ACADEMY OF ROMANIAN SCIENTISTS

A N N A L S

SERIES ON MATHEMATICS AND ITS APPLICATIONS

VOLUME 3 2011 NUMBER 1

ISSN 2066 – 6594

TOPICS:

- ORDINARY AND PARTIAL DIFFERENTIAL EQUATIONS
- OPTIMIZATION, OPTIMAL CONTROL AND DESIGN
- NUMERICAL ANALYSIS AND SCIENTIFIC COMPUTING
- ALGEBRAIC, TOPOLOGICAL AND DIFFERENTIAL STRUCTURES
- PROBABILITY AND STATISTICS
- ALGEBRAIC AND DIFFERENTIAL GEOMETRY
- MATHEMATICAL MODELLING IN MECHANICS ENGINEERING SCIENCES
- MATHEMATICAL ECONOMY AND GAME THEORY
- MATHEMATICAL PHYSICS AND APPLICATIONS

EDITURA
ACADEMIEI OAMENILOR DE ȘTIINȚĂ DIN ROMÂNIA

Annals of the Academy of Romanian Scientists

Series on Mathematics and its Applications

Secretariate: stiintematematice@gmail.com

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

CONTENTS

ADELINA GEORGESCU

25.04.1942 - 01.05.2010

I had the privilege of meeting Professor Adelina Georgescu almost 20 years ago during a series of courses that she held at the *University of Craiova*, Romania. What followed was more than a collaborative relationship; it was a long friendship, during which I discovered in her a gifted Romanian mathematician with a brilliant mind and a sharp spirit of righteousness.

Born on April 25, 1942 in Turnu Severin, Romania, she lost her mother at the age of two and was subsequently raised by her grandparents in Castranova. She began her schooling in Caracal and also graduated from an all-girls' high school in the same town. Between 1960 and 1965 she continued her studies in mathematics at the *University of Bucharest*. Here she benefited from the instruction of famous professors such as Victor Vâlcovici, Miron Nicolescu, Grigore Moisil, Gheorghe Vranceanu, Solomon Marcus, and Caius Iacob. Under the latter's guidance she completed her degree and earned her university diploma.

In 1965 she began working at the *Institute of Applied Mechanics Traian Vuia*, which later became the *Institute of Fluid Mechanics*. She then pursued a PhD at the *Institute of Mathematics*, studying hydrodynamic stability and corresponding with the best world specialists in the field. In 1970 she earned her doctorate degree in mathematics, under the supervision of academician Caius Iacob.

Between 1970 and 1975 she worked for the *Institute of Mathematics*, where she enjoyed a high-quality library, scientifically-advanced seminars, and elaborate discussions with the elite of Romanian mathematicians. After the closing of the *Institute of Mathematics* in 1975, she returned to the *Institute of Fluid Mechanics and Aerospatial Engineering*, where she remained until 1990. Here she worked on her first book in English, *Hydrodynamic Stability Theory*. The book was published by Kluwer in 1985 and has remained a highly appreciated reference in the field. At the same time she held her first courses at the *Faculty of Mathematics* from Bucharest.

The close relationship she had with her two sons, Andrei (born in 1971) and Sergiu Moroianu (born in 1973), well-known mathematicians today, and the relentless care for their education, enabled her to endure the difficulties of the communist regime and the stifling atmosphere maintained by it, until the Romanian Revolution in December 1989.

After the Revolution, returning to the re-established *Institute of Mathematics of the Romanian Academy*, she was often invited to hold conferences, seminars, and courses at foreign universities and research centers.

A three-month visit to the United States in 1990 gave her the strong conviction that mathematics needs to be more closely applied to real life problems and related to the universe of physics, economics and biology. This is why, after numerous complicated formalities, in 1991 she founded the *Institute of Applied Mathematics* (IMA), which she managed until 1995. Starting with a group of 25 researchers and modest working conditions, she succeeded to raise funds and to develop an institute of high scientific level, with an elegant headquarter and a rich library.

Still in 1990 she took the lead, with professors Cabiria Andreian Cazacu and Petre Osmatescu, to resume the organization of the *Congress of the Romanian Mathematicians* from all over the world, which had not been held since 1956. The idea finally materialized in 2003 with the organization of the *Fifth Congress of the Romanian Mathematicians* at Piteşti.

Promoting a rich scientific life of Romanian mathematicians through wide participation to conferences and constant exchange of research ideas, in 1992, in parallel with the IMA, she established the *Romanian Society of Applied and Industrial Mathematics* (ROMAI). This organization now has more than 150 members and an important branch in Basarabia. The organization has hosted annual *Conferences on Applied and Industrial Mathematics* (CAIM), the first one being held in Oradea in 1993. Professor Georgescu was involved in the organization of these conferences and in the improved rigor of their scientific content. Consequently, she published papers in volumes, starting with limited editions and then publishing them in the *Scientific Bulletins of the University of Pitesti*. Finally, the organization launched its own journal in 2005: *ROMAI Journal*.

In 1997, after her management position at IMA was terminated due to a controversial competition held while she was in Paris for a 6-month research stage, she moved to the University of Pitesti. Here she was elected head of the *Department of Applied Mathematics*. Following her dream to develop the field of applied mathematics, she organized at Pitesti the *Research Seminar Victor Vâlcovici*; she initiated and supported the hiring of talented young mathematicians and she edited a series of scientific monographs, which particularly reflected the activity of the ROMAI members. The first issue of the *Applied and Industrial Mathematics Series* appeared in 1999, and 29 issues

have been published up to this day. Disciplined, honest and dedicated, she contributed to the improved rigor and quality of education offered at the *University of Pitesti*. In permanent struggle with the established academic hierarchy, she fought tirelessly and often unsustained for a clean and high quality university.

During her last years, she extended her collaboration with Italian mathematicians such as Lidia Palese and Liliana Restuccia. They published tens of articles and attended many conferences and seminars at the universities of Bari, Messina and Catania.

Patriotically and tirelessly, she fought tooth and nail for the rapprochement of Basarabia. Thus, she collected and sent to Moldavia lots of books written in Romanian in a period when these books could hardly be found in that region. Additionally, with the support of a great number of Moldavian mathematicians, she expanded the organization of the CAIM conferences to Chisinau. Well-known mathematicians such as academician Mitrofan Ciobanu or professors Mefodie Rata, Mihail Popa, Dumitru Botnaru related their names to ROMAI and became regulars of CAIM conferences. Through all her actions, Adelina Georgescu showed that, in spite of all the existing obstacles, the idea of solidarity and spiritual unity with Moldavians is always present in the hearts of Romanians on the left side of the Prut River.

The main research contributions of Professor Adelina Georgescu were in hydrodynamics and their applications to complex fluid flows, hydrodynamic stability, turbulence, perturbative theories for differential equations, nonlinear dynamics, bifurcation theory, variational problems of mathematical physics, and synergetics. Her notable contributions comprise more than 200 scientific papers and 19 books, published both in Romania and abroad. She also hosted hundreds of conferences and other scientific meetings. A book of memoirs, published post-mortem, completes in a retrospective key her rich list of publications. With enthusiasm and persistence, she built a school of high academic quality, mainly focussed on applications of mathematical theories. The extraordinary personality of Professor Adelina Georgescu was felt by all her collaborators, friends and acquaintances. Her joy of sharing knowledge and science led her further to supervise and oversee 19 PhD theses, finalized between 1997 and 2009.

Of a rare honesty and a remarkable intelligence and generosity, she worked unwaveringly toward fulfilling her dream of advancing the field of applied

mathematics. In recognition of her scientific accomplishments, she was elected Doctor Honoris Causa of the *University of Tiraspol*, corresponding member of *Academia Peloritana dei Pericolanti* in Messina, member of the *Academy of Nonlinear Sciences* in Moscow, and member of the *Academy of Romanian Scientists.*

Professor Adelina Georgescu passed away on the First of May, 2010. She is resting near her father in the garden of a small church in her childhood village of Castranova. In her memory, at the annual conference organized at Iasi in October 2010, the ROMAI society awarded the first *Prize Adelina Georgescu for Applied Mathematics*, established for rewarding the most gifted young mathematicians from Romania and Republic of Moldavia.

I once heard Professor Adelina Georgescu citing her father's words at her graduation ceremony: "I gave you wings, now fly away!" And she did fly, throughout her subsequent life, higher and higher. And she gave wings to many disciples who found in her an energetic advisor always ready to share her knowledge and ideas. She was more than the founder of a research school: an unforgettable model of moral integrity, dignity and patriotism.

Farewell, dear Adelina Georgescu! You will always live in our hearts!

Carmen Rocşoreanu

Books published by Adelina Georgescu

1. Adelina Georgescu, Lucia Dragotescu, *Matematica si viata*, Seria Mat. Apl. Ind. 29, Ed. Pim, Iasi, 2010.

2. Adelina Georgescu, *Bifurcatie, fractali si haos determinist*, chapter 19 from Enciclopedia matematica, Editura AGIR, 2010, p. 945-994.

3. Adelina Georgescu, L. Palese, G. Raguso, *Biomatematica. Modelli, dinamica e biforcazione*, Cacucci Editore, Bari, 2009.

4. Adelina Georgescu, Lidia Palese, *Stability criteria of fluid flows*, Series on Advances in Mathematics for applied sciences 81, World Scientific, Singapore, 2009, 420 p.

5. Adelina Georgescu, George-Valentin Cârlig, Cătălin-Liviu Bichir, Ramona Radoveneanu, *Matematicieni români de pretutindeni*, second ed., Seria Mat. Apl. Ind. 24, Ed. Pamantul, Piteşti, 2006, ISBN 973-8280-90-7; 978-973-8280-90-8.

6. Adelina Georgescu, C.-L. Bichir, G.V. Cîrlig, *Matematicieni români de pretutindeni*, Seria de Mat. Apl. Ind. 18, E.d. Univ. Piteşti, Piteşti, 2004, ISBN 973-86901-6-1.

7 C. Rocşoreanu, Adelina Georgescu, N. Giurgiţeanu, *FitzHugh-Nagumo model: bifurcation and dynamics*, Kluwer, Dordrecht, 2000, 248 p., ISBN-7923-6427-9, MR 1779040 (2002a:34059).

8. B.-N. Nicolescu, N. Popa, Adelina Georgescu, M. Boloşteanu, *Mişcări ale fluidelor cavitante. Modelare şi soluţii*, Seria Mat. Apl. Ind, 3, Ed. Univ. Piteşti, Piteşti, 1999, 175 p., ISBN 973-9450-32-6. MR 1900777 (2003e:76014).

9. N. Popa, B.-N. Nicolescu, Adelina Georgescu, M. Boloşteanu, *Modelări matematice în teoria lubrificaţiei aplicate la etanşările frontale*, Seria Mat. Apl. Ind., 2, Ed. Univ. Piteşti, Piteşti, 1999, 145 p., ISBN 973-9450-20-2. MR 1897457.

10. Adelina Georgescu, M. Moroianu, I. Oprea, *Teoria bifurcaţiei. Principii şi aplicaţii*, Seria Mat. Apl. Ind. 1, Ed. Univ. Piteşti, Piteşti, 1999, 384 p., ISBN 973-9450-01-6. MR 1898530 (2003d:37067).

11. Adelina Georgescu, *Teoria stratului limită. Turbulenţă*, Ed. Gh. Asachi, Iaşi, 1997, 184 p., ISBN 973-9178-53-7.

12. Adelina Georgescu, *Asymptotic treatment of differential equations, Applied Mathematics and Mathematical Computation*, 9, Chapman and Hall, London, 1995, 268 p., MR 1316887 (96c:34107), ISBN 0-412-55860-2, Rev. Electr. 816.34002.

13. Adelina Georgescu, I. Oprea, *Bifurcation theory from application viewpoint*, Tipografia Univ. Timişoara, 1994, 281 p., (Monografii Matematice 51).

14. Adelina Georgescu, *Sinergetica. Solitoni. Fractali. Haos determinist.* Turbulenţă, Tipografia Univ. Timişoara, 1992, 338 p.

15. H. Dumitrescu, Adelina Georgescu, V. Ceangă, GH. Ghiţă, J. Popovici, B. Nicolescu, Al. Dumitrache, *Calculul elicei*, Ed.Academiei, Bucureşti, 1990, 675 p., ISBN 973-27-0053-X.

16. Adelina Georgescu, *Aproximaţii asimptotice*, Ed.Tehnică, Bucureşti, 1989, 172 p., ISBN 973-31-0142-7.

17. Adelina Georgescu, *Sinergetica-o nouă sinteză a ştiinţei*, Ed.Tehnică, Bucureşti, 1987, 148 p.

18. Adelina Georgescu, *Hydrodynamic stability theory*, Mechanics: Analysis, 9, Kluwer, Dordrecht, 1985, 307 p., ISBN 90-247-3120-8, MR 0850008 (87 k:76024).

19. St. N. Săvulescu, Adelina Georgescu, H.Dumitrescu, M.Bucur, *Cercetări matematice în teoria modernă a stratului limită*, Ed. Academiei, Bucureşti, 1981, 250 p.. MR 0632406 (84h: 76018); ZB 587.76001; Ref. Zh. Mat. 1 B 675K (1982); Ref. Zh. Mekh. 11 B 153K (1981).

20. Adelina Georgescu, *Teoria stabilităţii hidrodinamice*, Ed. ştiinţifică şi Enciclopedică, Bucureşti, 1976, 237 p., Ref. Zh. Mekh., 1977, 10 B 8K.

Ph. D theses supervised by Professor Adelina Georgescu

1. Rocşoreanu Carmen (University of Craiova, Romania) - *Dynamics and bifurcation in the FitzHugh-Nagumo equation*, 1997.

2. Giurgiteanu Nicolaie (University of Craiova, Romania, dead in 2010) - *Contributions to the study of systems of differential equations by numerical methods. Applications to biology*, 1997.

3. Suciu Nicolae (Friedrich-Alexander University Erlangen-Nuremberg, Germany and Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy, Cluj Napoca, Romania) - *On the Connection Between the Microscopic and the Macroscopic Modeling of the Thermodynamic Processes*, 1998.

4. Nicolescu Bogdan (University of Piteşti, Romania) - *Contributions to the mathematical study of cavitante fluid flows*, 1998.

5. Ion Anca Veronica (Institute of Applied Mathematics, Bucharest, Romania) - *Contributions to the study of material systems*, 2000.

6. Sterpu Mihaela (University of Craiova, Romania) - *Contributions at the study of codimensions of bifurcations for some three-dimensional dynamical systems*, 2001.

7. Băzăvan Petre (University of Craiova, Romania) - *Three-dimensional computation study of the attractors for Van der Pol - type equations*, 2001.

8. Bichir Cătălin-Liviu (ICEPRONAV, Galaţi, Romania) - *Contributions to the study of hydrodynamic stability*, 2002.

9. Curtu Rodica (Transilvania University of Braşov, Romania and University of Iowa, USA) - *Dynamics and bifurcation of the Gray-Scott model from enzymology*, 2002.

10. Ungureanu Laura (Spiru Haret University, Craiova, Romania) - *Structural stability and bifurcation in some mathematical models of economic dynamics*, 2002.

11. Georgescu Constantin (University of Piteşti, Romania) - *Contributions to the study of some ordinary differential equations from economy*, 2004.

12. Trifan Mariana (Chişinău, Moldavia) - *Contributions to the study of dynamics and bifurcation of some mathematical models for cancer*, 2005.

13. Codeci Elena (Vlaicu Vodă College, Curtea de Argeş, Romania) - *Perturbed bifurcation in models of microeconomic dynamics*, 2005.

14. Dragomirescu Florica-Ioana (University of Timişoara, Romania) - *Contributions to the study of spectral problems in hydrodynamic stability*, 2006.

15. Nistor Gheorghe (University of Piteşti, Romania) - *Contributions to asymptotic study of some nonlinear dynamics problems (including the FitzHugh-Nagumo model)*, 2007.

16. Georgescu Raluca-Mihaela (University of Piteşti, Romania) - *Group theory applications to the bifurcation study for some dynamical models*, 2007.

17. Nartea Simona Cristina (University of Piteşti, Romania) - *Inertial manifolds in concrete dynamical systems from biology and economics*, 2008.

18. Carlig George Valentin (University of Constanţa, Romania) - *Discrete dynamics and bifurcation in some economical and biological models*, 2009.

19. Gurgui Adriana (Constanţa, Romania) - *Applications of the theory of bifurcation when studying equilibriums and economic cycles*, 2009.

Applied and Industrial Mathematics Series, edited by Adelina Georgescu (published mainly at the Publishing House of the University of Pitesti)

1. Adelina Georgescu, Mihnea Moroianu, Iuliana Oprea, *Teoria bifurcaţiei; principii şi aplicaţii*, 1999.

2. Nicolae Popa, Bogdan Nicuşor Nicolescu, Adelina Georgescu, Mircea Boloşteanu, *Modelări matematice în teoria lubrificaţiei aplicate la etanşările frontale*, 1999.

3. Bogdan-Nicusor Nicolescu, Nicolae Popa, Adelina Georgescu, Mircea Bolosteanu, *Miscari ale fluidelor cavitante. Modelare şi soluţii*, 1999.

4. Anca-Veronica Ion, *Atractori globali şi varietăti inerţiale pentru două probleme din mecanica fluidelor*, 2000.

5. Nicolae Suciu, *Asupra relaţiei intre modelarea microscopică şi macroscopică a proceselor termodinamice*, 2001.

6. Laura Ungureanu, Liviu Ungureanu, *Elemente de dinamică economică*, 2000.

7. Marius-Florin Danca, *Funcţia logistică. Dinamică, bifurcaţie şi haos*, 2001.

8. Mihaela Sterpu, *Dinamică şi bifurcaţii pentru două modele Van der Pol generalizate*, 2001.

9. Iuliu Deac, *Dicţionar enciclopedic al matematicienilor*, Vol. I, 2001

10. Maria do Rosario de Pinho, Maria Margarida Ferreira, *Optimal control problems with constraints*, 2002.

11. Catalin-Liviu Bichir, *Stabilitate hidrodinamică teoretică şi numerică*, 2002

12. Iuliu Deac, *Dicţionar enciclopedic al matematicienilor*, Vol. II, 2002

13. Constanţa-Dana Constantinescu, *Haos, fractali şi aplicaţii*, 2003.

14. Danca Marius Florin, *Sisteme dinamice discontinue*, 2004.

15. Popa Mihail, *Metode cu algebre la sisteme diferenţiale*, 2004.

16. Raţă Mefodie, *Inexistenţa algoritmilor de recunoaştere a expresibilităţii sintactice în calcule logice*, 2004.

17. Emilia-Rodica Borşa, *Mişcări de fluide vâscoase generate de gradienţi de tensiune superficială*, 2004.

18. Adelina Georgescu, Cătălin-Liviu Bichir, George-Valentin Cârlig, *Matematicieni români de pretutindeni*, 2004.

19. Laura Ungureanu, *Stabilitate structurală şi bifurcaţie în două modele de dinamică economică*, 2004.

20. Cristian Grava, *Estimarea şi compensarea mişcării în secvenţe de imagini*, 2004.

21. Angela Muntean, *Introducere în studiul curenţilor marini*, 2005.

22. Mariana Trifan, *Dinamică şi bifurcaţie în studiul matematic al cancerului*, 2006.

23. Elena Codeci, *Bifurcaţie perturbată de dinamică economică*, 2006.

24. Adelina Georgescu, George-Valentin Cârlig, Cătălin-Liviu Bichir, Ramona Radoveneanu, *Matematicieni români de pretutindeni*, ed. a II-a, 2006.

25. Laurenţiu Calmutchi, *Metode algebrice şi funcţionale în teoria extensiilor spaţiilor topologice*, 2007.

26. Florica-Ioana Dragomirescu, *Probleme spectrale în stabilitatea hidrodinamică*, 2007.

27. Şerban E. Vlad, *Asynchronous Systems Theory*, 2007.

28. Raluca-Mihaela Georgescu, *Bifurcaţie în dinamica biologică cu metode de teoria grupurilor*, 2009

29. Adelina Georgescu, Lucia Dragotescu, *Matematică şi viaţă*, 2010.

Papers published by Adelina Georgescu

1966

1. Corecţii de compresibilitate pentru profile von Mises, St.Cerc.Mat., **18**, 2(1966), 301-308. AMR 20 # 3445, Ref. Zh. Mekh. **1** V, Ref. 260 (1967).

1969

2. Asupra soluţiilor asimptotice ale lui Heisenberg, St. Cerc. Mat., **21**, 5 (1969), 747-750. MR 0270616(**42**), 4, Ref. 5504; ZB **213**, p. 543; AMR 23 # 8876.

3. On a relationship between Heisenberg and Tollmien solutions, Rev.Roum. Math.Pures et Appl., 14, 7 (1969), 991-998. ZB **194**, p. 579; AMR 24 # 441; Ref. Zh. Mat. 7 B 230 (1969).

4. Improvement of one of Joseph's theorems and one of its applications, Rev. Roum. Math. Pures et Appl., **14**, 8 (1969), 1089-1092. ZB **197**, p. 253; AMR 23 # 9647.

5. Criterii de stabilitate liniară a mişcării plan paralele a unui fluid nenewtonian, St.Cerc.Mat., **21**, 7 (1969), 1027-1036. MR 0280059 (**43**), 4, Ref. 5780; ZB **197**, p. 239; Ref. Zh. Mekh. **7** V, Ref. 1155 (1970).

6. On the generalized Tollmien solutions of the Rayleigh equation for a general velocity profile, Bull.Math. de la Soc. Sci.Math. de la R.S. de Roumanie, **13** (**61**), 2 (1969), 147-158. ZB **216**, p. 529; Ref. Zh. Mat. **3** B 291 (1971).

1970

7. Note on Joseph's inequalities in stability theory, ZAMP, **21**, 2 (1970), 258-260. ZB **197**, p. 530; AMR 23 # 9646; Ref. Zh. Mekh. **11** A, Ref. 118 (1970).

8. Sur la stabilité linéaire des mouvements plans des fluides, Comptes Rendus, Paris, Série A, **271** (1970), 559-561. MR 0272253 (**42**), 5, Ref. 7134; ZB **215**, p. 584; Ref. Zh. Mekh. **3** V, Ref. 709 (1971).

9. Sufficient conditions for linear stability of two Ladyzhenskaya type fluids, Rev.Roum.Math.Pures et Appl., **15**, 6 (1970), 819-823. MR 0267825 (**42**), 2, Ref. 2727; ZB **211**, p. 297; Ref. Zh. Mekh. **3** V, Ref. 1147 (1971).

10. Contribuţii la studiul stabilităţii liniare a mişcării fluidelor, St. Cerc. Mat., **22**, 9 (1970), 1247-1333. MR 0337133 (**49**), 1, Ref. 1905; AMR 25 # 9144; Ref. Zh. Mekh. **6** V, Ref. 655 (1971), PhD thesis.

1971

11. Theorems of Joseph's type in hydrodynamic stability theory, Rev. Roum.Math.Pures et Appl., **16**, 3 (1971), 355-362. MR 0285183 (**44**), 2, Ref. 2406; ZB **216**, p. 528; AMR 25 # 1281; Ref. Zh. Mekh. **11** V, Ref. 542 (1971).

