x, y) + y^{-p} \varepsilon_2^2(x, y)\right) \|\xi\|^2, \forall \xi \in \mathbb{R}^n, \xi \neq 0;$$

$$(ii) \left\|A(x, y) - \widetilde{\alpha}(x, y) I_n\right\|_2 \leq 2\varepsilon_1(x, y), \left\|B(x, y) - \widetilde{\beta}(x, y) I_n\right\|_2 \leq 2\varepsilon_2(x, y),$$

then (5) holds.

Using Theorem 2 and choosing $\varepsilon_1 = \frac{1}{2}\sqrt{n}\gamma_F^A m_F^A$ and $\varepsilon_2 = \frac{1}{2}\sqrt{n}\gamma_F^B m_F^B$, the remark from above becomes:

Remark 3. *If:*

$$\xi C(x,y)\xi^* < -\frac{1}{4}n\left[\left(\gamma_F^A m_F^A\right)^2 + \left(\gamma_F^B m_F^B y^{-\frac{p}{2}}\right)^2\right]\|\xi\|^2, \forall \xi \in \mathbb{R}^n, \xi \neq 0, \forall (x,y) \in \Omega$$

then (5) holds.

Example 2. *Let us consider the system (2) with* $A = B = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_1 \end{pmatrix}$.

For $\widetilde{\alpha} = \widetilde{\beta} = a_1$ and $a_2, a_3 > 0$, we have $\varepsilon_1 = \varepsilon_2 = a_2 + a_3$, $A - a_1 I_2 = B - a_1 I_2 = \begin{pmatrix} 0 & a_2 \\ a_3 & 0 \end{pmatrix}$.

If $\xi C(x,y)\xi^* < -(a_2+a_3)^2(1+y^{-p})\|\xi\|^2$, then:

$$\xi \begin{pmatrix} -I_2 & 0 & 0 \\ 0 & -y^p I_2 & 0 \\ A - \widetilde{\alpha} I_2 & B - \widetilde{\beta} I_2 & C \end{pmatrix} \xi^* < \frac{1}{4}[a_2^2 + a_3^2 - 4(a_2+a_3)^2](1+y^{-p})\|\xi'\|^2 < 0$$

where $\xi = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6) \in \mathbb{R}^6, \xi \neq 0, \xi' = (\xi_5, \xi_6) \in \mathbb{R}^2, \xi' \neq 0$.

References

[1] C. Chifu-Oros, *Minimum value of a matrix norm with applications to maximum principles for second order differential systems*, (to appear).

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

[3] I.A. Rus, *Sur les proprietes des normes des solutions d'un system d'equations differentielles de second ordre*, Studia Universitatis Babes-Bolyai, Series Mathematica-Physica, Fascicula 1, p. 19-26, Cluj-Napoca, 1968.

Babeş-Bolyai University, Faculty of Business, Department of Business, 7^{th}, Horea Street, RO-3400, Cluj -Napoca, Romania

E-mail address: cochifu@tbs.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume XLVIII, Number 3, September 2003

SOME INTERPOLATION SCHEMES ON TRIANGLE

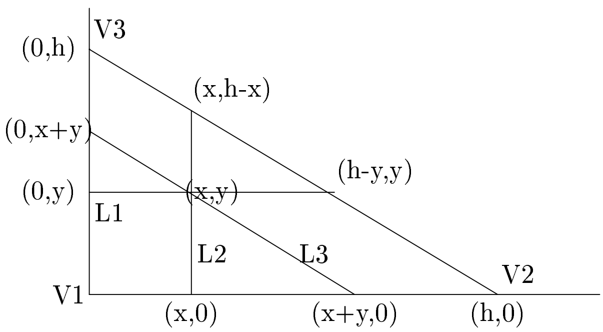
GHEORGHE COMAN AND IOANA POP

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. One of the quite simple procedures for constructing multidimensional approximation operators consists in the composition of univatiate approximation operators, using tensor product and boolean sum operations. In this paper, we will construct such interpolation operators for functions defined on a triangle, belonging to $B_{pq}^{r}(a,b)$ Sard spaces.

1. Introduction

Let $T_h = \left\{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, x+y \leq h\right\}$ be the standard triangle and $f : T_h \to \mathbb{R}$ a given function. For generating interpolation formulas on this triangle, we will use Lagrange, Hermite or Birkhoff univariate interpolation operators.



In the paper [1], using Lagrange interpolation operators defined by:

$$\begin{aligned}(L_1 f)(x,y) &= \tfrac{h-x-y}{h-y} f(0,y) + \tfrac{x}{h-y} f(h-y,y)\\ (L_2 f)(x,y) &= \tfrac{h-x-y}{h-x} f(x,0) + \tfrac{h}{h-x} f(x,h-x)\\ (L_3 f)(x,y) &= \tfrac{x}{x+y} f(x+y,0) + \tfrac{y}{x+y} f(0,x+y)\end{aligned} \qquad (1)$$

the authors studied some two dimensional discrete and blending interpolation operators. Thus, tensor product of the three operators, $P := L_1 L_2 L_3$, i.e.

$$(Pf)(x,y) = \tfrac{h-x-y}{h} f(0,0) + \tfrac{x}{h} f(h,0) + \tfrac{y}{h} f(0,h)$$

Received by the editors: 08.05.2003.

interpolates the function on the vertices of the triangle. For the remainder term of the interpolation formula:

$$f = Pf + Rf$$

was proved that for $f \in B_{11}(0,0)$

$$(Rf)(x,y) = \frac{x(x-h)}{2} f^{(2,0)}(\xi,0) + \frac{y(y-h)}{2} f^{(0,2)}(0,\eta) + xyf^{(1,1)}(\xi_1,\eta_1)$$

where $\xi,\eta \in [0,h], (\xi_1,\eta_1) \in T_h$.

Also, the boolean sum of every two operators L_i, $i = 1,2,3$ verifies the properties

$$L_i \oplus L_j f|_{\partial T_h} = f|_{\partial T_h}$$
$$L_i \oplus L_j g = g, g \in \mathbb{P}_2^2, \forall i,j = 1,2,3; i \neq j$$

That means that the operator $L_i \oplus L_j$ interpolates the function f on the boundary of T_h and its degree of exactness is 2 ($dex(L_i \oplus L_j) = 2$).

By appropriate select of interpolation operators, we can build interpolation formulas in which the values of the function are interpolated on certain sides and the normal derivatives on others.

For example, if B_1 is the Birkhoff interpolation operator defined by

$$(B_1 f)(x,y) = f(h-y,y) + (x+y-h) f^{(1,0)}(0,y)$$

which interpolates f on the ipotenuza of the triangle T_h and its normal derivative on the cathetus based on Ox, the operator

$$G = B_1 \oplus L_2$$

satisfies the interpolation properties:

$$(Gf)(x,0) = f(x,0), x \in [0,h]$$
$$(Gf)(h-y,y) = f(h-y,y), y \in [0,h]$$
$$(Gf)^{(1,0)}(0,y) = f^{(1,0)}(0,y), y \in [0,h]$$

and $dex(G) = 2$.

For the remainder of the formula $f = Gf + Rf$, it was proved that for $f \in C^{1,2}(T_h)$ and $f^{(0,3)}(0,y)$, $y \in [0,h]$ exist and is continuous, then

$$\|Rf\|_{L_\infty(T_h)} \leq \frac{h^3}{27}\left[\frac{2}{3}\left\|f^{(0,3)}(0,\cdot)\right\|_{L_\infty[0,h]} + \frac{1}{2}\left\|f^{(1,2)}\right\|_{L_\infty(T_h)}\right]$$

2. Next, we will build new interpolation operators for which we will determine the interpolation properties and degree of exactness. Also, the generated interpolation formulas will be studied.

2.1. Let us consider for the beginning the Taylor operator T_1^y defined by

$$(T_1^y f)(x,y) = f(x,0) + yf^{(0,1)}(x,0)$$

which interpolates the function f and its normal derivative $f^{(0,1)}$ with regard to the variable y on the Ox cathetus, respectively the operator L_1^x given in (1), i.e.

$$(L_1^x f)(x,y) = \frac{h-x-y}{h-y} f(0,y) + \frac{x}{h-y} f(h-y,y)$$

Let P_1 be

$$P_1 := L_1^x \oplus T_1^y$$

and

$$f = P_1 f + R_1 f \tag{2}$$

approximation formula generated by P_1.

Theorem 1. *Let consider $f : T_h \to \mathbb{R}$. If there exists $f^{(0,1)}(x,0)$, $x \in [0,h]$ and $f^{(1,0)}(h,0)$ then $P_1 f$ verifies the interpolation properties:*

$$(P_1 f)(0,y) = f(0,y), y \in [0,h]$$

$$(P_1 f)(h-y,y) = f(h-y,y), y \in [0,h]$$

$$(P_1 f)^{(0,1)}(x,0) = f^{(0,1)}(x,0), x \in [0,h]$$

and $dex(P_1) = 3$.

Proof.

$$\begin{aligned} (P_1 f)(x,y) &= \tfrac{h-x-y}{h-y} f(0,y) + \tfrac{x}{h-y} f(h-y,y) + f(x,0) + \\ &+ y f^{(0,1)}(x,0) - \tfrac{h-x-y}{h-y} f(0,0) - \tfrac{x}{h-y} f(h-y,0) - \\ &- \tfrac{y(h-x-y)}{h-y} f^{(0,1)}(0,0) - \tfrac{xy}{h-y} f^{(0,1)}(h-y,0) \end{aligned} \tag{3}$$

Now, the interpolation properties are easy verified, by direct computation.

So, $P_1 f$ coincides with f on a cathetus and the ipotenuza and the normal derivatives concides on the other cathetus.

We, also, have

$$P_1 e_{ij} = e_{ij} \text{ for } i,j \in \mathbb{N}, i+j \le 3 \text{ and } P_1 e_{22} \ne e_{22},$$

where $e_{ij}(x,y) = x^i y^j$. As P_1 is linear, it follows that $dex(P_1) = 3$.

Theorem 2. *If $f \in B_{2,2}(0,0)$ then*

$$(R_1 f)(x,y) = \int\int_{T_h} \varphi_{22}(x,y,s,t) f^{(2,2)}(s,t)\, dsdt$$

where $\varphi_{22}(x,y,s,t) = R_1\left[\frac{(x-s)_+^2}{2}\frac{(y-t)_+^2}{2}\right] = \frac{(x-s)_+^2}{2} \cdot \frac{(y-t)_+^2}{2}$

Furthermore, if $f^{(2,2)} \in C(T_h)$ then

$$(R_1 f)(x,y) = \frac{1}{36} x y^3 (x+y-h)(h+x-y) f^{(2,2)}(\xi,\eta), (\xi,\eta) \in T_h \tag{4}$$

Proof. As $dex(P_1) = 3$ it results, from the Peano's theorem, that

$$\begin{aligned} (R_1 f)(x,y) &= \int_0^h \varphi_{40}(x,y,s) f^{(4,0)}(s,0)\, ds + \int_0^h \varphi_{04}(x,y,t) f^{(0,4)}(0,t)\, dt + \\ &+ \int_0^h \varphi_{31}(x,y,s) f^{(3,1)}(s,0)\, ds + \int_0^h \varphi_{13}(x,y,t) f^{(1,3)}(0,t)\, dt + \\ &+ \int\int_{T_h} \varphi_{22}(x,y,s,t) f^{(2,2)}(s,t)\, dsdt \end{aligned}$$

Since $\varphi_{40}, \varphi_{31}, \varphi_{04}, \varphi_{13} = 0$ one obtain the first expression of the remainder term.

φ_{22} don't change the sign on T_h. Then, by the Mean theorem the expression (4) follows.

2.2. Now, let T_1 be defined by

$$\left(T_1^x f\right)(x,y) = f(0,y) + x f^{(1,0)}(0,y)$$

which interpolate f and its normal derivatives with regard to the variable x on the Oy cathetus.

Let be

$$P_2 = T_1^x \oplus T_1^y$$

and

$$f = P_2 f + R_2 f$$

the approximation formula generated by the operator P_2.

Theorem 3. *If $f : T_h \to \mathbb{R}$ and exist $f^{(1,0)}_{(x,0)}, f^{(0,1)}_{(0,y)}$, $x, y \in [0,h]$ then*

1. $P_2 f = f$ on ∂T_h.

2. $dex\,(P_2) = 3$.

Proof.

$$\begin{aligned}(P_2 f)(x,y) &= f(x,0) + y f^{(0,1)}(x,0) + f(0,y) + x f^{(1,0)}(0,y) - f(0,0) - \\ &- y f^{(0,1)}(0,0) - x\left[f^{(0,1)}(0,0) + y f^{(1,1)}(0,0)\right]\end{aligned}$$

The first statement results by a direct computation.

Also by direct computation, we obtain $P_2 e_{ij} = e_{ij}$ for $i + j \le 3$ and $P_2 e_{22} \neq e_{22}$, which implies that $dex\,(P_2) = 3$.

Theorem 4. *If $f \in B_{22}(0,0)$ then*

$$(R_2 f)(x,y) = \int\int_{T_h} \varphi_{22}(x,y,s,t)\, f^{(2,2)}(s,t)\, ds dt$$

where $\varphi_{22}(x,y,s,t) := R_2\left[\frac{(x-s)_+^2}{2}\frac{(y-t)_+^2}{2}\right] = \frac{(x-s)_+^2}{2} \cdot \frac{(y-t)_+^2}{2}$.

Furthermore, if $f^{(2,2)} \in C(T_h)$ then

$$(R_2 f)(x,y) = \frac{x^3 y^3}{36} f^{(2,2)}(\xi,\eta),\ (\xi,\eta) \in T_h.$$

Proof. Knowing that $dex\,(P_2) = 3$ it results, from the Peano's theorem, that

$$\begin{aligned}(R_2 f)(x,y) &= \int_0^h \varphi_{40}(x,y,s)\, f^{(4,0)}(s,0)\, ds + \int_0^h \varphi_{31}(x,y,s)\, f^{(3,1)}(s,0)\, ds + \\ &+ \int_0^h \varphi_{04}(x,y,t)\, f^{(0,4)}(0,t)\, dt + \int_0^h \varphi_{13}(x,y,t)\, f^{(1,3)}(0,t)\, dt + \\ &+ \int\int_{T_h} \varphi_{22}(x,y,s,t)\, f^{(2,2)}(s,t)\, ds dt\end{aligned}$$

Since $\varphi_{40}, \varphi_{31}, \varphi_{04}, \varphi_{13} = 0$ it results the first expression of the remainder term.

φ_{22} don't change the sign on T_h.Then, by the Mean theorem follows the second expression of the remainder term.

2.3. At last, let us consider the univariate operators B_1^x and B_1^y defined respectively by

$$(B_1^x)(x,y) = f(0,y) + xf^{(1,0)}(h-y,y)$$

and

$$(B_1^y)(x,y) = f(x,0) + yf^{(0,1)}(x,h-x)$$

The goal is to study the operator $P_3 := B_1^x \oplus B_1^y$ i.e.

$$\begin{aligned}(P_3f)(x,y) &= f(x,0) + f(0,y) + xf^{(1,0)}(h-y,y) + yf^{(0,1)}(x,h-x) - \\ &- f(0,0) - xf^{(1,0)}(h-y,0) - yf^{(0,1)}(0,h) - \\ &- xy\left(f^{(1,1)} - f^{(0,2)}\right)(h-y,y)\end{aligned}$$

Theorem 5. *If $f : T_h \to \mathbb{R}$ and there exists the derivatives $f^{(1,0)}(h-y,y)$, $f^{(0,1)}(x,h-x)$, $f^{(1,1)}(h-y,y)$, $f^{(0,2)}(h-y,y)$ and $f^{(1,0)}(h-y,0)$ for $x,y \in [0,h]$ than P_3 exists and*

$$(P_3f)(x,0) = f(x,0)$$

$$(P_3f)(0,y) = f(0,y)$$

$$(P_3f)^{(1,0)}(h-y,y) = f^{(1,0)}(h-y,y), x,y \in [0,h]$$

and

$$dex(P_3) = 2.$$

Proof. The first statement follows by a straightforward computation. Also, it is easy to verify that $P_3e_{ij} = e_{ij}$ for all $i,j \in \mathbb{N}$ with $i+j \leq 2$ and, for example $P_3e_{21} \neq e_{21}$. So, $dex(P_3) = 2$.

For the remainder term of the interpolation formula

$$f = P_3f + R_3f$$

we have:

Theorem 6. *If $f \in B_{12}(0,0)$ then*

$$\begin{aligned}(R_3f)(x,y) &= \frac{1}{6}y\left[y^2 + 6x(h-x-y)\right]f^{(0,3)}(0,\eta) - \\ &- \frac{1}{2}xy(2h+2x-y)f^{(1,2)}(\xi_1,\eta_1)\end{aligned}$$

Proof. As $dex(P_3) = 2$, using the Peano's theorem, one obtain

$$\begin{aligned}(R_3f)(x,y) &= \int_0^h \varphi_{30}(x,y,s)f^{(3,0)}(s,0)\,ds + \int_0^h \varphi_{21}(x,y,s)\cdot \\ &\cdot f^{(2,1)}(s,0)\,ds + \int_0^h \varphi_{03}(x,y,t)f^{(0,3)}(0,t)\,dt + \\ &+ \int\int_{T_h} \varphi_{12}(x,y,s,t)f^{(1,2)}(s,t)\,dsdt\end{aligned} \tag{5}$$

But, $\varphi_{30} = 0$, $\varphi_{21} = 0$, $\varphi_{03} \geq 0$ and $\varphi_{12} \leq 0$ on T_h. Using the mean theorem we have

$$\begin{aligned} (R_3 f)(x,y) &= f^{(0,3)}(0,\eta) \int_0^h \varphi_{03}(x,y,t)\,dt + \\ &+ f^{(1,2)}(\xi_1,\eta_1) \int\int_{T_h} \varphi_{12}(x,y,s,t)\,dsdt \end{aligned}$$

and (5) follows.

References

[1] Barnhill, R.E., Birkhoff, G., Gordon, W.J., *Smooth interpolation in triangle*, J. Approx. Theory 8, 1973, 114-128.

[2] Barnhill, R.E., Mansfield, *Sard Kernel theorems on triangular and rectangular domains with extensions and applications to finite element error*, Technical Report 11, Dep. of Math., Brunel Univ., 1972.

[3] Blaga, P., Coman, Gh., *Multivariate interpolation formulas of Birkhoff type*, Studia Univ. Babes-Bolyai, Mathematica 26, 1981, 14-22.

[4] Bohner, K., Coman, Gh., *Blending interpolation schemes on triangle with error bounds*, Springer Verlag, Berlin - Heidelberg - New York, 1977, 14-37.

[5] Coman, Gh., Tâmbulea, L., Gânscă, I., *Multivariate approximation*, Seminar on Numerical and Statistical Calcules, Preprint nr. 1, 1996, 29-60.

[6] Sard, A., *Linear Approximation*, American Mathematical Society, Providence, Rhode Island, 1963.

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

University of Agricultural Sciences and Veterinary Medicine,
Cluj-Napoca, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

A CHARACTERIZATION OF π-CLOSED SCHUNCK CLASSES

RODICA COVACI

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. A characterization of π-closed Schunck classes, followed by some consequences and applications in the formation theory of finite π-solvable groups are given.

1. Preliminaries

All groups considered in the paper are finite. Let π be a set of primes, π' the complement to π in the set of all primes and $O_{\pi'}(G)$ the largest normal π'-subgroup of a group G.

We first give some useful definitions.

Definition 1.1. ([9], [10], [12]) a) A class $\mathcal{X}$ of groups is a *homomorph* if $\mathcal{X}$ is epimorphically closed, i.e. if $G \in \mathcal{X}$ and N is a normal subgroup of G, then $G/N \in \mathcal{X}$.

b) A homomorph $\mathcal{X}$ is a *formation* if $G/N_1 \in \mathcal{X}$ and $G/N_2 \in \mathcal{X}$ imply $G/(N_1 \cap N_2) \in \mathcal{X}$.

c) A formation $\mathcal{X}$ is *saturated* if $\mathcal{X}$ is Frattini closed, i.e. if $G/\phi(G) \in \mathcal{X}$ implies $G \in \mathcal{X}$, where $\phi(G)$ denotes the Frattini subgroup of G.

d) A group G is *primitive* if G has a *stabilizer*, i.e. a maximal subgroup H with $core_G H = \{1\}$, where $core_G H = \cap\{H^g/g \in G\}$.

e) A homomorph $\mathcal{X}$ is a *Schunck class* if $\mathcal{X}$ is *primitively closed*, i.e. if any group G, all of whose primitive factor groups are in $\mathcal{X}$, is itself in $\mathcal{X}$.

Definition 1.2. a) ([8]) A group G is *π-solvable* if every chief factor of G is either a solvable π-group or a π'-group. For π the set of all primes, we obtain the notion of solvable group.

b) A class $\mathcal{X}$ of groups is said to be *π-closed* if

$$G/O_{\pi'} \in \mathcal{X} \;\Rightarrow\; G \in \mathcal{X}.$$

A π-closed homomorph, formation, respectively Schunck class is called *π-homomorph*, *π-formation*, respectively *π-Schunck class*.

Definition 1.3. ([9], [10]) Let $\mathcal{X}$ be a class of groups, G a group and H a subgroup of G.

Received by the editors: 07.12.2002.
2000 *Mathematics Subject Classification.* 20D10.
Key words and phrases. Key words and phrases: Schunck class, formation, π-solvable group.

a) H is *$\mathcal{X}$-maximal* in G if: (i) $H \in \mathcal{X}$; (ii) $H \leq H^* \leq G$, $H^* \in \mathcal{X}$ imply $H = H^*$.

b) H is an *$\mathcal{X}$-projector* of G if, for any normal subgroup N of G, HN/N is $\mathcal{X}$-maximal in G/N.

c) H is an *$\mathcal{X}$-covering subgroup* of G if: (i) $H \in \mathcal{X}$; (ii) $H \leq K \leq G$, $K_0 \lhd K$, $K/K_0 \in \mathcal{X}$ imply $K = HK_0$.

The following results will be used in the paper.

Theorem 1.4. ([4]) *Let $\mathcal{X}$ be a class of groups, G a group and H a subgroup of G.*

a) If H is an $\mathcal{X}$-covering subgroup or an $\mathcal{X}$-projector of G, then H is $\mathcal{X}$-maximal in G.

b) If $\mathcal{X}$ is a homomorph, any $\mathcal{X}$-covering subgroup of G is an $\mathcal{X}$-projector of G.

Theorem 1.5. ([9]) *If $\mathcal{X}$ is a homomorph, G a group, N a normal subgroup of G, K/N an $\mathcal{X}$-covering subgroup of G/N and H is an $\mathcal{X}$-covering subgroup of K, then H is an $\mathcal{X}$-covering subgroup of G.*

Theorem 1.6. ([1]) *A solvable minimal normal subgroup of a group is abelian.*

Theorem 1.7. ([1]) *If S is a maximal subgroup of G with $\mathrm{core}_G S = \{1\}$ and N is a minimal normal subgroup of G, then $G = SN$ and $S \cap N = \{1\}$.*

Theorem 1.8. ([10]) *Let $\mathcal{X}$ be a class of groups. $\mathcal{X}$ is a saturated formation if and only if $\mathcal{X}$ is both a Schunck class and a formation.*

Theorem 1.9. ([2], [3], [4]) *Let $\mathcal{X}$ be a π-homomorph. The following conditions are equivalent:*

(1) $\mathcal{X}$ is a Schunck class;

(2) any π-solvable group has $\mathcal{X}$-covering subgroups;

(3) any π-solvable group has $\mathcal{X}$-projectors.

2. **The main result**

In preparation for the main theorem of the paper, we give the following lemma.

Lemma 2.1. *Let $\mathcal{X}$ be a π-Schunck class, G a π-solvable group, such that $G \notin \mathcal{X}$, N a minimal normal subgroup of G with $G/N \in \mathcal{X}$ and H and $\mathcal{X}$-covering subgroup of G. Then H is a complement of N in G, i.e. $G = HN$ is $H \cap N = \{1\}$.*

Proof. Using that H is an $\mathcal{X}$-covering subgroup of G, from $H \leq G \leq G$, $N \lhd G$, $G/N \in \mathcal{X}$ follows that $G = HN$.

We prove now that $H \cap N = \{1\}$.

G is π-solvable group, hence the minimal normal subgroup N of G, being a chief factor of G, is either a solvable π-group or a π'-group. If we suppose that N is a π'-group, we obtain that $N \leq O_{\pi'}(G)$, hence

$$G/O_{\pi'} \cong (G/N)/(O_{\pi'}(G)/N).$$

But $G/N \in \mathcal{X}$ and $\mathcal{X}$ is a homomorph. So $G/O_{\pi'}(G) \in \mathcal{X}$, hence, $\mathcal{X}$ being π-closed, $G \in \mathcal{X}$, in contradiction with the hypothesis $G \notin \mathcal{X}$. It follows that N is a solvable π-group. By 1.6., N is abelian.

We prove that $H \cap N$ is a normal subgroup of G. Indeed, if $g \in G$ and $x \in H \cap N$, we have $g = nh$, with $n \in N$, $h \in H$ and

$$g^{-1}xg = (nh)^{-1}x(nh) = h^{-1}n^{-1}(xn)h = h^{-1}n^{-1}(nx)h = h^{-1}xh \in H \cap N,$$

where we used that N is abelian and that $H \cap N$ is normal in H.

Finally, N being a minimal normal subgroup of G and $H \cap N \lhd G$, $H \cap N \subseteq N$, we have $H \cap N = \{1\}$ or $H \cap N = N$. If we suppose that $H \cap N = N$, it follows that $N \subseteq H$, hence $G = HN = H$, a contradiction with $G \notin \mathcal{X}$ and $H \in \mathcal{X}$. So $H \cap N = \{1\}$. □

Theorem 2.2. *Let $\mathcal{X}$ be a π-homomorph. The following conditions are equivalent:*

(1) $\mathcal{X}$ is a Schunck class;

(2) if G is a π-solvable group, $G \notin \mathcal{X}$ and N is a minimal normal subgroup of G such that $G/N \in \mathcal{X}$, then N has a complement in G;

(3) any π-solvable group G has $\mathcal{X}$-covering subgroups;

(4) any π-solvable group G has $\mathcal{X}$-projectors.

Proof. (1) implies (2). Let G be a π-solvable group, $G \notin \mathcal{X}$ and N a minimal normal subgroup of G such that $G/N \in \mathcal{X}$. By (1) and 1.9., G has an $\mathcal{X}$-covering subgroup H. By Lemma 2.1., H is a complement of N in G.

(2) implies (3). We prove by induction on $|G|$ that any π-solvable group G has $\mathcal{X}$-covering subgroups.

Two cases are possible:

1. $G \in \mathcal{X}$. In this case, G is its own $\mathcal{X}$-covering subgroup.

2. $G \notin \mathcal{X}$. Let N be a minimal normal subgroup of G. By the induction, G/N has an $\mathcal{X}$-covering subgroup E/N. We consider two possibilities:

a) $G/N \in \mathcal{X}$. Then, by 1.4.a) and 1.3.a), $E/N = G/N$. Applying (2) for the π-solvable group G, $G \notin \mathcal{X}$ and for its minimal normal subgroup N with $G/N \in \mathcal{X}$, we obtain that N has a complement V in G, i.e. $G = NV$ and $N \cap V = \{1\}$.

We notice that $V \in \mathcal{X}$, because

$$V \cong V/(N \cap V) \cong NV/N = G/N \in \mathcal{X}.$$

By 1.2.a), N is either a solvable π-group or a π'-group. If we suppose that N is a π'-group, then $N \leq O_{\pi'}(G)$ and so

$$G/O_{\pi'}(G) \cong (G/N)/(O_{\pi'}(G)/N) \in \mathcal{X},$$

where we used that $G/N \in \mathcal{X}$ is a homomorph. Applying that $\mathcal{X}$ is π-closed, we get $G \in \mathcal{X}$, a contradiction. It follows that N is a solvable π-group, hence, by 1.6., N is abelian.

Let us consider two cases:

i) $core_G V \neq \{1\}$. By the induction, $G/core_G V$ has an $\mathcal{X}$-covering subgroup $H/core_G V$.

We notice that $H \neq G$, else $H = G$ implies $G/core_G V = H/core_G V \in \mathcal{X}$ and so $G/core_G V$ is its own $\mathcal{X}$-covering subgroup, hence, by 1.4.a), $G/core_G V$ is its own $\mathcal{X}$-maximal subgroup. But $V \in \mathcal{X}$ and $\mathcal{X}$ homomorph imply that $V/core_G V \in \mathcal{X}$. It follows that $V/core_G V = G/core_G V$ and so $V = G$, contradicting that $G \notin \mathcal{X}$ and $V \in \mathcal{X}$. Hence $H \neq G$.

The induction for H leads to the existence of an $\mathcal{X}$-covering subgroup L of H. Then $H/core_G V$ is an $\mathcal{X}$-covering subgroup of $G/core_G V$ and L is an $\mathcal{X}$-covering subgroup of H. Applying 1.5., we conclude that L is an $\mathcal{X}$-covering subgroup of G.

ii) $core_G V = \{1\}$. In this case, we prove that V is an $\mathcal{X}$-covering subgroup of G.

We proved that $V \in \mathcal{X}$.

Let now $V \leq K \leq G$, $K_0 \lhd K$ and $K/K_0 \in \mathcal{X}$. We shall prove that $K = VK_0$.

First, V is a maximal subgroup of G. Indeed, $V \neq G$, because $V \in \mathcal{X}$ and $G \notin \mathcal{X}$. Let now $V \leq V^* < G$. We show that $V = V^*$. Suppose $V < V^*$ and let $v^* \in V^* \setminus V \subset G = VN$ and put $v^* = vn$, where $v \in V$, $n \in N$. We have $n = v^{-1}v^* \in N \cap V^*$.

Let us prove that $N \cap V^* = \{1\}$. We notice that $G = NV \leq NV^* \leq G$ imply $G = NV^*$. Further, $N \cap V^*$ is a normal subgroup of G, because if $g \in G$, $x \in N \cap V^*$ we can prove that $g^{-1}xg \in N \cap V^*$. Indeed, if we take $g \in G = NV^*$ written as $g = mv^*$, with $m \in N$, $v^* \in V^*$, we have

$$g^{-1}xg = (mv^*)^{-1}(mv^*) = (v^*)^{-1}(m^{-1}x)mv^* =$$

$$= (v^*)^{-1}(xm^{-1})mv^* = (v^*)^{-1}xv^* \in N \cap V^*,$$

where we used that N is abelian and that $N \cap V^* \lhd V^*$. Hence $N \cap V^*$ is normal in G. N is a minimal normal subgroup of G and $N \cap V^* \subseteq N$. It follows that $N \cap V^* = \{1\}$ or $N \cap V^* = N$. But $N \cap V^* = N$ implies $N \subseteq V^*$ and so $G = NV^* = V^*$, in contradiction with the choice of V^*. Hence $N \cap V^* = \{1\}$.

From $n = v^{-1}v^* \in N \cap V^* = \{1\}$, we deduce $n = 1$ and so $v^{-1}v^* = 1$, which means $v^* = v \in V$, in contradiction with the choice of v^*. It follows that $V = V^*$. This completes the proof that V is a maximal subgroup of G.

By the above, we have for K with $V \leq K \leq G$ two possibilities: $K = V$ or $K = G$.

If $K = V$, we have $K_0 \lhd K = V$ and so $K = KK_0 = VK_0$.

If $K = G$, we reason as follows. Let us notice that $K_0 \neq \{1\}$, else

$$G = K \cong K/K_0 \in \mathcal{X},$$

a contradiction with $G \notin \mathcal{X}$. Let M be a minimal normal subgroup of G such that $M \subseteq K_0$. So we are in hypotheses of 1.7.: V is a maximal subgroup of G with $core_G V = \{1\}$ and M is a minimal normal subgroup of G. It follows that $G = VM$ and so

$$K = G = VM \leq VK_0 \leq G,$$

hence $K = G = VK_0$.

b) $G/N \notin \mathcal{X}$. In this case, we have $E/N \neq G/N$, because $E/N \in \mathcal{X}$. So $E \neq G$. By the induction, E has an $\mathcal{X}$-covering subgroup F. But E/N is an

$\mathcal{X}$-covering subgroup in G/N. Theorem 1.5. leads to the conclusion that F is an $\mathcal{X}$-covering subgroup of G.

(3) implies (4). Follows immediately from 1.9.

(4) implies (1). Follows also from 1.9. □

3. **Consequences**

Theorem 2.2. has some consequences on π-closed formations. In [5], we gave:

Theorem 3.1. ([5]) *Let $\mathcal{X}$ be a π-formation. The following conditions are equivalent:*

(1) $\mathcal{X}$ is saturated;

(2) if G is a π-solvable group and $G \notin \mathcal{X}$, but for the minimal normal subgroup N of G we have $G/N \in \mathcal{X}$, then N has a complement in G;

(3) any π-solvable group G has $\mathcal{X}$-covering subgroups.

From 2.2., 3.1. and 1.8., we obtain:

Corollary 3.2. *If $\mathcal{X}$ is a π-formation satisfying condition (2) from 2.2., then:*

a) $\mathcal{X}$ is a Schunck class;

b) $\mathcal{X}$ is Frattini closed, hence $\mathcal{X}$ is a saturated formation;

c) any π-solvable group G has $\mathcal{X}$-covering subgroups;

d) any π-solvable group G has $\mathcal{X}$-projectors.

4. **Some applications**

Finally, we give some applications of the main theorem of this paper, concerning to:

1. the existence and conjugacy given in [7] of $\mathcal{X}$-maximal subgroups in finite π-solvable groups, where $\mathcal{X}$ is a π-Schunck class;

2. the π-Schunck classes with the P property, introduced in [6].

4.1. In [7] we proved the following result:

Theorem 4.1.1. ([7]) *Let $\mathcal{X}$ be a π-Schunck class, G a π-solvable group and A an abelian normal subgroup of G with $G/A \in \mathcal{X}$. Then:*

(1) there is a subgroup S of G with $S \in \mathcal{X}$ and $AS = G$;

(2) there is an $\mathcal{X}$-maximal subgroup S of G with $AS = G$;

(3) if S_1 and S_2 are $\mathcal{X}$-maximal subgroups of G with $AS_1 = G = AS_2$, then S_1 and S_2 are conjugate in G.

Applying 4.1.1. and 2.2., we can prove the following theorem:

Theorem 4.1.2. *If $\mathcal{X}$ is a π-Schunck class, G is a π-solvable group, $G \notin \mathcal{X}$ and N is a minimal normal subgroup of G such that $G/N \in \mathcal{X}$, then:*

a) N has a complement H in G;

b) N is a solvable π-group, hence N is abelian;

c) H is $\mathcal{X}$-maximal in G;

d) H is conjugate to any $\mathcal{X}$-maximal subgroup S of G with $NS = G$.