12. On the Kelvin-Helmholtz instability in presence of porous media, Rev. Roum. Math. Pures et Appl., **16**, 1 (1971), 27-39 (with Ct. I. Gheorghiţă). MR 0281404 (**43**), 5, Ref. 7121; ZB **219**, p. 523; AMR 25 # 2434; Ref. Zh. Mat. **7** B 491 (1971).

13. On the neutral stability of the Couette flow between two rotating cylinders, Rev.Roum.Math.Pures et Appl. **16**, 4 (1971), 499-502. ZB **233**, Ref. 76084; AMR 26 # 5532; Ref. Zh. Mat. **1** B 520 (1972).

14. Instability of two superposed liquids in a circular tube in the presence of a porous medium, Rev.Roum.Math.Pures et Appl., **16**, 5 (1971), 677-680,

(with Şt. I. Gheorghiţă). MR 0286360 (**44**), 3, Ref. 3573; ZB **222**, Ref. 76049; AMR 25 # 2210; Ref. Zh. Mekh. **2**V, Ref. 472 (1972).

1972

15. Linear Couette flow stability for arbitrary gap between two rotating cylinders, Rev. Roum. Math. Pures et Appl., **17**, 4 (1972), 507-518. MR 0305717 (**46**), 3, Ref. 4847; ZB **245**, Ref. 76033; AMR 26 # 4681; Ref. Zh. Mekh. **10** V, Ref. 705 (1972).

16. Stability of spiral flow and of the flow in a curved channel, Rev. Roum. Math. Pures et Appl., **17**, 3 (1972), 353-357. ZB **245**, Ref. 76039; AMR 26 # 8194.

1973

17. Stability of the Couette flow of a viscoelastic fluid, Rev. Roum. Math. Pures et Appl., **18**, 9 (1973), 1371-1374. ZB **272**, Ref.: 76004; AMR 27 # 9399; Ref. Zh. Mekh. **3** V, Ref. 1096 (1974).

18. Teorema lui Squire pentru o mişcare într-un mediu poros, Petrol şi Gaze, **24**, 11 (1973), 676-678. Ref. Zh. Mekh. **7** V, Ref. 1105 (1974).

1976

19. Universal criteria of hydrodynamic stability, Rev. Roum. Math. Pures et Appl., **21**, 3 (1976), 287-302. MR 0443557 (**56)**, 1, Ref.: 1926; ZB **339**, Ref.: 76029; AMR, 29 # 10005; Ref. Zh. Mekh. **11** B 75 (1976).

1977

20. Stability of the Couette flow of a viscoelastic fluid. II, Rev. Roum. Math. Pures et Appl., **22**, 9 (1977), 1223-1233. (with O. Polotzka). ZB **372**, Ref.: 76008; Ref. Zh. Mekh. **6** B 964 (1978).

21. Metode analitice în studiul fenomenologic al stabilităţii mişcării fluidelor vâscoase incompresibile descrise de soluţii generalizate ale ecuaţiilor Navier-Stokes. St.Cerc.Mat., **29**, 6 (1977), 603-619. MR 0455867 (**56**), 5 , Ref.: 14101; ZB **406**, Ref. : 76033; Ref. Zh. Mekh. **5** B 98 (1978).

22. Variational formulation of some nonselfadjoint problems occuring in Bénard instability theory I, Preprint Series in Mathematics, **35**/1977, Institutul de Matematică, INCREST, Bucureşti, 1977.

1978

23 Bounds for linear characteristics of Couette and Poiseuille flows, Rev. Roum. Math. Pures et Appl., **23**, 5 (1978), 707-720 (with Tr. Bădoiu). MR 0506588 (80d:76040); ZB **383**, Ref. 76028; Ref. Zh. Mekh. **3** B 120 (1979).

1980

24. Neutral stability curves for a thermal convection problem, Acta Mechanica, **37** (1980), 165-168 (with V. Cardoş). ZB **441**, Ref.: 73136; Ref. Zh.Mekh. **3** B 515 (1981).

1981

25. On the nonexistence of regular solutions of a Blasius-like equation in the theory of the boundary layer of finite depth, Rev. Roum. Math. Pures Appl., **26**, 6 (1981), 849-854 (with M. Moroianu). MR 0627830 (83h:76025); ZB **471**, Ref. 76039; Ref. Zh. Mat. **2** B 482 (1982).

26. Recent results in fluid mechanics, Preprint **2**, Univ. "Babeş - Bolyai", Fac. Mat., Cluj-Napoca, 1981. MR 0655038 (84i:76001); ZB **517**, Ref. P76001.

27. On a Bénard convection in the presence of dielectrophoretic forces, J. Appl. Mech., **48** , 4 (1981), 980-981 (with O. Polotzka).

28. Catastrophe surface bounding the domain of linear hydromagnetic stability, Central Institute of Physics, National Institute of Scientific and Technical Creation, Bucharest, Romania, Preprint **FT-203**-1981.

1982

29. On a universal criterion of hydrodynamic stability, Univ. din Timişoara, Preprint **68**/1982.

30. Bifurcation (catastrophe) surfaces for a problem in hydromagnetic stability, Rev. Roum. Math. Pures Appl., **27**, 3 (1982), 335-337. MR 0669482 (84f:76033); ZB **495**, Ref. 76051.

31. Neutral stability curves for a thermal convection problem, Analele Univ. din Craiova, Secţia Mat. Fiz.-Chim., **X** (1982), 51-53 (with I. Oprea).

32. Characteristic equations for some eigenvalue problems in hydromagnetic stability theory, Mathematica, Cluj, **24** (**47**), 1-2 (1982), 31-41. MR 0692182 (84h:76023); ZB **521**, Ref. 76045.

1983

33. Neustanovivseesia ploscoe dvijenie tipa Puazeilia dlia jidkostei Rivlina-Eriksena, PMM, **47**, 2 (1983), 342-344. (with S.S. Chetti).

34. Stabilitatea şi ramificarea în contextul sinergeticii, St.Cerc.Mec.Apl., **42**, 2 (1983), 174-180

1984

35. Echilibrul plasmei în sisteme toroidale şi stabilitatea sa macroscopică, St.Cerc.Fiz., **36**, 1 (1984), 86-110.

1986

36. Proiectarea aerodinamică a elicei de randament maxim, St.Cerc. Mec.Apl., **45**, 2 (1986), 129-141 (with H. Dumitrecsu, Al. Dumitrache).

37. Bifurcaţia stratului limită, BITNAV, **3** (1986), 148-149.

38. Metode numerice în teoria bifurcaţiei, Stud. Cerc. Fiz., **38**, 10 (1986), 912-924 (with I. Oprea). MR 0873500 (88c: 58046)

1987

39. Metode de rezolvare a unor probleme de valori proprii care apar în stabilitatea hidrodinamică liniară, St.Cerc.Fiz., **39**, 1 (1987), 3-25 (with A. Setelecan). ZB **605**, Ref.: 76053.

40. Notă asupra unor probleme izoperimetrice în calculul elicei de randament maxim, St. Cerc. Mec. Apl., **46**, 5 (1987), 478-482.

41. Exact solutions for some instability of Bénard type, Rev. Roum. Phys., **32**, 4 (1987), 391-397.

1988

42. Metode numerice în teoria bifurcaţiei. II. Soluţii staţionare în cazul infinit dimensional, Stud. Cerc. Fiz., **40**, 1 (1988), 7-18 (with I. Oprea). MR 0949207 (89j : 65048)

43. Bifurcation (catastrophe) surfaces in multiparametric eigenvalue problems in hydromagnetic stability theory, Bull. Inst. Politehn. Bucureşti, Ser. Construc. Maş, **50** (1988), 9-12 (with I. Oprea), MR 0996532 (90e : 76077).

44. The bifurcation curve of characteristic equation provides the bifurcation point of the neutral curve of some elastic stability, Mathematica - Anal. Numér. Théor. Approx., **17**, 2 (1988), 141-145 (with I. Oprea). MR 1027220 (90i : 73061)

1989

45. Bifurcation manifolds in a multiparametric eigenvalue problem for linear hydromagnetic stability theory, Mathematica - Anal. Numér. Théor. Approx., **18**, 2 (1989), 123-138 (with I. Oprea, C. Oprea). MR 1089229 (92 i : 76044)

46. Model de aproximaţie asimptotică de ordinul patru pentru ecuaţiile meteorologice primitive când numărul Rossby tinde la zero, St. Cerc. Meteorolgie, **3** (1989), 13-21 (with C. Vamoş).

47. Filtred equations as an asymptotic approximation model, Meteorology and Hydrology, **19**, 2 (1989), 21-22 (with C. Vamoş).

48. Boundary layer separation I. Bubbles on leading edges, Rev. Roum. Sci. Tech.-Méc. Appl., **34**, 5 (1989), 509-525, (with H. Dumitrescu, Al. Dumitrache). MR 1054173 (91b:76043); Ref. Zh. Mekh., **5** B 145 (1990).

49. Stabilitatea mişcării lichidelor pe un plan înclinat, Stud. Cerc. Mec. Apl., **48**, 5 (1989), 471-479. MR 1050048

50. Comparative study of the analytic methods used to solve problems in hydromagnetic stability theory, An. Univ. Bucureşti, Mat., **38**, 1 (1989), 15-20 (with A. Setelecan). MR 1100332 (92a : 76042)

51. Lagrange and the calculus of variations, Noesis, **15** (1989), 29-35.

52. Fractalii şi unele aplicaţii ale lor, Stud. Cerc. Fiz., **41**, 3 (1989), 269-288. MR 1028540

53. Suprafeţe neutrale bifurcate într-o problemă de inhibiţie a convecţiei termice datorită unui câmp magnetic, Stud. Cerc. Mec. Apl., **48**, 3 (1989), 263-278. (with I. Oprea). MR 1023843 (90i :76086)

1990

54. Models of asymptotic approximation for synoptical flows, Zeitschrift f§r Meteorologie, **40**, 1 (1990), 14-20. (cu C. Vamoş).

55. Neutral stability curves for a thermal convection problem. II. The case of multiple solutions of the characteristic equation, Acta Mechanica, **81** (1990), 115-119. (with I. Oprea). MR 1059096 (91k : 76069)

56. Metode numerice în teoria bifurcaţiei. III. Soluţii periodice, Stud. Cerc. Fiz., **42**, 1 (1990), 117-125. (with I. Oprea). MR 1074148 (91j : 65106)

57. Scenarii de turbulenţă în cadrul haosului determinist, Stud. Cerc. Mec. Apl., **49**, 4 (1990), 413-417. (with C. Vamoş, N. Suciu). MR 1154192

58. Models of asymptotic approximation, IMA Preprint Series **724**, Minneapolis, Nov. 1990.

59. Asimptote oblice din punctul de vedere al aproximaţiei asimptotice, Gaz. Mat. M, **2** (1990), 58-60.

60. On a hydrodynamic-social analogy, Rev. Roum. Philos. Logique, **34**, 1-2 (1990), 100-102.

1991

61. Modelarea matematică în mecanica fluidelor, St. Cerc. Mec. Apl., **50**, 3-4 (1991), 295-298.

62. Efectul Toms, St. Cerc. Mec. Apl., **50**, 5-6 (1991), 305-321. (with C. Chiujdea).

63. Linear stability of a turbulent flow of Maxwell fluids in pipes, Université de Metz, **21**/1991, (with C. Chiujdea, R. Florea).

64. Evolution of the concept of asymptotic approximation, Noesis, **17** (1991), 45-50.

1992

65. Aspecte ale modelării stratului limită al atmosferei. I, Stud. Cerc. Mec. Apl., **51**, 1 (1992), 25-41. MR 1170345 (94b : 86001)

66. Studiul calitativ al ecuaţiilor diferenţiale, St. Cerc. Mec. Apl., **51**, 3 (1992), 317-326.

67.Models of asymptotic approximation governing the atmospheric motion over a low obstacle, Stud. Cerc. Mat., **44**, 3 (1992), 237-252. (with G. Marinoschi). MR 1182289 (93e : 86003)

68. Linear stability of a turbulent flow of Maxwell fluids in pipes, Rev. Roum. Math. Pures Appl., **37**, 7 (1992), 579-586 (with C. Chiujdea, R. Florea). MR 1188610 (93h : 76039)

69. Bifurcation problems in linear stability of continua, Quaderni. Dipto. di Mat. Univ. Bari, **1** (1992). (with I. Oprea).

70. Aspecte ale modelării stratului limită al atmosferei. II, Stud. Cerc. Mec. Apl., **51**, 2 (1992), 175-188. MR 1170347 (94b : 86002)

1993

71. Synergetics and synergetic method to study processes in hierarchical systems, Noesis, **18** (1993), 121-127.

72. The application of the shooting method to the hydrodynamic stability of the Poiseuille flow in channels and pipes, Computing, **4** (1993), 3-6. (with R. Florea).

73. Stability of a binary mixture in a porous medium with Hall and ion-slip effect and Soret-Dufour currents, Analele Univ. din Oradea, **3** (1993), 92-96. (with L. Palese, D. Paşca).

74. Metode de determinare a curbei neutrale în stabilitatea Bénard, St. Cerc. Mec. Apl., **52**, 4 (1993), 267-276. (with I. Oprea, D. Paşca).

75. Critical hydromagnetic stability of a thermodiffusive state, Rev. Roum. Math. Pures Appl., **38**, 10 (1993), 831-840. (with L. Palese, D. Paşca, M. Buican). MR 1264602 (95a : 76031)

76. Direcţii de cercetare principale în teoria sistemelor dinamice, Stud. Cerc. Mec .Apl., **52**, 2 (1993), 153-171. MR 1227549

77. Bifurcation manifolds in multiparametric linear stability of continua, ZAMM, **73**, 7-8 (1993), T831-T833. (with D. Paşca, S. Grădinaru, M. Gavrilescu).

78. Balance equations for the vector fields defined on orientable manifolds, Tensor (N. S.), **54** (1993), 88-90. (with C. Vamoş, N. Suciu). MR 1474041 (98i : 82036)

1994

79. Nonlinear stability criteria for l MHD flows. I. IsothermaI isotropic case, Rev. Roum. Math. Pures Appl., **39**, 2 (1994), 131-146, (with M. Maiellaro, L. Palese). MR 1298878 (95h:76054)

80. Extension of a Joseph's criterion to the nonlinear stability of mechanical equilibria in the presence of thermodiffusive conductivity, Rapp. Dipto. Mat., Univ. Bari, **12**/1994. (with L. Palese).

1995

81. Amélioration des estimations de Prodi pour le spectre, C. R. Acad. Sci. Paris Sér. I Math., **320**, 7 (1995), 891-896. (with L. Palese).

82. Sulla stabilitĞ globale del equilibrio meccanico per una miscela binaria in presenza di effetti Soret e Dufour, Rapp. Dipto. Mat., Univ. Bari, **8**/1995. (with L. Palese, A. Redaelli).

1996

83. Neutral stability hypersurfaces for an anisotropic MHD thermodiffusive mixture. III. Detection of false secular manifolds among the bifurcation characteristic manifolds, Rev. Roum. Math. Pures Appl.,**41**,1-2 (1996), 35-49.

84. Balance equations for physical systems with corpuscular structure, Physica **A**, **227** (1996), 81-92. (with C. Vamoş, N. Suciu, I. Turcu).

85. Balance equations for a finite number of material points, Stud. Cerc. Mat., **48**, 1-2 (1996), 115-127. (with C. Vamoş, N. Suciu). MR 1681175 (92m:82037)

86. A nonlinear stability criterion for a layer of a binary mixture, ZAMM Supplement **2**, **76** (1996), 529-530. (with L. Palese).

87. Extension of the Joseph's criterion on the nonlinear stability of mechanical equilibria in the presence of thermodiffusive conductivity, Theoret. Comput. Fluid Dyn., **8**, 6 (1996), 403-413.(with L. Palese) .)

88. Asymptotic analysis of nonlinear equilibrium solute transport in porous media, Water Ressources Research, **32**, 10, (1996), 3093-3098. (with U. Jaekel, H. Vereecken).

89. Nonlinear stability bounds for a binary mixture with chemical surface reactions, Rapp. Int. Dipto. Mat., Univ. Bari, **18**/1996. (with L. Palese).

90. Linearization principle for the stability of the mechanical equilibria of a binary mixture when the Soret and Dufour effects are present, Rapp. Int. Dipto. Mat., Univ. Bari, **14**/1996. (with L. Palese, A. Redaelli).

91. Coarse grained averages in porous media, KFA / ICG - 4 Internal Report No. 501296/1996. (with N. Suciu, C. Vamoş, U. Jaekel, H. Vereecken).

92. On Lagrangian passive transport in porous media, KFA/ICG-4 Internal Report No. 501196/1996 (with N. Suciu, H. Vereecken, C. Vamoş, U. Jaekel, O. Neuendorf).

1997

93. On the existence and on the fractal and Hausdorff dimensions of some global attractor, Nonlinear Anal., Theory, Methods & Applications, **30**, 8, (1997), 5527-5532. (with A. Ion). MR 1726057 (2000i:37142)

94. Stability spectrum estimates for confined fluids, Rev. Roum. Math. Pures Appl., **42**,1-2 (1997), 37-51.

95. Thermosolutal instability of a compresible Soret-Dufour mixture with Hall and ion-slip currents through a porous medium, Rev. Roum. Sci. Tech.-Méc. Appl., **42**, 3-4 (1997), 279-296, (with L. Palese, D. Paşca, D. Bonea).

96. Studiul portretului de fază. II. Punctele de inflexiune ale traiectoriilor de fază ale sistemului dinamic Van der Pol, St. Cerc. Mec. Apl., **56**, 1-2 (1997), 15-31. (with N. Giurgiţeanu).

97. Studiul portretului de fază. III. Influenţa liniarizării asupra sistemului dinamic neliniar, St. Cerc. Mec. Apl., **56**, 3 - 4 (1997), 141-153. (with N. Giurgiţeanu).

98. Hydrodynamical equations for one-dimensional systems of inelastic particles, Phys. Rev., **E** (3), **55**, 5 (1997), 6277-6280. (with C. Vamoş, N. Suciu). MR 1448402

99. Modelul continuu multiplicator-accelerator. Cazul liniar, Bul. Şt. Seria Mat.-Inform. Univ. Piteşti, **1** (1997), 95-104. (with C. Georgescu).

100. Bifurcation in the Goodwin model from economics. II, Bul. Şt. Seria Mat.-Inform. Univ. Piteşti, **1** (1997), 105-112. (with N. Giurgiţeanu, C. Rocşoreanu).

101. Studiul portretului de fază IV. Absenţa bifurcaţiei canard, St. Cerc. Mec. Apl., **56**, 5-6 (1997), 297-305. (with N. Giurgiţeanu, C. Rocşoreanu).

102. Degenerated Hopf bifurcation in the FitzHugh-Nagumo system. 1. Bogdanov-Takens bifurcation, Analele Univ. din Timişoara, **35**, 2 (1997), 285-298. (with C. Rocşoreanu, N. Giurgiţeanu). MR 1876887 (2002j:34085)

1998

103. Neutral thermal hydrodynamic and hydromagnetic stability hypersurfaces for a micropolar fluid layer, Indian J. Pure and Appl. Math.,**29**, 6 (1998), 575-582. (with M. Gavrilescu, L. Palese). MR 1636477 (99f:76056)

104. Coarse grained and stochastic averages. Applications to transport processes in porous media. ICG-4 Internal Report No. 500198/1998, J§lich. (with N. Suciu, C. Vamoş, U. Jaekel, H. Vereecken).

105. On the Misra-Prigogine-Courbage theory of irreversibility, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **2** (1998), 169-188. (with N. Suciu).

106. On the mechanism of drag reduction in Maxwell fluids, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **2** (1998), 107-114. (with C. Chiujdea).

107. Transport processes in porous media. 1. Continuous modelling, Romanian J. Hydrology Water Resources, **5**, 1-2 (1998) 39-56. (with N. Suciu, C. Vamoş, U. Jaekel, H. Vereecken).

108. Equilibria and relaxation oscillations of the nodal system of the heart. 2. Hopf bifurcation, Rev. Roum. Sci. Tech.-Méc. Appl., **43**, 3, (1998), 403-414, (with C. Rocşoreanu, N. Giurgiţeanu), MR 1830580.

109. Neutral surfaces for Soret - Dufour - driven convective instability, Rev. Roum. Sci. Tech. - Méc. Appl., **43**, 2 (1998), 251 - 260, (with L. Palese, L. Pascu).

1999

110. Set of attraction of certain initial data in a nonlinear diffusion problem, , Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 235-261.

111. Coincidence of the linear and nonlinear stability bounds in a horizontal thermal convection problem, Intern. J. Nonlin. Mech., **34**, 4 (1999), 603-613 (with D. Mansutti). MR 1688548 (2000a:76084)

112. New types of codimension-one and-two bifurcations in the plane, Inst. Matem. Acad. Rom, Preprint No. **12**/1999. (with C. Rocşoreanu, N. Giurgiţeanu).

113. Regimes with two or three limit cycles in the FitzHugh-Nagumo system, ZAMM **79**, Supplement **2** (1999), S293-S294 (with C. Rocşoreanu, N. Giurgiţeanu).

114. Asymptotic analysis of solute transport with linear nonequilibrium sorption in porous media, Transp. Porous Media, **36**, 2 (1999), 189-210 (with H. Vereecken, U. Jaekel). MR 1777016 (2001d:76132)

115. Symmetry of the solution of the nonlinear Reynolds equation describing mechanical face seals, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 333-343. (with B. Nicolescu, N. Popa).

116. Hopf bifurcation and canard phenomenon in the FitzHugh-Nagumo model, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 217-233. (with C. Rocşoreanu, N. Giurgiţeanu).

117. Convecţia termică cu efect Marangoni. Condiţii de echilibru, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 345-350. (with Gh. Nistor).

118. Applications of coarse-grained and stochastic averages to transport processes in porous media, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 435-445. (with N. Suciu, C. Vamos, U. Jaekel, H. Vereeken).

119. Modelul continuu multiplicator-accelerator. II. Cazul liniar pentru anumite valori negative ale parametrilor şi cazul neliniar, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 263-266. (with C. Georgescu).

120. Investigation of the normalized Gierer-Meinhardt system by center manifold method, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **3** (1999), 277-283. (with A. Ionescu).

121. Dynamics and bifurcations in a biological model, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **4** (1999), 137 - 153. (with N. Giurgiţeanu, C. Rocşoreanu).

122. On an inertial manifold in the dynamics of gas bubbles, Rev. Roum. Sci. Tech. - Méc. Appl., **44**, 6 (1999), 629 - 631. (with B. Nicolescu). MR 1872191

2000

123. Dynamics generated by the generalized Rayleigh equation. II. Periodic solutions, Mathematical Reports, **2**(**52**), 3 (2000), 367 - 378, 2001. (with M. Sterpu, P. Băzăvan), MR 1898619 (2003I:34111).

124. Neutral curves for the MHD Soret - Dufour driven convection, Rev. Roum. Sci. Tech. - Méc. Appl., **45**, 3 (2000), 265 - 275 (with S. Mitran, L. Palese).