Proof. a) Applying theorem 2.2., we obtain that N has a complement H in G, i.e. $HN = G$ and $H \cap N = \{1\}$.

b) N being a minimal normal subgroup of the π-solvable group G, N is either a solvable π-group, hence by 1.6. N is abelian, or N is a π'-group. We shall prove that the case N is π'-group is not possible in our hypotheses. Indeed, if we suppose that N is a π'-group, we have $N \leq O_{\pi'}(G)$ and

$$G/O_{\pi'}(G) \cong (G/N)/(O_{\pi'}(G)/N) \in \mathcal{X},$$

hence, by the π-closure of $\mathcal{X}$, $G \in \mathcal{X}$, a contradiction.

c) In order to prove that H is $\mathcal{X}$-maximal in G. let us first notice that $H \in \mathcal{X}$. Indeed, we have

$$H \cong H/\{1\} = H/(H \cap N) \cong HN/N = G/N \in \mathcal{X}.$$

Let now $H \leq H^* \leq G$ and $H^* \in \mathcal{X}$. We prove that $H = H^*$. Suppose that $H < H^*$. Then there is an element $h^* \in H^* \setminus H \subset G = HN$ and $h^* = hn$, with $h \in H$, $n \in N$. Then $n = h^{-1}h^* \in H^* \cap N = \{1\}$ and so $n = 1$ and $h^* = h \in H$, in contradiction with the choice of h^*. The fact that $H^* \cap N = \{1\}$ follows from $H^* \cap N \lhd G$ (since N is abelian and $H^* \cap N \lhd H^*$) and from the hypotheses that N is a minimal normal subgroup of G.

d) Since we are in the hypotheses of 4.1.1, there is an $\mathcal{X}$-maximal subgroup S of G with $NS = G$. Applying now 4.1.1.(3), we conclude that H is conjugate to S. □

4.2. In [6], we introduced the P property on a class $\mathcal{X}$ of groups. We say that $\mathcal{X}$ has the *P property* if, for any π-solvable group G, we have:

N minimal normal subgroup of G, N π'-group $\Rightarrow$ $G/N \in \mathcal{X}$.

Using theorem 2.2., we can prove the following result:

Theorem 4.2.1. *If $\mathcal{X}$ is a π-Schunck class with the P property and G is a π-solvable group, $G \notin \mathcal{X}$, then any minimal subgroup N of G which is a π'-group has a complement in G.*

Proof. By the P property, we have $G/N \in \mathcal{X}$. But $\mathcal{X}$ being a π-Schunck class, theorem 2.2. shows that $\mathcal{X}$ satisfies condition (2). Applying (2) for the π-solvable group G with $G \notin \mathcal{X}$ and for the minimal subgroup N of G with $G/N \in \mathcal{X}$, we conclude that N has a complement in G. □

References

[1] Baer, R., *Classes of finite groups and their properties*, Illinois J. Math., **1**, No. 2, 1957, 115-187.

[2] Covaci, R., *Projectors in finite π-solvable groups*, Studia Univ. Babeş-Bolyai, Math., **XXII**, No. 1, 1977, 3-5.

[3] Covaci, R., *Some properties of projectors in finite π-solvable groups*, Studia Univ. Babeş-Bolyai, Math., **XXVI**, No. 1, 1981, 5-8.

[4] Covaci, R., *Projectors and covering subgroups*, Studia Univ. Babeş-Bolyai, Math., **XXVII**, 1982, 33-36.

[5] Covaci, R., *On saturated π-formations*, Studia Univ. Babeş-Bolyai, Math., **XXXI**, No. 4, 1986, 70-72.

[6] Covaci, R., *On π-Schunck classes with the P property*, Babeş-Bolyai University Research Seminars, Preprint No. 5, 1988, 22-34.

[7] Covaci, R., *$\mathcal{X}$-maximal subgroups in finite π-solvable groups with respect to a Schunck class $\mathcal{X}$*, Studia Univ. Babeş-Bolyai, Math., **XLVII**, No. 3, 2002, 53-62.

[8] Čunihin, S. A., *O teoremah tipe Sylowa*, Dokl. Akad. Nauk SSSR, **66**, No. 2, 1949, 165-168.
[9] Gaschütz, W., *Zur Theorie der endlichen auflösbaren Gruppen*, Math. Z., **80**, No. 4, 1963, 300-305.
[10] Gaschütz, W., *Selected topics in the theory of soluble groups*, Australian National University, Canberra, 1969.
[11] Huppert, B., *Endliche Gruppen I*, Berlin - New York, Springer-Verlag, 1967.
[12] Schuck, H., *$\mathcal{H}$-Untergruppen in endlichen auflösbaren Gruppen*, Math. Z., **97**, No. 4, 1967, 326-330.

Faculty of Mathematics and Computer Science,
Babeş-Bolyai University, Cluj-Napoca, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

A NODAL SPLINE COLLOCATION METHOD FOR WEAKLY SINGULAR VOLTERRA INTEGRAL EQUATIONS

C. DAGNINO, V. DEMICHELIS, AND E. SANTI

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. A collocation method based on optimal nodal splines is presented for the numerical solution of linear Volterra integral equations of the second kind with weakly singular kernel. Since the considered spline operator is a bounded projector we can prove that, for sequences of locally uniform meshes, the approximate solution error converges to zero at exactly the same optimal rate as the spline approximation error. We consider in particular sequences of graded meshes, for which the local uniformity is proved. Finally, we give an upper bound for the condition number of the collocation system and we present some numerical examples.

1. **Introduction**

The Volterra integral equation of the second kind

$$y(x) = f(x) + \int_0^x k(x,s)y(s)ds, \quad x \in I \equiv [0, X] \tag{1}$$

with weakly singular kernel k provides mathematical model describing a wide variety of applicative problems. Particularly interesting kernels are the convolution ones, of the form $k(x-s)$, where $k(t) \in C(O, X] \cap L_1(O, X)$, but $k(t)$ may become unbounded as $t \to 0$. Examples of convolution kernels are

$$k(t) = \lambda |t|^{-\alpha} \quad , \quad 0 < \alpha < 1 \tag{2}$$

$$k(t) = \lambda \log |t| \, , \tag{3}$$

where $\lambda \in \mathbf{R}$.

If $f \in C(I)$, then (1) has a unique solution $y \in C(I)$. As f becomes smoother, y also becomes smoother, but only for $x > 0$. In general there will be no increase in smoothness of the solution at $x = 0$. At the same time, very special choices of f may force smoother behaviour at the origin [13].

In the recent literature, some collocation methods, based on piecewise polynomials for solving (1) with the above kernels, have been studied (cfr. [2,12] and

Received by the editors: 25.06.2003.
2000 *Mathematics Subject Classification.* 65R20, 65D30, 65D07.
Key words and phrases. spline approximation, collocation, Volterra equation, weakly singular kernel.

references therein). In order to find an approximate solution sufficiently smooth in $(0, X]$, one may use polynomial splines of order m, belonging to $C^{\nu}(I), 0 \le \nu \le m-2$.

In this context we propose a new product collocation method, for numerically solving (1), based on optimal nodal splines of order $m > 2$. We generate a sequence of spline approximations $\{y_n\}$ for the solution of (1) and we analyze its convergence to y. Since the constructed approximation operator is a bounded projection operator, it will be proved that $||y - y_n||$ converges to zero at exactly the same rate as the norm of nodal spline approximation error for sequences of locally uniform (l.u.) meshes.

In order to reflect the possible singular behaviour of the solution near to the initial point, we will resort to a sequence of graded meshes. Indeed in this context we will prove also that the above sequence is l.u..

The paper is organized as follows. In Section 2 we give some preliminaries relative to the nodal spline space of our interest, the construction and convergence properties of the approximating operator. In Section 3 we give our spline collocation method for the problem (1). Section 4 is devoted to the error analysis and in Section 5 we study the condition number for the collocation method. Finally, in Section 6 we present some numerical results; in one case, in particular, we will show the better performance of the sequence of graded partitions with respect to the uniform one, when the solution has the first derivative singular at $x = 0$.

2. **On optimal nodal splines**

We briefly review the definition and the main properties of the optimal nodal splines of interest in this context [5-8].

Let $I = [0, X]$ be a given finite interval of the real line $\mathbf{R}$, for a fixed integer $m \ge 3$ and $n \ge m - 1$, we define a partition Π_n of I by

$$\Pi_n : 0 = \tau_0 < \tau_1 < ... < \tau_n = X \ ,$$

generally called "primary partition". We insert $m - 2$ distinct points throughout $(\tau_\nu, \tau_{\nu+1})$, $\nu = 0, ..., n - 1$ obtaining a new partition of I

$$X_n : 0 = x_0 < x_1, < ... \qquad < x_{(m-1)n} = X,$$

where $x_{(m-1)i} = \tau_i$, $i = 0, ..., n$.

Let

$$R_n = \max_{\substack{0 \le k,j \le n-1 \\ |k-j|=1}} \frac{\tau_{k+1} - \tau_k}{\tau_{j+1} - \tau_j} \ , \tag{4}$$

we say that the sequence of primary partitions $\{\Pi_n; n = m - 1, m, ...\}$ is l.u. if, for all n, there exists a constant $A \ge 1$ such that $R_n \le A$, i.e.

$$\frac{1}{A} \le \frac{\tau_{k+1} - \tau_k}{\tau_{j+1} - \tau_j} \le A \ , \ k, j = 0, 1, ..., n - 1 \text{ and } |k - j| = 1 \ . \tag{5}$$

Since the convergence results of the nodal splines we shall consider are based on the local uniformity property of the primary partitions sequence and one of our objectives is the use of graded meshes, in the following proposition we shall prove that a sequence of primary graded partitions is l.u.

Proposition 1. *Let $[0, X]$ be a finite interval. The sequence of partitions $\{\Pi_n\}$, obtained by using graded meshes [3] of the form*

$$\tau_i = \left(\frac{i}{n}\right)^r X \quad , \quad 0 \le i \le n \,, \tag{6}$$

with grading exponent $r \in \mathbf{R}$ assumed ≥ 1, is l.u., i.e. it satisfies (5) with $A = 2^r - 1$.

Proof. For $r = 1$, the partition is uniform and (5) is satisfied with $A = 1$.

Consider now $r > 1$ and $k = j + 1$. We can write

$$f(j) = \frac{\tau_{j+2} - \tau_{j+1}}{\tau_{j+1} - \tau_j} = \frac{\left(1 + \frac{1}{j+1}\right)^r - 1}{1 - \left(1 - \frac{1}{j+1}\right)^r} \quad , \quad j = 0, 1, ..., n-2$$

and $f(0) = 2^r - 1$.

Consider now the function $f(x) = \frac{\left(1+\frac{1}{x+1}\right)^r - 1}{1 - \left(1 - \frac{1}{x+1}\right)^r}$, $x \in \mathbf{R}^+$. Then $f(j) = f(x), x \in N$. One can verify that $\lim_{x\to\infty} f(x) = 1$ and $f'(x) < 0$ for all x.

Then

$$1 \le f(j) \le 2^r - 1 \,. \tag{7}$$

If $k = j - 1$, for $j = 1, 2, ..., n-1$ we have

$$\frac{\tau_j - \tau_{j-1}}{\tau_{j+1} - \tau_j} = \frac{1}{\frac{(j+1)^r - j^r}{j^r - (j-1)^r}} = \frac{1}{f(j-1)}$$

and using (7), the thesis follows. ∎

Now, after introducing two integers [5]

$$i_0 = \begin{cases} \frac{1}{2}(m+1) & m \quad \text{odd} \\ \frac{1}{2}m + 1 & m \quad \text{even} \end{cases} \qquad \text{and} \quad i_1 = (m+1) - i_0$$

and two integer functions

$$p_\nu = \begin{cases} 0 & \nu = 0, 1, ..., i_1 - 2 \\ \nu - i_1 + 1 & \nu = i_1 - 1, ..., n - i_0 \\ n - (m-1) & \nu = n - i_0 + 1, ..., n-1 \end{cases}$$

$$q_\nu = \begin{cases} m - 1 & \nu = 0, 1, ..., i_1 - 2 \\ \nu + i_0 & \nu = i_1 - 1, ..., n - i_0 \\ n & \nu = n - i_0 + 1, ..., n-1 \end{cases}$$

consider the set $\{w_i(x); i = 0, 1, ..., n\}$ of functions defined as follows [6,7]

$$w_i(x) = \begin{cases} l_i(x) & x \in [\tau_0, \tau_{i_1-1}], & i \le m-1 \\ s_i(x) & x \in (\tau_{i_1-1}, \tau_{n-i_0+1}), & n \ge m \\ \bar{l}_i(x) & x \in [\tau_{n-i_0+1}, \tau_n], & i \ge n-(m-1) \end{cases}$$

where

$$l_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^{m-1} \frac{x - \tau_k}{\tau_i - \tau_k} \tag{8}$$

$$\bar{l}_i(x) = \prod_{\substack{k=0 \\ k \neq n-i}}^{m-1} \frac{x - \tau_{n-k}}{\tau_i - \tau_{n-k}} \tag{9}$$

$$s_i(x) = \sum_{r=0}^{m-2} \sum_{j=j_0}^{j_1} \alpha_{i,r,j} B_{(m-1)(i+j)+r}(x) \tag{10}$$

with $j_0 = \max\{-i_0, i_1-2-i\}$, $j_1 = \min\{-i_0+m-1, n-i_0-i\}$. The coefficients $\alpha_{i,r,j}$ are given in [5] and the B-splines sequence is constructed from the set of the normalized B-splines defined in [14] for $i = (m-1)(i_1-2), (m-1)(i_1-2)+1, ..., (m-1)(n-i_0+1)$. Then, the following locality property holds [6]

$$s_i(x) = 0 \quad , \quad x \notin [\tau_{i-i_0}, \tau_{i+i_1}]. \tag{11}$$

Each $w_i(x)$ is nodal with respect to Π_n, in the sense that

$$w_i(\tau_j) = \delta_{i,j} \quad , \quad i,j = 0,1,...,n\,. \tag{12}$$

Therefore, being $\det[w_i(\tau_j)] \neq 0$, the functions $w_i(x), i = 0,1,...,n$, are linearly independent. Let $S_{\Pi_n} = \text{span}\{w_i(x); i = 0,1,...,n\}$, it is proved in [7] that, for all $s \in S_{\Pi_n}$, one has $s \in C^{m-2}(I)$.

For all $g \in B(I)$, where $B(I)$ is the set of real-valued functions on I, we consider the spline operator $W_n : B(I) \to S_{\Pi_n}$, so defined

$$W_n g = \sum_{i=0}^{n} g(\tau_i) w_i(x) \quad , \quad x \in I\,.$$

By (12), for $0 \leq \nu < n$ we can write:

$$W_n g = \sum_{i=p_\nu}^{q_\nu} g(\tau_i) w_i(x), \quad x \in [\tau_\nu, \tau_{\nu+1}]\,. \tag{13}$$

It is proved in [6,7] that $W_n p = p$, for all $p \in \mathbf{P}_m$, where $\mathbf{P}_m$ denotes the set of polynomials of order m (degree $\leq m-1$), and $W_n g(\tau_i) = g(\tau_i)$, for $i = 0,1,...,n$, i.e. W_n is an interpolatory operator.

Using the results in [6,7,8] we deduce that, for l.u. $\{\Pi_n\}$, W_n is a bounded projection operator in S_{Π_n}. In fact, it is easy to show that

$$W_n s = s \quad , \quad \text{for all } s \in S_{\Pi_n}$$

and,if we denote:

$$||W_n|| = \sup\{||W_n h|| : h \in C(I), ||h|| < 1\},$$

with $||h|| = \max_{x \in I} |h(x)|$, considering that

$$||W_n|| \le (m+1)\left[\sum_{\lambda=1}^{m-1}(R_n)^\lambda\right]^{m-1} ,$$

where R_n is defined in (4), from (5) we obtain $||W_n|| < \infty$.

Finally, for all $g \in C^\nu(I)$, $0 \le \nu < m$, assuming that $\{\Pi_n\}$ is l.u., there results

$$||g - W_n g|| = O(H_n^\nu \omega(g^{(\nu)}; H_n; I)), \tag{14}$$

where $H_n = \max_{1 \le i \le n} (\tau_i - \tau_{i-1})$ and, for all $g \in C(I)$, $\omega(g; \delta; I) = \max_{\substack{x, x+h \in I \\ 0 < h \le \delta}} |g(x+h) - g(x)|$.

3. Spline collocation method

Consider now the linear integral equation (1) and a sequence of nodal spline spaces $\{S_{\Pi_n}; n = m-1, m, ...\}$ spanned by $\{w_i(x); i = 0, ..., n\}$ and based on a sequence of l.u. primary partitions $\{\Pi_n\}$.

For some fixed n we consider a spline $y_n \in S_{\Pi_n}$ written in the form

$$y_n(x) = \sum_{j=0}^{n} \alpha_j w_j(x), \alpha_j \in \mathbf{R} . \tag{15}$$

Substituting (16) in (1) we obtain

$$y_n(x) - \int_0^x k(x,s) y_n(s) ds + r_n(x) = f(x) ,$$

where $r_n(x)$ is the residual term obtained approximating y by y_n in (1).

The values α_j in (16), with $j = 0, 1, ..., n$, are choosen by requiring that

$$r_n(\tau_j) = 0, \quad j = 0, 1, ..., n . \tag{16}$$

This leads to determine $\alpha_0, \alpha_1, ..., \alpha_n$ as the solution of a linear system that, using (13), can be written in the form:

$$\alpha_j[1 - \mu_j(\tau_j)] - \sum_{\substack{i=0 \\ i \ne j}}^{n} \mu_i(\tau_j)\alpha_i = f(\tau_j), \quad j = 0, 1, ..., n , \tag{17}$$

where

$$\mu_i(\tau_j) = \int_0^{\tau_j} k(\tau_j, s) w_i(s) ds . \tag{18}$$

By (8)-(12) and (14) we can explicitly write each weight of the set $\{\mu_i(\tau_j); i, j = 0, 1, ..., n\}$ as follows.

For $i = 0, 1, ..., m-1$:

$$\mu_i(\tau_j) = \begin{cases} 0 & j = 0 \\ \int_{\tau_0}^{\tau_j} k(\tau_j, s) l_i(s) ds & 0 < j \leq i_1 - 1 \\ \int_{\tau_0}^{\tau_{i_1-1}} k(\tau_j, s) l_i(s) ds + \int_{\tau_{i_1-1}}^{\tau_j} k(\tau_j, s) s_i(s) ds \,, \\ i_1 - 1 < j \leq i + i_1 \\ \int_{\tau_0}^{\tau_{i_1-1}} k(\tau_j, s) l_i(s) ds + \int_{\tau_{i_1-1}}^{\tau_{i_1+i}} k(\tau_j, s) s_i(s) ds \,, \; i + i_1 < j \leq n \,. \end{cases}$$

For $i = m, ..., n-m$:

$$\mu_i(\tau_j) = \begin{cases} 0 & 0 \leq j \leq i - i_0 \\ \int_{\tau_{i-i_0}}^{\tau_j} k(\tau_j, s) s_i(s) ds & i - i_0 < j \leq i + i_1 \\ \int_{\tau_{i-i_0}}^{\tau_{i+i_1}} k(\tau_j, s) s_i(s) ds & i + i_1 < j \leq n \,. \end{cases}$$

For $i = n-m+1, ..., n$:

$$\mu_i(\tau_j) = \begin{cases} 0 \qquad 0 \leq j \leq i - i_0 \\ \int_{\tau_{i-i_0}}^{\tau_j} k(\tau_j, s) s_i(s) ds & i - i_0 < j \leq n - m + i_1 \\ \int_{\tau_{i-i_0}}^{\tau_{n-m+i_1}} k(\tau_j, s) s_i(s) ds + \int_{\tau_{n-m+i_1}}^{\tau_j} k(\tau_j, s) \bar{l}_i(s) ds, \\ n - m + i_1 < j \leq n \,. \end{cases}$$

We remark that writing the system (17) in the form $A\underline{\alpha} = \underline{f}$, where $A = \{a_{ji}\}_{j,i=0}^{n}$ is the coefficient matrix, $\underline{\alpha} = [\alpha_0 ... \alpha_n]^T$, $\underline{f} = [f(\tau_0) ... f(\tau_n)]^T$, the entries of A are as follows:

$$a_{jj} = \begin{cases} 1 & j = 0 \\ 1 - \mu_j(\tau_j) & j = 1, ..., n \end{cases} \tag{19}$$

and for $j \neq i$:

$$a_{ji} = \begin{cases} -\mu_i(\tau_j) & i = 0, ..., m-1; \quad j = 1, ..., n \\ & i = m, ..., n; \quad j = i - i_0 + 1, ..., n \\ & \\ 0 & i = 1, ..., m-1; \quad j = 0 \\ & i = m, ..., n; \quad j = 0, ..., i - i_0 \; . \end{cases} \tag{20}$$

The algorithm for the numerical evaluation of $\{\mu_i(\tau_j)\}$ is based on the procedure given in [4] and on the knowledge of integrals of the type

$$\int_{x_r}^{x_{r+1}} k(\tau_j, s) s^\nu ds \quad , \quad \nu = 0, 1, ..., m-1 \; . \tag{21}$$

For some kernels, as those ones given in (2) and (3), the integrals (22) can be easily evaluated in a closed form [10].

Once we have the solution $\underline{\alpha}$ of the system (18), by (16) we can obtain the approximation $y_n(x)$ of the solution $y(x)$ of (1).

4. **Error analysis**

In order to carry out the error analysis for the proposed method, we write the integral equation (1) in the operator form

$$(I - \widetilde{K})y = f \; , \tag{22}$$

where

$$\widetilde{K}y = \int_I \widetilde{k}(x, s) y(s) ds, \quad x \in I \tag{23}$$

and

$$\widetilde{k}(x, s) = \begin{cases} k(x, s) \quad , & 0 \le s \le x \\ 0, & s > x \end{cases} \tag{24}$$

We remark that, for the kernels $k(x, s)$ considered in Section 1, $\widetilde{k}(x, s)$ satisfies the following properties:

(i) $\widetilde{k}(x, s)$ is Riemann – integrable as a function of s , for all $x \in I$,

(ii) $\lim_{x \to x'} \int_I |\widetilde{k}(x', s) - \widetilde{k}(x, s)| ds = 0$, for $x', x \in I$,

(ii) $\max_{x \in I} \int_I |\widetilde{k}(x, s)| ds < \infty$.

Therefore, we conclude that the operator $\widetilde{K}$ is a bounded compact operator on $C(I)$.

In Section 2 it has been remarked that, considering a sequence of l.u. primary partitions $\{\Pi_n\}$, the spline operator W_n is a bounded interpolating projection operator, then the condition (17) can be rewritten as

$$W_n r_n = 0 \quad \text{or, equivalently,}$$

$$(I - W_n \widetilde{K}) y_n = W_n f \; . \tag{25}$$

Now, we will prove that the equation (26) has a unique solution y_n. Then we will study the convergence of y_n to y and we will give an upper bound for $||y - y_n||$.

In order to get the above results, we prove the following lemma.

Lemma 1. *Given a sequence of l.u. partitions $\{\Pi_n\}$, for the sequence of projections $\{W_n : C(I) \to S_{\Pi_n}\}$, there results*

$$||\widetilde{K} - W_n\widetilde{K}|| \to 0 \text{ as } n \to \infty \tag{26}$$

Proof. Being $\widetilde{K} : C(I) \to C(I)$ a compact operator and since (15) with $\nu = 0$ holds, we obtain the convergence result (27). ∎

Theorem 1. *Let $\{\Pi_n\}$ be a sequence of l.u. partitions. Consider the bounded projection operator W_n from $C(I)$ to S_{Π_n}.*

For all n sufficiently large, say $n \geq N$, the operator $(I - W_n\widetilde{K})^{-1}$ from $C(I)$ to $C(I)$ exists. Moreover it is uniformly bounded, i.e.:

$$\sup_{n \geq N} ||(I - W_n\widetilde{K})^{-1}|| \leq M < \infty \tag{27}$$

and

$$||y - y_n|| \leq ||(I - W_n\widetilde{K})^{-1}|| \, ||y - W_n y|| \,. \tag{28}$$

This leads to $||y - y_n||$ converging to zero exactly with the same rate of $||y - W_n y||$.

Proof. Adapting properly the results in [1], we write:

$$I - W_n\widetilde{K} = (I - \widetilde{K})[I - (I - \widetilde{K})^{-1}(W_n\widetilde{K} - \widetilde{K})] \,.$$

Using Lemma 1, we can find an integer N such that

$$\varepsilon_N = \sup_{n \geq N} ||\widetilde{K} - W_n\widetilde{K}|| < \frac{1}{||(I - \widetilde{K})^{-1}||} \,.$$

Then, for $n \geq N$, the inverse of $[I - (I - \widetilde{K})^{-1}(W_n\widetilde{K} - \widetilde{K})]$ exists and exploiting the geometric series theorem, there results

$$||[I - (I - \widetilde{K})^{-1}(W_n\widetilde{K} - \widetilde{K})]^{-1}|| \leq \frac{1}{1 - \varepsilon_N ||(I - \widetilde{K})^{-1}||} \,.$$

Therefore:

$$||(I - W_n\widetilde{K})^{-1}|| \leq \frac{||(I - \widetilde{K})^{-1}||}{1 - \varepsilon_N ||(I - \widetilde{K})^{-1}||} \equiv M < \infty \,. \tag{29}$$

In order to show (29) we multiply (23) by W_n and then rearrange to obtain

$$(I - W_n\widetilde{K})y = W_n f + (I - W_n)y \tag{30}$$

If we subtract (26) from (31) we obtain

$$y - y_n = (I - W_n\widetilde{K})^{-1}(y - W_n y) \,,$$

and using (30) the thesis follows. ∎

5. **Condition number of the collocation method**

We can also obtain an upper bound for the condition number of the linear system (18), by adapting some general results in [1].

For a given matrix $B \in \mathbf{R}^{d\times d}$ we will use the row norm so defined:

$$||B|| = \max_{0\le j\le (d-1)} \sum_{i=0}^{d-1} |B_{j,i}| \, .$$

If we denote by $\Gamma_n = [w_i(\tau_j)]_{i,j=0}^n$, using (13), there results $\Gamma_n = I$. Thus we can write

$$||A^{-1}|| \le ||W_n||\,||\Gamma_n^{-1}||\,||(I - W_n\widetilde{K})^{-1}|| = ||W_n||\,||(I - W_n\widetilde{K})^{-1}|| \, .$$

From (20), (21) we obtain:

$$\sum_{i=0}^{n} |a_{j,i}| \le \sum_{i=0}^{n} |\mu_i(\tau_j)| + 1 \, .$$

Therefore, setting $||\widetilde{K}|| = \max_{0\le t\le X} \int_0^X |\widetilde{k}(t,s)|ds$, there results:

$$||A|| \le \max_{0\le j\le n} \sum_{i=0}^{n} \int_0^{\tau_j} |k(\tau_j, s) w_i(s)| ds + 1 \le ||W_n||\,||\widetilde{K}|| + 1$$

and then

$$\mathrm{cond}(A) \le ||W_n||\,||(I - W_n\widetilde{K})^{-1}||\,(||W_n||\,||\widetilde{K}|| + 1) \, .$$

6. **Numerical examples**

In order to test the proposed method, we consider equations of the type (1) with

$$k(x,s) = \lambda(x-s)^{-\frac{1}{2}}, \quad x \in [0,1], \quad \lambda \in \mathbf{R}$$

In particular, we shall present some numerical results in the following cases:

$$\lambda = -\frac{1}{4}, \quad f(x) = \frac{1}{\sqrt{1+x}} + \frac{\pi}{8} - \frac{1}{4}\sin^{-1}\frac{1-x}{1+x} \, , \tag{31}$$

for which the exact solution is $\frac{1}{\sqrt{1+x}}$ and

$$\lambda = -1 \quad , \quad f(x) = \sqrt{x} + \frac{1}{2}\pi x \, , \tag{32}$$

for which the exact solution is $y(x) = \sqrt{x}$.

Referring to the equation defined by (32) we use our collocation method, based on cubic nodal splines ($m = 4$) with uniform primary partition Π_n, for increasing values of n. We report in Table 1 the corresponding absolute errors $|y(x) - y_n(x)|$ evaluated at the coinciding collocation points. In the last row of the table we also present the collocation matrix condition numbers.

Table 1

$|y(x) - y_n(x)|$ *for the equation (32)*, $m = 4$

x	$n = 10$	$n = 20$	$n = 40$
.1	0.10E-5	0.37E-7	0.15E-8
.2	0.55E-6	0.24E-7	0.94E-9
.3	0.39E-6	0.18E-7	0.12E-8
.4	0.30E-6	0.14E-7	0.47E-9
.5	0.24E-6	0.12E-7	0.76E-9
.6	0.21E-6	0.10E-7	0.11E-8
.7	0.18E-6	0.93E-8	0.11E-8
.8	0.15E-6	0.78E-8	0.57E-9
.9	0.58E-7	0.10E-7	0.37E-8
1.0	0.21E-6	0.80E-8	0.95E-9
condition number	1.35	1.35	1.34

Now we consider the equation defined by (33), whose exact solution $y(x) = \sqrt{x}$ has unbounded derivatives at $x = 0$. We use our collocation method and we remark that the knowledge of the behaviour of the solution suggests the use of a sequence of graded primary meshes of the form (6). Indeed we have proved in Section 2 that such a sequence of partitions Π_n is l.u., ensuring that the hypotheses of Theorem 1 are satisfied.

In Table 2, for increasing values on n, we compare absolute errors $|y(x) - y_n(x)|$, obtained using quadratic nodal splines and uniform partitions, with those ones resulting with the same splines and graded meshes of the form (6), with $r = 2$. As it was expected, the choice of graded primary partitions allows to obtain more accurate results in particular in a neighbouring of $x = 0$. In the last row of Table 2 we carry the condition number of collocation matrix.

Table 2

$|y(x) - y_n(x)|$ for the equation (33), $m = 3$

x	$n = 10$		$n = 20$		$n = 40$	
	$r = 1$	$r = 2$	$r = 1$	$r = 2$	$r = 1$	$r = 2$
0.01	0.58E-1	0.17E-2	0.43E-1	0.28E-3	0.24E-1	0.30E-4
0.51	0.99E-3	0.23E-4	0.33E-3	0.94E-5	0.11E-3	0.22E-5
1.	0.42E-3	0.12E-3	0.14E-3	0.12E-4	0.49E-4	0.24E-5
condition number	2.46	2.33	2.44	2.28	2.43	2.26

7. **Conclusions**

In this paper we have considered the numerical solution of linear Volterra integral equations of the second kind with weakly singular kernel of the form (2) and (3). In order to obtain a sufficiently smooth approximate solution in $(0, X]$, here we have proposed and analyzed a collocation method based on optimal nodal splines.

We remark that the above method could also be applied to obtain the starting values in $[0, T]$, with $T < X$, for another one based on piecewise polynomials on $[T, X]$. Such scheme has been used in [9], with a method based on quasi interpolatory splines defined in [11].

Finally, the generalization of the obtained results to the nonlinear equations would be interesting and its systematic study is under investigation.

References

[1] K.E. Atkinson, *The numerical solution of integral equations of the second kind*, Cambridge Monographs on Appl. and Comp. Math., Cambridge University Press, 1997.

[2] C.T.H.Baker, *A perspective on the numerical treatment of Volterra equations*, J. Comp. Appl. Math. **125** (2000), 217-249.

[3] H. Brunner, *The numerical solution of weakly singular Volterra integral equations by collocation on graded meshes*, Math. Comp. **45** (1985), 417-437.

[4] C. Dagnino, V. Demichelis, *Computational aspects of numerical integration based on optimal nodal splines*, Intern. J. Computer Math. **80** (2003), 243-255.

[5] J.M. De Villiers, *A convergence result in nodal spline interpolation*, J. Approx. Theory **74** (1993),6 266-279.

[6] J.M. De Villiers, C.H. Rohwer, *Optimal local spline interpolants*, J. Comput. Appl. Math. **18** (1987), 107-119.

[7] — , *A nodal spline generalization of the Lagrange interpolant*, in “Progress in Approximation Theory” (Eds. P. Nevai and A. Pinkus), Academic Press, 1991, pp. 201-211.

[8] —, *Sharp bounds for the Lebesgue constant in quadratic nodal spline interpolation*, International Series of Numerical Mathematics, vol. 115, Birkhäuser Verlag, 1994 (1-13).

[9] L.Gori, E. Santi, *A spline method for the numerical solution of Volterra integral equations of the second kind*. In “Integral and integro-differential equations” (Eds. R. Agarwal and D. O’Regan) vol. 2 Bookseries in Mathematical Analysis and Application (1999), 91-99.

[10] I.S. Gradshteyn, I.M. Ryzhik, *Tables of Integrals, Series and Products*, Academic Press, 2000.

[11] T. Lyche, L.L.Schumaker, *Local splines approximation methods*, J. Approx. Theory **15** (1975), 294-325.

[12] G. Micula, S. Micula, *Handbook of Splines*, Math. and its Appl. 462, Kluwer 1999.

[13] R.K. Miller, A. Feldstein, *Smoothness of solutions of Volterra integral equations with weakly singular kernels*, SIAM J. Math. Anal. **2** (1971), 242-

[14] L.L.Schumaker, *Spline functions: Basic Theory*, J. Wiley and Sons, New York, 1981.

Dipartimento di Matematica, Università di Torino,
Via Carlo Alberto 10, 10123 Torino, Italy
E-mail address: catterina.dagnino@unito.it

Dipartimento di Matematica, Università di Torino,
Via Carlo Alberto 10, 10123 Torino, Italy
E-mail address: vittoria.demichelis@unito.it

Dipartimento di Energetica, Università dell’Aquila
Monteluco di Roio, 67040 L-Aquila, Italy
E-mail address: esanti@dsiaq1.ing.univaq.it

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

PRICING DIGITAL CALL OPTION IN THE HESTON STOCHASTIC VOLATILITY MODEL

VASILE L. LAZAR

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. The aim of this paper is to analyze the problem of digital option pricing under a stochastic volatility model, namely the Heston model (1993). In this model the variance v, follows the same square-root process as the one used by Cox, Ingersoll and Ross (1985) from the short term interest rate. We present an analytical solution for this kind of options, based on S. Heston's original work [3].

1. **Introduction**

Options on stock were first traded in an organized way on The Chicago Board Option Exchange in 1973, but the theory of option pricing has its origin in 1900 in "Théorie de la Spéculation" of L.Bachelier. In the early 1970's, after the introduction of geometric Brownian motion, Fischer Black and Myron Scholes made a major breakthrough by deriving the Black-Scholes formula which is one of the most significant results in pricing financial instruments [1].

We begin by presenting some underlying knowledge about basic concepts of derivatives and pricing methods.

A financial derivative is a financial instrument whose payoff is based on other elementary financial instruments, such as bonds or stocks. The most popular financial derivatives are: forward contracts, futures, swaps and options.

Options are particular derivatives characterized by non-negative payoffs. There are two basic types of option contracts: call options and put options.

Definition 1.1. *A* ***call*** *option gives the holder the right to buy a prescribed asset, the underlying asset, with a specific price, called the exercise price or strike price, at a specified time in future, called expiry or expiration date.*

Definition 1.2. *A* ***put*** *option gives the holder the right to sell the underlying asset, with an agreed amount at a specified time in future.*

The options can also be classified based on the time in which they can be exercised:

- A European option can only be exercised at expiry;

Received by the editors: 02.04.2003.
Key words and phrases. option pricing, stochastic volatility, digital options, characteristic function.

- An American option can be exercised at any time up to and including the expiry

1.1. **Payoff Function.** Let S be the current price of the underlying asset and K be the strike price.Then, at expiry a European call option is worth:

$$max(S_T - K, 0) \tag{1.1}$$

This means that, the holder will exercise his right only if $S_T > K$ and than his gain is $S_T - K$. Otherwise, if $S_T \le K$, the holder will buy the underlying asset from the market and then the value of the option is zero.

The function (1.1) of the underlying asset is called **the payoff function**.

The payoff function from a European put option is:

$$max(K - S_T, 0) \tag{1.2}$$

Any option with a more complicated payoff structure than the usual put and call payoff structure is called an exotic option. In theory exists an unlimited number of possible exotic options but in practice there are only a few that have seen much use: digital or binary options, lookback options, barrier options, compound options, Asian options.

Digital options have a payoff that is discontinuous in the underlying asset price. For a digital call option with strike K at time T, the payoff is a Heaviside function:

$$DC(S\ ,\ T) = \mathcal{H}(S_T - K) = \begin{cases} 1 & \text{if } S_T \ge K \\ 0 & \text{if } S_T < K \end{cases} \tag{1.3}$$

and for a digital put option:

$$DP(S\ ,\ T) = \mathcal{H}(K - S_T) = \begin{cases} 1 & \text{if } S_T < K \\ 0 & \text{if } S_T \ge K \end{cases} \tag{1.4}$$

1.2. **Black-Scholes Formulae.** In 1973 Fischer Black and Myron Scholes derived a partial differential equation governing the price of an asset on which an option is based, and then solved it to obtain their formula for the price of the option, see [1].

We use the following notation:
S - the price of the underlying asset;
K - the exercise price;
t - current date;
T - the maturity date;
τ - time to maturity, $\tau = T - t$;
r - the risk free interest rate;
v - standard deviation of the underlying asset, i.e the volatility;
μ - the drift rate.

The assumptions used to derive the Black-Scholes partial differential equations are:

- the value of underlying asset is assumed to follow the log-normal distribution:

$$dS = \mu\, S\, dt + v\, S\, dW\ , \tag{1.5}$$

where the term $W(t)$ is a stochastic process with mean zero and variance t known as a Wiener process;
- the drift, μ, and the volatility, v, are constant throughout the option's life;
- there are no transaction costs or taxes;
- there are no dividends during the life of the option;
- no arbitrage opportunity;
- security trading is continuous;
- the risk-free rate of interest is constant during the life of the option.

Further, we give the most important result of stochastic calculus, Itô's lemma. Itô's lemma gives the rule for finding the differential of a function of one or more variables who follow a stochastic differential equation containing Wiener processes.

Lemma 1.1. *(One-dimensional Itô formula). Let the variable $x(t)$ follow the stochastic differential equation*

$$dx(t) = a(x, t)\, dt + b(x, t)\, dW.$$

Further, let $F(x(t), t) \in C^{2,1}$ be at least a twice differentiable function. Then the differential of $F(x, t)$ is given by:

$$dF = \left[\frac{\partial F}{\partial x} a(x,t) + \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2} b^2(x,t)\right] dt + \frac{\partial F}{\partial x} b(x,t)\, dW . \tag{1.6}$$

Proof: The proof of this lemma and the multi-dimensional case can be found in [4].

Using Itô's lemma and the foregoing assumptions, Black and Scholes have obtained the following partial differential equation for the option price $V(S, t)$:

$$\frac{\partial V}{\partial t} + \frac{1}{2} v^2 S^2 \frac{\partial^2 V}{\partial S^2} + r S \frac{\partial V}{\partial S} - r V = 0 . \tag{1.7}$$

In order to obtain a unique solution for the Black-Scholes equation we must consider final and boundary conditions. We will restrict our attention to a European call option, $C(S, t)$.

At maturity, $t = T$, a call option is worth:

$$C(S, T) = max(S_T - K, 0) \tag{1.8}$$

so this will be the final condition.

The asset price boundary conditions are applied at $S = 0$ and as $S \longrightarrow \infty$.

If $S = 0$ then dS is also zero and therefore S can never change. This implies on $S = 0$ we have:

$$C(0, t) = 0 . \tag{1.9}$$

Obviously, if the asset price increases without bound $S \longrightarrow \infty$, then the option will be exercised indifferently how big is the exercise price. Thus as $S \longrightarrow \infty$ the value of the option becomes that of the asset:

$$C(S, t) \approx S , S \longrightarrow \infty . \tag{1.10}$$

We have now the following final-boundary value problem:

$$\begin{cases} \frac{\partial C}{\partial t} + \frac{1}{2} v^2 S^2 \frac{\partial^2 C}{\partial S^2} + r S \frac{\partial C}{\partial S} - r C = 0 \\ C(0 , t) = 0 ; C(S , t) \approx S \text{ as } S \to \infty \\ C(S , T) = max(S_T - K , 0) \end{cases}$$

The analytical solution of this problem has the following functional form:

$$C(S , t) = S N(d_1) - K e^{-r(T-t)} N(d_2) \tag{1.11}$$

where

$$d_1 = \frac{log(S/K) + \left(r + \frac{1}{2} v^2\right) (T - t)}{v \sqrt{T - t}} \tag{1.12}$$

and

$$d_2 = \frac{log(S/K) + \left(r - \frac{1}{2} v^2\right) (T - t)}{v \sqrt{T - t}} \tag{1.13}$$

$N(x)$ is the *cumulative distribution function* for the standard normal distribution.

Similarly the price for a European put option is:

$$P(S , t) = - S N(- d_1) - K e^{-r(T-t)} N(- d_2) \tag{1.14}$$

In the digital option case, where we have the following final condition $DC(S , T) = \mathcal{H}(S_T - K)$, the solution for the option price equation is:

$$DC(S , t) = e^{-r(T-t)} N(d_2) \tag{1.15}$$

2. **Heston's Stochastic Volatility Model**

In the standard Black-Scholes model the volatility is assumed to be constant. Naturally the Black-Scholes assumption is incorrect and in reality volatility is not constant and it's not even predictable for timescales of more than a few months. This fact led to the development of stochastic volatility models, in which volatility itself is assumed to be a stochastic process.

We assume that S satisfies

$$dS = \mu S dt + v S dW_1 , \tag{2.1}$$

and, in addition the volatility follows the stochastic process:

$$dv = p(S , v , t) dt + q(S , v , t) dW_2 \tag{2.2}$$

where the two increments dW_1 and dW_2 have a correlation of ρ.

In this case the value V is not only a function of S and time t, it is also a function of the variance v, $V(S , v , t)$. The partial differential equation governing the option price is a generalization of Black-Scholes equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2} v^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho v S q \frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial v^2} + r S \frac{\partial V}{\partial S} + (p - \lambda q) \frac{\partial V}{\partial v} - r V = 0 . \tag{2.3}$$

where λ is the market price of volatility risk.

Examples of these models in continuous-time include Hull and White(1987), Johnson and Shanno(1987), Wiggins(1987), Stein and Stein(1991), Heston(1993), Bates(1996),and examples in discrete-time include Taylor(1986), Amin and Ng(1993) and Heston and Nandi(1993).

Among them, Heston's model is very popular because of its three main features:

- it does not allow negative volatility;
- it allows the correlation between asset return and volatility;
- it has a closed-form pricing formula.

Heston's option pricing formula is derived under the assumption that the stock price and its volatility follow the stochastic processes:

$$dS(t) \;=\; S(t)\;[\;\mu\; dt \;+\; \sqrt{v(t)}\; dW_1(t)\;] \tag{2.4}$$

and

$$dv(t) \;=\; k\;(\;\theta \;-\; v(t)\;)\; dt \;+\; \xi\; \sqrt{v(t)}\; dW_2(t)\;, \tag{2.5}$$

where:

$$\mathbf{Cov}[\; dW_1(t)\;,\; dW_2(t)\;] \;=\; \rho\; dt\;. \tag{2.6}$$

Finally, the market price of volatility risk is given by:

$$\lambda(S,v,t) \;=\; \lambda\; v\;. \tag{2.7}$$

According to the pricing equation (2.3) we have the following partial differential equation for the Heston model:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\, v\, S^2\, \frac{\partial^2 V}{\partial S^2} + \rho\,\sigma\, v\, S\, \frac{\partial^2 V}{\partial S\,\partial v} + \frac{1}{2}\,\sigma^2\, v\, \frac{\partial^2 V}{\partial v^2} + r\, S\, \frac{\partial V}{\partial S} + [\; k\,(\theta \;-\; v\;) \;-\lambda\, v\;]\, \frac{\partial V}{\partial v} \;-\; r\; V \;=\; 0\;. \tag{2.8}$$

The details of deriving the above equation and its closed-form solution, for a European call option, can be found in Heston's original work [3].

3. A Closed-Form Solution for a Digital Call Option in the Heston Model

In what follows we solve the partial differential equation (2.8) subject to the final condition:

$$DC(S\;,\;v\;,\;T) \;=\; \mathcal{H}(S_T \;-\; K) \;=\; \begin{cases} 1 & \text{if } S_T \;\geq\; K \\ 0 & \text{if } S_T \;<\; K \end{cases} \tag{3.1}$$

In order to simplify our work it is convenient to make the following substitution $x \;=\; \ln[S]$, $U(x\;,\;v\;,\;t) \;=\; V(S\;,\;v\;,\;t)$. Then the equation (2.8) is turn into

$$\frac{\partial U}{\partial t} + \frac{1}{2}\,\sigma^2\, v\, \frac{\partial^2 U}{\partial v^2} + \rho\,\sigma\, v\, \frac{\partial^2 U}{\partial x \partial v} + \frac{1}{2}\, v\, \frac{\partial^2 U}{\partial x^2} + \left(r \;-\; \frac{1}{2}\, v\right) \frac{\partial U}{\partial x} + [k\;(\theta \;-\; v) \;-\; v\;\lambda]\, \frac{\partial U}{\partial v} \;-\; r\; U \;=\; 0\;. \tag{3.2}$$

By analogy with the Black-Scholes formula (1.15), we guess a solution of the form:

$$DC(S\ ,\ v\ ,\ t)\ =\ e^{-r\ \tau}\ P \tag{3.3}$$

where the probability P correspond to $N(d_2)$ in the constant volatility case. P is the conditional probability that the option expires in-the-money:

$$P(x\ ,\ v\ ,\ T\ ;\ ln[K])\ =\ Pr[x(T)\ \geq\ ln[K]\ /\ x(t)\ =\ x\ ,\ v(t)\ =\ v]\ . \tag{3.4}$$

We now substitute the proposed value for $DC(S\ ,\ v\ ,\ t)$ into the pricing equation (3.2). We obtain:

$$e^{-r\tau}\frac{\partial P}{\partial t} + rPe^{-r\tau} + \frac{1}{2}\sigma^2 ve^{-r\tau}\frac{\partial^2 P}{\partial v^2} + \rho\sigma ve^{-r\tau}\frac{\partial^2 P}{\partial x\partial v} + \frac{1}{2}ve^{-r\tau}\frac{\partial^2 P}{\partial x^2} +$$
$$+\left(r - \frac{1}{2}v\right)e^{-r\tau}\frac{\partial P}{\partial x} + [k(\theta - v) - v\lambda]e^{-r\tau}\frac{\partial P}{\partial v} - rPe^{-r\tau} = 0\ . \tag{3.5}$$

This implies that P must satisfy the equation:

$$\frac{\partial P}{\partial t}\ +\ \frac{1}{2}\ \sigma^2\ v\ \frac{\partial^2 P}{\partial v^2}\ +\ \rho\ \sigma\ v\ \frac{\partial^2 P}{\partial x\partial v}\ +\ \frac{1}{2}\ v\ \frac{\partial^2 P}{\partial x^2}\ +\ \left(r\ -\ \frac{1}{2}\ v\right)\ \frac{\partial P}{\partial x}$$
$$+\ [k\ (\theta\ -\ v)\ -\ v\ \lambda]\ \frac{\partial P}{\partial v}\ =\ 0 \tag{3.6}$$

subject to the terminal condition:

$$P(x\ ,\ v\ ,\ T\ ;\ ln[K])\ =\ 1_{\{x\ \geq\ ln[K]\}}\ . \tag{3.7}$$

The probabilities are not immediately available in closed-form, but the next part shows that their characteristic function satisfy the same partial differential equation (3.6).

3.1. **The Characteristic Function.** Suppose that we have given the two processes

$$dx(t)\ =\ \left(r\ -\ \frac{1}{2}\ v(t)\right)\ dt\ +\ \sqrt{v(t)}\ dW_1(t) \tag{3.8}$$

$$dv(t)\ =\ [k\ (\theta\ -\ v(t))\ -\ \lambda\ v(t)]\ dt\ +\ \sigma\ \sqrt{v(t)}\ dW_2(t) \tag{3.9}$$

with

$$cov[dW_1(t)\ ,\ dW_2(t)]\ =\ \rho\ dt \tag{3.10}$$

and a twice-differentiable function

$$f(x(t)\ ,\ v(t)\ ,\ t)\ =\ E[g(x(T)\ ,\ v(T))\ /\ x(t)\ =\ x\ ,\ v(t)\ =\ v]\ . \tag{3.11}$$

From Itô's lemma we obtain:

$$df\ =\ \left(\frac{1}{2}\ \sigma^2\ v\ \frac{\partial^2 f}{\partial v^2}\ +\ \rho\ \sigma\ v\ \frac{\partial^2 f}{\partial x\partial v}\ +\ \frac{1}{2}\ v\ \frac{\partial^2 f}{\partial x^2}\ +\ \left(r\ -\ \frac{1}{2}\ v\right)\ \frac{\partial f}{\partial x}\right.$$
$$\left.+\ [k\ (\theta\ -\ v)\ -\ v\ \lambda]\ \frac{\partial f}{\partial v}\ +\ \frac{\partial f}{\partial t}\right)\ dt$$
$$+\ \left(r\ -\ \frac{1}{2}\ v\right)\ \frac{\partial f}{\partial x}\ dW_1\ +\ [k\ (\theta\ -\ v)\ -\ v\ \lambda]\ dW_2$$

In addition, by iterated expectations, we know that $f(x(t)\ ,\ v(t)\ ,\ t)$ is a martingale, therefore the df coefficient must vanisch, i.e.,

$$\frac{1}{2}\,\sigma^2\,v\,\frac{\partial^2 f}{\partial v^2}\;+\;\rho\,\sigma\,v\,\frac{\partial^2 f}{\partial x \partial v}\;+\;\frac{1}{2}\,v\,\frac{\partial^2 f}{\partial x^2}\;+ \\ \left(r\;-\;\frac{1}{2}\,v\right)\,\frac{\partial f}{\partial x}\;+\;[k\,(\theta\;-\;v)\;-\;v\,\lambda]\,\frac{\partial f}{\partial v}\;+\;\frac{\partial f}{\partial t}\;=\;0\;. \tag{3.12}$$

Equation (3.11) imposes the final condition

$$f(x\ ,\ v\ ,\ T)\;=\;g(x\ ,\ v) \tag{3.13}$$

Depending on the choice of g, the function f represents different objects. Choosing $g(x\ ,\ v)\;=\;e^{i\varphi x}$ the solution is the characteristic function, which is available in closed form. In order to solve the partial differential equation (3.12) with the above condition we invert the time direction: $\tau\;=\;T\;-\;t$. This mean that we must solve the following equation:

$$\frac{1}{2}\,\sigma^2\,v\,\frac{\partial^2 f}{\partial v^2}\;+\;\rho\,\sigma\,v\,\frac{\partial^2 f}{\partial x \partial v}\;+\;\frac{1}{2}\,v\,\frac{\partial^2 f}{\partial x^2}\;+ \\ \left(r\;-\;\frac{1}{2}\,v\right)\,\frac{\partial f}{\partial x}\;+\;[k\,(\theta\;-\;v)\;-\;v\,\lambda]\,\frac{\partial f}{\partial v}\;-\;\frac{\partial f}{\partial t}\;=\;0 \tag{3.14}$$

subject to the initial condition:

$$f(x\ ,\ v\ ,\ 0)\;=\;e^{i\varphi x} \tag{3.15}$$

We guess a solution, from this equation, of the form:

$$f(x\ ,\ v\ ,\ \tau)\;=\;e^{C(\tau)\;+\;D(\tau)\;v\;+\;i\varphi x} \tag{3.16}$$

with initial condition $C(0)\;=\;D(0)\;=\;0$.

This "guess" is due to the linearity of the coefficients.

Substituting the functional form (3.16) into equation (3.14) we find that:

$$\frac{1}{2}\,\sigma^2\,v\,D^2\,f\;+\;\rho\,\sigma\,v\,i\,\varphi\,D\,f\;-\;\frac{1}{2}\,v\,\varphi^2\,f\;+ \\ \left(r\;-\;\frac{1}{2}\,v\right)\,i\,\varphi\,f\;+\;[k\,(\theta\;-\;v)\;-\;v\,\lambda]\,D\,f\;-\;(C'\;+\;D'\,v)\,f\;=\;0$$

Therefore

$$v\,\left(\frac{1}{2}\,\sigma^2\,D^2\;+\;\rho\,\sigma\,i\,\varphi\,D\;-\;\frac{1}{2}\,\varphi^2\;-\;\frac{1}{2}\,i\,\varphi\;-\;(k\;+\;\lambda)\,D\;-\;D'\right)\;+ \\ +\;(r\,i\,\varphi\;+\;k\,\theta\,D\;-\;C')\;=\;0.$$

This can be reduce in two ordinary differential equations:

$$\text{a) } D'\;=\;\frac{1}{2}\,\sigma^2\,D^2\;+\;\rho\,\sigma\,i\,\varphi\,D\;-\;\frac{1}{2}\,\varphi^2\;-\;\frac{1}{2}\,i\,\varphi\;-\;(k\;+\;\lambda)\,D \tag{3.17}$$

and

$$\text{b) } C'\;=\;r\,i\,\varphi\;+\;k\,\theta\,D\;. \tag{3.18}$$

Basic theory on differential equation, including the Riccati equation, can be found in [7] a)We shall solve the Riccati differential equation

$$D' = \frac{1}{2}\sigma^2 D^2 + (\rho\,\sigma\,i\,\varphi - k - \lambda)\,D - \frac{1}{2}\varphi^2 - \frac{1}{2}\,i\,\varphi$$

using the substitution:

$$D = -\frac{E'}{\frac{\sigma^2}{2}\,E}$$

It follows that

$$E" - (\rho\,\sigma\,i\,\varphi - k - \lambda)\,E' + \frac{\sigma^2}{2}\left(-\frac{1}{2}\varphi^2 - \frac{1}{2}\,i\,\varphi\right) = 0 \tag{3.19}$$

Then the characteristic equation is

$$x^2 - (\rho\,\sigma\,i\,\varphi - k - \lambda)\,x + \frac{\sigma^2}{4}\left(-\varphi^2 - i\,\varphi\right) = 0\,.$$

Consequently, if we make the following notation

$$d = \sqrt{(\rho\,\sigma\,i\,\varphi - k - \lambda)^2 - \sigma^2\,(-\varphi^2 - i\,\varphi)}\,,$$

then the equation (3.19) has the general solution

$$E(\tau) = A\,e^{x_1\tau} + B\,e^{x_2\tau}\,,$$

where

$$x_{1,2} = \frac{\rho\,\sigma\,i\,\varphi - k - \lambda \pm d}{2}\,.$$

The boundary conditions

$$\begin{cases} E(0) = A + B \\ A\,x_1 + B\,x_2 = 0 \end{cases}$$

yield

$$A = \frac{g\,E(0)}{g - 1}$$

$$B = -\frac{E(0)}{g - 1}$$

where $g = \frac{x_1}{x_2}$. Hence we obtain

$$E(\tau) = \frac{E(0)}{g - 1}\,(g\,e^{x_1\tau} - e^{x_2\tau})$$

$$E'(\tau) = \frac{E(0)}{g - 1}\,(g\,x_1\,e^{x_1\tau} - x_2\,e^{x_2\tau})$$

and thus

$$D(\tau) = -\frac{2}{\sigma^2}\,\frac{E'}{E} = -\frac{2}{\sigma^2}\,x_2\,\frac{e^{x_2\tau} - e^{x_1\tau}}{e^{x_2\tau} - g\,e^{x_1\tau}}$$

Therefore our equation has the following solution:

$$D(\tau) = \frac{k + \lambda + d - \rho\,\sigma\,\varphi\,i}{\sigma^2}\left[\frac{1 - e^{d\tau}}{1 - g\,e^{d\tau}}\right] \tag{3.20}$$

where

$$d = \sqrt{(\rho\,\sigma\,\varphi\,i - k - \lambda)^2 - \sigma^2\,(-\varphi^2 - i\,\varphi)} \tag{3.21}$$

$$g \ = \ \frac{\rho\ \sigma\ \varphi\ i\ -\ k\ -\ \lambda\ -\ d}{\rho\ \sigma\ \varphi\ i\ -\ k\ -\ \lambda\ +\ d} \tag{3.22}$$

b)The second equation can be solved by mere integration:

$$\begin{aligned} C(\tau) \ &= \ r\ i\ \varphi\ \tau\ +\ k\ \theta \int_{\tau}^{0} -\ \frac{E'(s)}{\frac{\sigma^2}{2}\ E(s)}\ ds \\ &= \ r\ i\ \varphi\ \tau\ -\ \frac{2\ k\ \theta}{\sigma^2} \int_{\tau}^{0} \frac{E'(s)}{E(s)}\ ds \\ &= \ r\ i\ \varphi\ \tau\ -\ \frac{2\ k\ \theta}{\sigma^2}\ \ln \frac{E(\tau)}{E(0)}\ . \end{aligned}$$

It follows that

$$C(\tau) = ri\varphi\tau + \frac{k\theta}{\sigma^2}\left[(k+\lambda+d-\rho\sigma\varphi i)\tau - 2\ln\left(\frac{1-ge^{d\tau}}{1-e^{d\tau}}\right)\right]\ . \tag{3.23}$$

3.2. **Solution of the Digital Call Option.** We can invert the characteristic functions to get the desired probabilities, using a standard result in probability,that is, if F(x) is a one-dimensional distribution function and f its corresponding characteristic function, then the cumulative distribution function $F(x)$ and its corresponding density function $\phi(x) \ = \ F'(x)$ can be retrieved via:

$$\phi(x) \ = \ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz}\ f(t)\ dt \tag{3.24}$$

$$F(x) \ = \ \frac{1}{2} \ + \ \frac{1}{2\pi} \int_{0}^{\infty} \frac{e^{itx}\ f(-t)\ -\ e^{-itx}\ f(t)}{i\ t}\ dt \tag{3.25}$$

or

$$F(x) \ = \ \frac{1}{2} \ - \ \frac{1}{\pi} \int_{0}^{\infty} \mathcal{R}e \left[\frac{e^{itx}\ f(t)}{i\ t}\right]\ dt \tag{3.26}$$

This result is showed by J.Gil-Pelaez in [2].

Thus, we get the desired probability:

$$P(x\ ,\ v\ ,\ t\ ;\ \ln K) \ = \ \frac{1}{2} \ + \ \frac{1}{\pi} \int_{0}^{\infty} \mathcal{R}e \left[\frac{e^{-i\varphi \ln K}\ f(x\ ,\ v\ ,\ \tau\ ,\ \varphi)}{i\ \varphi}\right]\ d\varphi \tag{3.27}$$

We can summarize the above relations in the following Theorem:

Theorem 3.1. *Consider a Digital call option in the Heston model, with a strike price of K and a time to maturity of τ. Then the current price is given by the following formula:*

$$DC(S\ ,\ v\ ,\ t) \ = \ e^{-r\ \tau}\ P$$

where the probability function, P is given by:

$$P(x\ ,\ v\ ,\ t\ ;\ \ln K) \ = \ \frac{1}{2} \ + \ \frac{1}{\pi} \int_{0}^{\infty} \mathcal{R}e \left[\frac{e^{-i\varphi \ln K}\ f(x\ ,\ v\ ,\ \tau\ ,\ \varphi)}{i\ \varphi}\right]\ d\varphi$$

and the characteristic function is:

$$f(x\ ,\ v\ ,\ \tau) \ = \ e^{C(\tau)\ +\ D(\tau)\ v\ +\ i\varphi x}$$

where $C(\tau)$ and $D(\tau)$ are given by (3.23) and (3.20) respectively.

References

[1] Black, F & Scholes, M (1973) : *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 81:637-654;

[2] Gil-Pelaez, L (1951): *Note on the Inversion Theorem*, Biometrica 38, 481-482;

[3] Heston, S (1993): *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, 6:327-343;

[4] Korn, R & Korn, E (1999): *Option Pricing and Portfolio Optimization-Modern Methods of Financial Mathematics*, American Math.Soc.;

[5] Nagel, H. (1999): *Zur Optionsbewertung bei stochastischer Volatilität am Beispiel der DTB: Eine theoretische und empirische Analyse*, Inauguraldissertation zur Erlangung des Doktorgrades der Wirtschaftswissenschaftlichen Fakultät der Eberhard-Karls-Universität Tübingen;

[6] Precup, R (1997): *Ecuatii cu derivate partiale*, Transilvania Press;

[7] Rus, I.A. (1996): *Ecuatii diferentiale, ecuatii integrale si sisteme dinamice*, Transilvania Press;

[8] Shreve, S.E (1996): *Stochastic Calculus and Finance*, Lecture notes, Carnegie Mellon University;

[9] Wilmott, P (1999): *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley & Sons Ltd;

[10] Wilmott, P & Dewynne, J & Howison, S (1993): *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press;

[11] Wystup, U et al. MathFinance, URL: *http://www.mathfinance.de*

Fraunhofer-Institut für Techno- und Wirtschaftsmathematik,
Gottlieb-Daimler-Str.,Geb. 49, 67663, Kaiserslautern, Germany
E-mail address: vasilazar@yahoo.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

A CLASS OF EVEN DEGREE SPLINES OBTAINED THROUGH A MINIMUM CONDITION

GH. MICULA, E. SANTI, AND M. G. CIMORONI

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. A class of splines minimizing a special functional is investigated. This class is determined by the solution of quadratic programming problem. Convergence results and some numerical examples are given.

1. **Introduction**

The construction of splines, verifying minimum conditions has been proposed among others in [2], [5], [6]. In such papers the splines are interpolating the approximated function in the nodes and while in [5] the constructive method can be applied, in theory, for a spline of an arbitrary degree *m,* minimizing the integral $\int_I [g'(x)]^2 dx$, in [2] e [6] a cubic polynomial interpolating splines are considered satisfying some minimum conditions.

In particular, in [6], the considered splines have been applied for constructing quadrature sums approximating the Cauchy principal value integrals

$$I(wf;t) = ⨍_{-1}^{1} w(x)\frac{f(x)}{x-t}dx. \tag{1.1}$$

In this paper, utilizing the method proposed in [2], we construct the spline of even degree minimizing the functional

$$F(f) := \int_I [f^{(3)}(x)]^2 dx \qquad\qquad f \in W_2^3(I) \tag{1.2}$$

where, denoting $AC(I)$ the set of absolute continuous functions on I,

$$W_2^3(I) := \left\{ f : I \to I\!R, \quad f^{(0)} \in AC(I) \quad and \quad f^{(3)} \in L_2(I) \right\} . \tag{1.3}$$

This class of splines, called *interpolating-derivative* splines of degree $2m$, $m \geqslant 2$, has been determined in [3] by solving a linear system of $m+n+1$ equations, where n is the number of internal knots of the partition, and then the authors proved that the constructed spline solves problem (1.2). The convergence is proved by supposing $f \in W_2^{m+1}$.

Received by the editors: 14.05.2003.

In this paper, exploiting the different form that we use for defining the interpolating - derivative spline, we can obtain convergence results under weaker conditions on f, that gives more flexibility in the applications, as for example, when we consider the numerical evaluation of Cauchy singular integrals [8].

In Section 2 we give the details of the construction of the interpolating-derivative spline. In Section 3 we give some convergence results. Finally, in Section 4, some numerical experiments on test functions f are reported. In Appendix we prove some propositions whose results are necessary for proving theorem 2.5 and proposition 2.7 in Section 2 and theorems 3.2, 3.3 in Section 3.

2. **Construction of derivative-interpolating spline**

Let $m,\ n \geqslant m$ two given integer positive numbers and $Y \in I\!R^{n+1}, Y := \{y_0, y_1', ..., y_n'\}$ a given vector and

$$\Delta_n := \{a = x_0 < x_1 < \ldots < x_n < x_{n+1} = b\}$$

a given partition of $I \equiv [a, b]$ in $n+1$ subintervals $I_k := [x_k, x_{k+1}),\ k = 0, 1, ..., n$, limiting ourselves, for the sake of simplicity, to consider an uniform partition Δ_n, with $h = x_{i+1} - x_i\ ,\ i = 0, 1, ..., n$.

We denote by $I\!P_k$ the set of polynomials of degree $\leq k$. Consider the space of polynomial splines of degree $2m$

$$S_{2m}(\Delta_n) = \left\{ \begin{array}{l} s : s(x) = s_i(x) \in I\!P_{2m},\ x \in I_i,\ i = 0, 1, ..., n; \\ D^j s_{i-1}(x_i) = D^j s_i(x_i),\ j = 0, 1, ..., 2m-1,\ i = 1, 2, ..., n \end{array} \right\} \tag{2.1}$$

with simple knots $x_1, x_2, ..., x_n$. The space $S_{2m}(\Delta_n) \subset C^{2m-1}(I)$.

A function $s_f \in S_{2m}(\Delta_n)$ is called *derivative-interpolating* if

$$s_f(x_0) = y_0,\ s_f'(x_i) = y_i',\ i = 1, 2, ..., n;\ y_0 = f(x_0),\ y_i' = f(x_i). \tag{2.2}$$

Limiting ourselves to consider $m = 2$, if we set

$$M_i = s_f^{(2m-1)}(x_i),\ i = 0, 1, ..., n+1,$$

by successive integrations, we obtain

$$\begin{array}{rl} s_f(x)|_{I_i} = & [M_{i+1}(x - x_i)^4 - M_i(x - x_{i+1})^4]/(4!h) + \\ & +a_i(x - x_i)^2/2 + b_i(x - x_i) + c_i, \quad i = 0, 1, ..., n. \end{array} \tag{2.3}$$

By imposing the conditions (2.1) and (2.2), we obtain

$$\begin{cases} a_i &= \frac{y'_{i+1}-y'_i}{h} - \frac{h}{6}(M_{i+1}-M_i) & i=1,...,n-1 \\ b_0 &= y'_1 - \frac{h^2}{6}M_1 - a_0 h & \\ b_i &= y'_i - \frac{h^2}{6}M_i & i=1,...,n \\ c_0 &= y_0 + \frac{h^3}{24}M_0 & \\ c_1 &= c_0 + y'_1 h - \frac{h^3}{12}M_1 - \frac{h^2}{2}a_0 & \\ c_i &= c_{i-1} + (y'_i + y'_{i-1})\frac{h}{2} - \frac{h^3}{12}M_{i-1} & i=2,...,n \\ M_i h &= a_i - a_{i-1} & i=1,...,n. \end{cases} \tag{2.4}$$

Substituting the first equations of (2.4) in the last ones, we obtain a linear system

$$\widetilde{A}\widetilde{M} = \underline{b}^*(a_0, a_n) \tag{2.5}$$

where

$$\widetilde{A} = \begin{bmatrix} 5 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 5 \end{bmatrix}, \quad \widetilde{M} = \begin{bmatrix} M_1 \\ . \\ . \\ . \\ M_n \end{bmatrix},$$

$$\underline{b}^* = \frac{6}{h}\left[\frac{y'_2 - y'_1}{h} - a_0, \ldots, \frac{y'_{i+1} - y'_i}{h} - \frac{y'_i - y'_{i-1}}{h}, \ldots, -\frac{y'_n - y'_{n-1}}{h} + a_n\right]^T.$$

The spline function $s_f(x)$ will be determined by solving the following problem

$$\begin{cases} \min M^T \bar{A} M \\ \widetilde{A}\widetilde{M} = \underline{b}^*(a_0, a_n) \end{cases} \tag{2.6}$$

with $M = [M_{0,\ldots,}M_{n+1}]^T$,

$$\bar{A} = \left[\begin{array}{c|ccccc|c} 2 & 1 & & & & & 0 \\ \hline 1 & 4 & 1 & & & & \\ & 1 & \ddots & \ddots & & & \\ & & \ddots & \ddots & 1 & & \\ & & & 1 & 4 & & 1 \\ \hline 0 & & & & 1 & & 2 \end{array}\right] = \left[\begin{array}{c|c|c} 2 & \underline{e}_1^T & 0 \\ \hline \underline{e}_1 & A^* & \underline{e}_n \\ \hline 0 & \underline{e}_n^T & 2 \end{array}\right], \tag{2.7}$$

where

$$A^* = \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{bmatrix} \tag{2.8}$$

and $\underline{e}_1$, $\underline{e}_n$ are the vectors $[1, 0, ..., 0]^T$, $[0, 0, ..., 1]^T$ respectively.

We can write $\underline{b}^* = \underline{b} - \underline{e}_1 \widetilde{a}_0 + \underline{e}_n \widetilde{a}_n$ with

$$\underline{b} = \frac{6}{h} \left[\frac{y_2' - y_1'}{h}, \ldots, \frac{y_{i+1}' - y_i'}{h} - \frac{y_i' - y_{i-1}'}{h}, \ldots, -\frac{y_n' - y_{n-1}'}{h} \right]^T \tag{2.9}$$

and $\widetilde{a}_0 = \frac{6}{h} a_0$, $\widetilde{a}_n = \frac{6}{h} a_n$.

Considering that $\widetilde{A}$ is a symmetric positive definite and then, non singular matrix, from (2.5) we get

$$\widetilde{M} = \widetilde{A}^{-1} (\underline{b} - \underline{e}_1 \widetilde{a}_0 + \underline{e}_n \widetilde{a}_n) \tag{2.10}$$

thus:

$$\min M^T \bar{A} M = \min \left\{ \left[M_0 \widetilde{M}^T M_{n+1} \right] \begin{bmatrix} 2 & \underline{e}_1^T & 0 \\ \underline{e}_1 & A^* & \underline{e}_n \\ 0 & \underline{e}_n^T & 2 \end{bmatrix} \left[M_0 \widetilde{M}^T M_{n+1} \right]^T \right\}. \tag{2.11}$$

Using (2.10), the problem amounts to find out firstly the vector

$$N = [\widetilde{a}_0, -\widetilde{a}_n, -M_0, -M_{n+1}]^T,$$

solution of the linear system

$$BN = P \tag{2.12}$$

where, by setting $C = \widetilde{A}^{-1} A^* \widetilde{A}^{-1}$,

$$B = \begin{bmatrix} B_1 & B_2 \\ B_2 & 2I_2 \end{bmatrix}, B_1 = \begin{bmatrix} \underline{e}_1^T C \underline{e}_1 & \underline{e}_1^T C \underline{e}_n \\ \underline{e}_n^T C \underline{e}_1 & \underline{e}_n^T C \underline{e}_n \end{bmatrix}, B_2 = \begin{bmatrix} \underline{e}_1^T \widetilde{A}^{-1} \underline{e}_1 & \underline{e}_1^T \widetilde{A}^{-1} \underline{e}_n \\ \underline{e}_n^T \widetilde{A}^{-1} \underline{e}_1 & \underline{e}_n^T \widetilde{A}^{-1} \underline{e}_n \end{bmatrix}. \tag{2.13}$$

I_2 is the second order identity matrix and

$$P = \left[\underline{e}_1^T C \underline{b},\ \underline{e}_n^T C \underline{b},\ \underline{e}_1^T \widetilde{A}^{-1} \underline{b},\ \underline{e}_n^T \widetilde{A}^{-1} \underline{b} \right]^T. \tag{2.14}$$

Once determined N, we shall determine $s_f(x)$ by solving the system (2.5).