125. On a method in linear stability problems. Application to natural convection in a porous medium. J. of Ultrascientist of Physical Sciences, **12**, 3 (2000), 324 - 336. (with L. Palese).

126. On the Misra-Prigogine-Courbage theory of irreversibility. 2. The existence of the nonunitary similarity, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **6** (2000), 213 - 222. (with N. Suciu).

127. Codimension - three bifurcation for a FitzHugh-Nagumo like system, Bul. Şt. Univ. Piteşti, Seria Matematică şi Informatică, **6** (2000), 193 - 197. (with M. Sterpu).

128. Dynamics and bifurcation in the periodically forced FitzHugh-Nagumo system, Intern. J. of Chaos Theory and Applications, **5**, 2 (2000), 63 - 79. (with M. Sterpu) (invited paper)

129. On a new method in hydrodynamic stability theory, Math. Sciences Research Hot - Line, **4**, 7 (2000), 1 - 16. (with L. Palese, A. Redaelli). MR 1769518 (2001e:76054)

130. Hopf and homoclinic bifurcations in a biodynamical system, Bul. Şt. Univ. Baia Mare, Seria Mat-Inf., **16**, 1(2000), 131-142. (with C. Rocşoreanu, N. Giurgiţeanu). MR 1832131 (2002d:34067)

131. Dynamics of the cavitation spherical bubble. II. Linear and affine approximation, Rev. Roum. Sci. Tech. - Méc. Appl., **45**, 2 (2000), 163 - 175. (with B.-N. Nicolescu)

132. Degenerated Hopf bifurcation in the FitzHugh - Nagumo system. II. Bautin bifurcation, Mathematica- Anal. Numér. Theor. Approx., **29**, 1 (2000), 97 - 109. (with C. Rocşoreanu, N. Giurgiţeanu). MR 1928253 (2003h:34084)

2001

133. Concavity of the limit cycles in the FitzHugh-Nagumo model, Analele Univ. Iaşi, Seria I Matematica, **47**, 2 (2001), 287-298. (with C. Rocşoreanu, N. Giurgiţeanu). MR 1977388

134. Codimension-three bifurcations for the FitzHugh-Nagumo dynamical scheme, Mathema-tical Reports, **3** (**53**), 3 (2001), 287 - 292. (with M. Sterpu). MR 1929540 (2003j:34068).

135. Classes of solutions for a nonlinear diffusion PDE, J. of Comput. Appl. Math., **133**, 1-2 (2001), 373 - 381. (with H. Vereecken, H. Schwarze, U. Jaekel). MR 1858295 (2002h:76130)

136. On special solutions of the Reynolds equation from lubrication, J. of Comput. Appl. Math., **133**, 1-2 (2001), 367 - 372. (with B. Nicolescu, N. Popa, M. Boloşteanu). MR 1858294 (2002g:76046)

137. Conections between saddles for the FitzHugh-Nagumo system, Int. J. Bif. Chaos, **11**, 2 (2001), 533 - 540. (with C. Rocşoreanu, N. Giurgiţeanu). MR 1830350 (2002b:37077)

138. The complete form for the Joseph extended criterion, Ann. Univ. Ferrara, Sez. VII (N. S.), Sc. Mat., **47** (2001), 9 - 22 (with L. Palese, A. Redaelli). MR 1897556 (2003c:76058)

139. Determination of neutral stability curves for the dynamic boundary layer by splines, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **7** (2002), 15-22. (with L. Bichir).

140. Numerical integration of the Orr-Sommerfeld equation by wavelet methods, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **7** (2001), 9-14. (with L. Bichir)

141. Degenerated Bogdanov-Takens points in an advertising model, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **7** (2001), 173-177. (with L. Ungureanu)

2002

142. Codimention - one bifurcations for a Rayleigh model, Bul. Acad. Şt. Rep. Moldova, Seria Mat., **1** (**38**), (2002), 69 - 76. (with M Sterpu). MR 1954224 (2003m:34095)

143. Static bifurcation diagram for a microeconomic model, Bul. Acad. Şt. Rep. Moldova, **3** (**40**) (2002), 21-26 (with L. Ungureanu, M. Popescu), MR1991012

144. Improved criteria in convection problems in the presence of thermodiffusive conductivity, Analele Univ. Timişoara, **40**, 2 (2002), 49-66. (with L. Palese).

145. Existence and regularity of the solution of a problem modelling the Bénard problem, Mathematical Reports, **4** (**54**), 1 (2002), 87-102. (with A.-V. Ion). MR1994120

146. Stability criteria for quasigeostrophic forced zonal flows. I. Asymptotically vanishing linear perturbation energy, Magnetohydrodynamics: an International J. (with L. Palese) Rapp. Int. Dipto. Mat., Univ. Bari, **8**/1996

147. Domains of attraction for a model in enzimology, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **8** (2002), 49-58. (with R. Curtu).

148. Heteroclinic bifurcations for the FitzHugh - Nagumo system, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **8** (2002). (with C. Rocşoreanu, N. Giurgiţeanu).

149. Topological type of some nonhyperbolic equilibria in a problem of microeconomic dynamics, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **8** (2002). (with L. Ungureanu, M. Popescu).

150. On the Misra-Progogine-Courbage theory of irreversibility, Mathematica, **44** (**67**), 2 (2002), 215-231. (with N. Suciu)

151. Non - Newtonian solution and viscoelastic constitutive equations, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **5** (2000). (with C. Chiujdea).

152. Normal form for the degenerated Hopf bifurcation in an economic model, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **5** (2000). (with L. Ungureanu).

153. k > 3 order degenerated Bautin bifurcation and Hopf bifurcation in a mathematical model of economical dynamics, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., (2002). (with L. Ungureanu, M. Popescu).

154. Static bifurcation diagram for a mathematical model governing the capital of a firm, Bul. Şt. Univ. Piteşti, Seria Mat.- Inf, **8** (2002), 177-181. (with L. Ungureanu).

155. Concavity of the limit cycles in the FitzHugh-Nagumo model, An. Şt. Univ. Al.I.Cuza, Iaşi, Mat. (N. S.), **47**, 2 (2001), 287-298, (2002). (with C. Rocşoreanu, N. Giurgiţeanu) MR 1977388

2003

156. A Lie algebra of a differential generalized FitzHugh - Nagumo system, Bul. Acad. Şt., Rep. Moldova, Seria Mat. **1** (**41**) (2003), 18-30. (with M. Popa, C. Rocşoreanu) MR1992647

157. Approximation of pressure perturbations by FEM, Bul. Şt. Univ. Piteşti, Seria Mat. - Inf., **9** (2003), 31-36. (with C. - L. Bichir).

158. Global bifurcations for FitzHugh-Nagumo model, Dynamical Systems and Applications, Proc. of Conf. on Bifurcations, Symmetry and Patterns, (Porto, June 30 - July 4, 2000), in Trends in Mathematics: Bifurcations, Symmetry and Patterns, BirkhŁuser, Basel, 2003, 197-202, ISBN 3-7643-7020-3. (with C. Rocşoreanu, N. Giurgiţeanu).

159. Static and dynamic bifurcation of nonlinear oscillators, Bul. Şt. Univ. Piteşti, Seria Mec. Apl., **1**, **7** (2003), 133-138.

160. A Lorenz-like model for the horizontal convection flow, Int. J. Non-Linear Mech., **38** (2003), 629-644. (with E. Bucchignani şi D. Mansutti)

161. Bifurcation in biodynamics, Sci. Annals of UASVM Iasi, **46**, 2 (2003), 15-34.

162. Numerical integration of the Orr-Sommerfeld equation by wavelet methods, Bul. St. Univ. Pitesti, Seria Mat.-Inf., **9** (2003), 25-30. (with L. Bichir)

2004

163. Bifurcation in the Goodwin model I, Rev. Roum. Sci. Tech. - Méc. Appl., **49,** 1-6 (2004), 13-16. (with N. Giurgiţeanu, C. Rocşoreanu).

164. Curba valorilor de bifurcaţie Hopf pentru sisteme dinamice plane, Bul. St., Seria Mec. Apl., **10 (**2004), 55-62. (with E. Codeci)

165. Dynamic bifurcation diagrams for some models in economics and biology, Acta Universitatis Apulensis, Alba Iulia, Mathematics-Informatics, **8**

(2004), 156-161, Proceedings of the International Conference on Theory and Applications of Mathematics and Informatics - ICTAMI 2004, Thessaloniki, Greece.

166. On instability of the magnetic Bénard problem with Hall and ion-slip effects, Intern. J. Engng. Sci., **42** (2004), 1001-1012. (with L. Palese)

167. Liapunov method applied to the anisotropic Bénard problem, Math. Sci. Res. J., **8**, 7 (2004), 196-204. (with Lidia Palese).

2005

168. The static bifurcation in the Gray - Scott model, Rev. Roum. Sci. Tech. - Méc. Appl., **50**, 1-3 (2005), 3-13. (with R. Curtu).

169. A nonlinear hydromagnetic stability criterion derived by a generalized energy method, Bul. Acad. St. Rep. Moldova, Seria Mat., 1(47) (2005), 85-91. (with C.-L. Bichir, L. Palese)

170. A direct method and its application to a linear hydromagnetic stability problem, ROMAI J., **1**, 1 (2005), 67-76. (with L. Palese, A. Redaelli)

171. Sets governing the phase portrait (approximation of the asymptotic dynamics), ROMAI J., **1**, 1 (2005), 83-94. (with S.-C. Ion)

172. Application of the direct method to a microconvection model, Acta Universitatis Apulensis, Alba Iulia, Mathematics - Informatics**, 10** (2005), 131-142. ZB 1113.65076, MR 2240129. (with I. Dragomirescu)

173. Normal forms and unfoldings: a comparative study, Sci. Annals of UASVM Iaşi, **48,** 2 (2005), 15-26 (with E. Codeci)

174. Bifurcation of planar vector fields. I. Normal forms "at the point". One zero eigenvalue, Bul. Şt. Univ. Piteşti, Ser. Mec. Apl., **1 (12)** (2005), 41-47. (with D. Sârbu)

175. Bifurcation of planar vector fields. II. Normal forms "at the point". Hopf and cups cases, Bul. Şt. Univ. Piteşti, Ser. Mec. Apl., **1 (12)** (2005), 49-55. (with D. Sârbu)

176. Continuity of characteristics of a thin layer flow driven by a surface tension gradient, ROMAI J., **1**, 2 (2005), 11-16. (with E. Borsa)

2006

177. Some results on dynamics generated by Bazykin model, Atti Accad. Peloritana dei Pericolanti, Scienze FMN, **84,** C1A0601003 (2006), 1-10. (with R. - M. Georgescu)

178. A linear magnetic Bénard problem with tensorial electrical conductivity, Bolletttino U.M.I. (8) **9**-B (2006), 197-214. (with L. Palese, A. Redaelli)

179. Further study of a microconvection model using the direct method, ROMAI J., **2**, 1 (2006), 77-86. (with I. Dragomirescu)

180. Bifurcations and dynamics in the lymphocytes-tumor model, The 7th Congress of SIMAI 2004, Venezia, Italy, 2004. International Conference on Mathematical Models and Methods in Biology and Medicine -MMMBM 2005, Bedlewo, Poland, May 29-June 3, 2005, trimisa la Atti Accad. Peloritana dei Pericolanti, Scienze FMN. (with M. Trifan)

181. Asymptotic waves from the point of view of double-scale method, Atti Accad. Peloritana dei Pericolanti, Scienze FMN, **84,** C1A0601005 (2006), 1-9 (with L. Restuccia)

182. Linear stability bounds in a convection problem for variable gravity field, Bul. Acad. St. Rep. Moldova, Mat., 3(2006), 51-56. (with I. Dragomirescu)

183. Neutral manifolds in a penetrative convection problem. I. Expansions in Fourier series of the solutions, Sci. Annals of UASVM Iasi, **49**, 2 (2006), 19-32. MR 2300509 (with A. Labianca).

184. Mathematical models in biodynamics, Sci. Annals of UASVM Iasi, **49**, 2 (2006), 361-371.

2007

185. Analytical versus numerical results in a microconvection problem, Carpathian J. Math., **23**, 1-2 (2007), 81-88. MR 2305839 (with F.-I. Dragomirescu)

186. Relaxation oscillations and the "canard" phenomenon in the FitzHugh-Nagumo model, Cap. **4** in Recent Trends in Mechanics, **1,** Ed. Academiei Romane, Bucuresti, 2007, 82-106 (with M.-N. Popescu, Gh. Nistor, D. Popa)

187. Dynamical approach in biomathematics, ROMAI J., **2,** 2 (2007), 63-76. (with L. Palese, G. Raguso)

2008

188. A closed-form asymptotic solution of the FitzHugh-Nagumo model, Bul. Acad. St. Rep. Moldova, Seria Mat., **2** (2008), 24-34. (with Gh. Nistor, M.-N. Popescu, D. Popa)

189. Application of two spectral methods to a problem of convection with uniform internal heat source, Journal of Mathematics and Applications, 30 (2008), 43-52. (with F.-I. Dragomirescu)

190. A linear magnetic Bénard problem with Hall effect. Application of the Budiansky-DiPrima method, Trudy Srednevoljckogo Matematichescogo Obshchestva, Saransk, **10**, 1 (2008), 294-306. (with L. Palese).

191. Determination of asymptotic waves in Maxwell media by double-scale method, Technische Mechanik, **28**, 2 (2008), 140-151. (with L. Restuccia)

192. Approximate limit cycles for the Rayleigh model, ROMAI J. **4**, 2 (2008), 73-80. (with M. Sterpu, P. Băzăvan)

193. Lyapunov stability analysis in Taylor-Dean systems, ROMAI J. **4**, 2 (2008), 81-98. (with F.-I. Dragomirescu)

2009

194. Further results on approximate inertial manifolds for the FitzHugh-Nagumo model, Atti dellÕ Accad. Peloritana dei Pericolanti, Scienze FMN, vol LXXXVII, nr. 2. (2009). (with C.-S. Nartea)

195. On the stability bounds in a problem of convection with uniform internal heat source, arXiv. 0812.0517V1 [math-ph], Atti dell' Accad. Peloritana dei Pericolanti, Scienze FMN, accepted (with F.-I. Dragomirescu)

196. Degenerated Bogdanov-Takens bifurcations in an immuno-tumor model, Atti dell' Accad. Peloritana dei Pericolanti, Scienze FMN, vol LXXXVII, nr. 1. (2009). (with M. Trifan)

197. Stability criteria for quasigeostrofic forced zonal flows ; I. Asymptotically vanishing linear perturbation energy, ROMAI J. **5**, 1 (2009). 63-76 (with L. Palese)

198. Linear stability results in a magnetothermoconvection problem, An St. Univ. Ovidius Constanta, **17,** 3 (2009), 119-129. (with F.-I. Dragomirescu)

199. Polynomial based methods for linear nonconstant coefficients eigenvalue problems, Proceedings of the Middle Volga Mathematical Society, **11**, 2 (2009), 33-41. (with F.-I. Dragomirescu)

200. Thermodynamics of fluids, SIMAI e-Lecture Notes, Vol. **2** (2009), 1-26.

2011

201. On the nonlinear stability of a binary mixture with chemical surface reactions**,** Mathematics and its Applications, **3,** 1 (2011), 106-115 (with L. Palese)

202. An application of double-scale method to the study of non-linear dissipative waves in Jeffreys media, Mathematics and its Applications, **3,** 1 (2011), 116-134 (with L. Restuccia)

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

ATTRACTORS OF THE PERIODICALLY FORCED RAYLEIGH SYSTEM*

Petre Băzăvan†

Abstract

The autonomous second order nonlinear ordinary differential equation (ODE) introduced in 1883 by Lord Rayleigh, is the equation which appears to be the closest to the ODE of the harmonic oscillator with dumping.

In this paper we present a numerical study of the periodic and chaotic attractors in the dynamical system associated with the generalized Rayleigh equation. Transition between periodic and quasiperiodic motion is also studied. Numerical results describe the system dynamics changes (in particular bifurcations), when the forcing frequency is varied and thus, periodic, quasiperiodic or chaotic behaviour regions are predicted.

MSC: 34K28, 37C50, 37C70, 37G15, 37M20.

keywords: Periodic and chaotic attractors, bifurcations, Poincaré map and sections, Lyapunov exponents, periodic and quasiperiodic motion.

1 Introduction

The nonautonomous second order nonlinear ODE with time dependent sinusoidal forcing term, given by Diener [1979, 1],

$$\varepsilon\,\ddot{x} + \frac{\dot{x}^3}{3} - \dot{x} + ax = g\sin\omega t, \tag{1}$$

*Accepted for publication on December 16, 2010.

†bazavan@yahoo.com Department of Informatics, University of Craiova

is a generalisation of the Rayleigh equation $\ddot{x} + \frac{\dot{x}^3}{3} - \dot{x} + x = 0$ [Diener, 1979, 1]. Here, $x : \mathbf{R} \to \mathbf{R}$, $x = x(t)$ is the unknown function and the dot over x stands for the differentiation with respect to t. The control parameters are ε, a, g (forcing amplitude) and and ω (forcing frequency).

Some aspects concerning ščanardš bifurcations are analyzed in [Diener, 1979, 1] and [Diener, 1979, 2] for the periodically forced generalization of Rayleigh equation (1). From mathematical perspective the nonautonomous system of nonlinear ODEs associated with this equation is one of a class of periodically forced nonlinear oscillators, as the van der Pol (VP) and Bonhoeffer van der Pol (BVP) systems are. The behaviour of these systems was much numerically investigated in [Flaherty and Hoppensteadt, 1978], [Mettin et al., 1993] and [Barns and Grimshaw, 1997], due to their applications in electronics and physiology.

With (1), the two-dimensional non-linear non-autonomous system of ODEs

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -\frac{a}{\varepsilon} x_1 + \frac{1}{\varepsilon}\left(x_2 - \frac{x_2^3}{3}\right) + \frac{g}{\varepsilon}\sin\omega t, \end{cases} \tag{2}$$

and the three-dimensional nonlinear autonomous system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -\frac{a}{\varepsilon} x_1 + \frac{1}{\varepsilon}\left(x_2 - \frac{x_2^3}{3}\right) + \frac{g}{\varepsilon}\sin x_3, \\ \dot{x}_3 = \omega mod 2\pi, \end{cases} \tag{3}$$

are associated. A three-dimensional dynamical system with phase space $\mathbf{R^2} \times \mathbf{S^1}$ can be associated with (3). In [Sterpu et al., 2000], for the unforced case $g = 0$, the existence of a unique limit cycle for the dynamical system associated with the system,

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -\frac{a}{\varepsilon} x_1 + \frac{1}{\varepsilon}\left(x_2 - \frac{x_2^3}{3}\right), \end{cases} \tag{4}$$

for the case $a \cdot \varepsilon > 0$, is proved.

Therefore, the system (3) without periodic forcing ($g = 0$) exhibits a natural oscillation and we consider a sinusoidal forcing imposed on it ($g \neq 0$). Fixing the parameters ε, a, and g, as ω increases away from zero, the interaction between the frequencies of these two oscillations determines the resulting dynamics. Periodic as well as chaotic motion may occur.

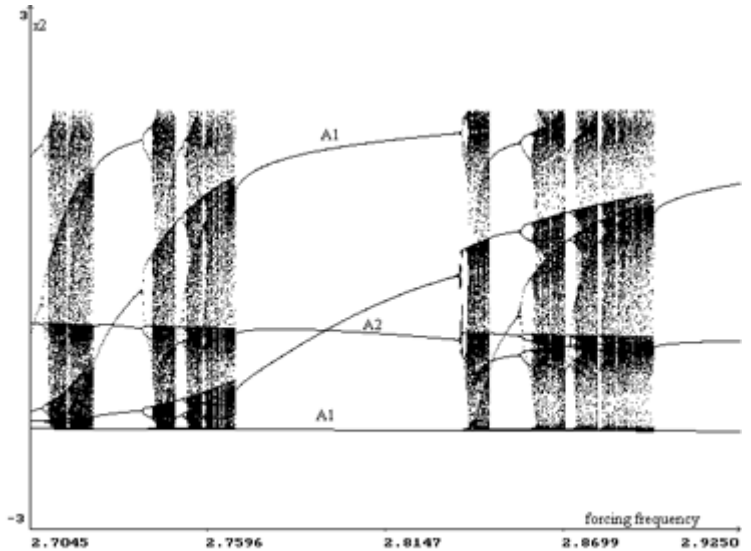


Figure 1: Bifurcation diagram for parameters ε=0.1250, a=0.5, g=0.6666 and $2.7045 \leq \omega \leq 2.9250$.

The lack of equilibria and the great number of parameters make the study of such a system difficult. Numerical methods often provide a useful and sometimes the only tool for study.

We intend to establish ω intervals for which specific behaviour concerning the attractors of the system (3) could be expected. By logistic reasons we investigated a region in the four-dimensional parameter space (ε, a, g, ω) given by $0 < \varepsilon \leq 1$, $0 < a \leq 1$, $0 < g \leq 1$, $2.7045 \leq \omega \leq 2.9250$ in case of Sec. 3 and $0 < \varepsilon \leq 1$, $0 < a \leq 1$, $1 < \omega \leq 3$, $0 < g \leq 2$ in case of Sec. 4.

The diagnostics used to establish structural changes of the system (3) involve representations of solutions in the phase space $\mathbf{R^2} \times \mathbf{S^1}$, time series, Poincaré sections at intervals of forcing period $\frac{2\pi}{\omega}$, bifurcation diagrams with $\omega - x_2$ coordinates, evaluations of the eigenvalues of the linearized Poincaré map-matrix, evaluations of the Lyapunov exponents. All the numerical computations were carried out through the application of a variable step-size four order Runge-Kutta method [Băzăvan, 1999]. The 3D-representation uses a centre projection [Băzăvan, 1994].

The bifurcation diagram plotted in Fig. 1, for the case $\varepsilon = 0.1250$, $a = 0.5$, $g = 0.6666$ and ω in the interval $2.7045 \leq \omega \leq 2.9250$ shows the typical system behaviour which will be interpreted in the next sections.

The mathematical model used in our numerical study is presented in Sec. 2. The Sec. 3 is concerned with the numerical study of alternating periodic and chaotic attractors in the behaviour of the system (3). Numerical results in Sec. 4 are concerned with the proof of the existence of the quasiperiodic motion and the study of the transition from quasiperiodic to periodic motion in the system (3).

2 The mathematical model

In order to present the mathematical model used in the numerical study from Secs. 3 and 4, we shortly write (3) in the form

$$\dot{x} = f(x), \tag{5}$$

where f is defined on the $\mathbf{R^2} \times \mathbf{S^1}$ cylinder.

We define the Poincaré map as follows. Let

$$\sum = \left\{ (x_1, x_2, x_3) \in \mathbf{R^2} \times \mathbf{S^1}, \mathbf{x_3} = \mathbf{0 mod} \frac{\mathbf{2}\pi}{\omega} \right\}$$

be a surface of section [Băzăvan, 2001], which is transversally crossed by the orbits of (5). The Poincaré map $P : \sum \to \sum$ is defined by

$$P(\mathbf{x}_0) = \mathbf{x}(t, \mathbf{x}_0) = \int_0^{\frac{2\pi}{\omega}} f(\mathbf{x}(t, \mathbf{x}_0))\, dt, \tag{6}$$

where $x_0 \in \sum$ and $x(t, x_0)$ is the solution of the Cauchy problem $x(0) = x_0$ for (5). We denote by P^n the n-times iterated map.