Before proving the below theorem 2.5, we need to investigate some properties of matrices $\widetilde{A}$, $\widetilde{A}^{-1}$and C.

Proposition 2.1. *The matrix $\widetilde{A} = (a_{ij})_{i,j=1}^n$, is:*

(a) *symmetric, positive definite;*
(b) *persymmetric, i.e. $a_{ij} = a_{n-i+1,n-j+1}$, $i, j = 1, ..., n$;*
(c) *totally positive (T.P.), i.e. all the minors are ≥ 0;*
(d) *oscillatory, then all the eigenvalues of $\widetilde{A}$ are distinct, real and positive.*

Proof. It is straightforward to verify (a), (b), (c). The property (d) follows by considering that a non singular T.P. matrix having the entries $a_{ik} \neq 0$, $|i - k| \leq 1$ is oscillatory [4]. □

Proposition 2.2. *The infinitive norm of $\widetilde{A}^{-1}$ satisfies the following relation*

$$\frac{1}{6} \leq \left\| \widetilde{A}^{-1} \right\|_{\infty} \leq \frac{1}{2}. \tag{2.15}$$

Proof. (For the proof, see Appendix). □

Proposition 2.3. *The entries a_{1j}^{-1}, $j = 1, ..., n$ of $\widetilde{A}^{-1}$ have decreasing absolute values, the sign of $(-1)^{j-1}$, in particular, the following inequalities:*

$$\frac{1}{5} \leq a_{11}^{-1} = \underline{e}_1^T \widetilde{A}^{-1} \underline{e}_1 \leq \frac{1}{4}, \tag{2.16}$$

$$\left| a_{1n}^{-1} \right| = \left| \underline{e}_1^T \widetilde{A}^{-1} \underline{e}_n \right| \leq \begin{cases} \frac{1}{24} & \text{if} \quad n = 2, \\ \frac{1}{90} & \text{if} \quad n \geq 3 \end{cases} \tag{2.17}$$

hold.

Proof. (For the proof, see Appendix). □

Proposition 2.4. *Let $C = \widetilde{A}^{-1} A^* \widetilde{A}^{-1}$. For the entries $c_{11} = \underline{e}_1^T C \underline{e}_1$, $c_{1n} = \underline{e}_1^T C \underline{e}_n$ we have:*

$$0 < c_{11} < 1 \tag{2.18}$$

$$|c_{1n}| < c_{11}. \tag{2.19}$$

Proof. (For the proof, see Appendix). □

Now we prove the following:

Theorem 2.5. *The system (2.12) is determined and the solution is*

$$N = \left[\left[\underline{e}_1^T \widetilde{A}^{-1} \underline{b}, \ \underline{e}_n^T \widetilde{A}^{-1} \underline{b} \right] B_2^{-1}, \ 0, \ 0 \right]^T. \tag{2.20}$$

Proof. Considering that $\underline{e}_1^T \widetilde{A}^{-1} \underline{e}_n = \underline{e}_n^T \widetilde{A}^{-1} \underline{e}_1$ and for the properties of $\widetilde{A}^{-1}$ and the definition of the symmetric positive matrix A^*, one has $\underline{e}_1^T C \underline{e}_n = \underline{e}_n^T C \underline{e}_1$, (2.11) can be written in the form

$$\begin{bmatrix} B_1 & B_2 \\ B_2 & 2I_2 \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} \underline{t}_1 \\ \underline{t}_2 \end{bmatrix}, \tag{2.21}$$

where

$\underline{x} = [\widetilde{a}_0, -\widetilde{a}_n]$, $\underline{y} = [-M_0, -M_{n+1}]$, $\underline{t}_1 = \left[\underline{e}_1^T C \underline{b}, \ \underline{e}_n^T C \underline{b} \right]^T$, $\underline{t}_2 = \left[\ \underline{e}_1^T \widetilde{A}^{-1} \underline{b}, \ \underline{e}_n^T \widetilde{A}^{-1} \underline{b} \right]^T$.

Since $A^* = \widetilde{A} - \left(\underline{e}_1 \underline{e}_1^T + \underline{e}_n \underline{e}_n^T \right)$, and then, $\widetilde{A}^{-1} A^* \widetilde{A}^{-1} = \widetilde{A}^{-1} - \widetilde{A}^{-1} \left(\underline{e}_1 \underline{e}_1^T + \underline{e}_n \underline{e}_n^T \right) \widetilde{A}^{-1}$ there results

$$B_1 = B_2 \left(I_2 - B_2 \right) \tag{2.22}$$

$$\underline{t}_1 = \left(I_2 - B_2 \right) \underline{t}_2, \tag{2.23}$$

and the system (2.21) reduces to:

$$\begin{bmatrix} B_2 \left(I_2 - B_2 \right) & B_2 \\ B_2 & 2I_2 \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} \left(I_2 - B_2 \right) \underline{t}_2 \\ \underline{t}_2 \end{bmatrix}. \tag{2.24}$$

The system (2.21) has unique solution.

In fact [4], since

$$2B_2I_2 = 2I_2B_2, \det\begin{bmatrix} B_2\,(I_2 - B_2) & B_2 \\ B_2 & 2I_2 \end{bmatrix} = \det(B_2)\det(2I_2 - 3B_2)$$

and using the results of propositions 2.2 and 2.3, it is straightforward to verify that $\det(B_2) \neq 0$ and $\det(2I_2 - 3B_2) \neq 0$. From the second equation of (2.24) we obtain $\underline{x} = B_2^{-1}(\underline{t}_2 - 2I_2\underline{y})$, and, substituting in the first one, there results:

$$B_2\,(I_2 - B_2)\,B_2^{-1}(\underline{t}_2 - 2\underline{y}) + B_2\underline{y} = (I_2 - B_2)\,\underline{t}_2,$$

that implicates

$$(3B_2 - 2I_2)\underline{y} = \underline{0}. \tag{2.25}$$

Therefore we obtain the solution $\underline{y} = 0$, $\underline{x} = B_2^{-1}\underline{t}_2$ and theorem 2.5 is proved. □

Corollary 2.6. *For the spline $s_f(x)$ the following property*

$$M_1 = s_f^{(2m-1)}(x_1) = M_n = s_f^{(2m-1)}(x_n) = 0 \tag{2.26}$$

holds.

Proof. Taking into account that, from the relation $B_2\underline{x} + 2\underline{y} = \underline{t}_2$, we obtain

$$\begin{bmatrix} M_0 \\ M_{n+1} \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} \underline{e}_1^T\widetilde{A}^{-1}\underline{b} \\ \underline{e}_n^T\widetilde{A}^{-1}\underline{b} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \underline{e}_1^T\widetilde{A}^{-1}\underline{e}_1 & -\,\underline{e}_1^T\widetilde{A}^{-1}\underline{e}_n \\ \underline{e}_n^T\widetilde{A}^{-1}\underline{e}_1 & -\,\underline{e}_n^T\widetilde{A}^{-1}\underline{e}_n \end{bmatrix}\begin{bmatrix} \widetilde{a}_0 \\ \widetilde{a}_n \end{bmatrix}, \tag{2.27}$$

we get the thesis considering the first and the last equation in (2.10), that is:

$$\begin{bmatrix} M_1 \\ M_n \end{bmatrix} = \begin{bmatrix} \underline{e}_1^T\widetilde{A}^{-1}\underline{b} \\ \underline{e}_n^T\widetilde{A}^{-1}\underline{b} \end{bmatrix} - \begin{bmatrix} \underline{e}_1^T\widetilde{A}^{-1}\underline{e}_1 & -\,\underline{e}_1^T\widetilde{A}^{-1}\underline{e}_n \\ \underline{e}_n^T\widetilde{A}^{-1}\underline{e}_1 & -\,\underline{e}_n^T\widetilde{A}^{-1}\underline{e}_n \end{bmatrix}\begin{bmatrix} \widetilde{a}_0 \\ \widetilde{a}_n \end{bmatrix}. \tag{2.28}$$

□

From theorem 2.5 and corollary 2.6 we can deduce that, as aspected, the obtained spline $s_f(x)$ reduces to a polynomial of second degree in the subintervals I_0 and I_n.

Remark 2.1. *In* [3, theorem 2] *the construction of such spline is obtained by solving a linear system of $m + n + 1$ equations that can have an increasing condition number when n increases. Our method is based on the solution of two linear systems, having matrix $\widetilde{A}$ and B respectively. By proposition 2.2, for each n, for the condition number of $\widetilde{A}$, we have $K_\infty\left(\widetilde{A}\right) \leq 3$; now we prove the following*

Proposition 2.7. *For the condition number $K_\infty(B)$ the inequality*

$$K_\infty(B) \leq \frac{4477}{219} \simeq 20.44, \ \textit{if } n \geq 3, \tag{2.29}$$

holds.

Proof. $B = \begin{bmatrix} B_1 & B_2 \\ B_2 & 2I_2 \end{bmatrix}$, then

$$B^{-1} = \begin{bmatrix} (2I_2 - 3B_2)^{-1} & 0 \\ 0 & (2I_2 - 3B_2)^{-1} \end{bmatrix} \begin{bmatrix} 2B_2^{-1} & -I_2 \\ -I_2 & I_2 - B_2 \end{bmatrix}.$$

Let $n \geq 3$. Using the results obtained in the propositions 2.3 and 2.4, one obtains

$$\|B\|_\infty \leq 2 + a_{11}^{-1} + \left|a_{1n}^{-1}\right| \leq \frac{407}{180}, \tag{2.30}$$

$$\|B_2\|_\infty = a_{11}^{-1} + \left|a_{1n}^{-1}\right| \leq \frac{47}{180} \tag{2.31}$$

and

$$\left\|B_2^{-1}\right\|_\infty = \frac{1}{\left|a_{11}^{-1}\right| - \left|a_{1n}^{-1}\right|} \leq 5.$$

Besides:

$$\left\|B^{-1}\right\|_\infty \leq \left\|(2I_2 - 3B_2)^{-1}\right\|_\infty \left(2\left\|B_2^{-1}\right\|_\infty + 1\right). \tag{2.32}$$

Using the results in [7], since for all vector $\underline{x}$ such that $\|\underline{x}\|_\infty = 1$, one has

$$\|(2I_2 - 3B_2)\,\underline{x}\|_\infty \geq 2 - 3\frac{47}{180} = \frac{73}{60}, \tag{2.33}$$

and then

$$\left\|B^{-1}\right\|_\infty \leq \frac{660}{73}. \tag{2.34}$$

Therefore

$$K_\infty(B) \leq \frac{407}{180}\frac{660}{73} \simeq 20.44, \text{ if } n \geq 3$$

and we get the thesis. □

3. **Convergence results**

Consider $I = [a, b]$ and the set W_2^3. In [3] has been proved the following

Theorem 3.1. *Let $f \in W_2^3(I)$ and let s_f be the derivative-interpolating spline, then*

$$\left\|f^{(k)} - s_f^{(k)}\right\|_\infty \leq \begin{cases} C_1 h^{2-\frac{1}{2}} \left\|f^{(3)}\right\|_2 & \text{if } k = 0, \\ C_2 h^{2-k+\frac{1}{2}} \left\|f^{(3)}\right\|_2 & \text{if } k = 1, 2, \end{cases} \tag{3.1}$$

where $C_1 = \sqrt{2}\,(b - a)$ and $C_2 = \sqrt{2}$.

We shall prove a new convergence theorem under weaker hypothesis on function f. For all $g \in C^1(I)$, we denote by

$$\omega(g'; h; I) = \max_{x, x+\delta \in I,\, 0 \leq \delta \leq h} |g'(x + \delta) - g'(x)|$$

the modulus of continuity of g'.

Supposing $f \in C^1(I)$, from (2.9), we obtain

$$\|\underline{b}\|_\infty \leq \frac{12}{h^2}\omega(f'; h; I) \tag{3.2}$$

then, using (2.15), (2.20):

$$\|a_0, a_n\|_\infty \le \frac{5}{h}\omega\left(f'; h; I\right). \tag{3.3}$$

and consequently, from (2.10), we obtain:

$$\|M\|_\infty = \left\|\widetilde{M}\right\|_\infty \le \frac{21}{h^2}\omega\left(f'; h; I\right). \tag{3.4}$$

If we consider that $f'(x_i) = y'_i$ and by (2.4):
$a_i \quad = \quad \frac{y'_{i+1}-y'_i}{h} - \frac{h}{6}(M_{i+1} - M_i) \quad i = 1, ..., n-1,$ we can write

$$|a_i| \le \frac{8}{h}\omega\left(f'; h; I\right). \tag{3.5}$$

Therefore,

$$\|\underline{a}\|_\infty \le \frac{8}{h}\omega\left(f'; h; I\right) \tag{3.6}$$

where $\underline{a} = [a_0, a_1, ..., a_n]^T$.

Theorem 3.2. *Let $f \in C^1(I)$ and $s_f(x)$ the interpolating-derivative spline quoted in Section 2 for a given partition Δ_n. Then*

$$\omega\left(s'_f; h; I\right) \le C\omega\left(f'; h; I\right) \tag{3.7}$$

where C is a constant independent of h.

Proof. It suffices to show that for $\forall u, v \in I$, $u < v$:

$$\left|s'_f(v) - s'_f(u)\right| \le \bar{C}\omega\left(f'; v-u; I\right).$$

Firstly consider $u, v \in [x_i, x_{i+1}]$, $i = 0, ..., n$; using the mean value theorem, we can derive:

$$\left|s'_f(v) - s'_f(u)\right| = \left|s''_f(\xi)\right| |v-u| \qquad \xi \in (u, v),$$

where $|v-u| \le h$.

Since for any $\xi \in (u, v)$, from (3.3), (3.4), (3.5), there results
$\left|s''_f(\xi)\right| \le \frac{29}{h}\omega\left(f'; h; I\right)$ if $u, v \in [x_i, x_{i+1}]\, t$, $i = 1, 2, ..., n-1$ and
$\left|s''_f(\xi)\right| \le \frac{5}{h}\omega\left(f'; h; I\right)$ if $u, v \in [x_0, x_1]$ or $u, v \in [x_n, x_{n+1}]$,
recalling that [9], $\frac{|v-u|}{h}\omega\left(f'; h; I\right) \le 2\omega\left(f'; |v-u|; I\right)$, we get

$$\left|s'_f(v) - s'_f(u)\right| \le C_1\omega\left(f'; |v-u|; I\right), \qquad C_1 = 58. \tag{3.8}$$

If $u \in [x_i, x_{i+1}]$, $v \in [x_j, x_{j+1}]$, $i+1 \le j$, then using (3.8) and the smoothness of modulus of continuity and, since being x_{i+1} and x_j internal nodes, $s'_f(x_{i+1}) = f'(x_{i+1})$, $s'_f(x_j) = f'(x_j)$:

$$\begin{aligned}
\left|s'_f(v) - s'_f(u)\right| &\le \left|s'_f(v) - s'_f(x_j)\right| + \left|s'_f(x_j) - s'_f(x_{i+1})\right| + \left|s'_f(x_{i+1}) - s'_f(u)\right| \\
&= \left|s'_f(v) - s'_f(x_j)\right| + \left|f'(x_j) - f'(x_{i+1})\right| + \left|s'_f(x_{i+1}) - s'_f(u)\right| \\
&\le (2C_1 + 1)\,\omega\left(f'; |v-u|; I\right).
\end{aligned}$$

This proves the theorem with $C = (2C_1 + 1)$. □

Supposing $f \in C^1(I)$, we define $r(x) = f(x) - s_f(x)$ and $r'(x) = f'(x) - s'_f(x)$, where $s_f(x)$ is the interpolating-derivative spline quoted in Section 2.

For $x \in I_i$, $i = 0, ..., n$ we can write

$$r'(x) = \begin{cases} r'(x_1) + (x - x_1)[x_1 x] r' & \text{if } x \in I_0, \\ r'(x_i) + (x - x_i)[x_i x] r' & \text{if } x \in I_i,\ i = 1, \dots, n. \end{cases} \tag{3.9}$$

where $[x_i x] r'$, $i = 1, \dots, n$, denotes the first divise difference of r'. Therefore, from (2.2) and theorem 3.2:

$$|r'(x)|_{I_i} \leq (C+1)\,\omega(f'; h; I),\ i = 0, 1, \dots, n. \tag{3.10}$$

We are ready to prove the following convergence result.

Theorem 3.3. *Let $f \in C^1(I)$ and $s_f(x)$ the interpolating-derivative spline. There results*

$$\|f - s_f\|_\infty \leq (b-a)(C+1)\,\omega(f'; h; I). \tag{3.11}$$

Proof. We can write, for $x \in I_i$, $i = 0, 1, ..., n$

$$|r(x)| = \left| \int_{x_0}^{x} r'(t)\,dt \right| \leq \max_{x \in I} |r'(x)|\,|x - x_{0\leq}| \leq (b-a)(C+1)\,\omega(f'; h; I) \tag{3.12}$$

and (3.11) is proved. □

We remark that the above theorems hold even when the partition Δ_n is quasi-uniform, i.e. such that: $\max\limits_{0 \leq i \leq n} \frac{h}{h_i}$ is bounded for $n \to \infty$ where $h_i = x_{i+1} - x_i$ and h is the norm of the partition. [E.Santi, M.G.Cimoroni: *Some new convergence results and applications of a class of interpolating-derivative splines.* In preparation].

We add now a property of the splines considered in this paper. The derivative-interpolating spline $s_f(x)$ defined in (2.3), considering a uniform partition Δ_n, reproduces any $f \in I\!P_2$. In fact, for $f = 1, x, x^2$ it is straightforward to verify that, $\underline{b}^*(a_0, a_n) = M = \underline{0}$. Therefore the coefficients of $s_f(x)|_{I_i}$ are:

$$\begin{cases} a_i = 0,\ b_i = 0, & c_i = 1, \quad \text{if} \quad f(x) = 1, \\ a_i = 0,\ b_i = 1, & c_i = x_i, \quad \text{if} \quad f(x) = x, \\ a_i = 2,\ b_i = 2x_i, & c_i = x_i^2, \quad \text{if} \quad f(x) = x^2, \end{cases}$$

and thus, for $x \in I_i$, $i = 0, \dots, n$:

$$\begin{cases} s_f(x) = 1, & \text{if} \quad f(x) = 1, \\ s_f(x) = (x - x_i) + x_i = x, & \text{if} \quad f(x) = x, \\ s_f(x) = \frac{2(x-x_i)^2}{2} + 2x_i(x - x_i) + x_i^2 = x^2, & \text{if} \quad f(x) = x^2. \end{cases}$$

4. **Numerical results**

We present now, some numerical results obtained by approximating some test functions by the spline considered in this paper. We denote $|r_n(x)| = |f(x) - s_f(x)|$ the error at x obtained by using a uniform partition of the interval $[-1, 1]$ in $n+1$ subintervals. In table 1 we report the results relative to a test function $f(x)$ having only $f'(x) \in C[-1, 1]$ and considering different uniform partition with

$n = 4, 19, 39, 99$. In table 2 we report, for confirming the polynomial reproducibility, the results relative to a function $f(x) \in I\!P_2$ considering a uniform partition with $n = 4$. Table 3 contains the results relative to a more regular function $f(x) \in C^\infty[-1, 1]$ with different uniform partition, taking $n = 4, 19, 39, 79$.

Table 1

$f(x) = sign(x)x^2/2 + e^x$				
x	$\|\mathbf{r}_4(x)\|$	$\|\mathbf{r}_{19}(x)\|$	$\|\mathbf{r}_{39}(x)\|$	$\|\mathbf{r}_{99}(x)\|$
-1	0.0	0.0	0.0	0.0
-0.6	5.1 (-3)	1.2 (-4)	1.4 (-5)	8.6 (-7)
-0.2	1.6 (-3)	1.8 (-4)	1.5 (-5)	8.6 (-7)
0.2	2.1 (-2)	1.7 (-3)	4.3 (-4)	6.8 (-5)
0.6	1.7 (-2)	1.8 (-3)	4.3 (-4)	6.8 (-5)
1	5.6 (-3)	2.5 (-3)	5.2 (-4)	7.4 (-5)

Table 2

$f(x) = x^2 + 2x - 5$	
x	$\|\mathbf{r}_4(x)\|$
-1	0.0
-0.6	0.0
-0.2	8.9 (-16)
0.2	1.8 (-15)
0.6	8.9 (-16)
1	8.9 (-16)

Table 3

$f(x) = 1/(x^2 + 25)$				
x	$\|\mathbf{r}_4(x)\|$	$\|\mathbf{r}_{19}(x)\|$	$\|\mathbf{r}_{39}(x)\|$	$\|\mathbf{r}_{79}(x)\|$
-1	0.0	0.0	0.0	0.0
-0.6	1.8 (-5)	3.4 (-7)	4.4 (-8)	5.6 (-9)
-0.2	1.7 (-5)	3.4 (-7)	4.4 (-8)	5.6 (-9)
0.2	1.7 (-5)	3.4 (-7)	4.4 (-8)	5.6 (-9)
0.6	1.8 (-5)	3.4 (-7)	4.4 (-8)	5.6 (-9)
1	0.0	0.0	3.5 (-17)	0.0

5. Appendix

Proposition 2.2. *The infinitive norm of* $\widetilde{A}^{-1}$ *satisfies the following relation*

$$\frac{1}{6} \leq \left\| \widetilde{A}^{-1} \right\|_\infty \leq \frac{1}{2}. \tag{5.1}$$

Proof. It is straightforward to verify that $\left\| \widetilde{A} \right\|_\infty = 6$ and then, being $\left\| \widetilde{A}^{-1} \right\|_\infty \left\| \widetilde{A} \right\|_\infty \geq 1$, we obtain the left inequality in (5.1). For proving that $\left\| \widetilde{A}^{-1} \right\|_\infty \leq \frac{1}{2}$, we write $\widetilde{A} = 4I + H$, where

$$H = \begin{bmatrix} 1 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 1 \end{bmatrix}.$$

Thus, for all $\underline{x} : \|\underline{x}\|_\infty = 1$, there results

$$\left\|\widetilde{A}\underline{x}\right\|_\infty = \|(4I+H)\underline{x}\|_\infty \geq \|4\underline{x}\|_\infty - \|H\underline{x}\|_\infty \geq 2,$$

and then [7], (5.1) is proved. □

Proposition 2.3. *The entries* a_{1j}^{-1}, $j = 1,...,n$ *of* $\widetilde{A}^{-1}$*have decreasing absolute values, the sign of* $(-1)^{j-1}$, *in particular, the following inequalities:*

$$\frac{1}{5} \leq a_{11}^{-1} = \underline{e}_1^T \widetilde{A}^{-1} \underline{e}_1 \leq \frac{1}{4}, \tag{5.2}$$

$$\left|a_{1n}^{-1}\right| = \left|\underline{e}_1^T \widetilde{A}^{-1} \underline{e}_n\right| \leq \begin{cases} \frac{1}{24} & \text{if} \quad n = 2, \\ \\ \frac{1}{90} & \text{if} \quad n \geq 3 \end{cases} \tag{5.3}$$

hold.

Proof. Using the results in [1], for the evaluation of the inverse matrix of a tridiagonal symmetric matrix, we can write

$$\widetilde{A}^{-1} = L + \underline{uv}^T \tag{5.4}$$

and then,

$$a_{ij}^{-1} = l_{ij} + u_i v_j \tag{5.5}$$

where $l_{ij} = 0$ for $i \leq j$.

In our case, there results

$$u_1 = 1, \quad u_2 = -5, \quad u_i = -4u_{i-1} - u_{i-2}, \quad i = 3,...,n \tag{5.6}$$

$$v_i = \alpha^{-1} u_{n-i+1}, \quad i = 1,...,n \tag{5.7}$$

with $\alpha = 5u_n + u_{n-1}$.The matrix $\widetilde{A}^{-1}$ is symmetric and $a_{1i}^{-1} = a_{i1}^{-1} = \alpha^{-1}u_{n-i+1}$, $a_{in}^{-1} = a_{ni}^{-1} = \alpha^{-1}u_i$, $i = 1, 2, ..., n$.

Therefore, using proposition 2.1 , we deduce that $a_{11}^{-1} = a_{nn}^{-1} = u_n/\alpha$, $a_{1j}^{-1} = a_{n,n-j+1}^{-1}$, $j = 1, 2, ..., n$ are decreasing in absolute value and have the sign of $(-1)^{j-1}$, and, in particular, $a_{1n}^{-1} = 1/\alpha$.

For proving (5.2) consider that $a_{11}^{-1} = u_1 v_1 = u_n/\alpha = \frac{1}{5}\frac{5u_n+u_{n-1}-u_{n-1}}{5u_n+u_{n-1}} = \frac{1}{5}\left(1 - \frac{u_{n-1}}{5u_n+u_{n-1}}\right)$, and then, using (5.6), $0 < -\frac{u_{n-1}}{5u_n+u_{n-1}} = \frac{1}{4}\frac{u_n+u_{n-2}}{5u_n+u_{n-1}} < \frac{1}{4}$, (5.3) follows.

We get the inequality (5.3) considering that $\left|a_{1n}^{-1}\right| = \left|\frac{1}{5u_n+u_{n-1}}\right|$ and then, from (5.6), $\left|a_{1n}^{-1}\right| = \frac{1}{24}$ if $n = 2$, $\left|a_{1n}^{-1}\right| = \frac{1}{90}$ for $n = 3$, and $\left|a_{1n}^{-1}\right|$ decreases when n increases. Therefore the proposition is completely proved. □

Proposition 2.4. *Let* $C = \widetilde{A}^{-1}A^*\widetilde{A}^{-1}$. *For the entries* $c_{11} = \underline{e}_1^T C \underline{e}_1$, $c_{1n} = \underline{e}_1^T C \underline{e}_n$ *we have:*

$$0 < c_{11} < 1 \tag{5.8}$$

$$|c_{1n}| < c_{11}. \tag{5.9}$$

Proof. For $n \geq 3$, by (2.22), using (5.2) and (5.3), it is straightforward to verify that $0 < c_{11} = a_{11}^{-1} - \left[\left(a_{11}^{-1}\right)^2 + \left(a_{1n}^{-1}\right)^2\right] < 1$ and $|c_{1n}| = \left|a_{1n}^{-1}\left(1 - 2a_{11}^{-1}\right)\right| < c_{11}$. □

References

[1] Bevilacqua, R., *Structural and computational properties of band matrices.* Complexity of structured computational problems. Applied Mathematics monographs. Comitato Nazionale per le Scienze Matematiche, C.N.R. (1991), 131-188.

[2] Bini, D., Capovani, M., *A class of cubic splines obtained through minimum conditions.* Mathematics of Computation **46**, n.173 (1986), 191-202.

[3] Blaga, P., Micula, G. *Natural spline functions of even degree.* Studia Univ. Babes-Bolyai Cluj-Napoca Mathematica, **38,** (1993) Nr. 2, 31-40.

[4] Gantmacher, F.R., *The theory of matrices*, Chelsea Publ. Company N.Y. 1974.

[5] Ghizzetti, A., *Interpolazione con splines verificanti un'opportuna condizione.* Calcolo **20** (1983), 53-65.

[6] Gori, L., *Splines and Cauchy principal value integrals.* Proc. Intern. Workshop on Advanced Math. Tools in Metrology (Ed. Ciarlini, Cox, Monaco, Pavese) (1994), 75-82.

[7] Kershaw, D., *A note on the convergence of interpolatory cubic splines.* Siam J.Numer. Anal. **8** (1971), 67-74.

[8] Rabinowitz, P., *Numerical integration based on approximating splines.* J. Comp. Appl. Math. **33** (1990), 73-83.

[9] Rabinowitz, P., *Application on approximating splines for the solution of Cauchy singular integral equations.* Appl. Numer. Math. **15** (1994), 285-297.

Dept. of Applied Mathematics, Babeş-Bolyai Univ. Cluj-Napoca, Romania
E-mail address: ghmicula@math.ubbcluj.ro

Dipartimento di Energetica, Universita degli Studi de l'Aquila, Italy
E-mail address: esanti@dsiaqi.ing.univaq.it

Dipartimento di Energetica, Universita degli Studi de l'Aquila, Italy

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

A COLLOCATION METHOD FOR SOLVING THE EXTERIOR NEUMANN PROBLEM

SANDA MICULA

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. In this paper we study the numerical solution of a boundary integral equation reformulation of the exterior Neumann problem. We give a brief outline of the problem and its solvability. Then, we propose a collocation method based on interpolation and give an error analysis. Numerical examples for the piecewise constant collocation method (centroid rule) conclude the paper.

1. The Exterior Neumann Problem

Let D denote a bounded open simply-connected region in $\mathbb{R}^3$, and let S denote its boundary. Let $\overline{D} = D \cup S$ and denote by $D_e = \mathbb{R}^3 - \overline{D}$ the region complementary to D. Let $\overline{D}_e = D_e \cup S$. At a point $P \in S$, let $\mathbf{n}_P$ denote the unit normal directed into D, provided that such a normal exists. Also assume that S is a piecewise smooth surface that can be decomposed into a finite union of smooth surfaces intersecting each other along common edges at most. In addition, assume that S has a triangulation $\mathcal{T}_n = \{\Delta_{n,k} \mid 1 \le k \le n\}$ with mesh size h (such a triangulation can be obtained as the image of a composition of bijections m_k from the unit simplex σ onto a planar triangle Δ_k and bijections F_j from a right triangle onto each smooth piece S_j of S; for details, see Micula [7, Chapter 3].

The Exterior Neumann Problem

Find $u \in C^1(\overline{D}_e) \cap C^2(D_e)$ that satisfies

$$\begin{aligned} \Delta u(P) &= 0,\ P \in D_e \\ \frac{\partial u(P)}{\partial \mathbf{n}_P} &= f(P),\ P \in S \end{aligned} \qquad (1)$$

$$u(P) = O(P^{-1}),\ \frac{\partial u(P)}{\partial r} = O(|P|^{-2}) \quad \text{, as} \quad r = |P| \to \infty \text{ uniformly in } \frac{P}{|P|}$$

with $f \in C(S)$ a given boundary function.

Received by the editors: 29.09.2003.
2000 *Mathematics Subject Classification.* 65N35, 65N30, 65P05, 35Q60.

The boundary value problem (1) has been studied extensively (see Mikhlin [8, Ch. 18], Günter [5, Ch. 3], Colton [4, Section 5.3]). Here we only give a very brief outlook at results on the solvability of the problem (1).

The Divergence Theorem (see Atkinson [2, Theorem 7.1.2]) can be used to obtain a representation formula for functions that are harmonic inside the region D_e. Let $u \in C^1(\overline{D}_e) \cap C^2(D_e)$ and assume that $\Delta u(P) = 0$ at all $P \in D_e$. Then

$$\int_S \frac{\partial u(Q)}{\partial n_Q} \frac{dS(Q)}{|P-Q|} - \int_S u(Q) \cdot \frac{\partial}{\partial n_Q} \left[\frac{1}{|P-Q|} \right] dS_Q = \begin{cases} [4\pi - \Omega(P)]u(P) & , P \in S \\ 4\pi u(P) & , P \in D_e \end{cases} \quad (2)$$

(see Atkinson [1].) In formula (2), $\Omega(P)$ denotes the *interior solid angle* at $P \in S$, defined in Atkinson [2, p. 430]. If S is smooth, then $\Omega(P) = 2\pi$. For a cube, the corners have interior solid angle of $\frac{1}{2}\pi$, and the edges have interior solid angles of π.

To study the solvability of (1), consider representing its solution as a *single layer potential*

$$u(A) = \int_S \frac{\rho(Q)}{|A-Q|}\, dS_Q, \quad A \in D_e \quad (3)$$

The function ρ in (3) is called a *single layer density* function. The function $u(A)$ in (3) is harmonic for all $A \notin S$. For well-behaved density functions and for $A \notin S$, the integrand in (3) is nonsingular. Even though for the case $A = P \in S$, the integrand in (3) becomes singular, it is relatively straightforward to show that the integral exists and moreover, if ρ is bounded on S, then

$$\sup_{A \in \mathbb{R}^3} |u(A)| \le c\|\rho\|_\infty \quad (4)$$

For a complete description of the properties of the single layer potential, see Günter [5, Chapter 2].

Now for the function u of (3), impose the boundary condition from (1) to get

$$\lim_{\substack{A \to P \\ A \in D_e}} \mathbf{n}_P \cdot \nabla \left[\int_S \frac{\rho(Q)}{|A-Q|}\, dS_Q \right] = f(P), \quad P \in S \quad (5)$$

for all $P \in S$ at which the normal $\mathbf{n}_P$ exists (which implies $\Omega(P) = 2\pi$). Using a limiting argument, we obtain the second kind integral equation

$$2\pi\rho(P) + \int_S \rho(Q) \cdot \frac{\partial}{\partial \mathbf{n}_P} \left[\frac{1}{|P-Q|} \right] dS_Q = f(P), \quad P \in S^* \quad (6)$$

The set S^* is to contain all points $P \in S$ at which a normal is defined. If S is a smooth surface, then $S^* = S$; otherwise, $S - S^*$ is a set of measure 0. The kernel function in (6) is given by

$$\frac{\partial}{\partial \mathbf{n}_P}\left[\frac{1}{|P-Q|}\right] = \frac{\mathbf{n}_P \cdot (P-Q)}{|P-Q|^3} = \frac{\cos\theta_P}{|P-Q|^2} \tag{7}$$

where θ_P denotes the angle between $\mathbf{n}_P$ and $(P-Q)$. Equation (6) can now be written as

$$\rho(P) + \frac{1}{2\pi}\int\limits_S \rho(P)\cdot\frac{\cos\theta_P}{|P-Q|^2}\,dS_Q = \hat{f}(P), \quad P \in S \tag{8}$$

where $\hat{f}(P) = \dfrac{1}{2\pi}f(P)$. For simplicity, we will write $f(P)$ instead of $\hat{f}(P)$.

Write the equation (8) in operator form:

$$(\mathcal{I} - K)\rho = f \tag{9}$$

The properties of the integral operator $\mathcal{K}$ and, implicitly, the solvability of equation (1) have been studied intensively in the literature, especially for the case that S is a smooth surface. For S sufficiently smooth, $\mathcal{K}$ is a compact operator from $C(S)$ to $C(S)$ and from $L^2(S)$ to $L^2(S)$. These results are contained in many textbooks, for example see Kress [6, Chapter 6], or Mikhlin [8, Chapters 12 and 16]. We will just state the following solvability result.

Theorem 1.1. *Let S be a C^2 surface. Then the equation* (9) *has a unique solution $\rho \in X$ for each given function $f \in X$, with $X = C(S)$ or $X = L^2(S)$.*

This theorem then leads to a solvability result for the Exterior Neumann Problem (1)

Theorem 1.2. *Let S be a smooth surface with $\overline{D_e}$ a region to which the Divergence Theorem can be applied. Assume the function $f \in C(S)$. Then, the Neumann problem* (1) *has a unique solution $u \in C^\infty(D_e)$.*

For the case when S is only piecewise smooth, the properties of $\mathcal{K}$ and the solvability of (8) are not yet fully understood. We will assume that Theorem 1.1 is true for the piecewise smooth surfaces that we will consider in our work.

2. A Collocation Method

We want to study the numerical solution of (8) using a an integral equation reformulation of (1) have been used before (see Atkinson and Chien [3] or Atkinson [2, Section 9.2]), but with the collocation nodes on the boundary of each triangular element. There are problems with defining the normal at the collocation points which are common to more than one triangular face, especially if the surface itself is approximated. This in turn means it is difficult to evaluate the kernel function in equation (8). For these reasons it makes sense to try collocation methods that use only interior collocation node points. We will use interpolation of order r (the collocation nodes will be the same as the interpolation nodes), of the form

$$q_{i,j} = \left(\frac{i+(r-3i)\alpha}{r}, \frac{j+(r-3j)\alpha}{r}\right), \quad i,j \geq 0, \quad i+j \leq r \tag{10}$$

for some $0 < \alpha < 1/3$ (these are points interior to the unit simplex, but they get mapped into points interior to each triangle in $\mathcal{T}_n$). For corresponding *Lagrange* functions (see Micula [7, pg. 7-11]), for $g \in C(S)$ define an operator $\mathcal{P}_n$ by

$$\mathcal{P}_n g(P) = \sum_{j=1}^{f_r} g\left(m_k(q_j)\right) l_j(s,t), \quad (s,t) \in \sigma, \quad P = m_k(s,t) \in \Delta_k \tag{11}$$

This interpolates $g(P)$ over each triangular element $\Delta_k \in S$, with the interpolating function polynomial in the parameterization variables s and t. Since $\mathcal{P}_n g$ is not continuous in general, we need to enlarge $C(S)$ to include the piecewise polynomial approximations $\mathcal{P}_n g$. To do this, we consider the equation (9) within the framework of the function space $L^\infty(S)$ with the uniform norm $\|\cdot\|_\infty$. Then, $\mathcal{P}_n : L^\infty(S) \longrightarrow L^\infty(S)$ is a bounded projection operator.

Define a collocation method with (10). Denote $v_{k,j} = m_k(q_i)$. Substitute

$$\begin{aligned} \rho_n(P) &= \sum_{j=1}^{f_r} \rho_n\left(v_{k,j}\right) l_j(s,t) \\ P &= m_k(s,t) \in \Delta_k, \quad k = 1, ..., n \end{aligned} \tag{12}$$

into (8). To determine the values $\{\rho_n(v_{k,j})\}$, force the equation resulting from the substitution to be true at the collocation nodes $\{v_1, ..., v_{nf_r}\}$. This leads to the linear system

$$\begin{aligned} \rho_n(v_i) \quad &- \quad \frac{1}{2\pi} \sum_{k=1}^{n} \sum_{j=1}^{f_r} \rho_n(v_{k,j}) \int_\sigma \frac{\cos\theta_{v_i}}{|v_i - m_k(s,t)|^2} \\ &\cdot \quad |(D_s m_k \times D_t m_k)(s,t)|\, d\sigma = f(v_i), \quad i = 1, ..., nf_r \end{aligned} \tag{13}$$

which we write abstractly as

$$(\mathcal{I} - P_n \mathcal{K})\rho_n = \mathcal{P}_n f \tag{14}$$

which will be compared to (9). We have the following result.

Theorem 2.1. *Let S be a C^2 surface as described earlier, with $F_j \in C^{r+2}$. Then for all sufficiently large n, say $n \geq n_0$, the operators $\mathcal{I} - P_n\mathcal{K}$ are invertible on $L^\infty(S)$ and have uniformly bounded inverses. For the solution ρ of* (9) *and the solution ρ_n of* (13)

$$\|\rho - \rho_n\|_\infty \leq \left\|(\mathcal{I} - P_n\mathcal{K})^{-1}\right\| \cdot \|\rho - \mathcal{P}_n\rho\|_\infty, \quad n \geq n_0 \tag{15}$$

Furthermore, if $f \in C^{r+1}(S)$, then

$$\|\rho - \rho_n\|_\infty = O(h^{r+1}), \quad n \geq n_0 \tag{16}$$

For the proof, see, for example, Atkinson [1]).

So interpolation of order r, leads to an error of order $O(h^{r+1})$. But superconvergent methods can be developed. Next, we want to explore in more detail the collocation method based on piecewise constant interpolation (the centroid method) and show that it is superconvergent at the collocation points. D efine the operator $\mathcal{P}_n$ by

$$\mathcal{P}_n g(P) = g(P_k),\ P \in \Delta_k,\ k = 1, ..., n \tag{17}$$

for $g \in C(S)$. Then, $\mathcal{P}_n$ is a bounded operator on $C(S)$ with $\|\mathcal{P}_n\| = 1$. Define a collocation method with (17). Substitute

$$\rho_n(P) = \rho_n(P_k),\ P = m_k(s,t) \in \Delta_k,\ k = 1, ..., n \tag{18}$$

into (8). To determine the values $\{\rho_n(P_k)\}$, force the equation resulting from the substitution to be true at the collocation nodes $\{P_k \mid k = 1, ..., n\}$. This leads to the linear system

$$\begin{aligned}\rho_n(P_i) \quad &+ \quad \frac{1}{2\pi}\sum_{k=1}^{n} \rho_n(P_k) \cdot \int_{\sigma} \frac{\cos\theta_{P_k}}{|P_k - m_k(s,t)|^2} \\ &\cdot \quad |(D_s m_k \times D_t m_k)(s,t)|\ d\sigma = f(P_k),\ i = 1, ..., n \end{aligned} \tag{19}$$

which can be rewritten abstractly as

$$(\mathcal{I} + P_n \mathcal{K})\, \rho_n = \mathcal{P}_n f \tag{20}$$

which will be compared to (9).

By Theorem 2.1., for the true solution ρ of (9) and the solution ρ_n of the collocation equation (20), we have

$$\|\rho - \rho_n\|_\infty = O(h), \quad n \geq n_0 \tag{21}$$

For $g \in C(\sigma)$, consider the interpolation formula (17), which has degree of precision 0. Integrating it over σ, we obtain

$$\int_\sigma g(s,t)\ d\sigma \approx \int_\sigma \mathcal{L}_\tau g(s,t)\ d\sigma = \frac{1}{2} g\left(\frac{1}{3}, \frac{1}{3}\right) \tag{22}$$

which has degree of precision 1.

For $\tau \subset \mathbb{R}^2$, a planar triangle and for a function $g \in C(\tau)$, the function

$$\mathcal{L}_\tau g(x,y) = g\left(m_\tau\left(\frac{1}{3}, \frac{1}{3}\right)\right) = g(P_\tau) \tag{23}$$

the constant polynomial interpolating g at the node $m_\tau\left(\frac{1}{3}, \frac{1}{3}\right) = P_\tau$ (the centroid of τ). We have the following.

Lemma 2.2. *Let τ be a planar right triangle and assume the two sides which form the right angle have length h. Let $g \in C^2(\tau)$. Let $\Phi \in L^1(\tau)$ be differentiable with the first derivatives $D_x\Phi,\ D_y\Phi \in L^1(\tau)$. Then*

$$\left|\int_\tau \Phi(x,y)\,(\mathcal{I} - L_\tau)\, g(x,y)\ d\tau\right| \leq ch^2 \left[\int_\tau (|\Phi| + |D\Phi|)\ d\tau\right] \cdot \max_\tau \left\{|Dg|, |D^2|g\right\} \tag{24}$$

For the proof, see Micula [7, pg 74-75].

This result can be extended to general triangles, provided

$$\sup_n \left[\max_{\Delta_{n,k} \in \mathcal{T}_n} r(\Delta_{n,k}) \right] < \infty \tag{25}$$

where

$$r(\tau) = \frac{h(\tau)}{h^*(\tau)} \tag{26}$$

with $h(\tau)$ and $h^*(\tau)$ denoting the diameter of τ and the radius of the circle inscribed in τ, respectively.

Corollary 2.3. *Let τ be a planar triangle of diameter h, let $g \in C^2(\tau)$, and let $\Phi \in L^1(\tau)$ with both first derivatives in $L^1(\tau)$. Then*

$$\begin{aligned} \left| \int_\tau \Phi(x,y)(\mathcal{I} - L_\tau) g(x,y) \right| &\leq c\,(r(\tau))\, h^2 \left[\int_\tau (|\Phi| + |D\Phi|)\ d\tau \right] \\ &\cdot \max_\tau \left\{ \|Dg\|_\infty, \|D^2 g\|_\infty \right\} \end{aligned} \tag{27}$$

where $c\,(r(\tau))$ is some multiple of $r(\tau)$ of (26).

Since formula (22) has degree of precision 1 (odd) over σ, extending it to a square would not improve the degree of precision, which means the same error bound as in Lemma 2.2 is true for a parallelogram formed by two symmetric triangles.

We want to apply the above results to the individual subintegrals in

$$\begin{aligned} \mathcal{K}g(P_i) &= \frac{1}{2\pi} \sum_{k=1}^{n} \int_\sigma \frac{\cos\theta_{P_k}}{|P_k - m_k(s,t)|^2} \rho\,(m_k(s,t)) \\ &\qquad \cdot |(D_s m_k \times D_t m_k)\,(s,t)|\ d\sigma \end{aligned} \tag{28}$$

with the role of g played by $\rho\,(m_k(s,t))\,|(D_s m_k \times D_t m_k)\,(s,t)|$, and the role of Φ played by $\dfrac{\cos\theta_{P_k}}{|P_k - m_k(s,t)|^2}$. For the derivatives of this last function, we have

Theorem 2.4. *Let i be an integer and S be a smooth C^{i+1} surface. Then*

$$\left| D_Q^i \left(\frac{\cos\theta_P}{|P-Q|^2} \right) \right| \leq \frac{c}{|P-Q|^{i+1}}, \quad P \neq Q \tag{29}$$

with c a generic constant independent of P and Q.

For details of the proof, see Micula [7, pg.76].

For the error at the collocation node points, we have the following.

Theorem 2.5. *Assume the hypotheses of Theorem 2.1, with each $F_j \in C^2$. Assume $\rho \in C^2$. Assume the triangulation $\mathcal{T}_n$ of S satisfies* (25) *and is symmetric. For those integrals in* (28) *for which $P_i \in \Delta_k$, assume that all such integrals are evaluated with an error of $O(h^2)$. Then*

$$\max_{1 \leq i \leq n} |\rho(P_i) - \widehat{\rho}_n(P_i)| \leq c h^2 \log h \tag{30}$$

Proof. We will bound

$$\max_{1\le i\le n} |\mathcal{K}(I-P_n)u(v_i))|$$

For a given node point v_i, denote Δ^* the triangle containing it and denote:

$$\mathcal{T}_n^* = \mathcal{T}_n - \{\Delta^*\}$$

By our assumption, the error in evaluating the integral of (28) over Δ^* will be $O(h^2)$.

Partition $\mathcal{T}_n^*$ into parallelograms to the maximum extent possible. Denote by $\mathcal{T}_n^{(1)}$ the set of all triangles making up such parallelograms and let $\mathcal{T}_n^{(2)}$ contain the remaining triangles. Then

$$\mathcal{T}_n^* = \mathcal{T}_n^{(1)} \cup \mathcal{T}_n^{(2)}.$$

It is easy to show that the number of triangles in $\mathcal{T}_n^{(1)}$ is $O(n) = O(h^{-2})$, and the number of triangles in $\mathcal{T}_n^{(2)}$ is $O(\sqrt{n}) = O(h^{-1})$.

It can be shown that all but a finite number of the triangles in $\mathcal{T}_n^{(2)}$, bounded independent of n, will be at a minimum distance from v_i. That means that the triangles in $\mathcal{T}_n^{(2)}$ are "far enough" from v_i, so that the function $G(v_i, Q)$ is uniformly bounded for Q being in a triangle in $\mathcal{T}_n^{(2)}$ (where we denote by $G(P,Q) = \dfrac{\cos\theta_P}{|P-Q|^2}$).

First, consider the contribution to the error coming from the triangles in $\mathcal{T}_n^{(2)}$. By Lemma 2.2. the error over each such triangle is $O\left(h^2\|D^2 g\|_\infty\right)$, since the area of each triangle is $O(h^2)$ and using our earlier observation. Having $O(h^{-1})$ such triangles in $\mathcal{T}_n^{(2)}$, the total error coming from triangles in $\mathcal{T}_n^{(2)}$ is $O\left(h^3\|D^2 g\|_\infty\right)$.

Next, consider the contribution to the error coming from triangles in $\mathcal{T}_n^{(1)}$. By Lemma 2.2., the error will be of size $O(h^2)$ multiplied times the integral over each such parallelogram of the maximum of the first derivatives of $G(v_i, Q)$ with respect to Q. Combining these we will have a bound

$$ch^2 \int\limits_{S-\Delta^*} (|G| + |DG|)\, dS_Q \tag{31}$$

By Theorem 2.4., the quantity in (31) is bounded by

$$ch^2 \int\limits_{S-\Delta^*} \left(\frac{1}{|P-Q|} + \frac{1}{|P-Q|^2}\right) dS_Q \tag{32}$$

Using a local representation of the surface and then using polar coordinates, the expression in (32) is of order

$$ch^2\,(h + \log h)$$

Thus, the error arising from the triangles in $\mathcal{T}_n^{(1)}$ is $O(h^2 \log h)$. Combining the error arising from the integrals over Δ^*, $\mathcal{T}_n^{(1)}$, and $\mathcal{T}_n^{(2)}$, we have (30). □

3. Numerical Examples

As a smooth surface consider the ellipsoid

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 \tag{33}$$

with $(a, b, c) = (1, 1, 1)$ (the surface $E1$), and $(a, b, c) = (2, 3, 5)$ (the surface $E2$). We solve the equation (1) with the function $f(P)$ so chosen that the true solution is

$$u = \frac{1}{\sqrt{x^2 + y^2 + z^2}} \tag{34}$$

In Tables 1 and 2 we give

$$|u(P) - u_n(P)| \tag{35}$$

where $P = P_{ij} = \tau_i \left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}\right) \in D_e(E_j)$ (the exterior of E_j), where $\tau_1 = 1.1$, $\tau_2 = 2$, and $\tau_3 = 10$ (points situated further and further away from the boundary of the ellipsoid). The results are consistent with a convergence rate of $O(h^2 \log h)$ predicted by Theorem 2.5. which illustrates the superconvergence.

As a simple piecewise smooth surface, we use again the unit cube

$$S = [0, 1] \times [0, 1] \times [0, 1] \tag{36}$$

	$P = P_{11}$		$P = P_{21}$		$P = P_{31}$	
n	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio
4	8.52 E-1		5.05 E-1		1.02 E-1	
16	9.29 E-2	9.16	6.05 E-2	8.35	1.20 E-2	8.53
64	1.10 E-2	8.44	8.32 E-3	7.27	1.63 E-3	7.36
256	2.67 E-3	4.12	1.88 E-3	4.40	3.71 E-4	4.39

TABLE 1. Errors in solving the Neumann Problem on $E1$

	$P = P_{12}$		$P = P_{22}$		$P = P_{32}$	
n	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio
4	2.87 E-1		1.56 E-1		2.70 E-2	
16	5.94 E-2	4.84	2.91 E-2	5.36	5.09 E-3	5.30
64	1.24 E-2	4.77	5.85 E-3	4.98	9.99 E-4	5.10
256	3.02 E-3	4.12	1.29 E-3	4.53	2.07 E-4	4.82

TABLE 2. Errors in solving the Neumann Problem on $E2$

The function f is chosen so that the true solution is

$$u = \frac{1}{\sqrt{(x - 0.5)^2 + (y - 0.5)^2 + (z - 0.5)^2}} \tag{37}$$

	$P = P_1$		$P = P_2$		$P = P_3$	
n	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio	$\|u(P) - u_n(P)\|$	Ratio
12	8.98 E-1		2.92 E-1		3.62 E-2	
48	4.11 E-1	2.17	9.18 E-2	3.18	3.01 E-3	12.01
192	1.96 E-1	2.02	3.61 E-2	2.54	3.61 E-4	8.33
768	9.89 E-2	1.98	1.68 E-2	2.14	1.18 E-4	3.05

TABLE 3. Errors in solving the Neumann Problem on the unit cube

In Table 3 we give the results for $|u(P) - u_n(P)|$ for $P = P_i = (\tau_i, \tau_i, \tau_i) \in D_e(S)$, $i = 1, 2, 3$. The ratios approach 2 as n increases, which is consistent with a rate of convergence of $O(h)$ as predicted by Theorem 2.1. (with $r = 0$). As shown in the table, the further away from the boundary of S the point P is, the better the approximation.

We conclude by noting that the ideas used in this paper to study the numerical solution of the exterior Neumann problem (1) apply very well to studying the numerical solutions of the interior Neumann problem and the (interior or exterior) Dirichlet problem as well. For the interior Neumann problem (analogous to (1), only with D instead of D_e), an auxiliary condition on $f(P)$ is needed for solvability (namely, $\int_S f(Q)\, dS = 0$). Also, this problem does not have a unique solution in the sense that two solutions differ by a constant, and the integral equation corresponding to (8) is no longer uniquely solvable.

The equation coming from the Dirichlet problem is similar, but the interest in solving it using collocation methods with only interior collocation points is not so great in this case, since the kernel does not involve the normal $\mathbf{n}_P$, but the normal $\mathbf{n}_Q$.

References

[1] Atkinson K.,*A survey of boundary integral equation methods for the numerical solution of Laplace's equation in three dimensions*, Plenum Publishing, 1990.
[2] Atkinson K., *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge University Press, 1997.
[3] Atkinson K. and Chien D.,*Piecewise polynomial collocation for boundary integral equations*, SIAM Journal of Scientific Computing, 16:651-681, 1995.
[4] Colton D., *Partial Differential Equations: An Introduction*, Random House, New York, 1988.
[5] Günter N., *Potential Theory*, Ungar publishing Company, New York, 1967.
[6] Kress R., *Linear Integral Equations*, Springer-Verlag, New York, 1989.
[7] Micula S., *Numerical Methods for the Radiosity Equation and Related Problems*, (Ph. D. Thesis), University of Iowa, Iowa City, 1997.
[8] Mikhlin S., *Mathematical Physics: An Advanced Course*, North - Holland Publishing, 1970.

DIVISION OF MATHEMATICS AND COMPUTER SCIENCE,
TRUMAN STATE UNIVERSITY, KIRKSVILLE MO (USA)
E-mail address: smicula@truman.edu

CARISTI TYPE OPERATORS AND APPLICATIONS

ADRIAN PETRUŞEL

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

1. Introduction

Caristi's fixed point theorem states that each operator f from a complete metric space (X, d) into itself satisfying the condition:

there exists a proper lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that:

$$d(x, f(x)) + \varphi(f(x)) \leq \varphi(x), \text{ for each } x \in X \tag{1.1}$$

has at least a fixed point $x^* \in X$, i. e. $x^* = f(x^*)$. (see Caristi [4]).

For the multi-valued case, there exist several results involving multi-valued Caristi type conditions. For example, if F is a multi-valued operator from a complete metric space (X, d) into itself and if there exists a proper, lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that

$$\text{for each } x \in X, \text{ there is } y \in F(x) \text{ so that } d(x, y) + \varphi(y) \leq \varphi(x), \tag{1.2}$$

then the multi-valued map F has at least a fixed point $x^* \in X$, i. e. $x^* \in F(x^*)$. (see Mizoguchi-Takahashi [11]).

Moreover, if F satisfies the stronger condition:

$$\text{for each } x \in X \text{ and each } y \in F(x) \text{ we have } d(x, y) + \varphi(y) \leq \varphi(x), \tag{1.3}$$

then the multi-valued map F has at least a strict fixed point $x^* \in X$, i. e. $\{x^*\} = F(x^*)$. (see Maschler-Peleg [10]).

Another result of this type was proved by L. van Hot, as follows.

If F is a multi-valued operator with nonempty closed values and $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ is a lower semi-continuous function such that the following condition holds:

$$\text{for each } x \in X, \inf \{ d(x, y) + \varphi(y) : y \in F(x) \} \leq \varphi(x), \tag{1.4}$$

then F has at least a fixed point. (see van Hot [6])

There are several extensions and generalizations of these important principles of nonlinear analysis (see the references list and also the bibliography therein)

The purpose of this paper is to present several new results and open problems for single-valued and multi-valued Caristi type operators between metric spaces. Also,

Received by the editors: 17.02.2003.
2000 *Mathematics Subject Classification.* 47H10, 54H25.
Key words and phrases. fixed point, w-distance, multifunction.

the case of metric spaces endowed with a weakly distance in the sense of Kada-Suzuki-Takahashi is considered.

2. **Preliminaries**

Throughout this paper (X, d) is a complete metric space, $f : X \to X$ is a single-valued operator and $F : X \multimap X$ denotes a multi-valued operator.

Definition 2.1. *A single-valued operator $f : X \to X$ is said to be:*

a) Caristi type operator if there exists a lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that

$$d(x, f(x)) + \varphi(f(x)) \leq \varphi(x), \text{ for each } x \in .$$

b) Kannan type operator if there exists $a \in [0, \frac{1}{2}[$ such that

$$d(f(x), f(y)) \leq a[d(x, f(x)) + d(y, f(y))], \text{ for each } x, y \in X.$$

c) Ciric-Reich-Rus type operator if there exist $a, b, c \in \mathbb{R}_+$, with $a+b+c < 1$ such that

$$d(f(x), f(y)) \leq ad(x, y) + bd(x, f(x)) + cd(y, f(y)), \text{ for each } x, y \in X.$$

If $b = c = 0$, then f is called an a-contraction.

Definition 2.2. *If $f : X \to X$ is a single-valued operator, then $x^* \in X$ is called a fixed point of f if $x^* = f(x^*)$. We will denote by $Fixf$ the fixed point set of f.*

If (X, d) is a metric space, then $\mathcal{P}(X)$ will denote the space of all subsets of X. Also, we denote by $P(X)$ the space of all nonempty subsets of X and by $P_p(X)$ the set of all nonempty subsets of X having the property "p", where "p" could be: cl = *closed*, b = *bounded*, cp = *compact*, cv = *convex* (for normed spaces X), etc.

We consider the following (generalized) functionals:

$$D : P(X) \times P(X) \to \mathbb{R}_+, \ D(A, B) = \inf \{ \ d(a, b) \mid a \in A, \ b \in B \ \}$$

$$H : P(X) \times P(X) \to \mathbb{R}_+ \cup \{+\infty\}, \ H(A, B) = \max \{ \ \sup_{a \in A} D(a, B), \ \sup_{b \in B} D(b, A) \ \}.$$

H is called the Pompeiu-Hausdorff generalized functional and it is well-known that if (X, d) is a complete metric space, then $(P_{cl}(X), H)$ is also a complete metric space.

Definition 2.3. *Let (X, d) be a metric space. Then a multi-valued operator $F : X \to P(X)$ is called:*

a) (M-T)- Caristi type multifunction if there exists a proper lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that

$$\text{for each } x \in X, \text{ there is } y \in F(x) \text{ so that } d(x, y) + \varphi(y) \leq \varphi(x)$$

(see Mizoguchi-Takahashi [11]*)*

b) (M-P)- Caristi type multifunction if there exists a proper lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that

$$\text{for each } x \in X \text{ and each } y \in F(x) \text{ we have } d(x, y) + \varphi(y) \leq \varphi(x).$$

(see Maschler-Peleg [10]*)*

c) (vH)- Caristi type multifunction if F has closed values and there exists a proper lower semicontinuous function $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ such that

$$\text{for each } x \in X,\ \inf \{\ d(x,y) + \varphi(y)\ :\ y \in F(x)\ \} \leq \varphi(x)$$

(see van Hot [6])

d) Kannan type multifunction if there exists $a \in [0, \frac{1}{2}[$ such that

$$H(F(x), F(y)) \leq a[D(x, F(x)) + D(y, F(y))],\ \text{for each } x, y \in X.$$

e) Reich type multifunction if there exist $a, b, c \in \mathbb{R}_+$, with $a + b + c < 1$ such that

$$H(F(x), F(y)) \leq ad(x,y) + bD(x, F(x)) + cD(y, F(y)),\ \text{for each } x, y \in X.$$

If $b = c = 0$, then F is called a multi-valued a-contraction.

Remark 2.1. *It is quite obviously that if F satisfies a (M-P)- Caristi type condition then F is a (M-T)- Caristi type multifunction and any (M-T)- Caristi type multifunction satisfies a (vH)- Caristi type condition.*

Definition 2.4. *Let (X, d) be a metric space and $F : X \to P(X)$ be a multi-valued map. Then an element $x^* \in X$ is called a fixed point of F if $x^* \in F(x^*)$ and it is called a strict fixed point of F if $\{x^*\} = F(x^*)$. We denote by $Fix(F)$ the fixed points set of F and by $SFix(F)$ the strict fixed points set of F.*

In 1996, Kada, Suzuki and Takahashi introduced the concept of w-distance on a metric space as follows.

Definition 2.5. *Let (X, d) be a metric space. Then a function $p : X \times X \to \mathbb{R}_+$ is called a w-distance on X if the following are satisfied:*

(1) $p(x,z) \leq p(x,y) + p(y,z)$, for any $x, y, z \in X$

(2) for any $x \in X$, $p(x, \cdot) : X \to \mathbb{R}_+$ is lower semicontinuous

(3) for any $\varepsilon > 0$, there exists $\delta > 0$ such that $p(z,x) \leq \delta$ and $p(z,y) \leq \delta$ imply $d(x,y) \leq \varepsilon$.

Some examples of w-distances are:

Example 2.1. *Let X be a metric space with metric d. Then $p = d$ is a w-distance on X.*

Example 2.2. *Let X be a normed space with norm $\|\cdot\|$. Then the function $p : X \times X \to \mathbb{R}_+$ defined by $p(x,y) = \|x\| + \|y\|$, for every $x, y \in X$ is a w-distance on X.*

Example 2.3. *Let X be a metric space with metric d and let T be a continuous mapping from X into itself. Then a function $p : X \times X \to \mathbb{R}_+$ defined by*

$$p(x,y) = max(d(Tx, y), d(Tx, Ty)),\ \text{for every } x, y \in X$$

is a w-distance on X.

Example 2.4. *Let $X = \mathbb{R}$ with the usual metric and $f : \mathbb{R} \to \mathbb{R}_+$ be a continuous function such that*

$$\inf_{x \in X} \int_x^{x+r} f(u)du > 0,\ \text{for each } r > 0.$$

Then a function $p : X \times X \to \mathbb{R}_+$*, defined by*

$$p(x,y) := |\int_x^y f(u)du|, \text{ for every } x, y \in X$$

is a w-distance on X.

For other examples and related results, see Kada, Suzuki and Takahashi [7].

Some important properties of the w-distance are contained in:

Lemma 2.1. *Let* (X,d) *be a metric space and* p *be a w-distance on* X*. Let* $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ *be sequences in* X*, let* $(\alpha_n)_{n\in\mathbb{N}}, (\beta_n)_{n\in\mathbb{N}}$ *be sequences in* $\mathbb{R}_+$ *converging to* 0 *and let* $y, z \in X$*. Then the following hold:*

(i) if $p(x_n, x_m) \leq \alpha_n$*, for any* $n, m \in \mathbb{N}$ *with* $m > n$*, then* (x_n) *is a Cauchy sequence.*

(ii) if $p(y, x_n) \leq \alpha_n$*, for any* $n \in \mathbb{N}$*, then* (x_n) *is a Cauchy sequence.*

(iii) if $p(x_n, y_n) \leq \alpha_n$ *and* $p(x_n, z) \leq \beta_n$*, for any* $n \in \mathbb{N}$ *then* (y_n) *converges to* z*.*

3. Single-valued Caristi type operators

If $f : X \to X$ is an a-contraction, then it is well-known (see for example Dugundji-Granas [5]) that f is a Caristi type operator with a function $\varphi(x) = \frac{1}{1-a}d(x, f(x))$. Also, Caristi type mappings include Reich type operators and in particular Kannan operators. Indeed, if f satisfies a Reich type condition with constants a, b, c, then f is a Caristi type operator with a function $\varphi(x) = \frac{1-c}{1-a-b-c}d(x, f(x))$.

Moreover, if $f : X \to X$ satisfies the following condition (see I. A. Rus (1972), [14]):

$$\text{there is } a \in [0,1[\text{ such that } d(f(x), f^2(x)) \leq ad(x, f(x)), \text{ for each } x \in X$$

then f is a Caristi operator with a function $\varphi(x) = \frac{1}{1-a}d(x, f(x))$.

Hence, the class of single-valued Caristi type operators is very large, including at least the above mentioned types of contractive mappings.

Some characterizations of metric completeness have been discussed by several authors such as Weston, Kirk, Suzuki, Suzuki and Takahashi, Shioji, Suzuki and Takahashi, etc. For example, Kirk [8] proved that a metric space is complete if it has the fixed point property for Caristi mappings. Moreover, Shioji, Suzuki and Takahashi proved in [15] that a metric space is complete if and only if it has the fixed point property for Kannan mappings. On the other hand, it is well-known that the fixed point property for a-contraction mappings does not characterize metric completeness, see for example Suzuki-Takahashi [16]. Thus, Kannan mappings and Caristi mappings characterize metric completeness, while contraction mappings do not. Regarding to the problem of characterizations of metric completeness by means of contraction mappings, Suzuki and Takahashi and independently M. C. Anisiu and V. Anisiu showed (see [16] respectively [1]) that a convex subset Y of a normed space is complete if and only if every contraction $f : Y \to Y$ has a fixed point in Y.

The following generalization of Caristi's theorem is proved in Kada-Suzuki-Takahashi [7]:

Theorem 3.1. *Let (X, d) be a complete metric space, let $\varphi : X \times X \to \mathbb{R}_+$ be a proper lower semicontinuous function and let $f : X \to X$ a mapping. Assume that there exists a w-distance p on X such that*

$$p(x, f(x)) + \varphi(f(x)) \leq \varphi(x), \text{ for each } x \in X.$$

Then there exists $x^* \in X$ such that $x^* = f(x^*)$ and $p(x^*, x^*) = 0$.

Open problem. The following result (see Zhong, Zhu and Zhao [18]) is a generalization of Caristi's fixed point principle:

Theorem 3.2. *If (X, d) is a complete metric space, $x_0 \in X$, $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ is proper lower semicontinuous and $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function such that $\int_0^\infty \frac{ds}{1+h(s)} = \infty$, then each single-valued operator f from X to itself satisfying the condition:*

$$\text{for each } x \in X, \frac{d(x, f(x))}{1 + h(d(x_0, x))} + \varphi(f(x)) \leq \varphi(x), \tag{3.5}$$

has at least a fixed point.

It is of interest to see if such a result, in terms of w-distances, is true.

4. **Multi-valued Caristi type operators**

It was proved by L. van Hot that any multi-valued a-contraction F on a metric space X is a (vH)-Caristi type multi-function with a function $\varphi(x) = \frac{1}{1-a} D(x, F(x))$. Moreover, if is a multi-valued a-contraction with nonempty and compact values then F satisfies a (M-T)- Caristi type condition with a same function $\varphi(x) = \frac{1}{1-a} D(x, F(x))$.

Let us remark now, that any Reich type multi-function (and hence in particular any Kannan multi-function) is a (vH)- Caristi type multi-function with a function φ given by $\varphi(x) = \frac{1-c}{1-a-b-c} D(x, F(x))$.

Definition 4.6. *If (X, d) is a metric space, then a multi-valued operator $F : X \to P(X)$ is said to be a Reich type graphic contraction if there exist $a, b, c \in \mathbb{R}_+$, with $a + b + c < 1$ such that*

$$H(F(x), F(y)) \leq ad(x, y) + bD(x, F(x)) + cD(y, F(y)),$$

for each $x \in X$ and each $y \in F(x)$.

A connection between multi-valued Reich type graphic contractions and multi-valued Caristi type operators is given in:

Lemma 4.2. *Let (X, d) be a metric space and let $F : X \to P(X)$ be a Reich type graphic contraction. Then F is a (vH)- Caristi type multi-function.*

Proof. Let $\varphi(x) = \frac{1-c}{1-a-b-c} D(x, f(x))$. Then, because for each $x \in X$ and each $y \in F(x)$ we have $D(y, F(y)) \leq H(F(x), F(y)) \leq ad(x, y) + bD(x, F(x)) + cD(y, F(y))$, we obtain $D(y, F(y)) \leq \frac{1}{1-c}(ad(x, y) + bD(x, F(x)))$.

Hence for $x \in X$ and $y \in F(x)$ we get $d(x, y) + \varphi(y) \leq d(x, y) + \frac{1}{1-a-b-c}(ad(x, y) + bD(x, F(x)))$ and so $inf\{d(x, y) + \varphi(y) | y \in F(x)\} \leq inf\{\frac{1-b-c}{1-a-b-c} d(x, y) | y \in F(x)\} + \frac{b}{1-a-b-c} D(x, F(x)) = \varphi(x)$. In conclusion, F is a (vH)- Caristi type multi-function and the proof is complete. ■

For the case of complete metric spaces endowed with a w-distance the following generalization of the Covitz-Nadler fixed point principle for multi-functions was proved by Suzuki and Takahashi in [16]. We need, first, a definition.

Definition 4.7. *Let (X,d) be a metric space. A multi-valued mapping $F : X \to P(X)$ is called p-contractive if there exist a w-distance p on X and a real number $a \in [0,1[$ such that for any $x_1, x_2 \in X$ and each $y_1 \in F(x_1)$ there exists $y_2 \in F(x_2)$ so that $p(y_1,y_2) \leq ap(x_1,x_2)$*

Theorem 4.3. *Let (X,d) be a complete metric space and $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ be a proper lower semicontinuous function. Let $F : X \to P_{cl}(X)$ be a p-contractive multi-function. Then there exists $x^* \in X$ a fixed point for F and $p(x^*,x^*) = 0$.*

Some extensions of the previous result are:

Theorem 4.4. *Let (X,d) be a complete metric space and $F : X \to P_{cl}(X)$ be a closed multi-valued operator such that the following assumption holds:*

there exist a w-distance p on X and a real number $a \in [0,1[$ so that for any $x \in X$ and any $y_1 \in F(x)$ there is $y_2 \in F(y_1)$ such that $p(y_1,y_2) \leq ap(x,y_1)$.

Then there exists $x^* \in X$ a fixed point for F and $p(x^*,x^*) = 0$.

Proof. Let $u_0 \in X$ and $u_1 \in F(u_0)$. Then there exists $u_2 \in F(u_1)$ such that $p(u_1,u_2) \leq ap(u_0,u_1)$. Then, $u_1 \in X$ and $u_2 \in F(u_1)$ we obtain that there is $u_3 \in F(u_2)$ with $p(u_2,u_3) \leq ap(u_1,u_2) \leq a^2p(u_0,u_1)$. Thus we can construct a sequence $(u_n)_{n\in\mathbb{N}}$ from X satisfying:

i) $u_{n+1} \in F(u_n)$, for each $n \in \mathbb{N}$.

ii) $p(u_n,u_{n+1}) \leq a^n p(u_0,u_1)$, for each $n \in \mathbb{N}$.

Hence, for any $n,m \in \mathbb{N}$ with $m > n$ we have: $p(u_n,u_m) \leq p(u_n,u_{n+1}) + ... + p(u_{m-1},u_m) \leq a^n p(u_0,u_1) + a^{n+1}p(u_0,u_1) + ... + a^{m-1}p(u_0,u_1) \leq \frac{a^n}{1-a}p(u_0,u_1)$. By Lemma 2.1 $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence $(u_n)_{n\in\mathbb{N}}$ converges to a point $u^* \in X$. Since F is closed we get that $u^* \in F(u^*)$.