Let $\xi(t, x_0)$ be a periodic solution of (5) with period $T = n \cdot \frac{2\pi}{\omega}$, lying on a closed orbit and consider the map P of the initial point x_0. Then, to this closed orbit an n-periodic orbit of P corresponds. Numerically, the period T (i.e. n from the expression of T) can be determined by integrating Eq. (5) with the initial condition x_0 and sampling the orbit points $x_k = P(x_{k-1})$, $k \geq 1$ at discrete times $t_k = k \cdot \frac{2\pi}{\omega}$, until $P^k(x_0) = x_0$. Then, $n = k$ [Băzăvan, 2001].

The stability discussion of the periodic orbit $\xi(t, x_0)$ is reduced to the stability discussion of the fixed point x_0 of P^n, i.e. $P^n(x_0) = x_0$. The linear stability of the n-periodic orbit of P is determined from the linearized-map matrix DP^n of P^n. Using the Floqet theory [Reithmeier, 1991], [Glendinning, 1995] the matrix DP^n of P^n can be obtained by integrating the linearized system (5) for a small perturbation $y \in \mathbf{R^2} \times \mathbf{S^1}$. The time history of the initial perturbation $y(0) = y_0$ is described by the linearized ODE around the periodic solution ξ.

The stability of the periodic solution $\xi(t, x_0)$ is determined by the eigenvalues of the matrix DP^n [Reithmeier, 1991], [Glendinning, 1995], [Kuznetsov, 1998]. We note that one of the eigenvalues of this matrix always equals 1 [Glendinning, 1995], and that the remained two eigenvalues, also called the Poincaré map multipliers, influence the stability. We denote these eigenvalues by λ_1 and λ_2.

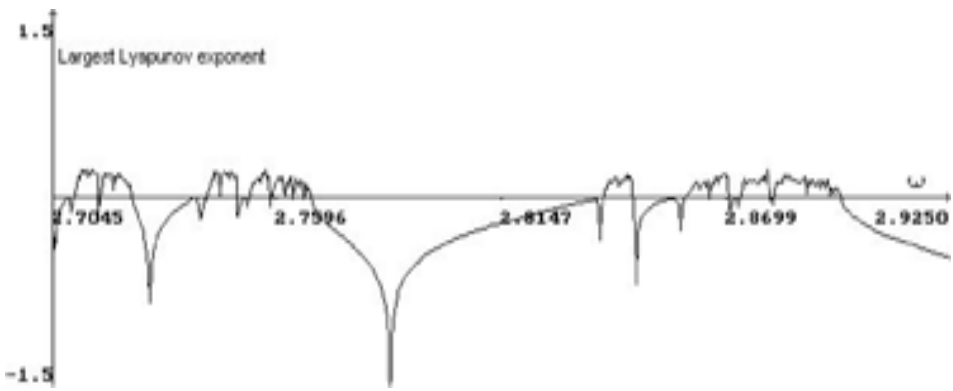


Figure 2: The largest Lyapunov exponent for (3), for parameter values ε=0.1250, a=0.5, g=0.6666 and $2.7045 \leq \omega \leq 2.9250$.

3 Periodic and chaotic attractors

In this section, by varying the parameter ω and keeping constant ε, a and g we study bifurcations associated with changes of stability in the periodically forced Rayleigh system (3).

The multipliers of the Poincaré map P^n, computed for $\varepsilon = 0.1250$, $a = 0.5$, $g = 0.6666$ and various ω values in the interval $2.7045 \leq \omega \leq 2.9250$, give information about the stability changes of an n-periodic orbit of (3) for which the map P is associated (see Sec. 2). Thus, the periodic orbit is stable only if $|\lambda_{1,2}| < 1$, [Reithmeier, 1991], [Glendinning, 1995], [Kuznetsov, 1998]. If, for a critical ω value, the multipliers satisfy $\lambda_1 = -1$, $-1 < \lambda_2 < 0$, [Reithmeier, 1991], [Glendinning, 1995], [Kuznetsov, 1998], the periodic orbit loses its stability through a period-doubling bifurcation. The motion becomes chaotic if, monotonically increasing ω, for sufficiently values, this process is repeated. This period doubling sequence leading to a chaotic state was reported in [Mettin, et al., 1993], [Barnes and Grimshaw, 1997] and [Sang-Yoon and Bumbi, 1998] for VP and BVP oscillators and inverted pendulum respectively. We also note that the reverse process can occur for the case of an unstable orbit. That is, when a multiplier λ of an unstable orbit increases through -1 the orbit becomes stable via period-doubling bifurcations.

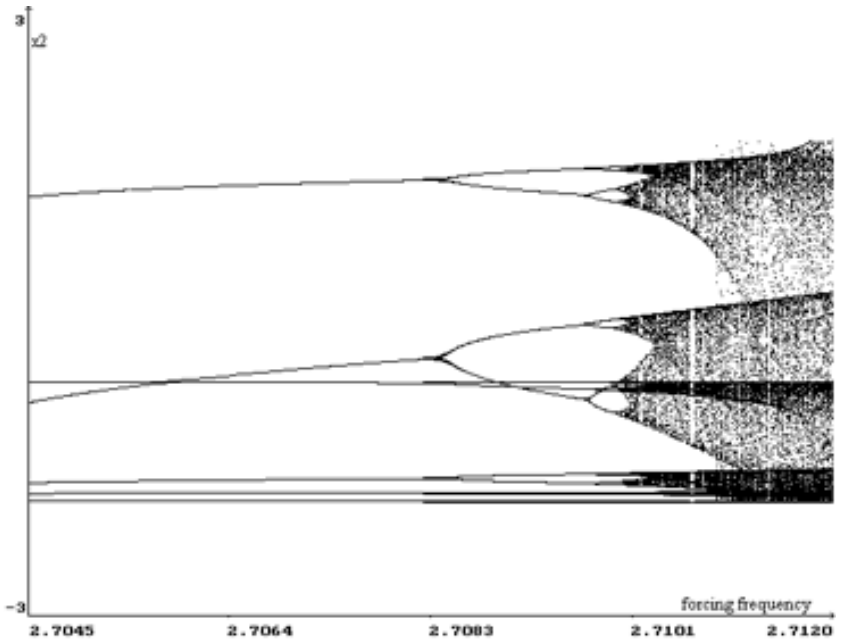


Figure 3: Bifurcation diagram for parameter values ε=0.1250, a=0.5, g=0.6666 and $2.7045 \leq \omega \leq 2.7120$.

As Fig. 1 shows, the system (3) exhibits the mentioned period-doubling sequences. Obvious chaotic regions interrupt periodic windows and then, chaotic attractors replace periodic attractors due to a destabilisation process through a period-doubling sequence. The reverse process, the stabilisation one, determines that periodic attractors replace chaotic attractors [Băzăvan, 2001].

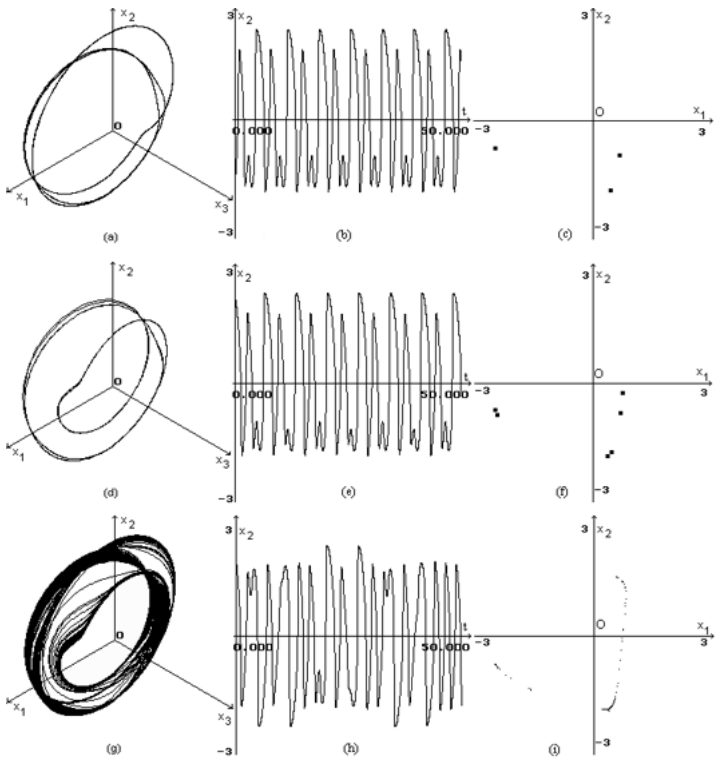


Figure 4: Closed trajectories, time series and Poincaré sections for system (3).

In order to ascertain these alternating regular and chaotic regions, the largest Lyapunov exponent measuring the convergence or divergence of neighbouring trajectories [Ott, 1993], [Barnes and Grimshaw, 1997] was plotted in Fig. 2 for the same parameter values as in Fig. 1. Negative values of this exponent correspond to periodic windows and positive values to chaotic regions.

In Fig. 3, which is a magnification of the bifurcation diagram in Fig. 1, for $2.7045 \leq \omega \leq 2.7120$, the typical route to chaotic state through a period-doubling sequence is more clearly seen. For $2.7045 \leq \omega < 2.7083$ two period-3 attractors are present.

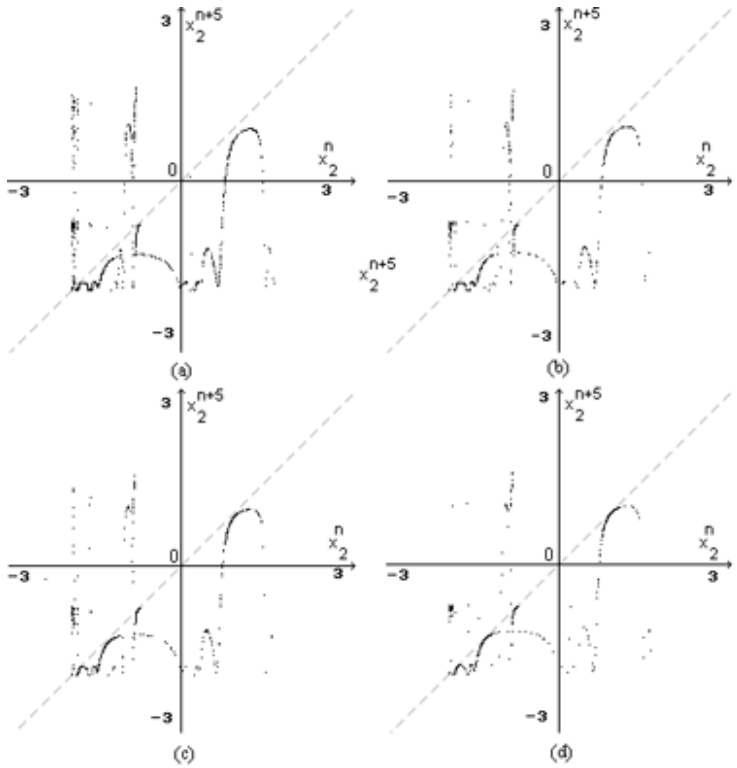


Figure 5: The points $\mathbf{X}_{n+5} = P^5(\mathbf{X}_n)$ for parameter values (a) ω=2.7225, (b) ω=2.7230, (c) ω=2.7235, (d) ω=2.7240.

The simultaneous presence of two attractors and the "jump" of the trajectories from one attractor to the other are characteristic to this system. Phase space with one of these period-3 solutions is represented on an invariant torus in Fig. 4a for $\omega = 2.7045$. For the solution in Fig. 4a, corresponding time series and Poincaré section with the three intersecting points are plotted in Figs. 4b-c. At $\omega \approx 2.7083$ the function curves split and the two solutions double their period as shows Fig. 3. The doubled periodic orbit, correspond-

ing to those from Fig. 4a, is represented in Fig. 4d for $\omega = 2.7090$. From the time series and the Poincaré section, plotted in Figs. 4e-f, the period six of the limit cycle is obvious.

The first period-doubling bifurcation at $\omega \approx 2.7083$ is followed by many subsequent period-doubling bifurcations. The length of the intervals of ω between these bifurcations decreases. Using magnifications of bifurcation diagram in Fig. 3, smaller ω step (i.e. 10^{-6}) and computing the $\lambda_{1,2}$ multipliers, for this period-doubling cascade the first five terms of the Feigenbaum progression $\frac{\omega_i - \omega_{i-1}}{\omega_{i+1} - \omega_i}$, [Kuznetsov, 1998], were estimated : 5.25, 5.18, 4.95, 4.81 and 4.72 [Băzăvan, 2001]. The convergence to the universal constant 4.6692 of this decreasing sequence is followed.

For $2.7106 < \omega < 2.7240$ the behaviour of the system is chaotic. The chaotic attractor, corresponding time series and Poincaré section are represented in Figs. 4g-i for $\omega = 2.7120$. At this ω value the largest Lyapunov exponent was computed to be 0.1812 [Băzăvan, 2001] providing the chaotic state of the system. As Fig. 1 shows, for $\omega \approx 2.7240$, the chaotic attractor is replaced by a period-5 attractor.

In order to illustrate this change from a chaotic attractor to a periodic attractor, the sequences of x_2 coordinates of the points $\mathbf{X}_{n+5} = P^5(\mathbf{X}_n)$ are plotted in Figs. 5a-d [Băzăvan, 2001]. For $\omega = 2.7225$ the diagonal $x_2^{n+5} = x_2^n$ is intersected in three separate locations. Here x_2^n represents the x_2 coordinate of the point $\mathbf{X}_n$. A channel between the diagonal and the return map curve is observed. As ω increases, the return map curve approaches the diagonal and at $\omega = 2.7240$ it is tangent in five distinct locations. A saddle-node bifurcation is encountered. The chaotic attractor is abruptly destroyed and replaced by a period-5 attractor. Note that, as the ω parameter increases, the density of the return points grows in the regions of the future attractor and diminishes in the other ones. This measure of the return points changes continuously with the continuous variation in the control parameter.

4 Transition between periodic and quasiperiodic motion

The dynamical system associated with (3) involves the interaction between two periodic motions, each with a different frequency. When the ratio of the frequencies is irrational the dynamical system behaves in a manner which is

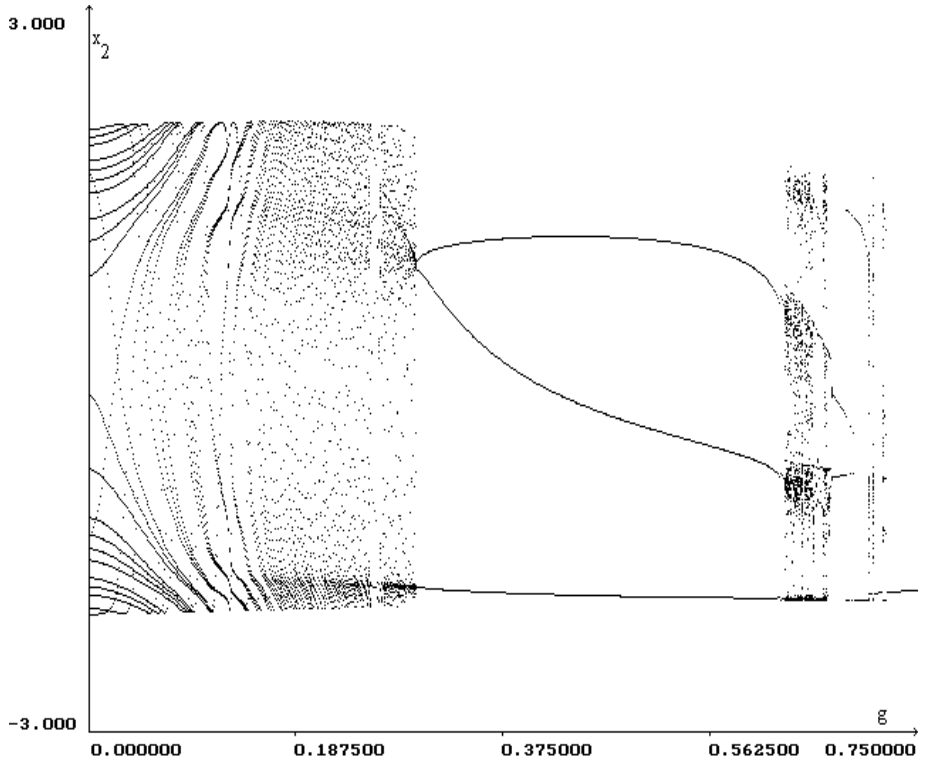


Figure 6: Bifurcation diagram for the dynamical system (3).

neither periodic or chaotic. This motion is called *quasiperiodic*. More precisely, the natural periodic motion, studied in [16] for the unforced case is modulated by a second periodic motion given by the sinusoidal term when $g > 0$. The system behaves in a manner with the motion never quite repeating any previous motion. This behaviour is generically followed by the system locking into a periodic motion, as the control parameter for the system is varied [18].

In our numerical study we investigated the region

$$\varepsilon = 0.125, \qquad a = 0.5, \qquad \omega = 2.84, \qquad 0 < g \leq 0.75. \tag{7}$$

An overview of the numerical results which typify the system is given by the bifurcation diagram in Fig. 6.

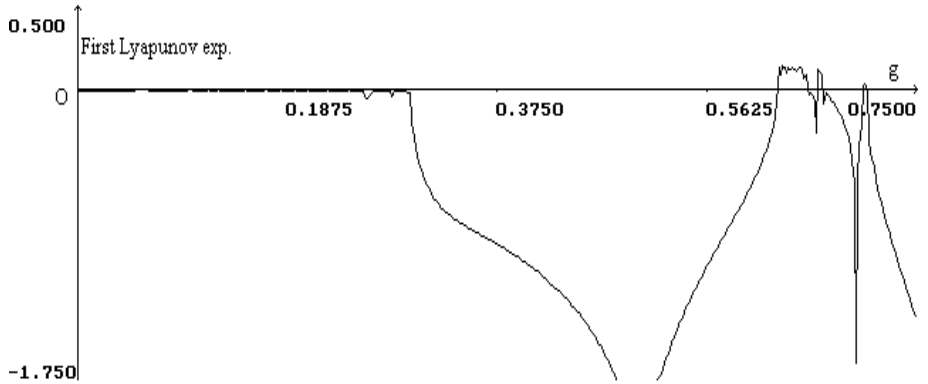


Figure 7: The first Lyapunov exponent for the dynamical system (3).

In the first part of the subinterval $0 < g < 0.3$ we observe an apparent regularity of the return points. This region which can indicate a quasiperiodic or chaotic behaviour is followed by a region with clear periodic motion. This last region is interrupted by short chaotic regions. We prove the existence of the quasiperiodic behaviour in two ways.

The first argument is the first Lyapunov exponent value. Recall that a leading Lyapunov exponent of zero verifies quasiperiodic behaviour [18].

Figure 7 is a graph of the control parameter (the forcing amplitude g) against the first Lyapunov exponent for the same parameter range as the bifurcation diagram of Fig. 6. In the interval $0 < g < 0.3$ the exponent was consistently within -0.01 of 0. This is the first numerical confirmation of the quasiperiodic behaviour.

The intersection points of the trajectories of the system (3) with the associated Poincaré section represent the second argument. At $g_1 = 0.07$ the section is represented in the Figure 8a.

The drift ring is associated with quasiperiodic motion. Integrating with a large period, the curve does not modify the shape. The fact that the points are situated on a closed curve and the constant shape related to the integration time confirm the quasiperiodic behaviour [18].

In proportion as g increases in the interval $0 < g < 0.3$ the return points remain on the same curve but the density increases markedly in some locations (Fig. 8b for $g_2 = 0.25$). At $g_3 = 0.3$ there are only three intersection points in the Poincaré section (Fig. 8c) and on the bifurcation diagram the

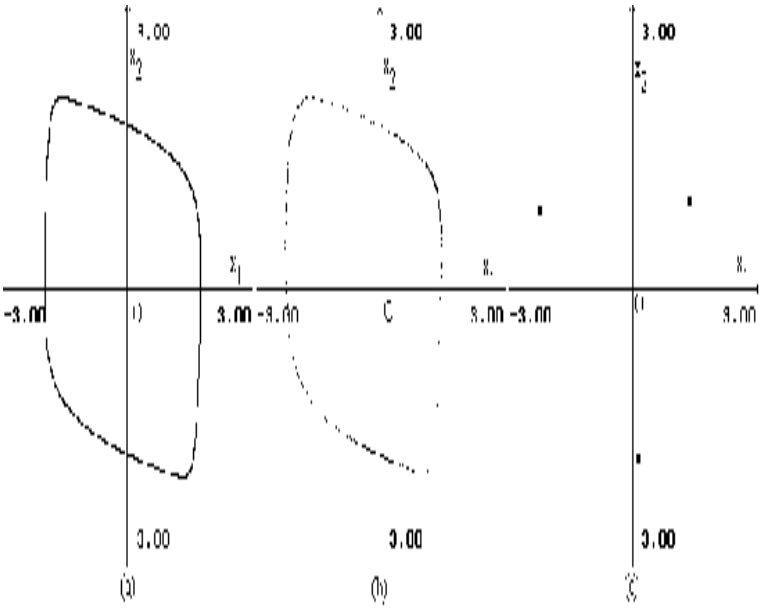


Figure 8: Poincaré sections for the dynamical system (3).

quasiperiodic region is replaced by a periodic window. The motion changes from quasiperiodic to periodic, with the emergence of a period-3 attractor. This is due to the saddle-node bifurcation of the Poincaré map P^3,

$$x_{n+3} = P^3(x_n), \quad x_0 \in \mathbf{R^2} \times \mathbf{S^1}, \quad \mathbf{n} \geq \mathbf{0}.$$

We numerically prove this fact. We use the projection of the graph of P^3 on the plane (y_n, y_{n+3}), $n \geq 0$, where we denote by y the x_2 coordinate of the point $x \in \mathbf{R^2} \times \mathbf{S^1}$.

In Figure 9a for $g_4 = 0.07$, when the motion is quasiperiodic, there are two intersection points of P^3 with the diagonal $y_n = y_{n+3}$. At the intersection the magnitude of the slope not equals 1. As g increases the curve approaches the diagonal in other locations (Fig. 9b for $g_5 = 0.28$). These locations suggest the imminent tangential intersections. At $g_6 = 0.2961$ there are three tangential intersections (Fig. 9c) and we have a saddle-node bifurcation of the map P^3. When $g_7 = 0.3$ (Fig. 9d) the graph of the map P^3 is a single point which is situated on the diagonal. This fact confirms the existence of the period-3 attractor.