Let us consider now a fixed $n \in \mathbb{N}$. Since $(u_m)_{m\in\mathbb{N}}$ converges to a $u^* \in X$ and $p(u_n,\cdot)$ is lower semicontinuous we have

$$p(u_n,u^*) \leq \liminf_{m\to\infty} p(u_n,u_m) \leq \frac{a^n}{1-a}p(u_0,u_1).$$

For $u^* \in F(u^*)$, there exists $w_1 \in F(u^*)$ such that $p(u^*,w_1) \leq ap(u^*,u^*)$. By a similar approach, we can construct a sequence $(w_n)_{n\in\mathbb{N}}$ in X such that $w_{n+1} \in F(w_n)$ and $p(u^*,w_{n+1}) \leq ap(u^*,w_n)$, for each $n \in \mathbb{N}$. So, as before we obtain

$$p(u^*,w_n) \leq ap(u^*,w_{n-1}) \leq ... \leq a^n p(u^*,u^*).$$

By Lemma 2.1, (w_n) is a Cauchy sequence and it converges to a point $x^* \in X$. Since $p(u^*,\cdot)$ is lower semicontinuous we have

$$p(u^*,x^*) \leq \liminf_{n\to\infty} p(u^*,w_n) \leq 0$$

and hence $p(u^*,x^*) = 0$. Then, for any $n \in \mathbb{N}$ we have

$$p(u_n,x^*) \leq p(u_n,u^*) + p(u^*,x^*) \leq \frac{a^n}{1-a}p(u_0,u_1).$$

Using again Lemma 2.1 we obtain $u^* = x^*$ and hence $p(x^*, x^*) = 0$. The proof is now complete. ■

Theorem 4.5. *Let (X, d) be a complete metric space and $\varphi : X \to \mathbb{R}_+ \cup \{+\infty\}$ be a proper lower semicontinuous function. Let $F : X \to P_{cl}(X)$ be a closed multi-valued operator having the following property:*

there exists a w-distance p on X so that for each $x \in X$ there is $y \in F(x)$ we have $p(x, y) + \varphi(y) \le \varphi(x)$.

Then $FixF \neq \emptyset$.

Moreover, if F satisfies the stronger condition:

there exists a w-distance p on X so that for each $x \in X$ and for each $y \in F(x)$ we have $p(x, y) + \varphi(y) \le \varphi(x)$,

then there exists $x^* \in X$ a fixed point for T and $p(x^*, x^*) = 0$.

Proof. From the hypothesis it follows that if p is a w-distance on (X, d), there exists a single-valued operator $f : X \to X$ such that f is a selection for F and satisfies the condition

$$p(x, f(x)) + \varphi(f(x)) \le \varphi(x), \text{ for each } x \in X.$$

The first conclusion follows now from Theorem 3.1. For the second conclusion of the theorem, we observe that for $x^* \in F(x^*)$ we have $p(x^*, x^*) + \varphi(x^*) \le \varphi(x^*)$ and hence $p(x^*, x^*) = 0$. The proof is now complete. ■

Following an idea from Kirk, Srivassan and Veeramani [9] we also have the following Caristi type theorem:

Theorem 4.6. *Let (X, d) be a complete metric space and $F : X \to P(X)$ be a multi-function. Let us suppose that there exist $A_1, A_2, ..., A_p$ closed subsets of X and $\varphi_i : A_i \to \mathbb{R}$, for $i \in \{1, 2, ..., p\}$ lower semicontinuous mappings, such that the following assumptions hold:*

i) $F(A_i) \subset A_{i+1}$, for $i \in \{1, 2, ..., p\}$, where $A_{p+1} = A_1$.

ii) for each $x \in A_i$ and each $y \in F(x)$ we have

$$d(x, y) + \varphi_{i+1}(y) \le \varphi_i(x), \text{ for } i \in \{1, 2, ..., p\}.$$

Then $FixF \neq \emptyset$.

Proof. Let $x_0 \in A_1$. Then for $x_1 \in F(x_0) \subset A_2$ we have $d(x_0, x_1) \le \varphi_1(x_0) - \varphi_2(x_1)$. Then for $x_1 \in A_2$ and $x_2 \in F(x_1) \subset A_3$ we have that $d(x_1, x_2) \le \varphi_2(x_1) - \varphi_3(x_2)$. So, we can construct a sequence $(x_n)_{n\in\mathbb{N}}$, having the following properties:

i) $x_{n+1} \in F(x_n)$, for $n \in \mathbb{N}$.

ii) $d(x_n, x_{n+1}) \le \varphi_{n+1}(x_n) - \varphi_{n+2}(x_{n+1})$, for $n \in \mathbb{N}$.

Let us observe that the sequence $(\varphi_{k+1}(x_k))_{k\in\mathbb{N}}$ converges to an element $a \in \mathbb{R}$. Then for $n, m \in \mathbb{N}$, with $n > m$ we get $d(x_n, x_m) \le d(x_n, x_{n+1}) + ... + d(x_{m-1}, x_m) \le \varphi_{n+1}(x_n) - \varphi_{m+1}(x_m)$. Therefore $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence that converges to a point $x \in \bigcap_{i=1}^{p} A_i$. So $A := \bigcap_{i=1}^{p} A_i \neq \emptyset$. Hence $F : A \to P(A)$. Moreover we can write:

for $x \in A \subset A_1$ and $y \in F(x)$ $d(x, y) \le \varphi_1(x) - \varphi_2(y)$

for $x \in A \subset A_2$ and $y \in F(x)$ $d(x,y) \leq \varphi_2(x) - \varphi_3(y)$
...
for $x \in A \subset A_p$ and $y \in F(x)$ $d(x,y) \leq \varphi_p(x) - \varphi_1(y)$

and therefore $pd(x,y) \leq \sum_{i=1}^{p}(\varphi_i(x) - \varphi_i(y))$. If we define $\varphi : A \to \mathbb{R}$ by $\varphi(x) = \frac{1}{p}\sum_{i=1}^{p}\varphi_i(x)$, then φ is lower semicontinuous and the following assertion holds:

for each $x \in A$ there exists $y \in F(x)$ such that $d(x,y) \leq \varphi(x) - \varphi(y)$.

Using Mizoguchi-Takahashi's theorem (see [11]) we get the conclusion. ■

Remark. If in the previous theorem F is a single-valued operator then Theorem 4.6 is Theorem 3.1 from Kirk, Srivassan and Veeramani [9].

Open problem. For other Caristi type conditions, results as the previous one, involving cyclical type conditions, can be established?

References

[1] Anisiu, M. C., Anisiu, V., *On the closedness of sets with the fixed point property for contractions*, Rev. Anal. Numér. Théor. Approx., 26 (1997), 13-17.

[2] Aubin, J.-P., *Optima and equilibria. An introduction to nonlinear analysis*, Springer-Verlag, Berlin, 1993.

[3] Aubin, J.-P., Siegel, J., *Fixed points and stationary points of dissipative multivalued maps*, Proc. A. M. S., 78(1980), 391-398.

[4] Caristi, J., *Fixed points theorems for mappings satisfying inwardness conditions*, Trans. A. M. S., 215(1976), 241-251.

[5] Dugundji, J. and Granas, A., *Fixed point theory*, Polish. Scientific Publishers, Warsaw, 1982.

[6] van Hot, L., *Fixed point theorems for multivalued mappings*, C. M. U. C., 23(1982), 137-145.

[7] Kada, O., Suzuki, T. and Takahashi, W., *Nonconvex minimization theorems and fixed point theorems in complete metric spaces*, Math. Japonica, 44(1996), 381-391.

[8] Kirk, W. A., *Caristi's fixed point theorem and metric convexity*, Colloquium Math., 36(1976), 81-86.

[9] Kirk, W. A., Srinivasan, P. S. and Veermani, P., *Fixed points for mappings satisfying cyclical contractive conditions*, Seminar on Fixed Point Theory Cluj-Napoca, 4(2003), to appear.

[10] Maschler, M., Peleg, B., *Stable sets and stable points of set-valued dynamic systems with applications to game theory*, SIAM J. Control Optimization, 14(1976), 985-995.

[11] Mizoguchi, N., Takahashi, W., *Fixed point theorems for multivalued mappings on complete metric spaces*, J. Math. Anal. Appl., 141(1989), 177-188.

[12] Petruşel, A., Sîntămărian, A., *Single-valued and multi-valued Caristi type operators*, Publ. Math. Debrecen, 60(2002), 167-177.

[13] Rus, I. A., *Generalized contractions and applications*, Cluj University Press, Cluj-Napoca, 2001.

[14] Rus, I. A., *The method of successive approximations*, Collection of articles dedicated to G. Călugăreanu on his 70-th birthday, Rev. Roumaine Math. Pures Appl., 17(1972), 1433-1437, (in Russian).

[15] Shioji, N., Suzuki, T., Takahashi, W., *Contractive mappings, Kannan mappings and metric completeness*, Proc. A. M. S., 126(1998), 3117-3124.

[16] Suzuki, T., Takahashi, W., *Fixed point theorems and characterizations of metric completeness*, Topological Meth. Nonlinear Anal., 8(1996), 371-382.

[17] Weston, J. D., *A characterization of metric completeness*, Proc. A. M. S., 64(1977), 186-188.
[18] Zhong, C.-K., Zhu, J., Zhao, P.-H., *An extension of multi-valued contraction mappings and fixed points*, Proc. A. M. S., (128)1999, 2439-2444.

Babeş-Bolyai University of Cluj-Napoca,
Department of Applied Mathematics, 3400 Cluj-Napoca, Romania,
E-mail address: petrusel@math.ubbcluj.ro

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

ON THE CONVERGENCE OF THE SOLUTION OF THE QUASI-STATIC CONTACT PROBLEMS WITH FRICTION USING THE UZAWA TYPE ALGORITHM

NICOLAE POP

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. The aim of the paper is to prove the convergence of a Uzawa type algorithm for a dual mixed variational formulation of a quasi-static contact problem with friction. This problem is considered as a saddle point problem which is approximated with the mixed finite element, where the stress, displacement and tangential displacement on the contact boundary will be simultaneously computed.

1. Introduction

The quasi-static model of the contact problems with friction, without the inertia effects, was proposed by [14] and consists of the formulation obtained through the approximation with finite differences of the variational inequality. The proof of the existence and uniqueness is based on the hypothesis that the displacements satisfy some conditions of regularity and the friction coefficient is small enough. The static contact problem with friction cannot describe the evolutive state of the contact conditions. For of this reason, the quasi-static formulation, of the contact problem with friction is preferred, which contains a dynamic formulation of the contact conditions and the inertial term is no longer used. Through the temporal discretization of the quasi-static contact problem, the so called incremental problem is obtained, equivalent with a sequence of static contact problems. Therefore, the quasi-static problem is solved step by step, at each time small deformations and displacements are calculated and are added at those calculated previously, as a result of a few small modifications of the applied forces, of the contact zone and of the contact conditions. Although, at each increment the dependence of the load-way is neglected, this hypothesis takes into account the way the applied forces change (modify themselves). From a mathematical point view, the problem obtained at each step is similar with a static problem.

This dual mixed variational formulation problem is descretized by the mixed finite element method and an Uzawa type algorithm is proposed. The iterative formulation of this algorithm is deduced and its convergence is proved.

Received by the editors: 28.05.2003.

Key words and phrases. unilateral quasi-static contact problem, dual mixed formulation, incremental scheme, Lagrange multipliers, mortar finite elements, mixed finite element, saddle point problem, Coulomb law.

The existence of solutions for the discrete problem by the mixed element method was obtained by Haslinger [7]. The contact problem has been recently studied by Andersen [11] and Rocca and Cocou [6] who proved that there exists a solution if the friction coefficient is small enough, and smooth and the contact functional is regular.

In this article is assumed that normal component of the stress vector and the contact zone is known.

2. Classical and variational formulation

Let $\Omega \subset \mathbb{R}^d, d = 2$ or 3, the polygonal domain occupied by a linear elastic body, and its boundary is denoted by Γ. Let Γ_1, Γ_2 and Γ_c be three open disjoint parts of Γ such that $\Gamma = \overline{\Gamma}_1 \cup \overline{\Gamma}_2 \cup \overline{\Gamma}_c$, $\overline{\Gamma}_1 \cap \overline{\Gamma}_c = \varnothing$ and mes $(\Gamma_1) > 0$. We assume for the simplicity that Γ_c is a segment for $d = 2$ and a polygon for $d = 3$. We denote by $\boldsymbol{u} = (u_1, \ldots, u_d)$ the displacement field, $\boldsymbol{\varepsilon} = (\varepsilon_{ij}(\mathbf{u})) = \left(\frac{1}{2}(u_{i,j} + u_{j,i})\right)$ the strain tensor and $\boldsymbol{\sigma} = (\sigma_{ij}(\mathbf{u})) = (a_{ijkl}\varepsilon_{kl}(\mathbf{u}))$ the stress tensor with the usual summation convention, where $i, j, k, l = 1, \ldots, d$. For the normal and tangential components of the displacement vector and stress vector, we use the following notation: $\mathbf{u}_N = u_i \cdot n_i$, $\boldsymbol{u}_T = \boldsymbol{u} - \boldsymbol{u}_N \cdot \mathbf{n}$, $\boldsymbol{\sigma}_N = \boldsymbol{\sigma}_{ij} u_i n_j$, $(\boldsymbol{\sigma}_T)_i = \boldsymbol{\sigma}_{ij} n_j - \boldsymbol{\sigma}_N \cdot n_i$, where $\mathbf{n} = (n_i)$ is the outward unit normal vector to $\partial\Omega$.

Lets us denote by $\boldsymbol{f}$ and $\boldsymbol{h}$ the density of body forces and traction forces, respectively. We assume that $a_{ijkl} \in L^\infty(\Omega)$, $l \le i, j, k, l \le d$, with usual condition of symmetry and elasticity, that is

$$a_{ijkl} = a_{jikl} = a_{klij}, \quad l < i, j, k, l \le d$$

$$\exists\, m_0 > 0,\ \forall\, \xi = (\xi_{ij}) \in \mathbb{R}^{d^2},\ \xi_{ij} = \xi_{ji},\ l \le i,\ j \le d,\ a_{ijkl}\, \xi_{ij}\, \xi_{kl} \ge m_0 |\xi|^2 .$$

In this conditions, the fourth-order tensor $\boldsymbol{a} = (a_{ijkl})$ is invertible a.e. on Ω and we denote its inverse $\boldsymbol{b} = (b_{ijkl})$, and $\boldsymbol{\varepsilon}_{ij}(\boldsymbol{u})) = (b_{ijkl}\sigma_{kl}(\boldsymbol{u}))$, $i, j, k, l = 1, \ldots, d$.

The classical contact problem with dry friction in elasticity is which the normal stress $\sigma_N(u)$ and Γ_c is assumed known, is follows: Find $\boldsymbol{u} = \boldsymbol{u}(x, t)$ such that $\boldsymbol{u}(0, \cdot) = \boldsymbol{u}^0(\cdot)$ in Ω and all $t \in [0, T]$,

$$-\operatorname{div} \boldsymbol{\sigma}(\boldsymbol{u}) = \boldsymbol{f}, \quad \text{in } \Omega \tag{2.1}$$

$$\boldsymbol{\sigma}_{ij}(\boldsymbol{u}) = a_{ijkl} \cdot \varepsilon_{kl}(\boldsymbol{u}), \quad \text{in } \Omega \tag{2.2}$$

$$\boldsymbol{u} = 0 \quad \text{on } \Gamma_1 \tag{2.3}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{h} \quad \text{on } \Gamma_2 \tag{2.4}$$

$$u_N \le 0,\ \boldsymbol{\sigma}_N(u) \le 0,\ u_N \boldsymbol{\sigma}_N(u) = 0 \quad \text{on } \Gamma_c \tag{2.5}$$

$$\mu_F |\boldsymbol{\sigma}_N(\boldsymbol{u})| = t, \quad t > 0$$

$$|\boldsymbol{\sigma}_T| < t \Rightarrow \dot{u}_T = 0;\ |\boldsymbol{\sigma}_T| = t \Rightarrow \exists\, \lambda \ge 0, \text{ s.t. } \dot{u}_T = -\lambda \boldsymbol{\sigma}_T \text{ on } \Gamma_c \tag{2.6}$$

where $\boldsymbol{u}^0$ is denoted the initial displacement of the body.

Condition (2.6) defines a form of Coulomb's law of friction for elastostatic problems: μ_F is the coefficient of friction $\mu_F \in L^\infty(\Gamma_c)$, $\mu_F \ge \mu_0$ a.e. on Γ_c.

The dual mixed variational formulation of the (2.1) - (2.6) in which stress, displacement and tangential displacement on contact zone are considerate unknown, it is shown the saddle-point problem with the form:

Find $(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\lambda}) \in S_t \times V \times \Lambda$ for all $t \in [0, T]$, such that

$$L(\boldsymbol{\sigma}, \boldsymbol{v}, \boldsymbol{\mu}) \leq L(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{\lambda}) \leq L(\boldsymbol{\tau}, \boldsymbol{u}, \boldsymbol{\lambda}) \; \forall \; (\boldsymbol{\tau}, \boldsymbol{v}, \boldsymbol{\mu}) \in S_0 \times V \times \Lambda, \tag{2.7}$$

where $\boldsymbol{u} \in W^{1,2}(0, T; V)$, $\boldsymbol{\sigma} \in W^{1,2}(0, T; \mathcal{S})$, $\boldsymbol{f} \in W^{1,2}(0, T; [L^2(\Omega)]^d)$,
$\boldsymbol{h} \in W^{1,2}(0, T; [L^2(\Gamma)]^d)$ with $supp(h(t)) \subset \Gamma_2$ for all $t \in [0, T]$.

$$L(\boldsymbol{\tau}, \boldsymbol{v}, \boldsymbol{\mu}) = J_0(\boldsymbol{\tau}) - (\operatorname{div} \boldsymbol{\tau}, \dot{\boldsymbol{v}}) - < \boldsymbol{t}, \; \boldsymbol{\mu} >_{\Gamma_c} \tag{2.8}$$

$$J_0(\boldsymbol{\tau}) = \frac{1}{2} \, a^*(\boldsymbol{\tau}, \boldsymbol{\tau}) + (\boldsymbol{f}, \operatorname{div} \boldsymbol{\sigma} + \dot{\boldsymbol{u}}) \tag{2.9}$$

$$\boldsymbol{t} = \mu_F |\boldsymbol{\sigma}_N(\boldsymbol{u})|, \quad \text{and} \quad \boldsymbol{\mu} = |\boldsymbol{u}_T| \text{ on } \Gamma_c \tag{2.10}$$

$$S_0 = \left\{ \boldsymbol{\tau} | \tau_{ij}, \tau_{ij,j} \in L^2(\Omega), \; \tau_{ij} = \tau_{ji}, \; \boldsymbol{\tau} \cdot \boldsymbol{n} = 0 \text{ a.e. on } \Gamma_2^f \right\} \tag{2.11}$$

$$S_t = \left\{ \boldsymbol{\tau} | \tau_{ij}, \tau_{ij,j} \in L^2(\Omega), \; \tau_{ij} = \tau_{ji}, \; \boldsymbol{\tau} \cdot \boldsymbol{n} = t \text{ a.e. on } \Gamma_2 \right\} \tag{2.12}$$

$$S = \left\{ \boldsymbol{\tau} | \tau_{ij} \in L^2(\Omega), \; \tau_{ij} = \tau_{ji}, \; \tau_{ij,j} \in L^2(\Omega) \right\}$$

endowed with inner product

$$(\boldsymbol{\sigma}, \boldsymbol{\tau})_{\mathcal{S}} = \int_\Omega \sigma_{ij} \, \tau_{ij} \, dx. \tag{2.13}$$

Norm $\| \cdot \|_{\mathcal{S}}$ is then

$$\|\boldsymbol{\tau}\|_{\mathcal{S}} = (\boldsymbol{\tau}, \boldsymbol{\tau})_{\mathcal{S}}^{1/2} \tag{2.14}$$

$$\text{and} \qquad a^*(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int_\Omega b_{ijkl} \, \sigma_{kl} \, dx \, . \tag{2.15}$$

Γ_2^f can be regarded as part of Γ_2 where $h \equiv 0$,

$$\Lambda = \{ \boldsymbol{\mu} \in H_{00}^{1/2}(\Gamma_c) | \, \boldsymbol{\mu} \geq 0 \; \text{ on } \Gamma_2^f \} \tag{2.16}$$

$$V = \{ \boldsymbol{v} \in H^1(\Omega) | \, \boldsymbol{v} / \Gamma_1 = 0 \} \tag{2.17}$$

$$H_{00}^{1/2}(\Gamma_c) = \{ \boldsymbol{\mu} \in H^{1/2}(\Gamma_c) | \rho^{-1/2} \boldsymbol{\mu} \in L^2(\Gamma_c) \} \, . \tag{2.18}$$

The norm of $H_{00}^{1/2}(\Gamma_c)$ is defined by

$${}_{00}\|\boldsymbol{\mu}\|_{1/2, \Gamma_c} = \left\{ \|\boldsymbol{\mu}\|_{1/2, \Gamma_c}^2 + \|d^{-1/2} \boldsymbol{\mu}\|_{0, \Gamma_c}^2 \right\}^{1/2} \, , \tag{2.19}$$

where d denotes the distance between the point on Γ_c and the end point of Γ_c see [4].

3. The time discretisation and the mixed finite element approximation of the saddle point problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded and $(T_h)_h$ a triangulation of Ω. We assume that each triangulation is compatible with the partition of Γ. i.e. each point where the boundary condition changes is a node of a set Ω_i, where $\overline{\Omega} = \cup_{i \in J_h} \overline{\Omega}_i$, with $\Omega_k \cup \Omega_l = \varnothing$ for all $k, l \in J_h, k \neq l$.

The finite element approximation to the saddle-point problem (2.7) is as follow:

Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\lambda}_h) \in S_t^h \times V_h \times \Lambda_h$ for all $t \in [0, T]$, such that

$$L(\boldsymbol{\sigma}_h, \boldsymbol{v}_h, \boldsymbol{\mu}_h) \leq L(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\lambda}_h) \leq L(\boldsymbol{\tau}_h, \boldsymbol{u}_h, \boldsymbol{\lambda}_h),\ \forall\, (\boldsymbol{\tau}_h, \boldsymbol{v}_h, \boldsymbol{\mu}_h) \in S_0^h \times V_h \times \Lambda_h \quad (3.1)$$

where $S_0^h = S_0 \cap S_h,\ S_t^h = S_h,\ \Lambda_h = M_h \cap \Lambda$ and S_h, V_h, M_h are subspaces of finite elements of S, V and $H_{00}^{1/2}(\Gamma_c)$, respectively. Let S_h be RT_1, Raviart-Thomas space, V_h the space of the piecewise constant and M_h piecewise continuous linear subspace of $H_{00}^{1/2}(\Gamma_c)$, is called the mortar space [10], as well.

We assume that the initial displacement field u satisfies the compatibility conditions, see ([8]).

The discrete Babuška-Brezzi condition should be satisfied for the dual mixed finite element method. It means to find an interpolation operator π_h from $\boldsymbol{S}$ to Ω^h, such that:

$$b(\boldsymbol{\tau} - \pi_h \boldsymbol{\tau}, \boldsymbol{v}_h, \boldsymbol{\lambda}_h) = 0 \quad (3.2)$$

$$\|\pi_h \boldsymbol{\tau}\|_s \leq c \|\boldsymbol{\tau}\|_s, \quad \forall\, \boldsymbol{\tau} \in S, \quad (3.3)$$

that means, for all $\pi_h \boldsymbol{\tau} \in S_h$ we have

$$\int_\Omega \operatorname{div}(\boldsymbol{\tau} - \pi_h \boldsymbol{\tau}) \boldsymbol{v}_h dx + \int_{\Gamma_c} (\boldsymbol{\tau}_N - \pi_h \boldsymbol{\tau}_N) \boldsymbol{\mu}_h ds = 0, \quad (\forall\, \boldsymbol{v}_h \in V_h, \boldsymbol{\mu}_h \in \Lambda_h). \quad (3.4)$$

Let

$$\int_\Omega \operatorname{div}(\boldsymbol{\tau} - \pi_h \boldsymbol{\tau}) \boldsymbol{v}_h dx = 0, \quad (\forall\, \boldsymbol{v}_h \in V_h) \quad (3.5)$$

$$\int_{\Gamma_c} (\boldsymbol{\tau}_N - (\pi_h \boldsymbol{\tau}_h)_N \boldsymbol{\mu}_h ds = 0, \quad \forall\, \boldsymbol{\mu}_h \in \Lambda_h\,. \quad (3.6)$$

Because $\boldsymbol{\sigma}_N(\boldsymbol{u})$ on Γ_c is regarded as given, applying Green's formula to equation (3.5) in the finite element discrete form, is clear that the elements of subspace S_h satisfies (3.2) and (3.3) and we finally obtain further

$$\|\boldsymbol{\tau}_{Nh}\|_{0,\Gamma_c} \leq \|\boldsymbol{\tau}_h\|_{0,\Omega} \leq \|\boldsymbol{\tau}_h\|_S, \quad (\forall\, \boldsymbol{\tau}_h \in \boldsymbol{S}_h). \quad (3.7)$$

The discretization of the saddle-point of the problem (3.1) by introduce a partition $(t_0, t_1, \ldots, t_N)$ of time interval $[0, T]$ and consider on incremental formulation obtained by using the backward finite difference approximation of the time derivative of u.

If we used $u_h^k = u_h(x, t_k)$, $\Delta u_h^k = u_h^{k+1} - u_h^k$, $\Delta t^k = t^{k+1} - t^k$, $\dot{u}_h(t^{k+1}) = \Delta u_h^k / \Delta t$, $f_h^k = f_h(k\Delta t)$, $h_h^k = h_h(k\Delta t)$, $\sigma_h^k = \sigma_h(u_h^k)$, $\lambda_h^k = |u_{Th}^k|$, for $k = 0, 1 \ldots, N$

where $\Delta t = \dfrac{T}{N}$.
In this case, we find $(\boldsymbol{\sigma}_h^k, \boldsymbol{u}_h^k, \boldsymbol{\lambda}_h^k) \in \mathcal{S}_t^h \times V_h \times \Lambda_h$ such that

$$L(\boldsymbol{\sigma}_h^k, \boldsymbol{v}_h^k, \boldsymbol{\mu}_h^k) \leq L(\boldsymbol{\sigma}_h^k, u_h^k, \lambda_h^k) \leq L(\boldsymbol{\tau}_h^k, u_h^k, \lambda_h^k),\ \forall\ (\boldsymbol{\tau}_h^k, \boldsymbol{v}_h^k, \boldsymbol{\mu}_h^k) \in S_0^h \times V_h \times \Lambda_h, \quad (3.8)$$

$k = 0, 1 \ldots, N$.

In this mode the quasi-static problem is approximated by a sequence of incremental problems (3.8).

Although, every problem (3.2) is a static one, it requires appropriate updating of the displacements and the loads after each increment.
The existence of the solution is guaranteed by the discrete Babuška-Brezzi condition should by satisfied for dual mixed element method, see ([4] and [14]).

4. **Convergence analysis of the Uzawa algorithm**

On the convergence (see [11]) with the finite element discrete problem (3.1) is following:

Proposition 4.1. *If* $(\boldsymbol{\sigma}_h^k, \boldsymbol{u}_h^k, \boldsymbol{\lambda}_h^k)$ *is the saddle-point of the problem* (3.8), *then*

(i) $\quad J_0(\boldsymbol{\sigma}_h^k, \boldsymbol{u}_h^k) - (\operatorname{div} \boldsymbol{\sigma}_h^k, \boldsymbol{u}_h^k) - < \mu_F \left|\boldsymbol{\sigma}_N^k\right|,\ \boldsymbol{\lambda}_h^k >_{\Gamma_c} \leq$

$\quad \leq J_0(\boldsymbol{\tau}_h^k, \boldsymbol{u}_h^k) - (\operatorname{div} \boldsymbol{\tau}_h^k, \boldsymbol{u}_h^k) - < \mu_F \left|\boldsymbol{\tau}_N^k\right|,\ \boldsymbol{\lambda}_h^k >_{\Gamma_c},\ (\forall\ \boldsymbol{\tau}_h^k \in S_0^h)$,

(ii) $\quad < \mu_F \left|\boldsymbol{\sigma}_N^k\right|,\ \boldsymbol{\mu}_h^k - \boldsymbol{\lambda}_h^k >_{\Gamma_c} + (\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}^k,\ \boldsymbol{v}_h^k - \boldsymbol{u}_h^k) \leq 0$,

$\quad (\forall\ \boldsymbol{\mu}_h^k \in \Lambda_h,\ \boldsymbol{v}_h^k \in V_h)$

where $\boldsymbol{\lambda}_h^k = |\boldsymbol{v}_{Th}^k|, \boldsymbol{\mu}_h = |\boldsymbol{u}_{Th}^k|$ *on* Γ_c, $k = 0, 1, \ldots, N$.

The proof can be deduced directly from the two inequalities showed at (3.8).

Proposition 4.2. *The variational problem*

$$(\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}^k, \boldsymbol{v}_h^k - \boldsymbol{u}_h^k) + < \mu_F \left|\boldsymbol{\sigma}_N^k\right|, \boldsymbol{\mu}_h^k - \boldsymbol{\lambda}_h^k >_{\Gamma_c} \leq 0\ (\forall\ \boldsymbol{\mu}_h^k \in \Lambda_h, \boldsymbol{v}_h^k \in V_h) \quad (4.1)$$

is equivalent to

$$\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k = 0, \quad \boldsymbol{\lambda}_h^k = P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k) \quad (4.2)$$

where P_Λ *is the projection operator from* $L^2(\Gamma_c)$ *to* Λ_h *is the convex subset of* $H^{1/2}(\Gamma_c), \rho > 0$, $\boldsymbol{s}_h^k = \mu_F \left|\boldsymbol{\sigma}_N^k\right|, k = 0, 1, \ldots, N$.

Proof. The inequation (4.1) is equivalent to

$$(\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}^k, \boldsymbol{u}_h^k - \boldsymbol{v}_h^k) + \boldsymbol{s}_h^k, \boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k >_{\Gamma_c} \geq 0\ (\forall\ \boldsymbol{\mu}_h^k \in \Lambda_h, \boldsymbol{v}_h^k \in V_h). \quad (4.3)$$

Multiplying the inequation (4.3) by ρ and adding $(\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \boldsymbol{u}_h^k)$ to the two sides of (4.3), we have

$$(\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \rho(\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}^k) + \boldsymbol{u}_h^k) + < \boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k, \rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k >_{\Gamma_c} \geq$$
$$\geq (\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \boldsymbol{u}_h^k) + < \boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k, \boldsymbol{\lambda}_h^k >_{\Gamma_c}. \quad (4.4)$$

But P_Λ is a projector operator,

$$(\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \rho(\operatorname{div} \boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k) + \boldsymbol{u}_h^k) + (\boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k, P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_{\boldsymbol{h}}^{\boldsymbol{k}}))_{0,\Gamma_c} \geq$$
$$\geq (\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \boldsymbol{u}_h^k) + < \boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k, \boldsymbol{\lambda}_h^k >_{\Gamma_c}.$$

Hence

$$(\boldsymbol{u}_h^k - \boldsymbol{v}_h^k, \rho(\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k)) + (\boldsymbol{\lambda}_h^k - \boldsymbol{\mu}_h^k, P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k) - \boldsymbol{\lambda}_h^k)_{0,\Gamma_c} \geq 0 \; . \tag{4.5}$$

Because V_h and Λ_h are convex sets, we can put $(0 < \alpha < 1)$:

$$\left. \begin{array}{lcl} \boldsymbol{v}_h^k & = & (1-\alpha)\boldsymbol{u}_h^k + \alpha(\rho(\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k) + \boldsymbol{u}_h^k) \\ \boldsymbol{\mu}_h^k & = & (1-\alpha)\boldsymbol{\lambda}_h^k + \alpha P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k) \end{array} \right\} . \tag{4.6}$$

Substituting (4.6) in (4.5) yields

$$\alpha(-\rho(\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k), \rho(\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k)) + \alpha(\boldsymbol{\lambda}_h^k - P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k), P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k) - \boldsymbol{\lambda}_h^k)_{0,\Gamma_c} \geq 0 \; ,$$

that is equivalent with

$$\alpha \|\rho(\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k)\|_{0,\Omega}^2 + \alpha \|\boldsymbol{\lambda}_h^k - P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k)\|_{0,\Gamma_c}^2 \leq 0 \; (0 < \alpha < 1, \; \rho > 0) \; ,$$

so we obtain

$$\operatorname{div}\boldsymbol{\sigma}_h^k + \boldsymbol{f}_h^k = 0 \quad \text{and} \quad \boldsymbol{\lambda}_h^k = P_\Lambda(\rho \boldsymbol{s}_h^k + \boldsymbol{\lambda}_h^k), \; \rho > 0, \; k = 0, 1, \ldots, N.$$

From this results we can define the following Uzawa algorithm type:

a) Given $\boldsymbol{u}_h^{nk} \in V_h, \boldsymbol{\lambda}_h^{nk} \in \Lambda_h$, we can define $\boldsymbol{\sigma}_h^{nk} \in S_t^h$ such that

$$\begin{aligned} J_0(\boldsymbol{\sigma}_h^{nk}) - (\operatorname{div}\boldsymbol{\sigma}_h^{nk}, \boldsymbol{u}_h^{nk}) - < \boldsymbol{s}_h^{nk}, \boldsymbol{\lambda}_h^{nk} >_{\Gamma_c} \leq \\ \leq J_0(\boldsymbol{\tau}_h^{nk}) - (\operatorname{div}\boldsymbol{\tau}_h^{nk}, \boldsymbol{u}_h^{nk}) + < \boldsymbol{t}_h^{nk}, \boldsymbol{\lambda}_h^{nk} >_{\Gamma_c}, \; \forall \, \boldsymbol{\tau}_h^{nk} \in S_0^h; \end{aligned} \tag{4.7}$$

b) Find $\boldsymbol{u}_h^{(n+1)k}$ and $\boldsymbol{\lambda}_h^{(n+1)k} = \left| \boldsymbol{v}_{Th}^{(n+1)k} \right|$ by using the following iterative method:

$$\boldsymbol{u}_h^{(n+1)k} = \boldsymbol{u}_h^{nk} + \rho_n(\operatorname{div}\boldsymbol{\sigma}_h^{nk} + \boldsymbol{f}^k) \tag{4.8}$$

$$\boldsymbol{\lambda}_n^{(n+1)k} = P_\Lambda(\boldsymbol{\rho}_n \boldsymbol{s}_h^{nk} + \boldsymbol{\lambda}_h^{nk}), \tag{4.9}$$

when $\rho_n > 0$ is chosen properly, $k = 0, 1, \ldots, N$. □

We define the following bounded linear operator: $g_\tau : S_h \to V \times L^2(\Gamma_c)$ by

$$g_\tau(\boldsymbol{v}, \boldsymbol{\mu}) = (\operatorname{div}\boldsymbol{\tau}, \boldsymbol{\nu}) + < \boldsymbol{s}, \boldsymbol{\mu} >_{\Gamma_c}, \quad s = \mu_F |\boldsymbol{\sigma}_N(\boldsymbol{v})|, \; \boldsymbol{\mu} = |\boldsymbol{v}_T| \, .$$

Proposition 4.3. *The operator $g_\tau : S_h \to V \times L^2(\Gamma_c)$ is Lipschitz continuous, i.e. there exists a constant $c > 0$, such that*

$$\|g_\tau(\boldsymbol{\tau}_1) - g_\tau(\boldsymbol{\tau}_2)\|_{V \times L^2(\Gamma_c} \leq c \|\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2\|_s, \quad \forall \, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in S_h,$$

where $\| \cdot \|_{V \times L^2(\Gamma_c)}$ denotes the norm of product space $V \times L^2(\Gamma_c)$.

Proof is obtained from definition of g_r and from (3.7).

Theorem 4.4. *There exists the constant α_0 and α_1, with $0 < \alpha_0 \leq \rho_n \leq \alpha_1$, such that, the Uzawa type algorithm a)-b), is convergent in sense that $\boldsymbol{\sigma}_h^{nk} \to \boldsymbol{\sigma}_h^k$ strongly in S.*

Proof. We denote $\boldsymbol{r}_1^{nk} = \boldsymbol{u}_h^{nk} - \boldsymbol{u}_h^k$, $\boldsymbol{r}_2^{nk} = \boldsymbol{\lambda}_h^{nk} - \boldsymbol{\lambda}_h^k$, and from (4.7)-(4.9) we can deduce:

$$\begin{aligned}
&\left\|\boldsymbol{r}_1^{(n+1)k}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{(n+1)k}\right\|_{0,\Gamma_c}^2 = \left\|\boldsymbol{u}_h^{(n+1)k} - \boldsymbol{u}_h^k\right\|_{0,\Omega}^2 + \left\|\boldsymbol{\lambda}_h^{(n+1)k} - \boldsymbol{\lambda}_h^k\right\|_{0,\Gamma_c}^2 = \\
&= \left\|\boldsymbol{u}_h^{nk} + \rho_n(\operatorname{div}\boldsymbol{\sigma}_h^{nk} + \boldsymbol{f}^k) - \boldsymbol{u}_h^{nk} - \rho_n(\operatorname{div}\boldsymbol{\sigma}_h^{nk} + \boldsymbol{f}^k)\right\|_{0,\Omega}^2 + \\
&+ \left\|P_\Lambda(\rho_n \boldsymbol{s}_h^{nk} + \boldsymbol{\lambda}_h^{nk}) - P_\Lambda(\rho_n \boldsymbol{s}_h^{nk} + \boldsymbol{\lambda}_h^{nk}\right\|_{0,\Gamma_c}^2 \le \\
&\le \left\|\boldsymbol{r}_1^{nk} + \rho_n \operatorname{div}(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k)\right\|_{0,\Omega}^2 + \left\|\rho_n(\boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k) + (\boldsymbol{\lambda}_h^{nk} - \boldsymbol{\lambda}_h^k)\right\|_{0,\Gamma_c}^2 = \\
&= \left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + 2\rho_n(\boldsymbol{r}_1^{nk}, \operatorname{div}(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k)) + \rho_n^2\left\|\operatorname{div}(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k)\right\|_{0,\Omega}^2 + \\
&+ \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2 + 2\rho_n(\boldsymbol{r}_2^{nk}, \boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k)_{0,\Gamma_c} + \boldsymbol{\rho}_n^2\left\|\boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k\right\|_{0,\Omega}^2 = \\
&= \left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2 + 2\rho_n(\boldsymbol{r}_1^{nk},\ \operatorname{div}(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k)) + (\boldsymbol{r}_2^{nk},\ (\boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k))_{0,\Gamma_c} + \\
&+ \rho_n^2\left\|\operatorname{div}(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k)\right\|_{0,\Omega}^2 + \left\|\boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k\right\|_{0,\Gamma_c}^2 . \qquad (4.10)
\end{aligned}$$

With the Proposition 4.3 and (4.10) can be regarded as positive algebraic equations with degree two in ρ, we get

$$a(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k, \boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k) + (\boldsymbol{r}_1^{nk}, \operatorname{div}(\boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k)) + < \boldsymbol{r}_2^{nk}, \boldsymbol{s}_h^{nk} - \boldsymbol{s}_h^k >_{\Gamma_c} \le 0\,,$$

where a is a linear symmetric form $a : S \times S \to \mathbb{R}$, which with (4.10) implying:

$$\begin{aligned}
&\left\|\mathbf{r}_1^{(n+1)k}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{(n+1)k}\right\|_{0,\Gamma_c}^2 \le \left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2 - \\
&-2\rho_n a(\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_{\boldsymbol{h}}^{\boldsymbol{k}}) + 2\rho_n^2\left\|\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k\right\|_S^2 \le \\
&\le \left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2 - (2\rho_n - \rho_n^2)\left\|\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k\right\|_S^2 .
\end{aligned}$$

For this inequation, we suppose $2\rho_n - 2\rho_n^2 \ge \beta > 0$, and we choose $\alpha_0 = \dfrac{1-\sqrt{1-2\beta}}{2}$, $\alpha_1 = \dfrac{1+\sqrt{1-2\beta}}{2}$ such that for $\rho_n \in [\alpha_0, \alpha_1]$, then we have:

$$\left\|\boldsymbol{r}_1^{(n+1)k}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{(n+1)k}\right\|_{0,\Gamma_c}^2 + \beta\left\|\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k\right\|_S^2 \le \left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2 \qquad (4.11)$$

From (4.11) results that the sequence $\left(\left\|\boldsymbol{r}_1^{nk}\right\|_{0,\Omega}^2 + \left\|\boldsymbol{r}_2^{nk}\right\|_{0,\Gamma_c}^2\right)_n$ is decreasing and has a finite limit, so that $\beta\left\|\boldsymbol{\sigma}_h^{nk} - \boldsymbol{\sigma}_h^k\right\|_S^2 \to 0$ for $n \to \infty$, and Theorem 4.4 is proved. □

The solution $\boldsymbol{\sigma}_h^k$ of (3.8) is a fixed point of function $M_h : S_h \to S_h$, so that $\boldsymbol{\sigma}_h^k$ is the limit of a sequence $(\boldsymbol{\sigma}_h^{nk})_n$, defined by $\boldsymbol{\sigma}_h^{nk} = M_h \boldsymbol{\sigma}_h^{(n-1)k}$, (see [13]).

Theorem 4.5. *In the conditions of Theorem 4.4, if $\alpha_0 < \rho_n < \alpha_1$ is true (α_1 are chosen according to Theorem 4.4, then for the sequences $\left\{\boldsymbol{u}_h^{nk}\right\}_n$, $\left\{\boldsymbol{\lambda}_h^{nk}\right\}_n$ defined by (4.8) – (4.9) we have:*

a) $\lim\limits_{n\to\infty}\left\|\boldsymbol{u}_h^{(n+1)k} - \boldsymbol{u}_h^k\right\|_{0,\Omega} = 0,\ \lim\limits_{n\to\infty}\left\|\boldsymbol{\lambda}_h^{(n+1)k} - \boldsymbol{\lambda}_h^k\right\|_{0,\Gamma_c} = 0;$

b) $\left\{\boldsymbol{u}_h^{nk}, \boldsymbol{\lambda}_h^{nk}\right\}_n \to \{u_h, \lambda_h\}$ *weakly in* $V_h \times \Lambda_h$ *where* $\left\{\boldsymbol{u}_h^k, \boldsymbol{\lambda}_h^k\right\}$ *is such that* $\boldsymbol{\sigma}_h^k, \boldsymbol{u}_h^k, \boldsymbol{\lambda}_h^k$ *is a saddle-point of* $L(\boldsymbol{\tau}_h^k, \boldsymbol{v}_h^k, \boldsymbol{\mu}_h^k)$ *on* $S_t^h \times V_h \times \Lambda_h$.

The proof is similar to that of Theorem 4.4, see [3].

5. Conclusions

We have analyzed, with Uzawa type algorithm of dual mixed variational formulation of the reduced version of a contact problem with friction in which it is assumed that the normal contact component of stress vector is known. For a more general contact problem, the existence solution is proved, but in very special cases.

References

[1] Brezzi, F., Fortin., M., *Mixed and Hybrid Finite Element Method*, New York, Springer-Verlag, 1991.

[2] Ciarlet, P.G., *The Finite Element Method for Elliptic Problems*, Amsterdam, New York, Oxford, North Holland, 1978.

[3] Glowinski,R., Lions, J., Tremolieres, R., *Numerical Analysis of Variational Inequalities*, Amsterdam, New York, Oxford, North Holland, 1981.

[4] Kiruchi, N., Oden, J.T., *Contact problem in Elasticity: A Study of Variational Inequalities and Finite Element Method*, SIAM Philadelphia, 1988.

[5] Nečas, J., Jarušek, J., Haslinger, J., *On the solution of the variational inequality to the Signorini problem with small friction*, Boll.Un.Mat.Ital., B(6), 17(1980), pp.796-881.

[6] Rocca, R., Cocou, M., *Numerical analysis of quasi-static unilateral contact problems with local friction*, SIAM J. NUMER. ANAL., vol.39, No.4, pp.1324-1342, 2001.

[7] Haslinger, J., *Approximation of the Signorini problem with friction obeying the Coulomb law*, Math. Methods Appl. Sci. 5(1983), pp.422-437.

[8] Rocca, R., Cocou, M., *Existence and approximation of a solution to quasi-static Signorini problem with local friction*, Internat J. Engrg.Sci. 39(2001), pp. 1253-1258.

[9] Pop, N., *On inexact Uzawa methods for saddle point problems arising from contact problem*, Bul.Ştiinţ. Univ. Baia Mare, Ser.B, Matematică-Informatică, Vol.XV, Nr.1-2(1999), pp.45-54.

[10] Wohlmuth, I.B., Krause, H.R., *A multigrid method based on the unconstrained product space for mortar finite element discretetizations*, SIAM J.NUMER. ANAL. Vol.39, No.1, pp.192-213(2001).

[11] Wang, G., Wang, L., *Uzawa type algorithm based on dual mixed variational formulation*, Applied Mathematics and Mechanics, Vol.23, No.7, (2002), pp.765-772.

[12] Anderson, L.E., *Approximation of the Signorini problem with friction obeying the Coulomb law*, Math. Methods Appl. Sci. 5(1983), pp.422-437.

[13] Cocou, M., *Existence of solutions of Signorini problems with friction*, Int. J. Engrg. Sci., 22(5), 567-575, 1984.

[14] Klarbring, A., Mikelic, A., Shillor, M., *Global existence result for the quasistatic frictional contact problem with normal compliance*, In "Unilateral Problems in Structural Analysis IV (Capri 1989), 85-111, Birkhäuser, 1991

North University of Baia Mare, Faculty of Science,
Department of Mathematics and Computer Science,
Victoriei 76, 4800 Baia Mare, Romania
E-mail address: nic_pop2002@yahoo.com

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

THE NUMERICAL ANALYSIS OF SOME SYSTEMS OF DIFFERENTIAL EQUATIONS ARISING FROM MOLECULAR DYNAMICS

ADRIAN REVNIC, RONALD H. W. HOPPE, AND GUSTAVO SIBONA

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. Molecular dynamics (MD) has become an important tool in the study of molecules/atoms interaction. In the classical MD, the motion of the atoms is described by the Newton equations, the quantum effects being either neglected or incorporated implicitly in the potential function. In this paper we study the application of MD in the formation of thin films, by an appropriate choice of interacting potentials of Tersof type. The numerical integration is performed by a (parallel) version of the Störmer–Verlet scheme using a particle-in-cell method and nearest neighbor concept. Higher order methods based on composition are also considered.

1. Introduction

Molecular dynamics is a modern computational technique used in condensed matter physics, materials science, chemistry, and other fields, consisting of following the temporal evolution of a system of N particles, interacting with each other by means of a certain law. In classical molecular dynamics, the evolution is based on the Newton's equations of motion and the forces are obtained as gradients of a certain potential which is function of all the particle coordinates.

The MD simulation of coating processes must give both insights into the dynamics of the absorption and growth procedures at the surface layer, and information about the structure of the developing crystal layers. In the first case, one is thus interested in the short time dynamics of the atomic reciprocal effects, while in the second case in the temporal average values of the atomic positions and on variables that depend on it.

Received by the editors: 12.03.2003.
2000 *Mathematics Subject Classification.* 65C20, 70F10, 81V99.
Key words and phrases. Molecular dynamics, Tersoff potentials, Störmer-Verlet method, composition methods.

2. Modeling and simulation by molecular dynamics

In molecular dynamics, the behavior of a system of N particles is modeled by Newton's equations of motion

$$\begin{aligned} m_i \, \dot{\mathbf{v}}_i &= - \nabla_i V(\mathbf{r}_1, ..., \mathbf{r}_N) \ , \\ \dot{\mathbf{r}}_i &= \mathbf{v}_i \ , \ 1 \le i \le N \ , \end{aligned} \tag{1}$$

where $\mathbf{r}_i, \mathbf{v}_i$, and m_i stand for the position vector, the velocity vector, and the mass of the i-th particle, and $V(\mathbf{r}_1, ..., \mathbf{r}_N)$ refers to the potential energy of the system as a function of the position vectors of all particles. The negative gradient $-\nabla V = -(\frac{\partial V}{\partial x_i}, \frac{\partial V}{\partial y_i}, \frac{\partial V}{\partial z_i})^T$ corresponds to the force $\mathbf{F}_i$ acting on the i-th particle.

We note that (1) represents a Hamiltonian system with respect to the Hamiltonian

$$H(\mathbf{r}_1, ..., \mathbf{r}_N, \mathbf{v}_1, ..., \mathbf{v}_N) \ = \ \frac{1}{2} \sum_{i=1}^N m_i \ \mathbf{v}_i^2 \ + \ V(\mathbf{r}_1, ..., \mathbf{r}_N) \tag{2}$$

which describes the total energy of the system.

The physics of the system is completely determined by the function V which comprises all exterior and interatomar potentials.

For the particular coating processes, suitable choices are Brenner- and Tersoff-type potentials (cf., e.g., [6, 7]) which are of the form

$$V \ = \ \frac{1}{2} \sum_{i \neq j} V_{ij} \ = \ \frac{1}{2} \sum_{i \neq j} f_C(r_{ij}) \ [\ f_R(r_{ij}) \ + \ b_{ij} f_A(r_{ij}) \ ] \ . \tag{3}$$

Here, $r_{ij} := \mid \mathbf{r}_i - \mathbf{r}_j \mid \ , \ 1 \le i \neq j \le N$, and $f_C(\cdot)$ is a cut-off function

$$f_C(r) \ := \ \begin{cases} 1 \ , & r < R - D \\ \frac{1}{2} \ - \ \frac{1}{2} \sin \left(\frac{\pi}{2} \frac{r-R}{D} \right) \ , & R - D \le r \le R + D \\ 0 \ , & R + D < r \end{cases} \ , \tag{4}$$

whereas $f_A(\cdot), f_R(\cdot)$ denote attractive and repulsive potentials, respectively,

$$f_A(R) \ := \ - \ A \ \exp \ (- \ \lambda_1 \ r) \quad , \quad f_R(R) \ := \ B \ \exp \ (- \ \lambda_2 \ r) \ . \tag{5}$$

Moreover, the bond-order parameter b_{ij} is chosen as a monotonically decreasing function of the number of neighbors of the atoms i and j according to the bond-order-concept which states that the more neighbors an atom has, the weaker the bond to each neighbor:

$$b_{ij} \ := \ (\ 1 \ + \ \beta^m \ \xi_{ij}^m \ )^{-\frac{1}{2m}} \ . \tag{6}$$

Here, ξ_{ij} is the effective coordination number given by

$$\begin{aligned} \xi_{ij} \ &:= \ \sum_{k \neq i,j} f_C(r_{ik}) \ g(\theta_{ijk}) \ \exp \ (\lambda_3^3 \ (r_{ij} - r_{ik})^3) \ , \\ g(\theta) \ &:= \ 1 \ + \ \frac{c^2}{d^2} \ - \ \frac{c^2}{d^2 + (h - \cos \theta)^2} \ , \end{aligned} \tag{7}$$

where θ_{ijk} represents the bond angle formed by the bond between atom i and atom j and the bond between atom j and atom k.

Note that the weighting factor exp $(\lambda_3^3(r_{ij} - r_{ik})^3)$ takes into account the relative distance between different neighbors: a weaker bond (longer distance r_{ik}) will be considerably more weakened by a stronger bond (shorter distance r_{ij}) than vice versa. Furthermore, the function g, depending on the bond angle, is another weighting factor which is chosen such that it stabilizes the crystallographic structure with regard to shear forces.

Note that the weighting factors do not occur in the classical Tersoff potentials but have been introduced to improve the quality of the model for the specific BN-system under consideration.

The parameters A, B, c, d, h, m, β, and $\lambda_i, 1 \le i \le 3$, in (5),(6),(7) are fitted both by using experimentally obtained data such as elasticity modules and lattice specific constants as well as with regard to structural energies (e.g., surface and defect energies) and interatomar forces computed by means of ab-initio quantum mechanical calculations.

Ab-initio methods consider every atom as a many particle system consisting of the atomic nucleus and the surrounding electrons. The many particle system is then solved by self consistent pseudopotential calculations based on the density functional theory. However, such computations require an enormous amount of work and therefore, they have been carried out for less particles than are used in the molecular dynamics approach.

For the numerical integration of the Hamiltonian system (1), symplectic integrators are well suited due to conservation of energy [1]. In fact, a backward analysis [2] shows that the total energy is well preserved for exponentially long time horizons $T = \Delta t \exp(C/2\Delta t)$:

$$\begin{aligned} & H(\mathbf{r}_1(k\Delta t), ..., \mathbf{r}_N(k\Delta t), \mathbf{v}_1(k\Delta t), ..., \mathbf{v}_N(k\Delta t)) = \\ = \ & H(\mathbf{r}_1(0), ..., \mathbf{r}_N(0), \mathbf{v}_1(0), ..., \mathbf{v}_N(0)) + O((\Delta t)^p) \quad , \quad k\Delta t < T \ , \end{aligned}$$

where Δt is the time stepsize and p refers to the order of consistency of the integrator.

We have used a standard symplectic integrator of order $p = 2$, commonly used in molecular dynamics, namely the Störmer–Verlet scheme

$$\mathbf{r}_i(t + \Delta t) - 2\mathbf{r}_i(t) + \mathbf{r}_i(t - \Delta t) = \Delta t^2 \frac{1}{m_i} \mathbf{F}_i(t) \tag{8}$$

which is in fact the most natural discretization of the Newton's equations $m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i$.

The two-term recursion (1.8) can now be easily modified to the classical formulation of the Störmer–Verlet schema in molecular dynamics.

$$\begin{aligned} \mathbf{r}_i(t + \Delta t) &= \mathbf{r}_i(t) + \Delta t \ \mathbf{v}_i(t) + \frac{1}{2} \frac{(\Delta t)^2}{m_i} \mathbf{F}_i(t) \ , \qquad (9) \\ \mathbf{v}_i(t + \Delta t) &= \mathbf{v}_i(t) + \frac{1}{2} \frac{\Delta t}{m_i} (\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)) \quad , \quad 1 \le i \le N \ . \end{aligned}$$

Thereby, a time loop of the Störmer–Verlet algorithm implementation looks as shown in Figure 1.

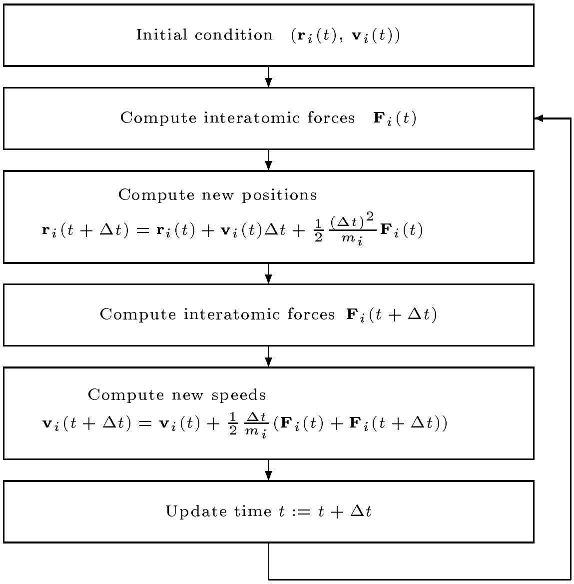


FIGURE 1. Time loop of the Störmer–Verlet algorithm

The Störmer–Verlet method admits an interesting one-step formulation, which is usefull for numerical computations. Introducing the velocity approximation at the midpoint

$$\mathbf{v}_i(t+\frac{\Delta t}{2}) := \mathbf{v}_i(t) + \frac{\Delta t}{2}\frac{1}{m_i}\mathbf{F}_i(t)$$

we get

$$\begin{aligned} \mathbf{v}_i(t+\frac{\Delta t}{2}) &= \mathbf{v}_i(t) + \frac{\Delta t}{2}\frac{1}{m_i}\mathbf{F}_i(t), \\ \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \Delta t\mathbf{v}_i(t+\frac{\Delta t}{2}), \\ \mathbf{v}_i(t+\Delta t) &= \mathbf{v}_i(t+\frac{\Delta t}{2}) + \frac{\Delta t}{2}\frac{1}{m_i}\mathbf{F}_i(t+\Delta t), \end{aligned} \tag{10}$$

which is an explicit one-step numerical method

$$\Phi_{\Delta t}^{SV} : (\mathbf{r}_i(t), \mathbf{v}_i(t)) \rightarrow (\mathbf{r}_i(t+\Delta t), \mathbf{v}_i(t+\Delta t)) \tag{11}$$

It is interesting to notice that for the implementation of the Störmer–Verlet method, the one-step formulation (1.10) is numerically more stable than the two-term recursion (1.8).

A specific feature, to be dealt with in the following section, is that we have implemented the Verlet algorithm in a parallel setting.

3. *Cell-Partition* Method

The simulation volume is divided into ("*cells* "), congruent subsections filling all the space, in such a way that all neighbours j of a particle i within a distance $r_{ij} < r_c$ are in the same subsection as i or in one of the directly neighbouring subsections. In this way one can limit the computation of reciprocal forces to particles that are in the same and neighbouring subsections. In addition, cubic subsections of edge length r_c can be usually used, the adjustment for the simulation of solids with crystal structure can require other geometry. Figure 2 shows the partitioning for the two-dimensional case.

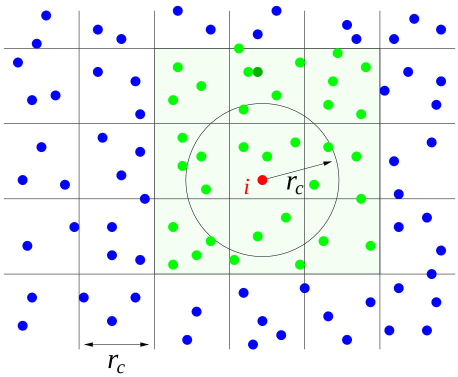


FIGURE 2. of the simulation space (2D)

Neighbour lists

From Figure 2 is evident that particles with a distance $r_{ij} > r_c$ can still be in neighbouring subsections, for which the reciprocal forces were then unnecessarily computed using the *Cell Partitioning* method alone. Therefore, additionally neighbour lists are provided, in which for each particle all next neighbours are seized. For the efficient looking for of the neighbours a appropriately *Cell Partitioning* procedure is used. To make worthwhile the production of the neighbour lists, we must use the same list over several time steps, i.e. we may extend the neighbourhood seize radius not only to neighbours with in a distance $r_{ij} < r_c$, but must increase it introducing a safety distance δ_s (see figure 3). The neighbour lists are valid only during a period $\delta_s/2v_{max}$, for a maximum particle speed v_{max}, because two particles can reduce their maximal distance with the speed of $2v_{max}$), i.e. it must be recalculated again after $[\delta_s/2v_{max}\Delta t]$ time steps.

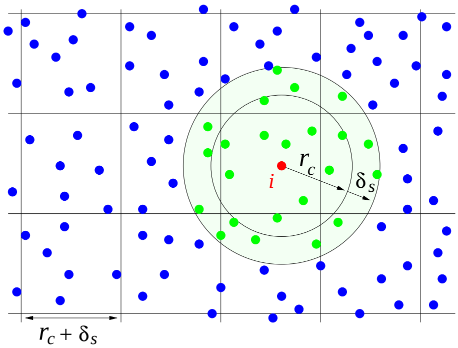


FIGURE 3. Range of the next neighbours (2D)

4. **Composition Methods**

An interesting procedure for constructing integration methods of higher order is by composition of simple methods. The aim is to increase the order of a simple underlying method, while preserving its desirable properties as symplecticity, symmetry and straightforward implementation. In the following we will use the ideas in [3] and [4], in obtaining a method of order 4 based by the composition of three Störmer–Verlet methods, mainly

$$\Phi_{\Delta t}^{com} := \Phi_{\alpha_1 \Delta t}^{SV} \circ \Phi_{\alpha_2 \Delta t}^{SV} \circ \Phi_{\alpha_3 \Delta t}^{SV} \tag{12}$$

As the Störmer–Verlet method is of order 2, this means $\Phi_{\Delta t}^{SV}$ satisfies (componentwise)

$$\Phi_{\Delta t}^{SV}(\mathbf{w}_0) = \Phi_{\Delta t}^{exact}(\mathbf{w}_0) + C(\mathbf{w}_0)(\Delta t)^3 + \mathcal{O}((\Delta t)^4), \tag{13}$$

where $\Phi_{\Delta t}^{exact}(\mathbf{w}_0)$ denotes the exact flux of the problem. Consequently,

$$\begin{aligned} \Phi_{\Delta t}^{com} &:= \Phi_{\alpha_1 \Delta t}^{SV} \circ \Phi_{\alpha_2 \Delta t}^{SV} \circ \Phi_{\alpha_3 \Delta t}^{SV} \\ &= \Phi_{(\alpha_1+\alpha_2+\alpha_3)\Delta t}^{exact}(\mathbf{w}_0) + \left(\alpha_1^3 + \alpha_2^3 + \alpha_3^3\right) C(\mathbf{w}_0)(\Delta t)^3 + \mathcal{O}((\Delta t)^4). \end{aligned} \tag{14}$$

so, by imposing

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 1 \\ \alpha_1^3 + \alpha_2^3 + \alpha_3^3 &= 0, \end{aligned} \tag{15}$$

the method $\Phi_{\Delta t}^{com}$ has at least order 3. As the Störmer–Verlet method is symmetric, i.e. $\Phi_{\Delta t}^{SV} = \left(\Phi_{-\Delta t}^{SV}\right)^{-1}$, the composition method will be symmetric if

$$\alpha_1 = \alpha_3. \tag{16}$$

But the order of a symmetric method is always an even number (see [5]), so the method must have at least order 4. Solving the system (15)–(16), we obtain a solution

$$\begin{aligned} \alpha_1 &= 1.3512072 \\ \alpha_2 &= -1.7024145 \\ \alpha_3 &= 1.3512072. \end{aligned} \tag{17}$$

The implementation of this method is straightforward. Once a Stöermer–Verlat (SV) routine is implemented, the composition method consists of calling this routine three times, with different time steps given using the scaling parameters given by (17). This means the method takes two positive intermediate steps $1.3512072 \times \Delta t$ and one negative intermediate step $-1.17024145 \times \Delta t$, as it can be seen from Figure 4.

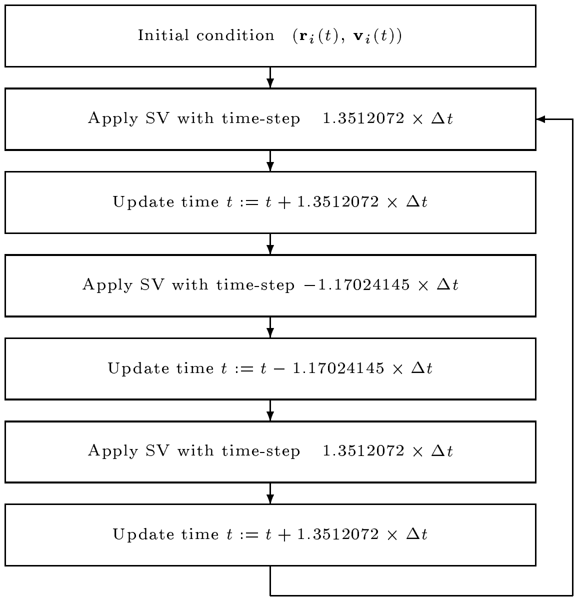


FIGURE 4. Implementation of the composition method

References

[1] Sanz Serna, J.M., Calvo, M.P, *Numerical Hamiltonian Problems*, Chapman & Hall, 1994.
[2] Hairer, E., *Backward analysis of numerical integrators and symplectic methods*, Annals of Numerical Mathematics,**1**, 1994, 107–132 1994.
[3] Yoshida, H., *Construction of higher order symplectic integrators*, Phys. Lett. A **150**, 262–268 (1990).
[4] McLachlan, R.I., Quispel, G.R.W., *Splitting methods*, Acta Numerica 2002, 341–434.

[5] Hairer, E., Lubich, Ch., Wanner, G., *Geometric Numerical Integration. Structure–Preserving Algorithms for ODEs*, Springer Verlag, 2002.
[6] Tersoff, J., *Chemical order in amorphous silicon carbide*, Phys. Rev. B, **49**, 16349 (1994).
[7] Brenner, D. W., *Empirical potential for hydrocarbond for use in simulating the chemical vapor deposition of diamond films*, Phys. Rev. B, **42**, 9458 (1990).

Institute of Mathematics, University of Augsburg,
Univesitätsstr. 14, D–86159 Augsburg
E-mail address: {revnic, hoppe, sibona}@math.uni-augsburg.de

MAXIMAL FIXED POINT STRUCTURES

IOAN A. RUS, SORIN MUREŞAN, AND EDITH MIKLOS

Dedicated to Professor Gheorghe Micula at his 60^{th} anniversary

Abstract. Examples, counterexamples and properties of the maximal fixed point structures are given.

1. Introduction

Let X be a nonempty set and $P(X) := \{Y \subseteq X \mid Y \neq \emptyset\}$. For $A, B \in P(X)$ we denote
$\mathbb{M}(A,B) := \{f : A \to B \mid F \text{ is an operator}\}$, $\mathbb{M}(A) := \mathbb{M}(A,A)$.

Definition 1.1. *(Rus* [39], [40], [41]*). A triple $(X, S(X), M)$ is a fixed point structure (briefly FPS) iff*
(i) $S(X) \subseteq P(X)$, $S(X) \neq \emptyset$;
(ii) M is an operator which attaches to each pair (A,B), $A, B \in P(X)$, a nonempty subset of $\mathbb{M}(A,B)$ such that, for any $Y \in P(X)$, if $Z \subseteq Y$, $Z \neq \emptyset$, $f(Z) \subseteq Z$, then $f|_Z \in M(Z)$, for all $f \in M(Y)$;
(iii) every $Y \in S(X)$ has the fixed point property (briefly FPP) with respect to $M(Y)$.

Definition 1.2. *(Rus* [43]*). The triple $(X, S(X), M)$ which satisfies (i) and (iii) in Definition 1.1 is called weak fixed point structure (briefly $WFPS$).*

Let $(X, S(X), M)$ be a FPS and $S_1(X) \subseteq P(X)$ such that $S_1(X) \subseteq S(X)$.

Definition 1.3. *(Rus [45]). The FPS $(X, S(X), M)$ is maximal in $S_1(X)$ iff we have*
$S(X) = \{A \in S_1(X) \mid f \in M(A) \text{ implies that } F_f \neq \emptyset\}$.

The aim of this paper is to give some examples of maximal FPS and to study the maximal FPS. Some open problems are formulated. Throughout the paper we follow terminologies and notations in [45] (see also [41], [42]).

Received by the editors: 14.10.2002.
2000 *Mathematics Subject Classification.* 47H10, 06A05, 54H25.
Key words and phrases. fixed point structure, maximal fixed point structure, ordered sets with fixed point property, topological spaces with fixed point property.

2. Examples and counterexamples

Example 2.1. *The trivial* FPS *is maximal in* $P(X)$*. In this case* X *is a nonempty set,* $S(X) := \{\{x\} \mid x \in X\}$ *and* $M(Y) := \mathbb{M}(Y)$*. We remark that if* $cardY \geq 2$ *there exists an operator* $f : Y \to Y$ *such that* $F_f = \emptyset$*.*

Example 2.2. *The Tarski* FPS *isn't maximal in* $P(X)$*. In this case* $(X, \leq)$ *is a partial ordered set,* $S(X) :=\{ Y \in P(X) \mid (Y, \leq)$ *is a complete lattice* $\}$ *and* $M(Y) := \{f : Y \to Y \mid f$ *is an increasing operator*$\}$*. To prove this assertion we consider* $X := \mathbb{R}^2$ *which is partial ordered by*

$$(x_1, x_2) \leq (y_1, y_2) \Leftrightarrow x_1 \leq y_1 \text{ and } x_2 \leq y_2.$$

We consider $Y = \{(1,1), (1,5), (2,4)\}$ *and we remark that* $(Y, \leq)$ *has the* FPP *with respect to increasing operators but* $(Y, \leq)$ *isn't a lattice.*

Remark 2.1. *For other results see:* [7], [30], [30], [34], [46].

Example 2.3. *The Tarski* FPS*,* $(X, S(X), M)$ *is maximal in* $S_1(X)$*, for all ordered set* $(X, \leq)$*, where* $S_1(X) := \{Y \in P(X) \mid (Y, \leq)$ *is a lattice*$\}$*.By a theorem of Davies (*[14], [34]*) it follows that* $Y \in S(X)$*.*

Example 2.4. *The Schauder* FPS *isn't, in general, maximal in* $P(X)$*. In this example* X *is a Banach space,* $S(X) := P_{cp,cv}(X)$ *and* $M(A, B) := C(A, B)$*. For* $Y \notin P_{cp,cv}(X)$ *with topological* FPP *see* [4], [18], [24], [35] *and* [37].

We have

Theorem 2.1. *The Schauder* FPS *is maximal in* $P_{b,cl,cv}(X)$*.*

3. FPS of contractions

Let (X, d) be a complete metric space, $S(X) := P_{cl}(X)$ and $M(Y) := \{f : Y \to Y \mid f \text{ is a contraction}\}$. By definition $(X, S(X), M)$ is the FPS of contractions. It is clear that the FPS of contractions is maximal iff

$$(Y \in P(X), f \in M(Y) \Rightarrow F_f \neq \emptyset) \Rightarrow Y \in P_{cl}(X).$$

This problem is studied by M-C. Anisiu and V. Anisiu [6]. The main results are the following

Theorem 3.1. *(*[6], [12]*) There exists a complete metric space and a nonclosed subset with* FPP *with respect to contractions.*

Theorem 3.2. *(*[6]*) Let* X *be a Banach space and* $Y \in P(X)$ *a convex set with* $IntY \neq \emptyset$*. If each contraction* $f : Y \to Y$ *has a fixed point, then* Y *is closed.*

Remark 3.1. *For other results see* [13], [22], [26], [28].

4. Some properties of the maximal FPS

Let $\mathcal{C}$ be the class of structured sets (the class of sets, the class of all partial ordered sets, the class of Banach spaces, the class of Hausdorff topological spaces,...). Let S be an operator which attaches to each $X \in \mathcal{C}$ a nonempty set set $S(X) \subseteq$

$P(X)$. By M we denote an operator which attaches to each pair (A, B), $A \in P(X)$, $B \in P(Y)$, $X, Y \in \mathcal{C}$, a subset $M(A, B) \subseteq \mathbb{M}(A, B)$.

We have

Lemma 4.1. *Let $X \in \mathcal{C}$ and $(X, S(X), M)$ be a maximal FPS. Let $A \in S(X)$ and $B \in P(A)$. If there exists a retraction $r \in M(A, B)$ of A onto B such that*

$$f \in M(B) \Rightarrow f \circ r \in A$$

then $B \in S(X)$.