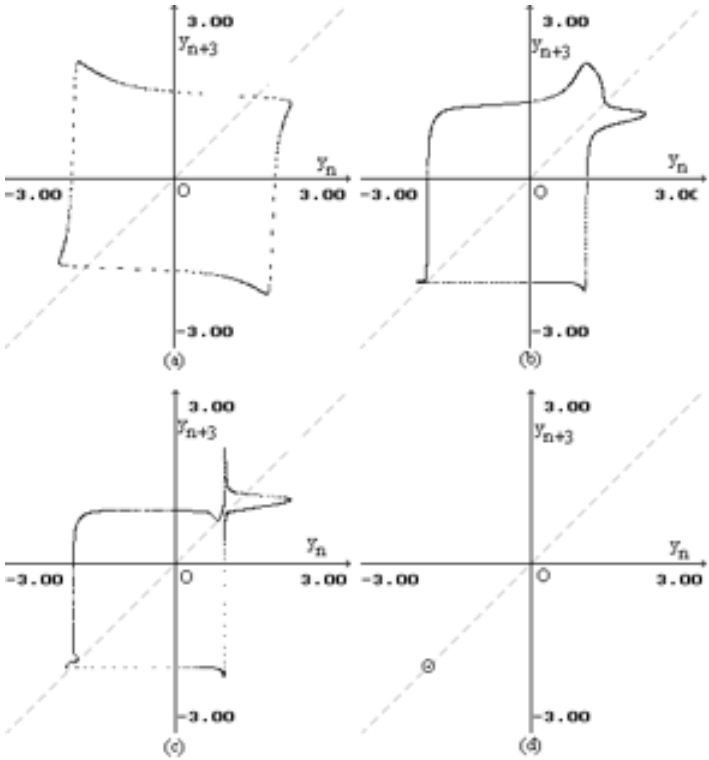


Figure 9: The Poincaré map P^3 associated with the dynamical system (3).

Conclusions

The numerical study in this paper shows that the periodically forced Rayleigh system possesses a lot of phenomena encountered in many other nonlinear systems. Some of them as period-doubling and saddle-node bifurcations, alternating periodic and chaotic attractors, alternating periodic and quasiperiodic motion, simultaneous presence of more than one periodic attractors were outlined here.

References

[1] B. Barnes, R. Grimshaw. Numerical studies of the periodically forced Bonhoeffer van der Pol oscillator, Int. J. Bifurcation and Chaos, 7(12), 2653–2689, 1997.

[2] P. Băzăvan. *Problems of Computational Geometry in "Ray-tracing"*, Annals of University of Craiova, Romania, Math. Comp. Sci. Series XX, 80–88, 1994.

[3] P. Băzăvan. *The dynamical system generated by a variable step size algorithm for Runge-Kutta methods*, Int. J. Chaos and Theory and Applications, 4(4), 21–28, 1999.

[4] P. Băzăvan. *Tridimensional Computation Study of Attractors for van der Pol - type Equations*, Ph.D. thesis, Institute of Mathematics of Roumanian Academy, Bucharest, Romania, 2001.

[5] M. Diener. *Nessie et les canards*, IRMA, Strasbourg, 1979.

[6] M. Diener. *Quelques examples de bifurcations et les canards*, Publ. IRMA, 37, 1979.

[7] J.E. Flaherty, F.C. Hoppensteadt. *Frequency entrainment of a forced van der Pol oscillator*, Studies Appl. Math, 58, 5–15, 1978.

[8] P. Glendinning. *Stability, Instability and Chaos*, Cambridge, New York, 1995.

[9] E. Grebogi, E. Ott, J. Yorke. *Metamorphoses of basin boundaries in nonlinear dynamical systems*, Phys. Rev. Lett., 56(10), 1011–1014, 1986.

[10] J. Guckenheimer, P. Holmes. *Nonlinear oscillations, dynamical systems and bifurcations of vector fields*, Springer, New York, p.168, 1983.

[11] Y. Kuznetsov. *Elements of applied bifurcation theory*, Springer Verlag, New York, 1998.

[12] R. Mettin, U. Parlitz, W. Lauterborn *Bifurcation structure of the driven van der Pol oscillator*, Int. J. Bifurcation and Chaos, 3(6), 1529–1555, 1993.

[13] E. Ott. *Chaos in dynamical systems*, Cambridge University Press, 1993.

[14] E. Reithmeier. *Periodic Solutions of Nonlinear Dynamical Systems*, Springer Verlag, 1991.

[15] Kim Sang-Yoon, Uh. Bumbi. Bifurcations and Transitions to Chaos in an Inverted Pendulum, Electronic paper, 1998.

[16] M. Sterpu, P. Băzăvan. *Study on a Rayleigh equation*, Annals of University of Pitesti, Romania, Math. Comp. Sci. Series, 3, 429–434, 1999.

[17] M. Sterpu, A. Georgescu, P. Băzăvan. *Dynamics generated by the generalized Rayleigh equation II. Periodic solutions*, Mathematical Reports, 2000, 2(52), 367–378.

[18] G.L. Baker, J.P. Gollub. *Chaotic dynamics an introduction*, Cambridge University Press, 193-195, 1996.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

THE FLOW OF A PARTICULAR CLASS OF OLDROYD-B FLUIDS*

Ilie Burdujan†

Abstract

This paper deals with Taylor-Couette flow formation in a particular class of Oldroyd-B fluids filling the annular region between two infinitely long coaxial circular cylinders, due to a time-dependent axial shear applied on the outer surface of the inner cylinder. The obtained solution is presented as the sum of a related Newtonian solution and the specific non-Newtonian contribution. Afterwards, it was specialized to give the solution for second grade fluids and Maxwell fluids, as well. Some exact solutions for particular classes of Oldroyd-B fluids arise as limiting cases of our solution. These results were established as limiting cases of the solution of an initial-boundary problem in fractional derivatives which was obtained, in its turn, by using the Laplace and Hankel transformations.

MSC: 76A05

keywords: Taylor-Couette flow, Oldroyd-B fluid, Maxwell fluid, second grade fluid.

1 Introduction

Oldroyd-B model provides a simple linear viscoelastic model for dilute polymer solutions, based on the dumbbell model. A wide class of fluids, such as

*Accepted for publication on December 27, 2010.

†burdujan_ilie@yahoo.com Department of Mathematics, University of Agricultural and Veterinary Medicine "Ion Ionescu de la Brad" Iaşi, 700490, Romania

polymer solutions, petroleum products, oils, blood, etc., are non-Newtonian. Moreover, the non-Newtonian fluids arise in a large variety of industrial applications - such as chemical processes (e.g. the processing of synthetic fibres, foams), food industries, construction engineering and so on, what motivates the great interest in their study. Certainly, the analysis of the behavior of the fluid motion for non-Newtonian fluids is essentially more complex in comparison with that of Newtonian fluids. It is well known that for a wide class of flows of Newtonian fluids it is possible to give a closed form for their analytical solutions, while for non-Newtonian fluids such solutions are rarely found. On the other hand, some of the mathematical models do not fit well with experimental data. That is why some mathematical objects, obtained by placing some fractional derivatives instead of some time derivatives into the rheological constitutive equations that describe the rheological properties of some classes of materials, were tested. On this line we can quote the papers of Bagley [1], Friedrich [7], Makris and Constantinou [17], Glökle and Nonnenmacher [9], Mainardi [15], Mainardi and Gorenflo [16], Rossikhin and Y. A., Shitikova [19], [20] and so on; they had obtained results which are in a good agreement with experimental data. Unfortunately, as it was already remarked in [5], an initial-boundary problem for an equation with fractional derivatives (shortly, IBPEFD) is not necessarily the mathematical model for a real dynamical system, because the fractional derivatives have no always a tensorial character. Nevertheless, some of its limiting cases are the mathematical models for real phenomena. Therefore, it becomes important to solve such an IBPEFD because its solution gives the possibility to find the solutions for all its limiting cases, among them being the solutions of problems modelling real dynamical systems. For example, this is the case of limits for parameters which allow to avoid the presence of fractional derivative. In fact, the mathematical models for Newtonian fluids, ordinary Maxwell fluids, ordinary second grade fluids, ordinary Oldroyd-B fluids are limiting cases for the before mentioned IBPEFD.

Two important situations may arise in the limiting processes. A result of such a limit can be the disappearance of all fractional derivatives. As example, in the problem under consideration in the present paper (i.e., the IBPEFD [(7), (9), (10), (12)]), the Newtonian solution is obtained by making the relaxation time λ (and, necessarily, the retardation time λ_r) tends to zero. The second kind of results corresponds to the case when the orders of all fractional derivatives, here α or/and β, tend to 1; in this case the obtained

equation contains ordinary or partial derivatives only. This time the limiting process is considered in the sense of Schwartz's distribution theory with respect to some appropriately classes of testing functions. For example, the solution for ordinary Maxwell fluids is obtained from the before mentioned IBPEFD when $\lambda_r \to 0$ and $\alpha \to 1$.

These remarks will be used in what follows in order to find the exact solution for Taylor-Couette flow of an incompressible Oldroyd-B fluid in a circular pipe. More exactly, the main purpose of this paper is to provide exact solutions for the velocity field and the shear stress corresponding to the large class of unsteady flows of incompressible Oldroyd-B fluids between two infinite coaxial circular cylinders, one of them being subject to a time-dependent rotational shear stress. More exactly, by the suggestion given in [13], we study the case when in the boundary condition (8) we put $a = 2$, so that this paper can be considered as a continuation of paper [5].

To this end, into the governing equations, corresponding to an Oldroyd-B fluid in the absence of body forces and a pressure gradient in the flow direction, some time derivatives are replaced by fractional derivatives. The obtained mathematical object was named by Tong and Liu [22] the governing equations of an incompressible "generalized" Oldroyd-B fluid. After making the similar replacement in the initial-boundary conditions, an IBPEFD is obtained. The governing equations for an incompressible "generalized" Maxwell fluid or for a "generalized" second grade fluid are similarly obtained. The attribute "generalized" will be used here for designing the hypothetical fluids that would be characterized by such IBPEFDs.

The solution of IBPEFD [(7), (9), (10), (12)] is presented as a sum of the Newtonian solution and the corresponding non-Newtonian contribution. It can be easily specialized to give the similar solutions for the second grade and Maxwell fluids. As it was already remarked, the Newtonian solutions can be also obtained as limiting cases of general solutions. Furthermore, the non-Newtonian contributions to the general solutions have been expressed in terms of the time derivatives of a Newtonian solution. The exact expressions for the fluid velocity and the shear stress are obtained by the successive use of the methods of Hankel and Laplace transforms.

In the particular cases $a = 0$ and $a = 1$, this problem was already solved in [5]. The present paper solve this problem in case $a = 2$. That is why this paper is really a continuation of [5]. We try to make its reading as selfcontained as possible.

2 Model and basic equations

Recall that the Oldroyd-B model is a classical model for dilute solutions of polymers suspended in a viscous incompressible solvent. The Oldroyd-B model can be derived from microscopic principles by assuming a linear Hook's Law for the restoring force under distention of immersed polymer coils. In the recent years the Oldroyd-B fluids has gained a special place among the fluids of rate types. They contain as special cases the classical Newtonian fluids and the Maxwell fluids as well as the second grade fluids.

Let us consider an incompressible Oldroyd-B fluid at rest filling the annular region between two infinitely long coaxial circular cylinders of radii $R_1,\ R_2$ $(0 < R_1 < R_2)$. The outer cylinder is always at rest, while at time $t = 0^+$ the inner cylinder is suddenly set in rotation around its axis by a time-dependent shear stress.

The equations governing the unsteady motion of an incompressible fluid are

$$div\ \mathbf{V} = 0, \qquad \rho\ \frac{d\mathbf{V}}{dt} = div\ \mathbf{T},$$

where $\mathbf{V}$ is the velocity field, ρ the density, $\mathbf{T}$ the Cauchy *stress tensor* and d/dt the material time derivative.

The Cauchy *stress tensor* $\mathbf{T}$ for an incompressible Oldroyd-B fluid, is given by

$$\mathbf{T} = -p\mathbf{I} + \mathbf{S}, \qquad \mathbf{S} + \lambda\frac{D\mathbf{S}}{Dt} = \mu\left(\mathbf{A} + \lambda_r\frac{D\mathbf{A}}{Dt}\right), \tag{1}$$

where $-p\mathbf{I}$ is the *indeterminate spherical stress* (p is the *isotropic pressure*), $\mathbf{S}$ is the *extra-stress tensor*, $\mathbf{A}$ is the *first* Rivlin-Ericksen *tensor*, μ is the *dynamic viscosity* of the fluid, λ and $\lambda_r (< \lambda)$ are *material constants* (namely, the *relaxation time* and the *retardation time*, respectively), and the upper convected time derivatives are defined by

$$\frac{D\mathbf{S}}{Dt} = \frac{d\mathbf{S}}{dt} + (\mathbf{V}\cdot\nabla)\mathbf{S} - \mathbf{L}\mathbf{S} - \mathbf{S}\mathbf{L}^T, \qquad \frac{D\mathbf{A}}{Dt} = \frac{d\mathbf{A}}{dt} + (\mathbf{V}\cdot\nabla)\mathbf{A} - \mathbf{L}\mathbf{A} - \mathbf{A}\mathbf{L}^T. \tag{2}$$

Into above equation (2), ∇ is the gradient operator, $\mathbf{L}$ denotes the velocity gradient and the superscript T indicates the transpose operation. This time, the body forces have been neglected.

Since the motion is axial symmetric we shall use the cylindrical coordinates $(r,\ \theta,\ z)$. That is why, for the problem under consideration, we assume a velocity field $\mathbf{V}$ and an extra-stress tensor $\mathbf{S}$ of the form

$$\mathbf{V} = \mathbf{V}(r,t) = \omega(r,t)\mathbf{e}_\theta, \qquad \mathbf{S} = \mathbf{S}(r,t) \tag{3}$$

where $\mathbf{e}_\theta$ is the unit vector along the θ-direction of the cylindrical coordinate system. For such flows the constraint of incompressibility is automatically satisfied. Furthermore, if the fluid is at rest up to the moment $t = 0$, i.e.

$$\mathbf{V}(r,t) = \mathbf{0}, \quad \mathbf{S}(r,t) = \mathbf{0} \quad for\ t \leq 0, \tag{4}$$

then the governing equations for an Oldroyd-B fluid, in the absence of body forces and a pressure gradient in the flow direction, are given for $r \in (R_1, R_2)$, $t > 0$ by

$$\lambda\frac{\partial^2\omega(r,t)}{\partial t^2} + \frac{\partial\omega(r,t)}{\partial t} = \nu\left(1 + \lambda_r\frac{\partial}{\partial t}\right)\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right)\omega(r,t), \tag{5}$$

$$\left(1 + \lambda\frac{\partial}{\partial t}\right)\tau(r,t) = \mu\left(1 + \lambda_r\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial r} - \frac{1}{r}\right)\omega(r,t), \tag{6}$$

where $\tau(r,t) = S_{r\theta}(r,t)$ is the nonzero *shear stress*, $\nu = \mu/\rho$ is the *kinematic viscosity*, ρ is its constant *density*, while λ and λ_r are respectively the *relaxation* and *retardation times*. The system of equations (5)-(6) must be solved subject to the initial and boundary conditions

$$\omega(r,0) = \frac{\partial\omega(r,0)}{\partial t} = 0, \quad \tau(r,0) = 0, \tag{7}$$

respectively,

$$\left(1 + \lambda\frac{\partial}{\partial t}\right)\tau(R_1,t) = \mu\left(1 + \lambda_r\frac{\partial}{\partial t}\right)\left(\frac{\partial\omega(R_1,t)}{\partial r} - \frac{1}{R_1}\omega(R_1,t)\right) = ft^a, \tag{8}$$

for $r \in (R_1, R_2)$, $t > 0$ and $a \geq 0$.

By replacing some inner time derivatives by the fractional differential operators D_t^α and D_t^β $(0 < \beta \leq \alpha < 1)$, the governing equations (5) and (6) of an incompressible Oldroyd-B fluid become (see [22])

$$(1 + \lambda D_t^\alpha)\frac{\partial\omega(r,t)}{\partial t} = \nu(1 + \lambda_r D_t^\beta)\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right)\omega(r,t), \tag{9}$$

$$(1 + \lambda D_t^\alpha)\tau(r,t) = \mu(1 + \lambda_r D_t^\beta)\left(\frac{\partial}{\partial r} - \frac{1}{r}\right)\omega(r,t), \tag{10}$$

for $r \in (R_1, R_2),\ t > 0$; here the fractional derivatives are defined by [18]

$$D_t^p[f(t)] = \frac{1}{\Gamma(1-p)} \frac{d}{dt} \int_0^t \frac{f(\tau)}{(t-\tau)^p} d\tau, \qquad 0 < p < 1 \tag{11}$$

(where $\Gamma(\cdot)$ is EULER's Gamma function).

As before, by replacing in (8) the inner derivations with respect to t by the fractional differential operators D_t^α and D_t^β $(\beta \leq \alpha)$, we get

$$(1 + \lambda D_t^\alpha)\tau(R_1, t) = \mu(1 + \lambda_r D_t^\beta) \left(\frac{\partial \omega(R_1, t)}{\partial r} - \frac{1}{R_1} \omega(R_1, t) \right) = f t^a \tag{12}$$

for $t > 0,\ a \geq 0$. Consequently, an IBPEFD, consisting of equations [(9), (7), (12)], is associated with the model of the Taylor-Couette flow of an Oldroyd-B fluid in an annulus due to a time depending couple and characterized by simultaneously equations [(5), (7), (8)]. It will be solved by using the integral transform techniques. More exactly, the Laplace and finite Hankel transforms are used to change the IBPEFD [(9), (7), (12)] into an algebraic system.

Moreover, the equations (9) and (10) contain as limiting cases the governing equations of the so called (see [22]) "generalized" second grade and Maxwell models (i.e. the models obtained by replacing some inner time derivatives by some fractional differential operators in the governing equations of a second grade or a Maxwell fluid), as well as the ordinary Oldroyd-B, Maxwell and second grade models.

In this paper, we are especially interested in the case when the boundary condition corresponds to $a = 2$; the cases $a = 0$ and $a = 1$ were already analyzed in [5].

COMMENT. In order to ensure the dimensional consistency of equations (7) and (8), the material constants λ and λ_r must have necessarily the dimensions of t^α and t^β, respectively. Into several papers (e.g. [13]) the authors (correctly) used λ^α and λ_r^β instead of λ and λ_r. However, for simplicity, we shall keep the notations λ and λ_r (like [11] or [22]) having in mind their correct significations.

3 Exact solutions for the velocity field

In what follows, we shall use the modified Hankel transform, with respect to r, defined by means of the Bessel functions of index 1

$$B(r, r_n) = J_1(rr_n)Y_2(R_1r_n) - J_2(R_1r_n)Y_1(rr_n),$$

where $J_1(\cdot)$, $J_2(\cdot)$, $Y_1(\cdot)$ and $Y_2(\cdot)$ are Bessel functions (of index 1 and 2), $(r_n)_{n\in\mathbf{N}^*}$ is the increasing sequence of the positive roots of the transcedental equation $J_1(R_2x)Y_2(R_1x) - J_2(R_1x)Y_1(R_2x) = 0$ (i.e. $B(R_2, r_n) = 0$ for all $n \in \mathbf{N}^*$). We shall denote by $\omega_H(r_n, t)$ the image of $\omega(r, t)$ by the modified Hankel transform, defined by

$$\omega_H(r_n, t) = \int_{R_1}^{R_2} r\ \omega(r, t)\ B(r, r_n)\ dr\ . \tag{13}$$

Recall that the inverse of the modified Hankel transform (13) is defined by

$$f(r) = \frac{\pi^2}{2}\sum_{n=1}^{\infty}\frac{r_n^2J_1^2(R_2r_n)B(r, r_n)}{J_2^2(R_1r_n) - J_1^2(R_2r_n)}f_H(r_n) = \frac{\pi^2}{2}\sum_{n=1}^{\infty}r_n^2C_{Fn}f_H(r_n), \tag{14}$$

where $f_H(r_n)$ denotes the image of $f(r)$ by Hankel transform (13) and

$$C_{Fn} = \frac{J_1^2(R_2r_n)B(r, r_n)}{J_2^2(R_1r_n) - J_1^2(R_2r_n)} = D_{Fn}B(r, r_n). \tag{15}$$

By applying successively the Hankel transform (13) and the Laplace transform, the following expression for the velocity field $\omega(r, t)$ was obtained in [5]:

$$\begin{aligned}\omega(r, t) &= \omega_{N,a}(r, t) - \\ &-\frac{\pi f}{\rho}\Gamma(1 + a)\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\left(-\frac{\nu r_n^2}{\lambda}\right)^k r_nC_{Fn}\int_0^t F_k(s)e^{-\nu r_n^2(t-s)}ds,\end{aligned} \tag{16}$$

where

$$\begin{aligned}F_k(t) = \sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\Bigg[&G_{\alpha,\alpha_m-a,k+1}\left(-\frac{1}{\lambda}, t\right) + \\ &+\nu r_n^2\frac{\lambda_r}{\lambda}G_{\alpha,\beta_m-a,k+1}\left(-\frac{1}{\lambda}, t\right)\Bigg]\end{aligned} \tag{17}$$

with (see [14])

$$G_{a,b,c}(d,t)=\sum_{j=0}^{\infty}\frac{\Gamma(j+c)t^{a(j+c)-b-1}}{\Gamma(j+1)\Gamma(c)\Gamma[a(j+c)-b]}d^j=\mathcal{L}^{-1}\left\{\frac{q^b}{(q^a-d)^c}\right\}, \quad (18)$$

for $Re\ (ac-b)>0,\ \left|\frac{d}{q^a}\right|<1$, and

$$\omega_{N,a}(r,t)=\frac{\pi f}{\rho}\sum_{n=1}^{\infty}r_nC_{Fn}\int_0^t s^a e^{-\nu r_n^2(t-s)}ds, \quad (19)$$

represents the velocity field corresponding to a Newtonian fluid performing the same motion. More exactly, $\omega_{N,a}(r,t)$ for $R_1<r<R_2,\ t>0$ is the solution of the problem:

$$\begin{cases}\dfrac{\partial\omega_{N,a}(r,t)}{\partial t}=\nu\left[\dfrac{\partial^2\omega_{N,a}(r,t)}{\partial r^2}+\dfrac{1}{r}\dfrac{\partial\omega_{N,a}(r,t)}{\partial r}-\dfrac{1}{r^2}\ \omega_{N,a}(r,t)\right],\\ \omega_{N,a}(r,0)=0,\ \ \omega_{N,a}(R_2,t)=0,\\ \tau_{N,a}(R_1,t)=\mu\left[\dfrac{\partial\omega_{N,a}(R_1,t)}{\partial r}-\dfrac{1}{R_1}\omega_{N,a}(R_1,t)\right]=ft^a;\end{cases}$$

in last boundary condition, the presence of $\tau_{N,a}(R_1,t)$ can be ignored; it is just for helping us to motivate the form of the boundary condition. A special interest is for velocity field $\omega_N(r,t)=\omega_{N,0}(r,t)$. Recall that, it was proved in [5] that

$$\omega_N(r,t)=\omega_{N,0}(r,t)=\varphi_0(r)-\frac{\pi f}{\mu}\sum_{n=1}^{\infty}\frac{1}{r_n}C_{Fn}e^{-\nu r_n^2t}, \quad (20)$$

where

$$\varphi_0(r)=-\frac{f}{2\mu}\left(\frac{R_1}{R_2}\right)^2\left(\frac{R_2^2}{r}-r\right). \quad (21)$$

Further, Eq. (20) gives

$$\frac{\partial\omega_N(r,t)}{\partial t}=\frac{\pi f}{\rho}\sum_{n=1}^{\infty}r_nC_{Fn}e^{-\nu r_n^2t}$$

and consequently (19) becomes

$$\omega_{N,a}(r,t)=t^a*\partial_t\omega_N(r,t). \quad (22)$$

Then, for $a = 1$ we get

$$\omega_{N,1}(r,t) = \varphi_0(r)t + \varphi_1(r) + \frac{\pi f}{\mu\nu}\sum_{n=1}^{\infty}\frac{1}{r_n^3}C_{Fn}e^{-\nu r_n^2 t}, \tag{23}$$

where

$$\varphi_1(r) = \frac{A}{r} + Br + Cr^3 + E\ r\ ln\ r, \tag{24}$$

with

$$\begin{aligned} A &= -\frac{fR_1^4}{8R_2^2\mu\nu}(2R_2^2 - R_1^2), & C &= \frac{f}{8\mu\nu}\left(\frac{R_1}{R_2}\right)^2, \\ B &= \frac{fR_1^2}{8R_2^2\mu\nu}\left[4R_2^4\ ln\ R_2 - (R_2^2 - R_1^2)^2\right], & E &= -\frac{fR_1^2}{2\mu\nu}. \end{aligned} \tag{25}$$

Similarly, for $a = 2$, we get

$$\omega_{N,2}(r,t) = \varphi_0(r)t^2 + \varphi_1(r)t + \varphi_2(r) - \frac{2\pi f}{\mu\nu^2}\sum_{n=1}^{\infty}\frac{1}{r_n^5}D_{Fn}B(r,r_n)e^{-\nu r_n^2 t}, \tag{26}$$

where $\varphi_2(r)$ is the solution of the boundary problem

$$\begin{cases} \varphi_2''(r) + \frac{1}{r}\varphi_2'(r) - \frac{1}{r^2}\varphi_2(r) = \frac{1}{\nu}\varphi_1(r), \quad R_1 < r < R_2, \\ \varphi_2(R_2) = 0, \\ \varphi_2'(R_1) - \frac{1}{R_1}\varphi_2(R_1) = 0. \end{cases}$$

The solution of this last problem is:

$$\begin{aligned} \varphi_2(r) = &\frac{C_1}{r} + C_2 r + \\ &+\frac{1}{96}\left(48A\ ln\ r + 12B\ r^2 + 4C\ r^4 + 12E\ r^2 ln\ r - 9E\ r^2\right) r, \end{aligned} \tag{27}$$

where

$$\begin{aligned} C_1 = &-\frac{fR_1^6}{192\mu\nu^2 R_2^4}\left[12\ R_2^4 - \right. \\ &\left. -14R_1^2R_2^2 + 12R_2^4\ ln\ R_1 - 12R_2^4\ lnR_2 + 3R_1^4\right], \\ C_2 = &\frac{fR_1^2}{192\mu\nu^2 R_2^6}\ \left[15R_1^4R_2^4 - 14R_1^6R_2^2 + 12R_1^4R_2^4\ ln\ R_1 - \right. \\ &\left. -24R_1^4R_2^4\ ln\ R_2 + 3R_1^8 - 7R_2^8 + 24R_1^2R_2^6\ ln\ R_2 - 6R_1^2R_2^6\right]. \end{aligned} \tag{28}$$

Finally, having in mind the expression (20) of the Newtonian solution $\omega_N(r,t)$, it is easy to show that the general solution $\omega(r,t)$ can be written in a suitable form in terms of its time derivatives, namely

$$\begin{aligned}\omega(r,t) &= t^a * \partial_t \omega_N(r,t) - \\ &-\Gamma(1+a)\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+1}\omega_N(r,t) * G_{\alpha,\alpha_m-a,k+1}\left(-\frac{1}{\lambda},t\right) + \\ &+\frac{\lambda_r}{\lambda}\Gamma(1+a)\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+2}\omega_N(r,t) * G_{\alpha,\beta_m-a,k+1}\left(-\frac{1}{\lambda},t\right),\end{aligned} \tag{29}$$

where $\alpha_m = m\beta + \alpha - k - 1,\ \beta_m = m\beta + \beta - k - 2$. In case $a = 2$, it seem natural to express the general solution $\omega(r,t)$ in terms of $\omega_{N,2}(r,t)$ and its time derivatives. As

$$\begin{aligned}\frac{\partial^3\omega_{N,2}(r,t)}{\partial t^3} &= \frac{2\pi f}{\rho}\sum_{n=1}^{\infty} r_n D_{Fn} B(r,r_n) e^{-\nu r_n^2 t} = \\ &= \frac{2\pi f}{\rho}\sum_{n=1}^{\infty} r_n C_{Fn} e^{-\nu r_n^2 t} = 2\frac{\partial\omega_N(r,t)}{\partial t}\end{aligned}$$

it results

$$\begin{aligned}\omega(r,t) &= \omega_{N,2}(r,t) - \\ &-\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+3}\omega_{N,2}(r,t) * G_{\alpha,\alpha_m-2,k+1}\left(-\frac{1}{\lambda},t\right) + \\ &+\frac{\lambda_r}{\lambda}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+4}\omega_{N,2}(r,t) * G_{\alpha,\beta_m-2,k+1}\left(-\frac{1}{\lambda},t\right).\end{aligned} \tag{30}$$

4 Calculations of the shear stress

By using Eq. (29), the following form of the shear stress was already obtained in [5]:

$$\tau(r,t) = \tau_{N,a}(r,t) + \mu\Gamma(1+a)A\left(-\frac{1}{\lambda},t\right) * \partial_t\Omega_N(r,t) -$$

$$-\mu\Gamma(1+a)\sum_{k=0}^{\infty}\left[B_k\left(-\frac{1}{\lambda},t\right)*\partial_t^{k+1}\Omega_N(r,t)-\right.$$
$$\left.-\frac{\lambda_r}{\lambda}C_k\left(-\frac{1}{\lambda},t\right)*\partial_t^{k+2}\Omega_N(r,t)\right] \tag{31}$$

where

$$\Omega_N(r,t)=\frac{\partial\omega_N(r,t)}{\partial r}-\frac{1}{r}\omega_N(r,t), \tag{32}$$

$$\tau_{N,a}(r,t)=\mu t^a*\partial_t\Omega_N(r,t)=$$
$$=-\pi f\nu\sum_{k=0}^{\infty}\frac{r_n^2[J_2(rr_n)Y_2(R_1r_n)-J_2(R_1r_n)Y_2(rr_n)]}{J_2^2(R_1r_n)-J_1^2(R_2r_n)}\int_0^t s^a e^{-\nu r_n^2(t-s)}\,ds \tag{33}$$

represents the shear stress corresponding to a Newtonian fluid and

$$A\left(-\frac{1}{\lambda},t\right)=\frac{\lambda_r}{\lambda}R_{\alpha,\beta-a-1}\left(-\frac{1}{\lambda},t\right)-R_{\alpha,\alpha-a-1}\left(-\frac{1}{\lambda},t\right),$$

$$R_{\alpha,\beta}(a,t)=\sum_{k=0}^{\infty}\frac{a^k t^{(k+1)\alpha-\beta-1}}{\Gamma((k+1)\alpha-\beta)},\quad Re\ (\alpha-\beta)>0,\quad |at^\alpha|<1,$$

$$B_k\left(-\frac{1}{\lambda},t\right)=\frac{1}{\lambda^k}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\left[G_{\alpha,\alpha_m-a,k+1}\left(-\frac{1}{\lambda},t\right)+\right.$$
$$\left.+\frac{\lambda_r}{\lambda}G_{\alpha,\alpha_m+\beta-a,k+2}\left(-\frac{1}{\lambda},t\right)-G_{\alpha,\alpha_m+\alpha-a,k+2}\left(-\frac{1}{\lambda},t\right)\right],$$

$$C_k\left(-\frac{1}{\lambda},t\right)=\frac{1}{\lambda^k}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\left[G_{\alpha,\beta_m-a,k+1}\left(-\frac{1}{\lambda},t\right)+\right.$$
$$\left.+\frac{\lambda_r}{\lambda}G_{\alpha,\beta_m+\beta-a,k+2}\left(-\frac{1}{\lambda},t\right)-G_{\alpha,\beta_m+\alpha-a,k+2}\left(-\frac{1}{\lambda},t\right)\right].$$

Starting from Eq. (16), the following equivalent form of the shear stress is obtained

$$\tau(r,t)=\tau_{N,a}(r,t)+\mu\Gamma(1+a)A\left(-\frac{1}{\lambda},t\right)*\partial_t\Omega_N(r,t)+$$

$$+\pi f\nu\ \Gamma(1+a)\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}r_n\widetilde{C}_{Fn}(-\nu r_n^2)^k\left[B_k\left(-\frac{1}{\lambda},t\right)+\right.$$
$$\left.+\nu r_n^2\frac{\lambda_r}{\lambda}C\left(-\frac{1}{\lambda},t\right)\right]*e^{-\nu r_n^2 t}, \tag{34}$$

where

$$\widetilde{C}_{Fn}=\frac{J_2(rr_n)Y_2(R_1r_n)-J_2(R_1r_n)Y_2(rr_n)}{J_2^2(R_1r_n)-J_1^2(R_2r_n)}J_1^2(R_2r_n).$$

Making $a=0,\ 1$ and 2 into (31) and (34), the shear stresses corresponding to $f,\ ft$ and ft^2 into (7) are obtained. For instance, the shear stresses for the corresponding Newtonian solutions are

$$\tau_N(r,t)=\tau_{N,0}(r,t)=\tau_0(r)+\pi f\sum_{n=1}^{\infty}\widetilde{C}_{Fn}\ e^{-\nu r_n^2 t}, \tag{35}$$

$$\tau_{N,1}(r,t)=t\tau_0(r)+\tau_1(r)-\frac{\pi f}{\nu}\sum_{n=1}^{\infty}\widetilde{C}_{Fn}e^{-\nu r_n^2 t}, \tag{36}$$

$$\tau_{N,2}(r,t)=t^2\tau_0(r)+t\tau_1(r)+\tau_2(r)-\frac{\pi f}{\nu}\sum_{n=1}^{\infty}\widetilde{C}_{Fn}e^{-\nu r_n^2 t}, \tag{37}$$

where

$$\tau_0(r)=\frac{fR_1^2}{r^2},$$
$$\tau_1(r)=\frac{fR_1^2\left[(R_2^2-r^2)^2-(R_2^2-R_1^2)^2\right]}{8\nu R_2^2r^2},$$
$$\tau_2(r)=\mu[\varphi_2'(r)-\frac{1}{r}\varphi_2'(r)]=$$
$$=\frac{\mu}{48r^2}(-96C_1+24Ar^2+12Br^4+8Cr^6+12Er^4\ ln\ r-3Er^4).$$

5 Limiting cases

1. Making the limit of Eqs. (16), (29), (31) and (34) as $\lambda_r \to 0$, we get the similar solutions corresponding to the so-called "generalized" Maxwell fluids, namely

$$\begin{aligned}
&\omega(r,t) = \omega_{N,2}(r,t) - \\
&-2\frac{\pi f}{\rho}\sum_{n=1}^{\infty} r_n C_{Fn} \sum_{k=0}^{\infty}\left(-\frac{\nu r_n^2}{\lambda}\right)^k \int_0^t G_{\alpha,\gamma_k-2,k+1}\left(-\frac{1}{\lambda},t\right) e^{-\nu r_n^2(t-s)}\, ds = \\
&= \omega_{N,2}(r,t) - 2\sum_{k=0}^{\infty}\frac{1}{\lambda^k} G_{\alpha,\gamma_k-2,k+1}\left(-\frac{1}{\lambda},t\right) * \partial_t^{k+1}\omega_{N,2}(r,t),
\end{aligned} \tag{38}$$

$$\begin{aligned}
\tau(r,t) &= \tau_{N,2}(r,t) - 2\mu\int_0^t \partial_s\Omega_N(r,s) R_{\alpha,\alpha-31}\left(-\frac{1}{\lambda},t-s\right)\, ds + \\
&+2\pi f\nu \sum_{n=1}^{\infty}\sum_{k=0}^{\infty} r_n \widetilde{C}_{Fn}\left(\frac{-\nu r_n^2}{\lambda}\right)^k \left[G_{\alpha,\gamma_k-2,k+1}\left(-\frac{1}{\lambda},t\right) - \right. \\
&\left. -G_{\alpha,\gamma_k+\alpha-2,k+2}\left(-\frac{1}{\lambda},t\right)\right] * e^{-\nu r_n^2 t} = \\
&= \tau_{N,2}(r,t) - 2\mu\partial_t\Omega_N(r,t) * R_{\alpha,\alpha-3}\left(-\frac{1}{\lambda},t\right) - \\
&-2\mu\sum_{k=0}^{\infty}\frac{1}{\lambda^k}\partial_t^{k+1}\Omega_N(r,t) * \left[G_{\alpha,\gamma_k-2,k+1}\left(-\frac{1}{\lambda},t\right) - \right. \\
&\left. -G_{\alpha,\gamma_k+\alpha-2,k+2}\left(-\frac{1}{\lambda},t\right)\right],
\end{aligned} \tag{39}$$

where $\gamma_k = \alpha - k - 1$. Furthermore, making $\lambda \to 0$ in (38) and (39) and taking into account of

$$\lim_{\lambda\to 0}\frac{1}{\lambda^k} G_{a,b,k}(-1/\lambda,t) = \frac{t^{-b-1}}{\Gamma(-b)}, \quad \lim_{\lambda\to 0}\frac{1}{\lambda} R_{a,b}(-1/\lambda,t) = \frac{t^{-b-1}}{\Gamma(-b)}, \tag{40}$$

the Newtonian solutions

$$\omega_{N,2}(r,t)=t^2*\partial_t\omega_N(r,t),\qquad \tau_{N,2}(r,t)=\mu t^2*\partial_t\Omega_N(r,t) \tag{41}$$

are recovered.

2. By making now $\alpha\to 1$ into (38) and (39), the solutions for ordinary Maxwell fluids are obtained, namely

$$\begin{aligned}&\omega(r,t)=\omega_{N,2}(r,t)-\\&-2\frac{\pi f}{\rho}\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\left(-\frac{\nu r_n^2}{\lambda}\right)^k r_nC_{Fn}\int_0^t G_{1,-k-2,k+1}\left(-\frac{1}{\lambda},t\right)e^{-\nu r_n^2(t-s)}\,ds=\\&=\omega_{N,2}(r,t)-2\sum_{k=0}^{\infty}\frac{1}{\lambda^k}\partial_t^{k+1}\omega_N(r,t)*G_{1,-k-2,k+1}\left(-\frac{1}{\lambda},t\right),\end{aligned} \tag{42}$$

$$\begin{aligned}&\tau(r,t)=\tau_{N,2}(r,t)-2\mu\int_0^t\partial_s\Omega_N(r,s)R_{1,-2}\left(-\frac{1}{\lambda},t-s\right)\,ds+\\&+2\pi f\nu\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}r_n\widetilde{C}_{Fn}\left(\frac{-\nu r_n^2}{\lambda}\right)^k\times\\&\times\int_0^t e^{-\nu r_n^2t}\left[G_{1,-k-2,k+1}\left(-\frac{1}{\lambda},t-s\right)-G_{1,-k-1,k+2}\left(-\frac{1}{\lambda},t-s\right)\right]\,ds=\\&=\tau_{N,2}(r,t)-2\mu\partial_t\Omega_N(r,t)*R_{1,-2}\left(-\frac{1}{\lambda},t\right)-\\&-2\mu\sum_{k=0}^{\infty}\frac{1}{\lambda^k}\partial_t^{k+1}\Omega_N(r,t)*\left[G_{1,-k-2,k+1}\left(-\frac{1}{\lambda},t\right)-G_{1,-k-1,k+2}\left(-\frac{1}{\lambda},t\right)\right].\end{aligned} \tag{43}$$

Indeed, direct computations implying suitable grouping of terms and the use of equation

$$\sum_{k=0}^{\infty}\left(-\frac{\nu r_n^2}{\lambda}\right)^k G_{1,-k-a,k+1}\left(-\frac{1}{\lambda},t\right)=\lambda\mathcal{L}^{-1}\left(\frac{1}{q^{a-1}}\frac{1}{\lambda q^2+q+\nu r_n^2}\right) \tag{44}$$

shows that Eq. (42) can be respectively written in the form

$$\omega(r,t) = \omega_{N,2}(r,t) - \\ -\frac{4f\lambda}{R\rho}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{J_1(Rr_n)}e^{-\nu r_n^2 t} * \mathcal{L}^{-1}\left(\frac{1}{q}\cdot\frac{1}{\lambda q^2+q+\nu r_n^2}\right), \tag{45}$$

Then, by using the formula

$$e^{-\nu r_n^2 t} * \mathcal{L}^{-1}\left(\frac{1}{q}\frac{1}{\lambda q^2+q+\nu r_n^2}\right) = \\ = \frac{1}{(\nu r_n^2)^3}\left[e^{-\nu r_n^2 t} + \lambda^2\frac{q_{n2}^3 e^{q_{n1}t} - q_{n1}^3 e^{q_{n2}t}}{q_{n1}-q_{n2}} - \lambda\nu r_n^2\right], \tag{46}$$

one obtains

$$\omega(r,t) = \omega_{N,2}(r,t) - \frac{4f\lambda}{R\rho}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{J_1(Rr_n)}\left\{\frac{1}{(\nu r_n^2)^2} - \frac{1}{\nu r_n^2(\lambda-\lambda_r)}\left[e^{-\nu r_n^2 t} + \right.\right. \\ \left.\left. + \frac{1}{q_{n1}-q_{n2}}\left(\frac{e^{q_{n1}t}}{q_{n1}(1-\nu r_n^2\lambda)+\nu r_n^2} - \frac{e^{q_{n1}t}}{q_{n2}(1-\nu r_n^2\lambda)+\nu r_n^2}\right)\right]\right\}, \tag{47}$$

where q_{n1}, q_{n2} are the roots of equation $\lambda q^2 + q + \nu r_n^2 = 0$. Taking into account Eq. (20), the solution (47) can be written in the following simpler form:

$$\omega(r,t) = \omega_{N,2}(r,t) + \frac{f\lambda}{\mu\nu}\frac{r^3(2R^2-r^2)}{24R^2} - \frac{4f}{R\mu\nu^2}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^6 J_1(Rr_n)}\left[e^{-\nu r_n^2 t} + \right. \\ \left. + \frac{[\nu r_n^2 + (1-\nu r_n^2\lambda)q_{n2}]e^{q_{n1}t} - [\nu r_n^2 + (1-\nu r_n^2\lambda)q_{n1}]e^{q_{n2}t}}{q_{n1}-q_{n2}}\right]. \tag{48}$$

A similar procedure, applied to Eq. (43), yields

$$\tau(r,t) = \tau_{N,2}(r,t) - 2\mu\left(\lambda^2 e^{-\frac{t}{\lambda}} - \lambda^2 + \lambda t\right) * \partial\Omega_N(r,t) + \\ + \mu\frac{4f}{R\rho}\sum_{n=1}^{\infty}\frac{r_n J_2(rr_n)}{J_1(Rr_n)}\, e^{-\frac{t}{\lambda}} * e^{-\nu r_n^2 t} * \mathcal{L}^{-1}\left(\frac{1}{q}\frac{1}{\lambda q^2+q+\nu r_n^2}\right). \tag{49}$$

After a straightforward computation we get

$$\begin{aligned}\tau(r,t)=\tau_{N,2}(r,t)-2\mu\lambda\left(\lambda^2e^{-\frac{t}{\lambda}}-\lambda^2+\lambda t\right)*\partial\Omega_N(r,t)+\\+\frac{4f}{R}\sum_{n=1}^{\infty}\frac{J_2(rr_n)}{r_nJ_1(Rr_n)}\left(\frac{\lambda}{(\nu r_n^2)^2}\frac{q_{n2}^2e^{q_{n1}t}-q_{n1}^2e^{q_{n2}t}}{q_{n1}-q_{n2}}+\right.\\\left.+\frac{e^{-\nu r_n^2t}}{(\nu r_n^2)^2(\lambda\nu r_n^2-1)}-\frac{\lambda^2e^{-\frac{t}{\lambda}}}{\lambda\nu r_n^1-1}+\frac{\lambda^2}{\nu r_n^2}\right).\end{aligned}\tag{50}$$

3. In the special case when $\lambda\to 0$ into (16) and (29), the solutions

$$\begin{aligned}&\omega(r,t)=\omega_{N,2}(r,t)+\\&+2\frac{\pi f}{\rho}\lambda_r\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{k}C_{Fn}\frac{k!\lambda_r^m(-\nu r_n^2)^{k+1}}{m!(k-m)!\Gamma(2-\beta_m)}\int_0^t s^{1-\beta_m}e^{-\nu r_n^2(t-s)}\,ds=\\&=\omega_{N,2}(r,t)+2\lambda_r\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\int_0^t\partial_s^{k+2}\omega_N(r,s)\,\frac{(t-s)^{1-\beta_m}}{\Gamma(2-\beta_m)}\,ds\end{aligned}\tag{51}$$

and

$$\begin{aligned}\tau(r,t)=\tau_{N,2}(r,t)+2\mu\lambda_r\int_0^t\frac{(t-s)^{2-\beta}}{\Gamma(3-\beta)}\partial_s\Omega_N(r,t)\,ds-\\-2\pi f\nu\lambda_r\sum_{n=1}^{\infty}r_n\widetilde{C}_{Fn}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m(-\nu r_n^2)^{k+1}}{m!(k-m)!}\times\\\times\int_0^t\left[\frac{s^{1-\beta_m}}{\Gamma(2-\beta_m)}+\lambda_r\frac{s^{2-\beta_m-\beta-1}}{\Gamma(2-\beta_m-\beta)}\right]e^{-\nu r_2^2(t-s)}\,ds=\\=\tau_{N,2}(r,t)+2\mu\lambda_r\int_0^t\frac{(t-s)^{2-\beta}}{\Gamma(3-\beta)}\,\partial_s\Omega_N(r,s)\,ds+2\mu\lambda_r\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^k}{m!(k-m)!}\times\end{aligned}$$

$$\times \int_0^t \left[\frac{(t-s)^{1-\beta_m}}{\Gamma(2-\beta_m)} + \lambda_r \frac{(t-s)^{1-\beta_m-\beta}}{\Gamma(2-\beta_m-\beta)} \right] \partial_s^{k+2} \Omega_N(r,s)\, ds \tag{52}$$

corresponding to a "generalized" second grade fluid are obtained.