Proof. Let $f \in M(B)$. Then $f \circ r \in M(A)$. ¿From $A \in S(X)$ it follows that $F_{f \circ r} \neq \emptyset$. Let $x^* \in F_{f \circ r}$. We have $f(r(x^*)) = x^*$. We remark that $x^* \in B$ and so we have $f(x^*) = x^*$. By the maximality of $(X, S(X), M)$ it follows that $B \in S(X)$. □

Lemma 4.2. *Let $X, Y \in \mathcal{C}$. Let $(X, S(X), M)$ and $(Y, S(Y), M)$ be two FPS. Let $A \in S(X)$ and $B \in S(Y)$. We suppose that:*
i) $(Y, S(Y), M)$ is a maximal FPS;
ii) there exists a bijection $\varphi \in M(A, B)$ such that $\varphi^{-1} \circ g \circ \varphi \in M(A)$, for all $g \in M(B)$.
Then $B \in S(Y)$.

Proof. Let $f \in M(B)$. Then, from ii), it follows that $F_{\varphi^{-1} \circ f \circ \varphi} \neq \emptyset$. Let $x^* \in F_{\varphi^{-1} \circ f \circ \varphi}$. We remark that $\varphi(x^*) \in F_f$. So, by the maximality of $(Y, S(Y), M)$, we have $B \in S(Y)$. □

5. Open problems

The above considerations give rise to the following open problems.

Problem 1 Characterize the partial ordered sets with FPP with respect to increasing operator.

References: K. Baclavski and A. Bjőrner [7], A.C. Davies [14], G. Markowsky [32], J.D. Mashburn [33], I.A Rus [34], L.E. Ward [46].

Problem 2. Characterize the metric space with the FPP with respect to isometric operators.

References: K. Goebel and W.A. Kirk [20], W.A. Kirk and B. Sims [28], A.T.-M. Lau [29].

Problem 3.Characterise the metric space with the FPP with respect to contractions.

References: R.P. Agarwal, M. Meehan and D.O'Regan [4], M.C. Anisiu and V. Anisiu [6], V. Conserva and S. Rizzo [13], T.K. Hu [22], J. Jachymski [26], W.A. Kirk and B. Sims [28], I.A.Rus [45], H. Cohen [12].

Problem 4. Characterize the topological spaces with FPP with respect to continuous operators.

References: V.N. Akis [5], R.F. Brown [9], E.H. Connel [12], J. Dugundji and A.

Granas [18], A.A. Fora [19], W. Hans [21], S.Y. Husseini [23], E. de Pascale, G. Trombetta and H. Weber [16], I.A. Rus [35], [37].

Problem 5. Characterize the categories (S. MacLane [31]) with the FPP (I.A.Rus [38]).

References: J. Adàmek and V. Koubek [1], J. Adàmek, V. Koubek and J. Reitermann [2], A. Bjōrner [8], J. Isbel and B. Mitchel [25], J. Lambek [30], I.A. Rus [34], [38] and [43].

References

[1] Adàmek, J., Koubek, V., *Remark on fixed points of functors*, Fundamental of computation theory, Proc. Intern. Conf. Poznan-Kornik, 1977.
[2] Adàmek, J., Koubek, V., Reitermann, J., *Embeddings into categories with fixed points in representations*, Czech. Math. J., 31(1981), 368-389.
[3] Adàmek, J., Merzenich, W., *Fixed points as equations and solutions*, Canad. J. Math., 36(1984), 495-519.
[4] Agarwal, R.P., Meehan, M., O'Regan, D., *Fixed point theory and applications*, Cambridge University Press, 2001.
[5] Akis, V.N., *Fixed point theorems and almost continuity*, Fund. Math., 121(1983), 43-52.
[6] Anisiu, M.-C., Anisiu, V., *On the closedeness of sets with the fixed point property for contractions*, Revue d'Analyse Num. Theorie de l'Approx., 26(1997), 13-17.
[7] Baclawski, K., Björner, A., *Fixed points in partially ordered sets*, Adv. Math., 31(1979), 263-287.
[8] Björner, A., *Reflexive domains and fixed points*, Acta Appl. Math. 4(1985), 99-100.
[9] Brown, R.F., *On some old problems of fixed point theory*, Roky Mounntain J. Math., 4(1974), 3-14.
[10] ________, *Retraction method in Nielsen fiexed point theory*, Pacific J. Math., 115(1984), 277-297;
[11] Cohen, H., *A fixed point problem for products of metric spaces*, Nieuw Archief voor Wiskunde, 21(1973), 59-63.
[12] Connel, E.H., *Properties of fixed point spaces*, Proc. AMS, 10(1959), 974-979.
[13] Conserva, V., Rizzo, S., *Fixed points and completeness*, Math. Japonica, 38(1993), 901-903.
[14] Davies, A.C., *A characterization of complete lattices*, Pacific J. Math., 5(1955), 311-319.
[15] de Pascale, E., *A finite dimensional reduction of the Schauder conjecture*, Comm. Math. Univ. Carolinae, 34(1993), 401-404.
[16] de Pascale, E., Trombetta, G., Weber, H., *Convexly totally bounded and strongly totally bounded sets*, Ann. Scuola. Nor. Sup. di Pisa, 20(1993), 341-355.
[17] Downing, D.J., Ray, W.O., *Some remarks on set-valued mappings*, Nonlinear Analysis, 5(1981), 1367-1377.
[18] Dugundji, J., Granas, A., *Fixed point theory*, P.W.N., Warsawa, 1982.
[19] Fora, A.A., *A fixed point theorem theorem for product spaces*, Pacific J. Math., 99(1982), 327-335.
[20] Goebel, K., Kirk, W.A., *Topics in metric fixed point theory*, Cambridge Univ. Press, 1990.
[21] Hans, W., *Compact convex sets in nonlocally convex linear spaces*, Schauder-Tyhonoff fixed point theorem, Topology, measure and fractals, Math. Res., 66, Akademie-Verlag, Berlin, 1992.
[22] Hu, T.K., *On a fixed point theorem for metric space*, Amer. Math. Monthly, 74(1967), 441-442.

[23] Husseini, S.Y., *The products of manifolds with the fixed point property*, Bull. Amer. Math. Soc., 81(1975), 441-442.
[24] Istratescu, V.I., *Fixed point theory*, D. Reidel, Dordrecht, 1981.
[25] Isbel, J., Mitchel, B., *Exact colimits and fixed points*, Trans. Amer. Math. Soc., 220(1976), 289-298.
[26] Jachimski, J., *A short proof of the converse to the contraction principle and some related results*, Topological Methods in Nonlinear Analysis, 15(2000), 179-186.
[27] Kirk, W.A., *Fixed point theorems for nonexpansive mappings*, Proc. AMS Symp. in Pure Math., 18(1970), 162-168.
[28] Kirk, W.A., Sims, B. (eds), *Handbook of metric fixed point theory*, Kluwer, Dordrecht, 2001.
[29] Lau, A.T.-M., *Sets with fixed point property for isometric mappings*, Proc. Amer. Math. Soc., 78(1980), 388-392.
[30] Lambek, J., *A fixpoint theorem for complete categories*, Math. Z., 103(1968), 151-161.
[31] MacLane, S., *Categories for the working mathematician*, Springer, New York, 1971.
[32] Markowsky, G., *Chain-complete posets with applications*, Algebra Univ., 6(1976), 53-68.
[33] Mashburn, J.D., *Three counterexamples concerning w-chain completness and fixed point properties*, Proc. Edinburg Math. Soc., 24(1981), 141-146.
[34] Rus, I.A., *Teoria punctului fix in structuri algebrice*, Univ. Babes-Bolyai, Cluj-Napoca, 1971.
[35] ________, *Teoria punctului fix in analiza functionala*, Univ. Babes-Bolyai, Cluj-Napoca, 1973.
[36] ________, *Metrical fixed point theorems*, Babes-Bolyai Univ., Cluj-Napoca, 1979.
[37] ________, *Principii si aplicatii ale teoriei punctului fix*, Ed. Dacia, Cluj-Napoca, 1979.
[38] ________, *Punct de vedere categorial in teoria punctului fix*, Seminar itinerant, Timisoara, 1980, 205-209.
[39] ________, *Fixed point structures*, Mathematica, 28(1986), 59-64.
[40] ________, *Further remarks on the fixed point structures*, Studia Univ. Babes-Bolyai, 31(1986), No.4, 41-43.
[41] ________, *Technique of the fixed point structures*, Bull. Appl. Math., 737(1991), 3-16.
[42] ________, *Technique of the fixed point structures for multivalued mappings*, Math. Japonica, 38(1993), 289-296.
[43] ________, *Some open problems in fixed point theory by means of fixed point structures*, Libertas Mathem., 14(1994), 65-84.
[44] ________, *Maximal fixed point structures*, Zilele Acad. Clujene, 18-23 Nov. 1996.
[45] ________, *Generalized contractions*, Cluj Univ. Press, 2001.
[46] Ward, L.E., *Completness in semilattices*, Can. J. Math., 9(1975), 578-582.

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

University of Oradea, Romania

Faculty of Mathematics and Computer Science,
Babeş-Bolyai University, Cluj-Napoca, Romania

STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **XLVIII**, Number 3, September 2003

WEAKLY SINGULAR VOLTERRA AND FREDHOLM-VOLTERRA INTEGRAL EQUATIONS

SZILÁRD ANDRÁS

Dedicated to Professor Gheorghe Micula at his 60^{th} *anniversary*

Abstract. Some existence and uniqueness theorems are established for weakly singular Volterra and Fredholm-Volterra integral equations in $C[a,b]$. Our method is based on fixed point theorems which are applied to the iterated operator and we apply the fiber Picard operator theorem to establish differentiability with respect to parameter. This method can be applied only for linear equations because otherwise we can't compute the iterated equation.

1. Introduction

The integral equation

$$u(x) = f(x) + \int_a^x K_1(x,s)u(s)ds, \tag{1}$$

with $f \in C[a,b]$ is weakly singular if there exists $L_1 \in C([a,b] \times [a,b])$ and $\alpha \in (0,1)$ such that $K_1(x,s) = \frac{L_1(x,s)}{|x-s|^\alpha}$ $\forall\, x,s \in [a,b]$ with $x \neq s$. In this case the kernel function K_1 is called *weakly singular*. The integral equation

$$u(x) = f(x) + \int_a^x K_1(x,s)u(s)ds + \int_a^b K_2(x,s)u(s)ds, \tag{2}$$

with $f \in C[a,b]$ is called weakly singular if at least one of the kernel functions K_1 and K_2 is weakly singular. In this paper we give an existence and uniqueness theorem for the equation 1 by using fixed point approach and we obtain the continuous dependence and differentiability with respect to a parameter. For equation 2 we study two different cases, in the first case K_1 is weakly singular and K_2 is continuous and in the second case both kernels are weakly singular. In both cases we obtain existence, uniqueness, continuous dependence and differentiability with respect to the parameter. We'll use the following theorems

Received by the editors: 14.04.2003.
2000 *Mathematics Subject Classification.* 47H10, 45G15.
Key words and phrases. Picard operators, fiber Picard operators, iterated equation.

Theorem 1.1. *If (X, d) is a complete metric space and $T : X \to X$ is an operator with*

$$d(Tu, Tv) \leq L \cdot d(u, v)\ \forall\ u, v \in X,\ \text{where } 0 < L < 1,$$

then

1. *T has an unique fixed point u^*.*
2. *The sequence $u_{n+1} = Tu_n$, $\forall\ n \in \mathbb{N}$ is convergent to u^* for all $u_0 \in X$.*
3. *$d(u_n, u^*) \leq \frac{L^n}{1-L} \cdot d(u_1, u_0)\ \forall\ n \in \mathbb{N}$.*

Theorem 1.2. *(Fiber Picard operator's) [8] Let (V, d) be a generalized metric space with $d(v_1, v_2) \in \mathbb{R}^p_+$, and (W, ρ) a complete generalized metric space with $\rho(w_1, w_2) \in \mathbb{R}^m_+$. Let $A : V \times W \to V \times W$ be a continuous operator. If we suppose that:*

a) *$A(v, w) = (B(v), C(v, w))$ for all $v \in V$ and $w \in W$;*
b) *the operator $B : V \to V$ is a weakly Picard operator;*
c) *there exists a matrix $Q \in M_m(\mathbb{R}_+)$ convergent to zero, such that the operator $C(v, \cdot) : W \to W$ is a Q contraction for all $v \in V$,*

then the operator A is a weakly Picard operator. Moreover, if B is a Picard operator, then the operator A is a Picard operator.

Theorem 1.3. *If X is a set and $T : X \to X$ is a function such that the equation $T^n(u) = u$ has an unique solution u^*, than u^* is the unique solution of the equation $Tu = u$*

Theorem 1.4. *If (X, d) is a generalized complete metric space and $T : X \to X$ is an operator such that T^k is a contraction, then the sequence $u_{n+1} = Tu_n\ \forall\ n \in \mathbb{N}$ is convergent to the unique fixed point of T^k.*

In order to apply these theorems to weakly singular integral equations we need the following properties of the weakly singular kernels.

Theorem 1.5. *If $K(x, s) = \frac{L(x,s)}{|x-s|^\alpha}$ with $0 < \alpha < 1$ and $L \in C([a, b] \times [a, b])$, then the operator $T : C[a, b] \to C[a, b]$,*

$$(Tu)(x) = \int_a^x K(x, s)u(s)ds$$

is well defined ($Tu \in C[a, b]$).

Proof. If $a \leq x < x' \leq b$ and $\delta_1 > 0$ we have

$$|(Tu)(x') - (Tu)(x)| \leq \int_a^{x-\delta_1} |K(x', s) - K(x, s)||u(s)|ds +$$

$$+ \int_{x-\delta_1}^{x'-\delta_1} |K(x', s)||u(s)|ds + \int_{x-\delta_1}^{x} |K(x, s)||u(s)|ds +$$

$$+ \int_{x'-\delta_1}^{x'} |K(x', s)||u(s)|ds.$$

$u \in C[a,b]$, implies that there exists $M = \max_{s\in[a,b]} |u(s)|$.

$K : [x - \frac{\delta_1}{2}, b] \times [a, x - \delta_1] \to \mathbb{R}$ is continuous so it is uniform continuous and $\forall\, \epsilon > 0$ there exists $\delta_2 > 0$ such that

$$|K(x', s) - K(x, s)| < \tfrac{\epsilon}{2M(b-a)} \text{ if } |x - x'| < \delta_2 \text{ and } s \le x - \delta_1.$$

This implies

$$|(Tu)(x') - (Tu)(x)| \le \frac{\epsilon}{2} + M \cdot \int_{x-\delta_1}^{x'-\delta_1} |K(x', s)|ds +$$

$$+ M \cdot \int_{x-\delta_1}^{x} |K(x, s)| + M \cdot \int_{x'-\delta_1}^{x'} |K(x', s)|ds,$$

if $|x - x'| < \delta_2$. On the other hand we have the following inequalities:

$$\int_{x-\delta_1}^{x'-\delta_1} |K(x', s)|ds \le P \cdot \int_{x-\delta_1}^{x'-\delta_1} \frac{ds}{(x'-s)^\alpha} = P \cdot \left(- \frac{(x'-s)^{1-\alpha}}{1-\alpha} \Big|_{x-\delta}^{x'-\delta_1} \right)$$

$$= \frac{P}{1-\alpha} \left((x' - x + \delta_1)^{1-\alpha} - \delta_1^{1-\alpha}) \right) \le \frac{P}{1-\alpha} \cdot (2(x'-x))^{1-\alpha} < \frac{\epsilon}{6M}$$

where $|x' - x| < \delta_3$, and $P = \max_{x,s\in[a,b]} |L(x, s)|$.

$$\int_{x-\delta_1}^{x} |K(x, s)| \le P \cdot \int_{x-\delta_1}^{x} \frac{ds}{(x-s)^\alpha} = \frac{P}{1-\alpha} \left(-(x-s)^{1-\alpha} \Big|_{x-\delta_1}^{x} \right) =$$

$$= \frac{P}{1-\alpha} \cdot \delta_1^{1-\alpha} < \frac{\epsilon}{6M}$$

for $\delta_1 \le \delta_4$.

$$\int_{x'-\delta_1}^{x'} |K(x', s)| \le \frac{P}{1-\alpha} \delta_1^{1-\alpha} < \frac{\epsilon}{6M}$$

for $\delta_1 \le \delta_3$. From these inequalities we deduce

$$|(Tu)(x') - (Tu)(x)| < \epsilon$$

if $|x - x'| < \min(\delta_1, \delta_2, \delta_3, \delta_4)$, so the operator T is well defined.

Theorem 1.6. *If K_1 or K_2 is weakly singular kernel, then the operator $T : C[a,b] \to C[a,b]$,*

$$(Tu)(x) = \int_a^x K_1(x, s)u(s)ds + \int_a^b K_2(x, s)u(s)ds$$

is well defined ($Tu \in C[a,b]$).

Proof. As in theorem 1.1 we can prove that the operator $T_2 : C[a,b] \to C[a,b]$,

$$(T_2u)(x) = \int_a^b K_2(x,s)u(s)ds$$

is well defined if K_2 is a weakly singular kernel, so T is well defined because it is the sum of two well defined operators .

Theorem 1.7. [6] *If K_1 and K_2 are weakly singular kernels and*

$$|K_1(x,s)| \le \frac{P_1}{|x-s|_1^{\alpha}}, \quad |K_2(x,s)| \le \frac{P_2}{|x-s|_2^{\alpha}},$$

where $P_1, P_2 \in \mathbb{R}$, $0 \le \alpha_1 < 1$, $0 \le \alpha_2 < 1$, then the function

$$K_3(x,s) = \int_a^b K_1(x,t)K_2(t,s)dt$$

satisfies the following conditions:

1. *If $\alpha_1 + \alpha_2 > 1$, the function $K_3(x,s)$ is a weakly singular kernel and*
$$|K_3(x,s)| < \frac{P_3}{|x-s|^{\alpha_1+\alpha_2-1}},$$
where $P_3 \in \mathbb{R}$.
2. *If $\alpha_1 + \alpha_2 = 1$, the function $K_3(x,s)$ is continuous for $x \ne s$ and*
$$|K_3(x,s)| < P_3 + P_4 \ln|x-s|,$$
where $P_3, P_4 \in \mathbb{R}$.
3. *If $\alpha_1 + \alpha_2 < 1$, the function $K_3(x,s)$ is continuous in $D = [a,b] \times [a,b]$.*

The proof can be found in [6] at pp. 374. An analogous theorem can be proved for the Volterra integral operator.

Theorem 1.8. *If the functions K_1 and K_2 are weakly singular kernels and*

$$|K_1(x,s)| \le \frac{P_1}{(x-s)^{\alpha_1}},$$

$$|K_2(x,s)| \le \frac{P_2}{(x-s)^{\alpha_2}},$$

for $x \ge s$, then the function

$$K_3(x,s) = \int_s^x K_1(x,t)K_2(t,s)dt$$

satisfies the following properties

1. *If $\alpha_1 + \alpha_2 > 1$, then K_3 is a weakly singular kernel and*
$$|K_3(x,s)| \le \frac{P_3}{(x-s)^{\alpha_1+\alpha_2-1}}.$$
2. *If $\alpha_1 + \alpha_2 = 1$, then K_3 is continuous and $|K_3(x,s)| \le P_4$.*

3. *If* $\alpha_1 + \alpha_2 < 1$, *then* K_3 *is continuous and*

$$|K_3(x,s)| \leq P_4 \cdot (x-s)^{1-\alpha_1-\alpha_2}.$$

2. The main results

2.1. The Volterra integral equation.

Theorem 2.1. *If* $K(x,s,\lambda) = \frac{L(x,s,\lambda)}{(x-s)^\alpha}$ *with* $L \in C([a,b] \times [a,b] \times [\lambda_1, \lambda_2])$ *and* $0 < \alpha < 1$, *then the equation*

$$u(x) = f(x) + \int_a^x K(x,s,\lambda)u(s)ds \tag{3}$$

with $f \in C[a,b]$ *and* $\lambda \in [\lambda_1, \lambda_2]$ *has a unique solution in* $C([a,b])$ *and this solution can be obtained by successive approximation. This solution depends continuously on* λ *and if* K *is continuously differentiable with respect to* λ, *the solution is also continuously differentiable with respect to* λ.

Proof. Due to theorem 1.5 the operator

$$T : C[a,b] \to C[a,b],\ (Tu)(x) = f(x) + \int_a^x K(x,s,\lambda)u(s)ds$$

is well defined. Theorem 1.8 implies that there exists $n \in \mathbb{N}^*$ such that the iterated kernel $K^{(n)}$ defined by the following relations $K^{(1)}(x,s,\lambda)=K(x,s,\lambda)$ and $K^{(j+1)}(x,s,\lambda) = \int_s^x K(x,t,\lambda) \cdot K^{(j)}(t,s,\lambda)dt\ \forall\ j \geq 1$ is continuous. But any solution of the equation 3 satisfies the iterated equation

$$u(x) = f(x) + \sum_{i=1}^{n-1} \int_a^x K^{(i)}(x,s,\lambda)f(s)ds + \int_a^x K^{(n)}(x,s,\lambda)u(s)ds. \tag{4}$$

We apply theorem 1.1 to the operator $\bar{T} : C[a,b] \to C[a,b]$

$$(\bar{T}u)(x) = f(x) + \sum_{i=1}^{n-1} \int_a^x K^{(i)}(x,s,\lambda)f(s)ds + \int_a^x K^{(n)}(x,s,\lambda)u(s)ds. \tag{5}$$

which has a continuous kernel, so by choosing a Bielecki metric in $C[a,b]$ $\bar{T}$ is a contraction. This implies that the equation $\bar{T}u = u$ has an unique solution u^* in $C[a,b]$. By the other hand from theorem 1.3 we obtain that u^* is the unique solution of the equation $Tu = u$, because $\bar{T} = T^{(n)}$. From theorem 1.4 we deduce that the sequence of successive approximation $u_{n+1} = Tu_n$ is convergent to u^* for every $u_0 \in C[a,b]$. This implies that equation 3 has an unique continuous solution, and this can be approximated by successive approximation. By applying the same technique to the equation

$$u(x,\lambda) = f(x) + \int_a^x K(x,s,\lambda)u(s,\lambda)ds \tag{6}$$

we obtain that u^* is the unique solution in $C([a,b]\times[\lambda_1,\lambda_2])$, so the solution is depending continuously on the parameter λ. To study the differentiability of the solution we apply theorem 1.2 with the following spaces and operators:

a) $V = C([a,b]\times[\lambda_1,\lambda_2])$ and $B = \bar{T}$

b) $W = C([a,b]\times[\lambda_1,\lambda_2])$ and

$$C(v,w)(x,\lambda) = g(x,\lambda) + \int_a^x K^{(n)}(x,s,\lambda)\cdot w(s,\lambda)ds + \int_a^x \frac{\partial K^{(n)}}{\partial \lambda}\cdot v(s,\lambda)ds$$

where $g(x,\lambda) = \sum_{i=1}^{n-1}\int_a^x \frac{\partial K^{(i)}(x,s)}{\partial\lambda} f(s)ds$

The operator $A=(B,C)$ satisfies the conditions of theorem 1.2 because in $C([a,b]\times[\lambda_1,\lambda_2])$ we use a Bielecki metric and $K^{(n)}$ is a continuous function. This implies the uniform convergence of the sequence $v_{n+1} = V(v_n)$ to the unique solution u^* of equation 6 and the uniform convergence of the sequence $w_{n+1} = C(v_n,w_n)$ to a function w^*. If we choose $v_0 \in C^1[a,b]\times[\lambda_1,\lambda_2]$ and $w_0 = \frac{\partial v_0}{\partial\lambda}$ due to the operator C (which was obtained by a formal differentiation of the operator B)we have $w_n = \frac{\partial v_n}{\partial \lambda}$ $\forall\ n\in\mathbb{N}$. The Weierstrass's theorem implies that w^* is continuous and $w^*(x,\lambda) = \frac{\partial u^*(x,\lambda)}{\partial\lambda}$. So the solution u^* is continuously differentiable with respect to the parameter λ.

Remark 2.1. *We can use a direct proof (without the iterated operators) if we use the following inequality:*

$$|Tu(x)-Tv(x)| \le \int_a^x \frac{\max\limits_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]}|L(x,s,\lambda)|}{|x-s|^\alpha}\cdot|u(s)-v(s)|ds \le$$

$$\le L^*||u-v||\cdot\int_a^x \frac{e^{\tau(s-a)}}{(x-s)^\alpha}ds \le \left(\int_a^x \frac{ds}{(x-s)^{\alpha p}}\right)^{\frac{1}{p}}\cdot\left(\int_a^x e^{\tau(s-a)q}ds\right)^{\frac{1}{q}} \le$$

$$\le \left(\frac{(b-a)^{1-\alpha\cdot p}}{1-\alpha\cdot p}\right)^{\frac{1}{p}}\cdot\frac{e^{\tau(x-a)}}{(\tau\cdot q)^{\frac{1}{q}}},$$

where $\alpha\cdot p<1$, $\frac{1}{p}+\frac{1}{q}=1$, $L^* = \max\limits_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]}|L(x,s,\lambda)|$ *and*

$$||u-v|| = \max_{x\in[a,b],\lambda\in[\lambda_1,\lambda_2]}|u(x,\lambda)-v(x,\lambda)|\cdot e^{-\tau(x-a)}.$$

So we can choose τ such that the operator T be a contraction with the corresponding Bielecki metric.

2.2. **The Fredholm-Volterra integral equation.**

Theorem 2.2. *For the equation*

$$u(x) = f(x) + \int_a^x K_1(x,s,\lambda)y(s)ds + \int_a^b K_2(x,s,\lambda)y(s)ds \tag{7}$$

with

$$L_1 = \max_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]} |K_1(x,s,\lambda)|$$

and

$$L_2 = \frac{2 \cdot \max\limits_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]} |L(x,s,\lambda)|}{1-\alpha} \cdot (b-a)^{1-\alpha}$$

where $K_1, L \in C([a,b]\times[a,b]\times[\lambda_1,\lambda_2])$ *and* K_2 *is a weakly singular kernel* ($K_2(x,s,\lambda) = \frac{L(x,s,\lambda)}{|x-s|^\alpha}$, $0<\alpha<1$) *the iterated kernels are*

$$K_1^{(n+1)}(x,s,\lambda) = \int_s^x K_1(x,t,\lambda)K_1^{(n)}(t,s,\lambda)dt + \int_a^b K_2(x,t,\lambda)K_1^{(n)}(x,t,\lambda)dt \quad (8)$$

and

$$K_2^{(n+1)}(x,s,\lambda) = \int_a^x K_1(x,t,\lambda)K_2^{(n)}(t,s,\lambda)dt + \int_a^b K_2(x,t,\lambda)K_2^{(n)}(x,t,\lambda)dt \quad (9)$$

and the resolvent kernels are

$$R_1(x,s,\lambda) = \sum_{j=1}^{\infty} K_1^{(j)}(x,s,\lambda), \quad (10)$$

$$R_2(x,s,\lambda) = \sum_{j=1}^{\infty} K_2^{(j)}(x,s,\lambda). \quad (11)$$

If L_1 *and* L_2 *satisfies condition a) or b), there exist an unique continuous solution to the equation 7, this solution depends continuously on* λ *and if the functions* K_1 *and* L *are continuously differentiable with respect to* λ*, then the solution is also continuously differentiable with respect to* λ*. The solution of the equation 7 can be represented in the form*

$$u(x) = f(x) + \int_a^x R_1(x,s,\lambda)f(s)ds + \int_a^b R_2(x,s,\lambda)f(s)ds.$$

The series (10) and (11) are convergent if L_1 *and* L_2 *satisfy the condition a) or b)*

a) $\frac{L_1}{2-L_2(b-a)} + \left(e^{\frac{L_1(b-a)}{2-L_2(b-a)}} - 2\right) L_1L_2(b-a) < 0$;

b) $\frac{1}{b-a}\ln\frac{1-L_2(b-a)}{(b-a)^2L_1L_2} + \left(\frac{1-L_2(b-a)}{(b-a)^2}L_1L_2 - 2\right)(b-a)L_1L_2 > 0$ *and*

$$\frac{1}{b-a}\ln\frac{1-L_2(b-a)}{(b-a)^2L_1L_2}(1-L_2(b-a)) +$$

$$+(b-a)L_1L_2\left(2 - \frac{1-L_2(b-a)}{(b-a)^2L_1L_2}\right) - L_1 > 0.$$

Proof. Due to theorem 1.2 we can apply the same reasoning as in [1] theorem 2.2.

Remark 2.2. *By applying the Fiber Picard operator theorem ([8]) as in* [1] *we can prove that the solution is differentiable with respect to the parameter* λ

Theorem 2.3. *If in the equation 7 both kernels are singular,* $K_1(x,s,\lambda) = \frac{L_1^*(x,s,\lambda)}{|x-s|^{\alpha_1}}$ *and* $K_2(x,s,\lambda) = \frac{L_2^*(x,s,\lambda)}{|x-s|^{\alpha_2}}$ *with* $L_1^*, L_2^* \in C([a,b]\times[a,b]\times[\lambda_1,\lambda_2])$, $0 < \alpha_1 < 1$, $0 < \alpha_2 < 1$ *and the numbers*

$$L_1 = \max_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]} |K_1^{(n)}(x,s,\lambda)| \tag{12}$$

and

$$L_2 = \max_{x,s\in[a,b],\lambda\in[\lambda_1,\lambda_2]} |K_2^{(n)}(x,s,\lambda)| \tag{13}$$

satisfies condition a) or b) from theorem 2.2 then equation 7 has an unique solution in $C[a,b]\times[\lambda_1,\lambda 2]$. *If in addition the functions* L_1^* *and* L_2^* *are continuously differentiable with respect to the parameter* λ, *the solution is also continuously differentiable with respect to* λ.

Proof. The iterated equation is

$$u(x) = f(x) + \sum_{j=1}^{n-1}\int_a^x K_1^{(j)}(x,s,\lambda)\cdot f(s)ds + \sum_{j=1}^{n-1}\int_a^b K_2^{(j)}(x,s,\lambda)\cdot f(s)ds +$$

$$+\int_a^x K_1^{(n)}(x,s,\lambda)u(s)ds + \int_a^b K_1^{(n)}(x,s,\lambda)u(s)ds$$

where the iterated kernels are defined by the relations 8 and 9. Due to theorem 1.5 and 1.6 the function

$$g_1(x,\lambda) = f(x) + \sum_{j=1}^{n-1}\int_a^x K_1^{(j)}(x,s,\lambda)\cdot f(s)ds + \sum_{j=1}^{n-1}\int_a^b K_2^{(j)}(x,s,\lambda)\cdot f(s)ds$$

is a continuous function. From theorem 1.7 and 1.8 we deduce that if $\max(\alpha_1,\alpha_2) < \frac{n-1}{n}$ and $\max\left(\frac{\alpha_2}{1-\alpha_1}, \frac{\alpha_1}{1-\alpha_2}\right) < n$ than $K_1^{(n)}$ and $K_2^{(n)}$ are continuous kernels so we can apply theorem 1.2 from [1] (because L_1 and L_2 satisfy a) or b)). From this theorem we deduce that the equation 7 has an unique solution u^* in $C[a,b]\times[\lambda_1,\lambda_2]$. This u^* is also the unique solution of the equation 7 because of theorem 1.3 and can be approximated by successive approximation due to theorem 1.4. To study the differentiability of the solution we apply theorem 1.2 again with the following spaces and operators:

a) $V = C([a,b]\times[\lambda_1,\lambda_2])$ and

$$(Bu)(x) = g_1(x,\lambda) + \int_a^x K_1^{(n)}(x,s,\lambda)u(s)ds + \int_a^b K_1^{(n)}(x,s,\lambda)u(s)ds$$

b) $W = C([a,b] \times [\lambda_1, \lambda_2])$ and

$$C(v,w)(x,\lambda) = \frac{\partial g_1(x,\lambda)}{\partial \lambda} + \int_a^x K_1^{(n)}(x,s,\lambda) \cdot w(s,\lambda)ds +$$

$$+ \int_a^x \frac{\partial K_1^{(n)}}{\partial \lambda} \cdot v(s,\lambda)ds + \int_a^b K_2^{(n)}(x,s,\lambda) \cdot w(s,\lambda)ds + \int_a^b \frac{\partial K_2^{(n)}}{\partial \lambda} \cdot v(s,\lambda)ds,$$

where $\frac{\partial g_1(x,\lambda)}{\partial \lambda} = \sum_{j=1}^{n-1} \int_a^x \frac{\partial K_1^{(j)}(x,s,\lambda)}{\partial \lambda} \cdot f(s)ds + \sum_{j=1}^{n-1} \int_a^b \frac{\partial K_2^{(j)}(x,s,\lambda)}{\partial \lambda} \cdot f(s)ds$

The operator $A = (B, C)$ satisfies the conditions of theorem 1.2 because in $C([a,b] \times [\lambda_1, \lambda_2])$ we use a Bielecki metric and $K^{(n)}$ is a continuous function. This implies the uniform convergence of the sequence $v_{n+1} = V(v_n)$ to the unique solution u^* of equation 7 and the uniform convergence of the sequence $w_{n+1} = C(v_n, w_n)$ to a function w^*. If we choose $v_0 \in C^1[a,b] \times [\lambda_1, \lambda_2]$ and $w_0 = \frac{\partial v_0}{\partial \lambda}$ due to the operator C (which was obtained by a formal differentiation of the operator B)we have $w_n = \frac{\partial v_n}{\partial \lambda}$ $\forall\, n \in \mathbb{N}$. The Weierstrass's theorem implies that w^* is continuous and $w^*(x,\lambda) = \frac{\partial u^*(x,\lambda)}{\partial \lambda}$. So the solution u^* is continuously differentiable with respect to the parameter λ.

Remark 2.3. 1. *Conditions 12 and 13 can be transferred inductively to the original kernels, but the conditions obtained are much more technical.*
2. *By using the same inequalities as in remark 2.1 we can avoid the use of the iterated kernels to obtain existence and uniqueness.*

References

[1] András, Sz., *Fredholm-Volterra equations*, 4th Joint Conference on Matematics and Computer Science, 2001, Felix, Romania.
[2] Anselone, P. M., *Nonlinear integral equations*, The University of Wisconsin Press, 1964.
[3] Chakrabarti, A., Vanden Berge, G., *Numerical solution of singular integral equations*, Elsevier Preprint, 2002.
[4] Estrada, R., Kanwal, R. P., *Singular integral equations*, Birkhauser, 2000.
[5] Fitt, A. D., Stefanidis, V., *Film cooling effectivness for subsonic slot injection into a cross flow.*
[6] Ionescu, D. V., *Ecuaţii differenţiale şi integrale*, ediţia a II-a, Ed. Didactică şi Pedagogică, Bucureşti.
[7] Pogorzelski, W., *Integral equations and their applications*, Pergamon Press, 1966.
[8] Rus, I. A., *Fiber Picard operators and applications*, Mathematica, Cluj Napoca, 1999.
[9] Srivastava, H. M., Buschman, R. G., *Theory and applications of convolution integral equations*, Klemer Academic Publishers, 1992.

Department of Applied Mathematics, Babeş-Bolyai University,
Cluj-Napoca, M. Kogălniceanu, No. 1, Romania
E-mail address: andrasz@math.ubbcluj.ro