Of course, making $\lambda_r \to 0$ into Eqs. (51) and (52), we again attain to the Newtonian solutions given by Eq. (41). Moreover, in the special case when $\beta \to 1$, Eqs. (51) and (52) reduce to the solutions for an ordinary second grade fluid, namely

$$\begin{aligned}
&\omega(r,t) = \omega_{N,2}(r,t) + \\
&+2\frac{\pi f \lambda_r}{\rho} \lambda_r \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{k} C_{Fn} \frac{k! \lambda_r^m (-\nu r_n^2)^{k+1}}{m!(k-m)!\Gamma(k-m+3)} \int_0^t s^{k-m+2} e^{-\nu r_n^2 (t-s)}\, ds = \\
&= \omega_{N,2}(r,t) + 2\lambda_r \sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{k! \lambda_r^m}{m!(k-m)!} \int_0^t \partial_s^{k+2} \omega_N(r,s)\, \frac{(t-s)^{k-m+2}}{\Gamma(k-m+3)}\, ds
\end{aligned} \tag{53}$$

and

$$\begin{aligned}
&\tau(r,t) = \tau_{N,2}(r,t) + 2\mu\lambda_r \int_0^t (t-s) \partial_s \Omega_N(r,t)\, ds - \\
&-2\pi f \nu \lambda_r \sum_{n=1}^{\infty} r_n\, \widetilde{C}_{Fn} \sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{k! \lambda_r^m}{m!(k-m)!} (-\nu r_2^2)^{k+1} \times \\
&\times \int_0^t \left[\frac{s^{k-m+2}}{\Gamma(ka-m+3)} + \lambda_r \frac{s^{k-m+1}}{\Gamma(k-m+2)} \right] e^{-\nu r_n^2 (t-s)}\, ds = \\
&= \tau_{N,2}(r,t) + 2\mu\lambda_r t * \partial_t \Omega_N(r,t) + \\
&+2\mu\lambda_r \sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{k! \lambda_r^k}{m!(k-m)!} \left[\frac{t^{k-m+2}}{\Gamma(k-m+3)} + \right. \\
&\left. + \lambda_r \frac{t^{k-m+1}}{\Gamma(k-m+2)} \right] * \partial_t^{k+2} \Omega_N(r,t)
\end{aligned} \tag{54}$$

i.e.

$$\omega(r,t)=\omega_{N,2}(r,t)-\frac{f\lambda_r r^3}{\mu R^2}t-\frac{f\lambda_r}{\mu\nu}\frac{r^3(r^2-2R^2)}{12R^2}+\frac{f\lambda_r^2 r^3}{\mu R^2}-$$
$$-\frac{4f}{\mu\nu^2 R}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^6 J_1(Rr_n)}\left(e^{-\frac{\nu r_n^2 t}{1+\nu r_n^2\lambda_r}}-e^{-\nu r_n^2 t}\right)-$$
$$-\frac{8f\lambda_r}{\mu\nu R}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^2 J_1(Rr_n)}e^{-\frac{\nu r_n^2 t}{1+\nu r_n^2\lambda_r}}-\frac{4f\lambda_r^2}{\mu R}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^2 J_1(Rr_n)}e^{-\frac{\nu r_n^2 t}{1+\nu r_n^2\lambda_r}},$$

$$\tau(r,t)=\tau_{N,2}(r,t)+2\mu\lambda_r t*\partial_t\Omega(r,t)+$$
$$+\frac{2f\lambda_r}{R}\sum_{n=1}^{\infty}\frac{r_n^3 J_2(rr_n)}{r_n J_1(Rr_n)}\left(1-\frac{2}{\nu r_n^2}-\right.$$
$$\left.-\frac{1+\lambda_r\nu r_n^2}{\lambda_r\nu^2 r_n^4}e^{-\frac{\nu r_n^2 t}{1+\lambda_r\nu r_n^2}}+\frac{\lambda_r\nu r_n^2-1}{\lambda_r\nu^2 r_n^4}e^{-\nu r_n^2 t}\right). \tag{55}$$

4. In the special case when $\alpha\to 1$ and $\beta\to 1$ into (16), (29), (32) and (34), the solutions for an Oldroyd-B fluid are obtained, namely:

$$\omega(r,t)=\omega_{N,2}(r,t)-$$
$$-\Gamma(1+a)\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+1}\omega_N(r,t)*G_{1,m-k-2,k+1}\left(-\frac{1}{\lambda},t\right)+$$
$$+2\frac{\lambda_r}{\lambda}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\partial_t^{k+2}\omega_N(r,t)*G_{1,m-k-3,k+1}=$$
$$=\omega_{N,2}(r,t)-2\frac{\pi f}{\rho}\sum_{n=1}^{\infty}r_n\ C_{Fn}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\left(-\frac{\nu r_n^2}{\lambda}\right)^k\times$$
$$\times\left[G_{1,m-k-2,k+1}\left(-\frac{1}{\lambda},t\right)+\nu r_n^2\frac{\lambda_r}{\lambda}G_{1,m-k-3,k+1}\left(-\frac{1}{\lambda},t\right)\right], \tag{56}$$

$$\tau(r,t)=\tau_{N,2}(r,t)+2\mu\frac{\lambda_r-\lambda}{\lambda}R_{1,-2}\left(-\frac{1}{\lambda},t\right)*\partial_t\Omega_N(r,t)-$$
$$-2\mu\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\left[G_{1,m-k-2,k+1}\left(-1/\lambda,t\right)+\right.$$
$$\left.+\frac{\lambda_r}{\lambda}G_{1,m-k-1,k+2}\left(-1/\lambda,t\right)-G_{1,m-k-1,k+2}\left(-\frac{1}{\lambda},t\right)\right]*\partial_t^{k+1}\Omega_N(r,t)+$$
$$+2\mu\frac{\lambda_r}{\lambda}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!\lambda^k}\left[G_{1,m-k-3,k+1}\left(-1/\lambda,t\right)+\right.$$
$$\left.+\frac{\lambda_r}{\lambda}G_{1,m-k-2,k+2}\left(-1/\lambda,t\right)-G_{1,m-k-2,k+2}\left(-1/\lambda,t\right)\right]*\partial_t^{k+2}\Omega_N(r,t)=$$
$$=\tau_{N,2}(r,t)+2\mu\frac{\lambda_r-\lambda}{\lambda}R_{1,-2}\left(-\frac{1}{\lambda},t\right)*\partial_t\Omega_N(r,t)+$$
$$+2\pi f\nu\sum_{n=1}^{\infty}r_n\widetilde{C}_{Fn}\sum_{k=0}^{\infty}\sum_{m=0}^{k}\frac{k!\lambda_r^m}{m!(k-m)!}\left(-\frac{\nu r_n^2}{\lambda}\right)^k\times$$
$$\times\left\{\left[G_{1,m-k-2,k+1}\left(-\frac{1}{\lambda},t\right)+\frac{\lambda_r-\lambda}{\lambda}G_{1,m-k-1,k+2}\left(-\frac{1}{\lambda},t\right)\right]+\right.$$
$$\left.\nu r_n^2\frac{\lambda_r}{\lambda}\left[G_{1,m-k-3,k+1}\left(-\frac{1}{\lambda},t\right)+\frac{\lambda_r-\lambda}{\lambda}G_{1,m-k-1,k+2}\left(-\frac{1}{\lambda},t\right)\right]\right\}*e^{-\nu r_n^2 t} \tag{57}$$

As before we get

$$\omega(r,t)=\omega_{N,2}(r,t)-\frac{4fR}{\mu}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^2J_1(Rr_n)}\left[\lambda_r\frac{t^2}{2}+\frac{\lambda-2\lambda_r-\lambda_r^2\nu r_n^2}{\nu r_n^2}t+\right.$$
$$+\frac{\lambda_r^2(\nu r_n^2)^2+(1+\lambda_r\nu r_n^2)(3\lambda_r-2\lambda)}{(\nu r_n^2)^2}+\frac{1}{(\nu r_n^2)^3}e^{-\nu r_n^2 t}+$$
$$\left.\lambda^3\frac{q_{n2}^3(1+\lambda q_{n1}+\lambda_r q_{n2}+\lambda_r\nu r_n^2)e^{q_{n1}t}-q_{n1}^3(1+\lambda q_{n2}+\lambda_r q_{n1}+\lambda_r\nu r_n^2)e^{q_{n2}t}}{(\lambda-\lambda_r)(\nu r_n^2)^4(q_{n1}-q_{n2})}\right]$$
$$=\frac{4fR}{\mu}\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^2J_1(Rr_n)}\left[\frac{t^2}{2}-\frac{1-\lambda_r\nu r_n^2}{\nu r_n^2}t+\frac{1+(2\lambda_r-\lambda)\nu r_n^2+\lambda_r^2(\nu r_n^2)^2}{(\nu r_n^2)^2}+\right.$$
$$\left.+\frac{1+(2\lambda_r-\lambda)\nu r_n^2}{(\nu r_n^2)^2}\frac{q_{n2}e^{q_{n1}t}-q_{n1}e^{q_{n2}t}}{q_{n1}-q_{n2}}-(1+\lambda_r\nu r_n^2)\frac{e^{q_{n1}t}-e^{q_{n2}t}}{q_{n1}-q_{n2}}\right], \tag{58}$$

$$\tau(r,t)=\tau_{N,2}(r,t)+2\mu(\lambda\lambda_r-\lambda)\left(t-\lambda+\lambda e^{-\frac{t}{\lambda}}\right)*\partial_t\Omega_N(r,t)+$$
$$+\frac{4f\nu}{R}\sum_{n=1}^{\infty}\frac{r_nJ_2(rr_n)}{J_1(Rr_n)}\left[\frac{\lambda_r}{\nu r_n^2}t+\frac{\lambda-2\lambda_r-\nu r_n^2\lambda\lambda_r}{(\nu r_n^2)^2}+\right.$$
$$+\frac{\lambda^2}{\nu r_n^2}\frac{1-\nu r_n^2\lambda_r}{1-\nu r_n^2\lambda}e^{-\frac{t}{\lambda}}+\frac{\lambda^2}{(\nu r_n^2)^3}\frac{1-\nu r_n^2\lambda_r}{1-\nu r_n^2\lambda}e^{-\nu r_n^2t}+$$
$$\left.+\frac{\lambda(\lambda-\lambda_r)}{q_{n1}-q_{n2}}\left(\frac{q_{n1}e^{q_{n1}t}}{A_nq_{n1}+B_n}-\frac{q_{n2}e^{q_{n2}t}}{A_nq_{n2}+B_n}\right)\right]=$$
$$=4f\sum_{n=1}^{\infty}\frac{J_1(rr_n)}{r_n^2J_1(Rr_n)}\left[-\frac{t^2}{2}+\frac{1+\lambda_r\nu r_n^2+\lambda-\lambda_r}{\nu r_n^2}+\right.$$
$$+\frac{(\lambda+\lambda_r)(1+\lambda\nu r_n^2)}{\nu r_n^2}-\frac{\lambda-\lambda_r}{\lambda}\nu r_n^2e^{-\frac{t}{\lambda}}+$$
$$\left.+\frac{\lambda^2}{(\nu r_n^2)^3}\frac{(1+\lambda_rq_{n1})q_{n2}^3(1+\lambda q_{n2})e^{q_{n1}t}-(1+\lambda_rq_{n2})q_{n1}^3(1+\lambda q_{n1})e^{q_{n2}t}}{q_{n1}-q_{n2}}\right], \tag{59}$$

where q_{n1} and q_{n12} are the real roots of $\lambda q^2+(1+\nu r_n^2\lambda_r)q+\nu r_n^2=0$ (they are real negative numbers because $(1+\lambda_r\nu r_n^2)^2-4\lambda\nu r_n^2>0$, $q_{n1}\cdot q_{n2}\lambda\nu r_n^2>0$ and $q_{n1}+q_{n2}=-\dfrac{1+\nu r_n^2\lambda_r}{\lambda}<0$).

6 Conclusions

The main purpose of this paper is to provide exact solution for the unsteady flow of an incompressible Oldroyd-B fluid filling the annular region between two infinitely long co-axial cylinders subject to a particular time-dependent shear stress. Such solutions, obtained by using the Hankel and Laplace transforms, are presented as sums between the Newtonian solutions and the corresponding non-Newtonian contributions. Furthermore, the non-Newtonian contributions of the general solutions are also presented in equivalent forms, under series form in terms of the time derivative of the (simplest) Newtonian

solution ω_N and $\omega_{N,2}$ as well. For $\lambda \to 0$ (and, consequently, $\lambda_r \to 0$) these contributions tend to zero, such that the general solutions become Newtonian solutions corresponding to the given initial-boundary conditions.

It is remarkable that the general solutions can be easily specialized to give both the similar solutions for "generalized" second grade and Maxwell fluids and the solutions for all ordinary fluids (Oldroyd-B, Maxwell and second grade) performing the same motions. Direct computations shows that the solutions which have been obtained certainly satisfy both the governing equations and all imposed initial and boundary conditions. Furthermore, the solutions corresponding to ordinary Maxwell and second grade fluids can be also obtained as limiting cases of those for ordinary Oldroyd-B fluids. As regard the Newtonian solutions, given under simple forms (20), (22), and (37), they can be obtained as limiting cases of the previous solutions.

From our general solutions, corresponding to non-Newtonian fluids, it clearly results that the non-Newtonian contributions of these solutions exponentially decrease in time, the motion of the non-Newtonian fluids being well approximated, for large values of t, by the motion of the corresponding Newtonian fluid.

References

[1] R.L. Bagley. A theoretical basis for the application of fractional calculus to viscoelasticity. *J.Rheology* 27: 201–210, 1983.

[2] R. Bandelli, K.R. Rajagopal. Start-up flows of second grade fluids in domains with one finite dimension. *Int. J. Non-Linear Mech.* 30 (1995), 817-839.

[3] R. Bandelli, K.R. Rajagopal, G.P. Galdi. On some unsteady motions of fluids of second grade. *Arch. Mech.* 47 : 661-676, 1995.

[4] G.K. Batchelor. *An Introduction to Fluid Dynamics*. Cambridge University Press, Cambridge, 1967.

[5] I. Burdujan. Some Accelerated Flows for an Oldroyd-B Fluid. *ROMAI J.* 5.2: 29–48, (2009).

[6] L. Debnath, D. Bhatta. *Integral Transforms and their Applications*(second ed.). Chapman and Hall/CRC Press, Boca-Raton-London-New York, 2007.

[7] Ch. Friedrich. Relaxation and retardation functions of the Maxwell model with fractional derivatives. *Rheol. Acta* 30: 151–158, 1991.

[8] A. Georgescu, L. Palese, A. Redaelli. A direct method and its application to a linear hydromagnetic stability problem. *ROMAI J.* 1.1: 67–76, 2005.

[9] W.G. Glökle,T.F. Nonnenmacher. Fractional relaxation and the time-temperature superposition principle. *Rheol. Acta* 33: 337–343, 1994.

[10] T. Hayat, M. Khan, T. Wang. Non-Newtonian flow between concentric cylinders. *Comm. Non-Linear Sci. Numer. Simm.* 11: 297-305, 2006.

[11] T. Hayat, M. Hussain, M. Khan. Hall effect on flows of an Oldroyd-B fluid through porous medium for cylindrical geometries. *Computers and Mathematics with Applications* 52: 269-282, 2006.

[12] A. Heibig, L.I. Palade. On the rest state stability of an objective fractional derivative viscoelastic fluid model. *J. Math. Phys.* 49: 043101-22, 2008.

[13] M. Khan, S. Hyder Ali, H. Qi. Some accelerated flows for a generalized Oldroyd-B fluid. *Nonlinear Analysis: Real world Applications* (2007), doi: 10.1016/j.nonrwa.2007.11.017.

[14] C.F.Lorenzo, T.T. Hartley. Generalized Functions for the Fractional Calculus. NASA/TP-1999-209427, 1999.

[15] F. Mainardi. Fractional relaxation-oscillation and fractional diffusion-wave phenomena. *Chaos, Solitons&Fractals* 7(9): 1461–1477, 1996.

[16] F. Mainardi, R. Gorenflo. On Mittag-Leffler-type functions in fractional evolution processes. *J. Comput. Appl. Math.* 116(2): 283–299, 2000.

[17] N. Makris, M. C. Constantinou. Fractional derivative Maxwell model for viscous dampers. *J. Struct. ASCE* 117(9): 2708–2724, 1991.

[18] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, 1999.

[19] Y.A. Rossikhin,M.V. Shitikova. A new method for solving dynamic problems of fractional derivative viscoelasticity. *Int. J. Engng Sci.* 39: 149–176, 2000.

[20] Y.A. Rossikhin, M.V. Shitikova. Analysis of dynamic behavior of viscoelastic rods whose rheological models contain fractional derivatives of two different orders. *ZAMP* 81(6): 363–376, 2001.

[21] P.N. Srivastava. Non-steady helical flow of a viscoelastic liquid. *Arch. Mech. Stos.* 18: 145-150, 1966.

[22] D. Tong, Y. Liu, Exact solutions for the unsteady rotational flow of non-Newtonian fluid in an annular pipe. *Int. J. Eng. Sci.* 43: 281-289, 2005.

[23] D. Tong, Y. Ruihe,W. Heshan. Exact solutions for the flow of non-Newtonian fluid with fractional derivative in an annular pipe. *Science in China Ser. G Physics, Mechanics & Astronomy.* 48: 485-495, 2005.

[24] W. P. Wood. Transient viscoelastic helical flows in pipes of circular and annular cross-section. *J. Non-Newtonian Fluid Mech.* 100: 115-126, 2001.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

FIXED POINTS THEOREMS IN MULTI-METRIC SPACES*

Mitrofan M. Choban† Laurenţiu I. Calmuţchi‡

Abstract

The aim of the present article is to give some general methods in the fixed point theory for mappings of general topological spaces. Using the notions of the multi-metric space and of the E-metric space, we proved the analogous of several classical theorems: Banach fixed point principle, Theorems of Edelstein, Meyers, Janos etc.

MSC: 54H25, 54E15, 54H13, 12J17, 54E40.

keywords: fixed point, m-scale, semifield, multi-metric space, E-metric space, pseudo-metric.

1 Introduction

Any space is considered to be Tychonoff and non-empty. We use the terminology from [12, 13].

Let $\mathbb{R}$ be the space of real numbers.

A pseudo-metric on a set X is a function $\rho : X \times X \longrightarrow \mathbb{R}$ satisfying the following conditions: $\rho(x,x)=0$, $\rho(x,y)=\rho(y,x)$ and $\rho(x,z)\leq\rho(x,y)+\rho(y,z)$ for all $x,y,z \in X$. The number $\rho(x,y)$ is called the ρ-distance between the points x,y.

*Accepted for publication on November 12, 2010.

†Department of Mathematics, Tiraspol State University, Kishinev, Republic of Moldova, MD 2069, mmchoban@gmail.com

‡Department of Mathematics, Tiraspol State University, Kishinev, Republic of Moldova, MD 2069

For any $x, y \in X$ we have $\rho(x,y) \geq 0$. If $A \subseteq B$, $B \subseteq X$, $x \in X$ and $r \geq 0$, then $\rho(A,B) = inf\{\rho(x,y) : x \in A, y \in B\}$ and $B(x,\rho,r) = \{y \in X : \rho(x,y) < r\}$.

The pseudo-metric ρ generates on X the canonical equivalence relation: $x \sim y$ iff $\rho(x,y) = 0$. Let X/ρ be the quotient set with the canonical projection $\pi_\rho : X \longrightarrow X/\rho$ and the metric $\bar{\rho}(u,v) = \rho(\pi_\rho^{-1}(u), \pi_\rho^{-1}(v))$.

Definition 1.1. *A multi-metric space is a pair $(X, \mathcal{P})$, where X is a set and $\mathcal{P}$ is a non-empty family of pseudo-metrics on X satisfying the condition: $x = y$ if and only if $\rho(x,y) = 0$ for each $\rho \in \mathcal{P}$.*

Fix a multi-metric space $(X, \mathcal{P})$. A subset $U \subseteq X$ is called *$\mathcal{P}$-open* if for any $x \in U$ there exist a number $\epsilon = \epsilon(x,U) > 0$ and a finite set $A = A(x,U) \subseteq \mathcal{P}$ such that $B(x,A,\epsilon) = \cap\{B(x,\rho,\epsilon) : \rho \in A\} \subseteq U$. The family $\mathcal{T}(\mathcal{P})$ of all $\mathcal{P}$-open subsets is a completely regular Hausdorff topology on X. If $\mathcal{T}$ is a completely regular Hausdorff topology on X, then $\mathcal{T} = \mathcal{T}(\mathcal{P})$ for some family $\mathcal{P}$ of pseudo-metrics on X (see [12]).

A $\mathcal{P}$-number is the set $\lambda_\mathcal{P} = (\lambda_\rho : \rho \in \mathcal{P})$, where $\lambda_\rho \in \mathbb{R}$ for any $\rho \in \mathcal{P}$. If $\alpha, \beta \in \mathbb{R}$ and $\lambda_\mathcal{P} = (\lambda_\rho : \rho \in \mathcal{P})$ is a $\mathcal{P}$-number, then $\lambda \leq \lambda_\mathcal{P} \ll \beta$ if $\lambda \leq \lambda_\rho < \beta$ for any $\rho \in \mathcal{P}$.

In [1, 2, 3] there were introduced the metric spaces over topological semifields. The general conception of the metrizability of spaces is contained in [21]. Every multi-metric space can be considered as a metric space $(X, d, \mathbb{R}^\mathcal{P})$, where $d(x,y) = (\rho(x,y) : \rho \in \mathcal{P})$ for all $x, y \in X$, over the Tichonoff semifield $\mathbb{R}^\mathcal{P}$ (see [1, 2, 3]).

The notion of a topological semifield may be generalized in the following way. We say that E is *a metric scale* or, briefly, *an m-scale* if:

1. E is a topological algebra over the field of reals $\mathbb{R}$;
2. E is a commutative ring with the unit $1 \neq 0$;
3. E is a vector lattice and $0 \leq xy$ provided $0 \leq x$ and $0 \leq y$;
4. For any neighborhood U of 0 in E there exists a neighborhood V of 0 in E such that if $x \in V$ and $0 \leq y \leq x$, then $y \in U$.

From the condition 4 it follows:

5. If $0 \leq y_n \leq x_n$ and $lim_{n\to\infty} x_n = 0$, then $lim_{n\to\infty} y_n = 0$ too.

Any topological semifield is an m-scale.

Let E be an m-scale.

Denote by $E^{-1} = \{x : x \cdot y = 1 \text{ for some } y \in E\}$ the set of all invertible elements of E.

By $N(0,E)$ we denote some base of the space E at the point 0.

We consider that $0 \leq x \ll 1$ if $0 \leq x < 1$, $1 - x$ is invertible and $lim_{n\to\infty} x^n = 0$. We put $E^{(+,1)} = \{x \in E : 0 \leq x \ll 1\}$. If $t \in \mathbb{R}$ and $0 \leq t < 1$, then $t \cdot 1 \in E^{(+,1)}$. We identify t with $t \cdot 1 \in E$ for each $t \in \mathbb{R}$.

A mapping $d : X \times X \longrightarrow E$ is called *a metric over m-scale E* or *an E-metric* if it is satisfying the following axioms of the metric:

- $d(x, y) = 0$ if and only if $x = y$;
- $d(x, y) = d(y, x)$;
- $d(x, y) \leq d(x, z) + d(yz, y)$.

Every E-metric is non-negative, i.e. $d(x, y) \geq 0$ for all $x, y \in X$.

The ordered triple (X, d, E) is called *a metric space over m-scale E* or *an E-metric* space if d is an E-metric on X.

Let (X, d, E) be an E-metric space. If $U \in N(0, E)$, then we put $B(x, d, U) = \{y \in X : d(x, y) \in U\}$ for any $x \in X$. The family $\{B(x, d, U) : x \in X, U \in N(0, E)\}$ is the base of the topology $\mathcal{T}(d)$ of the E-metric space. The space $(X, \mathcal{T}(d))$ is a Tychonoff space.

Let X be a space and $f : X \longrightarrow X$ be a mapping. By $Fix(f) = \{x \in X : f(x) = x\}$ we denote the set of all fixed points of the mapping f. The excellence book [13] contains the fixed point theory for metric spaces with the important applications in distinct domains. Several results for general topological spaces with interesting applications contain the surveys [13, 17, 23]. Distinct generalizations of the Banach fixed point principle were proposed in [6, 11, 13, 17, 23, 26]. In [24] it was arisen the following general problem: to find topological analogies of the Banach fixed point principle. Some solutions of this problem were proposed in [8, 25]. This article is a continuation of the works [8, 24, 25, 16].

Definition 1.2 (see [13]). *Let X be a space and $\mathcal{M}$ be a class of mappings $f : X \longrightarrow X$. If $Fix(f) \neq \varnothing$ for each $f \in \mathcal{M}$, then X is called a fixed point space relative to $\mathcal{M}$.*

If X is a fixed point space relative to all continuous mappings $f : X \longrightarrow X$, then X is called a fixed point space.

For each family γ of subsets of a set X and any $x \in X$ the set $St(x, \gamma) = \cup\{U \in \gamma : x \in U\}$ is the star of the point x relative to γ. We put $St(A, \gamma) = \cup\{St(x, \gamma) : x \in A\}$.

Theorem 1.3 (see [13], Theorem 0.4.4, for a compact metric space X). *For a compact space X the following assertions are equivalent:*

1. X is a fixed point space.

2. There exists a family $\mathcal{P}$ of continuous pseudo-metrics on X such that:

- for any two pseudo-metrics $\rho_1, \rho_2 \in \mathcal{P}$ there exists $\rho \in \mathcal{P}$ such that $sup\{\rho_1(x,y), \rho_2(x,y)\} \leq \rho(x,y)$ for all $x, y \in X$;

- $\mathcal{T}(\mathcal{P})$ is the topology of the space X;

- for any $\rho \in \mathcal{P}$ and $\epsilon > 0$ there exist a compact fixed point subspace $X_{(\rho,\epsilon)}$ of X and two continuous mappings $\alpha_{(\rho,\epsilon)} : X \longrightarrow X_{(\rho,\epsilon)}$ and $\beta_{(\rho,\epsilon)} : X_{(\rho,\epsilon)} \longrightarrow X$ such that $\rho(x, \beta_{(\rho,\epsilon)}(\alpha_{(\rho,\epsilon)}(x))) < \epsilon$ for any $x \in X$.

3. There exists a family $\mathcal{U}$ of open covers of X such that:

- for any open cover ξ of X there exists a refinement $\gamma \in \mathcal{U}$ of ξ;

- for any $\gamma \in \mathcal{U}$ there exist a compact fixed point subspace X_γ of X and two continuous mappings $\alpha_\gamma : X \longrightarrow X_\gamma$ and $\beta_\gamma : X_\gamma \longrightarrow X$ such that $\beta_\gamma(\alpha_\gamma(x)) \in St(x, \gamma)$ for any $x \in X$.

Proof. Since any compact space has a unique uniform structure and any uniform structure is generated by a family of continuous pseudometrics, the assertions 2 and 3 are equivalent.

If X is a fixed point space, then we put $X = X_\gamma = X_{(\rho,\epsilon)}$ and consider that $\alpha_\gamma, \beta_\gamma, \alpha_{(\rho,\epsilon)}, \beta_{(\rho,\epsilon)}$ are the identical mapping. Thus the implications $1 \to 2$ and $1 \to 3$ are obvious.

Assume that the assertion 3 is true. If $\xi, \gamma \in \mathcal{U}$ and γ is a refinement of ξ, then we put $\xi \leq \gamma$. Fix a continuous mapping $f : X \longrightarrow X$. We put $g_\gamma = \beta_\gamma \circ f \circ \alpha_\gamma : X_\gamma \longrightarrow X_\gamma$ and $f_\gamma = \beta_\gamma \circ \alpha_\gamma \circ f : X \longrightarrow X$ for any $\gamma \in \mathcal{U}$. Since X_γ is a fixed point space, then we can fix a point $z_\gamma \in X_\gamma$ such that $g_\gamma(z_\gamma) = z_\gamma$. We put $x_\gamma = \beta_\gamma(z_\gamma)$. Obviously, $f_\gamma(x_\gamma) = x_\gamma$ and $f(x_\gamma) \in St(x_\gamma, \gamma)$.

The set $\mathcal{U}$ is directed. Thus $\{x_\gamma : \gamma \in \mathcal{U}\}$ is a net in the space X. Since X is a compact space, the net $\{x_\gamma : \gamma \in \mathcal{U}\}$ has a cluster point x_0, i.e. for each neighborhood V of the point x_0 in X and any $\xi \in \mathcal{U}$ there exists $\gamma \in \mathcal{U}$ such that $\xi \leq \gamma$ and $x_\gamma \in V$ (see [12], Theorem 3.1.23).

We affirm that $f(x_0) = x_0$.

Assume that $f(x_0) \neq x_0$. Since f is a continuous mapping, there exist $\xi \in \mathcal{U}$ and an open subset V of X such that $St(x_0, \xi) \subseteq V$, $St(V, \xi) \cap St(f(x_0), \xi) = \varnothing$ and $f(St(V, \xi)) \cap St(f(x_0), \xi)$. By construction, $f(x) \notin St(V, \xi)$ for any $x \in St(V, \xi)$.

There exists $\gamma \in \mathcal{U}$ such that $\xi \leq \gamma$ and $x_\gamma \in St(x_0, \xi) \subseteq V$. By construction, $x_\gamma \in V$, $f(x_\gamma) \in St(x_\gamma, \gamma) \subseteq St(V, \gamma)$ and $f(x_\gamma) \in St(V, \xi)$, a contradiction with the condition that $f(x) \notin St(V, \xi)$ for any $x \in St(V, \xi)$. The proof is complete.

As Theorem 0.4.5 from ([13], p. 8), using Tychonoff cube I^A, one can prove the next assertion.

Theorem 1.4. *Let $f; X \to X$ be a continuous mapping and there exist a compact absolute retract space Y and the continuous mapping $\alpha : X \longrightarrow Y$, $\beta : Y \longrightarrow X$ such that $f = \beta \circ \alpha$. Then f has a fixed point.*

2 Complete multi-metric spaces

Fix an m-scale E and an E-metric space (X, d, E).

A set $\{x_\mu : \mu \in M\}$ is a net or a generalized sequence if M is a directed set.

A point $x \in X$ is a limit of the net $\{x_\mu : \mu \in M\}$ and we put $lim_{m\in M} x_\mu = x$ if for any $U \in N(0, E)$ there exists $\lambda \in M$ such that $d(x, x_\mu) \in U$ for any $\mu \geq \lambda$.

A net $\{x_\mu : \mu \in M\}$ is called *fundamental* if for any $U \in N(0, E)$ there exists $\lambda \in M$ such that $d(x_\mu, x_\eta) \in U$ for any $\mu, \eta \geq \lambda$. Any convergent net is fundamental. The limit of a fundamental sequence is unique (if the limit exists).

The space (X, d, E) is called *complete* if any fundamental net is convergent (see [12, 1, 2, 3]).

The space (X, d, E) is called *sequentially complete* if any fundamental sequence $\{x_n \in X : n \in \mathbb{N}\}$ is convergent.

Let $\{x_\mu : \mu \in M\}$ be a net and $U \in N(0, E)$. We assume that $x \in L(U, \{x_\mu : \mu \in M\})$ if there exists $\lambda \in M$ such that $d(x, x_\mu) \in U$ for any $\mu \geq \lambda$.

The space (X, d, E) is called *conditionally complete* if $\cap\{L(U, \{x_n : n \in \mathbb{N}\}) : U \in \gamma\} \neq \varnothing$ for any fundamental sequence $\{x_n \in X : n \in \mathbb{N}\}$ and any countable non-empty family $\gamma \subseteq N(0, E)$.

Any complete metric space is sequentially complete and any sequentially complete space is conditionally complete.

Example 2.1. Let (X, d, E) be an E-metric space. On X consider the topology $\mathcal{T}(d)$. Assume that the space X is pseudocompact. Then the metric space (X, d, E) is conditionally complete. If the space X is countably compact, then the metric space (X, d, E) is sequentially complete. The space (X, d, E) is complete if and only if the space X is compact. Thus if X is a countably compact non-compact space, then the space (X, d, E) is sequentially complete and non-complete.

Example 2.2. Let A be an uncountable set. We put $I_\alpha = I = [0,1]$ for any $\alpha \in A$. Let $Y = I^A = \Pi\{I_\alpha : \alpha \in A\}$, $0_A = (0_\alpha : \alpha \in A) \in Y$ and $X = Y \setminus \{0_A\}$. The space $E = \mathbb{R}^A$ is a topological field and an m-scale. By construction, $X \subseteq I^A \subseteq E$. We put $d(x,y) = (|x_\alpha - y_\alpha| : \alpha \in A)$ and $d_\alpha(x,y) = |x_\alpha - y_\alpha|$ for any pair of points $x = (x_\alpha : \alpha \in A) \in Y$ and $x = (x_\alpha : \alpha \in A) \in Y$. Let $\mathcal{P} = \{d_\alpha : \alpha \in A\}$. Then (X, d, E) is an E-metric space and $(X, \mathcal{P})$ is a multi-metric space.

Obviously, $\mathcal{T}(d) = \mathcal{T}(\mathcal{P})$. The space Y is compact and the space X is pseudocompact. Thus the spaces (X, d, E) and $(X, \mathcal{P})$ are conditionally complete. We put $x_n = (2_\alpha^{-n} : \alpha \in A)$. Then $\{x_n : n \in \mathbb{N}\}$ is a fundamental sequence. We have $lim_{\mathcal{P}} x_n = 0_A$. Thus the spaces (X, d, E) and $(X, \mathcal{P})$ are not sequentially complete. In particular, the spaces (X, d, E) and $(X, \mathcal{P})$ are not complete.

Remark 2.3. Let (X, d, E) be an E-metric space and the space E be first countable, i.e. it is metrizable in the usual sense. Then the space X is metrizable in the usual sense too. Moreover, if the space (X, d, E) is conditionally complete, then the space (X, d, E) is complete.

Remark 2.4. Let (X, d, E) be an E-metric space and X be an infinite extremally disconnected countably compact space. Then each fundamental sequence $\{x_n : n \in \mathbb{N}\}$ is trivial, i.e. there exists $m \in \mathbb{N}$ such that $x_n = x_m$ for any $n \geq m$.

Now we fix a multi-metric space $(X, \mathcal{P})$.

In this case we put $E = \mathbb{R}^{\mathcal{P}}$ and $d(x,y) = (\rho(x,y) : \rho \in \mathcal{P})$ for any $x, y \in X$. Obviously, (X, d, E) is an E-metric space and $\mathcal{T}(d) = \mathcal{T}(\mathcal{P})$.

A sequence $\{x_n \in X : n \in \mathbb{N} = \{1, 2, ...\}\}$ is called ρ-fundamental, where $\rho \in \mathcal{P}$, if for any $\epsilon > 0$ there exists $m \in \mathbb{N}$ such that $\rho(x_n, x_k) < \epsilon$ for all $n, k \geq m$. A point $x \in X$ is a ρ-limit of the sequence $\{x_n : n \in \mathbb{N}\}$ if $lim\rho(x, x_n) = 0$. Let $L(\{x_n : n \in \mathbb{N}\}, \rho) = \{x \in X : lim\rho(x, x_n) = 0\}$. A sequence $\{x_n \in X : n \in \mathbb{N}\}$ is fundamental if and only if it is ρ-fundamental for any $\rho \in \mathcal{P}$.

Obviously, the space $(X, \mathcal{P})$ is conditionally complete if and only if $\cap\{L(\{x_n : n \in \mathbb{N}\}, \rho) : \rho \in \mathcal{F}\} \neq \varnothing$ for any countable family $\mathcal{F} \subseteq \mathcal{P}$ and any fundamental sequence $\{x_n : n \in \mathbb{N}\}$.

There exists a canonical embedding $e_{\mathcal{P}} : X \longrightarrow \Pi\{X/\rho : \rho \in \mathcal{P}\}$, where $e_{\mathcal{P}}(x) = (\pi_\rho(x) : \rho \in \mathcal{P})$. If $(Y_\rho, \bar{\rho})$ is the Hausdorff completion of the metric space $(X/\rho, \bar{\rho})$ and the subspace $e_{\mathcal{P}}(X)$ is closed in $\Pi\{Y_\rho : \rho \in \mathcal{P}\}$, then $(X, \mathcal{P})$ is a complete multi-metric space.

Obviously, a multi-metric space $(X, \mathcal{P})$ is a complete multi-metric space if and only if the E-metric space (X, d, E) is complete.

3 Banach fixed point theorem for E-metric spaces

Fix an m-scale E.

Let (X, d, E) be an E-metric space.

A mapping $f : X \longrightarrow X$ is *a contractive mapping* if there exists an element $k \in E^{(+,1)}$ such that $d(f(x), f(y)) \leq kd(x, y)$ for all $x, y \in X$. The element k is called the Lipschitz constant of the mapping f. Every contractive mapping is uniformly continuous.

For any mapping $f : X \longrightarrow X$ and any point $x \in X$ we put $0(x, f) = x$ and $n(x, f) = f((n-1)(x, f))$ for any $n \in \mathbb{N}$. We put $\omega = \{0, 1, 2, ...\}$. Then the set $T(f, x) = \{n(x, f) : n \in \omega\}$ is the Picard orbit of the point x relative to the mapping f.

Theorem 3.1. *Let $f : X \longrightarrow X$ be a contractive mapping with the Lipschitz constant $k \in E^{(+,1)}$. Then for each point $x \in X$ the Picard orbit $T(f, x) = \{n(x, f) : n \in \omega\}$ is a fundamental sequence of the metric space (X, d, E).*

Proof. Fix $x = x_0$. We put $x_n = f(x_{n-1})$ for any $n \in \mathcal{N}$.

There exists $b \in E$ such that $b \cdot (1 - k) = 1$.

Then $d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq k \cdot d(x_{n-1}, x_n)$ for any $n \in \mathbb{N}$. Hence $d(x_n, x_{n+1}) \leq k^n \cdot d(x_0, x_1)$ for any $n \in \mathbb{N}$. Obviously, $d(x_n, x_m) \leq d(x_n, x_{n+1} + ... + d(x_{m-1}, x_m) \leq (k^n + k^{n+1} + ... + k^{m-1} + k^m) \cdot d(x_0, x_1)$. Therefore $d(x_n, x_m) \leq (k^n - k^{n+m}) \cdot b \cdot d(x_0, x_1)$ provided $n, m \in \mathbb{N}$ and $n \leq m$. Since $lim_{n\to\infty} k^n = 0$, the sequence $\{x_n : n \in \omega\}$ is fundamental. The proof is complete.

Corollary 3.2. *Let (X, d, E) be a sequentially complete E-metric space and $f : X \longrightarrow Y$ be a contraction mapping with the Lipschitz constant $k \in E^{(+,1)}$. Then the mapping f admits one and only one fixed point $b \in X$. Moreover, $b = lim_{n\to\infty} n(x, f)$ for any point $x \in X$.*

Remark 3.3. For a topological semifield E the assertion of Corollary 3.2 was proved by K.Iseki in [16].

4 Banach fixed point theorem for a multi-metric space

Let $(X, \mathcal{P})$ be a multi-metric space.

A mapping $f : X \longrightarrow X$ is *a Lipschitz mapping* if there exists a $\mathcal{P}$-number $k = (k_\rho : \rho \in \mathcal{P})$ such that $\rho(f(x), f(y)) \leq k_\rho$ for all $x, y \in X$ and $\rho \in \mathcal{P}$. The $\mathcal{P}$-number k is called the Lipschitz constant of the mapping f. Every Lipschitz mapping is uniformly continuous.

A mapping $f : X \longrightarrow X$ is called *a contraction mapping* on $(X, \mathcal{P})$ if there exists a non-negative $\mathcal{P}$-number $k = (k_\rho : \rho \in \mathcal{P}) \ll 1$ such that $\rho(f(x), f(y)) \leq k_\rho \rho(x, y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$.

Theorem 4.1. *Let $(X, \mathcal{P})$ be a sequentially complete multi-metric space and $f : X \longrightarrow Y$ be a contraction mapping on $(X, \mathcal{P})$. Then the mapping f admits one and only one fixed point $b \in X$.*

Proof. Let $f : X \longrightarrow X$ be a mapping. Then $Fix(f) = \{x \in X : f(x) = x\}$ is the set of all fixed points of the mapping f.

Suppose that f is a contraction mapping with the Lipschitz constant $k = (k_\rho : \rho \in \mathcal{P}) \ll 1$, i.e. $o \leq k_\rho < 1$ for any $\rho \in \mathcal{P}$.

We put $E = \mathbb{R}^{\mathcal{P}}$ and $d(x, y) = (\rho(x, y) : \rho \in \mathcal{P})$ for any $x, y \in X$. Obviously, (X, d, E) is a sequentially complete E-metric space and f is a contraction mapping with the Lipschitz constant $k \in E^{(+,1)}$. Corollary 3.2 completes the proof.

5 Representations of mappings of multi-metric spaces

Let $(X, \mathcal{P})$ be a multi-metric space and the set $\mathcal{P}$ be non-empty.

We say that a mapping $f : X \longrightarrow X$ admits *a pointwise identification* if there exists a family of non-negative functions $\{\varphi_{(f,\rho)}, \psi_{(f,\rho)} : X \times X \longrightarrow \mathcal{R} : \rho \in \mathcal{P}\}$ such that $\varphi_{(f,\rho)}(x, y) \cdot \rho(x, y) \leq \rho(f(x), f(y)) \leq \psi_{(f,\rho)}(x, y) \cdot \rho(x, y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$. The functions $\{\varphi_{(f,\rho)}, \psi_{(f,\rho)} : X \times X \longrightarrow \mathcal{R} : \rho \in \mathcal{P}\}$ are called the *pointwise identification functions* of the mapping f.

Theorem 5.1. *Let $(X, \mathcal{P})$ be a multi-metric space and $f : X \longrightarrow X$ be a mapping with the pointwise identification functions $\{\varphi_{(f,\rho)}, \psi_{(f,\rho)} : X \times X \longrightarrow \mathcal{R} : \rho \in \mathcal{P}\}$. Then for any $\rho \in \mathcal{P}$ there exists a mapping $f_\rho : X/\rho \longrightarrow X/\rho$ such that $\pi_\rho(f(x)) = f_\rho(\pi_\rho(x))$ and $\varphi_{(f,\rho)}(x, y) \cdot \rho(x, y) \leq \hat{\rho}(f_\rho(\pi_\rho(x)), f_\rho(\pi_\rho(y))) \leq \psi_{(f,\rho)}(x, y) \cdot \rho(x, y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$.*

Proof. Fix $u, v \in X/\rho$. Let $x, y \in \pi_{\rho^{-1}}(u)$. Then $\rho(x, y) = 0$ and $0 \leq \rho(f(x),\ f(y)) \leq \psi_{(f,\rho)}(x, y)\rho(x, y) = 0$. Hence $f(\pi_\rho^{-1}(u))$ is a singleton and we put $f_\rho(u) = \pi_\rho(f(\pi_{\rho^{-1}}(u)))$. The mapping $f_\rho : X/\rho \longrightarrow X/\rho$ is well defined. The proof is complete.

Theorem 5.2. *Let $(X, \mathcal{P})$ be a sequentially complete multi-metric space, $f : X \longrightarrow X$ be a mapping and for any $\rho \in \mathcal{P}$ there exists a non-negative number $k_\rho \neq 1$ such that:*

1. If $k_\rho < 1$, then $\rho(f(x), f(y)) \leq k_\rho \cdot \rho(x, y)$ for all $x, y \in X$.

2. If $k_\rho > 1$, then $\rho(f(x), f(y)) \geq k_\rho \cdot \rho(x, y)$ for all $x, y \in X$.

Then the mapping f admits one and only one fixed point $b \in X$.

Proof. By virtue of Theorem 5.1, for any $\rho \in \mathcal{P}$ there exists a mapping $f_\rho : X/\rho \longrightarrow X/\rho$ such that:

1. If $k_\rho < 1$, then $\bar{\rho}(f_\rho(x), f_\rho(y)) \leq k_\rho \cdot \bar{\rho}(x, y)$ for all $x, y \in X/\rho$.
2. If $k_\rho > 1$, then $\bar{\rho}(f(x), f(y)) \geq k_\rho \cdot \bar{\rho}(x, y)$ for all $x, y \in X/\rho$.
3. $\pi_\rho(f(x)) = f_\rho(\pi_\rho(x))$ for any $x \in X$.

We put $q_\rho = k_\rho$ and $C_{(f,\rho)}(x) = f_\rho(x)$ for any $x \in X/\rho$ if $k_\rho < 1$. If $k_\rho > 1$, then for the mapping f_ρ there exists the inverse mapping $C_{(f,\rho)} = f_\rho^{-1}$ and we put $q_\rho = 1/k_\rho$. Let $C_f(x) = (C_{(f,\rho)}(\pi_\rho(x)) : \rho \in \mathcal{P})$ for any $x \in X$. By construction, C_f is a contraction of the multi-metric space $(X, \mathcal{P})$ with the constant $q = (q_\rho : \rho \in \mathcal{P}) \ll 1$. By virtue of Theorem 4.1, there exists a unique fixed point for C_f. Since $Fix(f) = Fix(C_f)$, the proof is complete.

Theorem 5.3. *Let $(X, \mathcal{P})$ be a multi-metric space, $\mathcal{P} = \{\rho_\alpha : \alpha \in A\}$ and $f : X \longrightarrow Y$ be a contraction mapping. Then for any number $q \in (0, 1)$ there exists a family of pseudo-metrics $\mathcal{D} = \{d_\alpha : \alpha \in A\}$ such that:*

1. $d_\alpha(f(x), f(y)) \leq q \cdot d_\alpha(x, y)$ for all $\alpha \in A$ and $x, y \in X$.

2. $\mathcal{T}(d_\alpha) = \mathcal{T}(\rho_\alpha)$ for any $\alpha \in A$.

3. $\mathcal{T}(\mathcal{D}) = \mathcal{T}(\mathcal{P})$.

4. If the space $(X, \mathcal{P})$ is complete, then the space $(X, \mathcal{D})$ is complete too.

5. If the space $(X, \mathcal{P})$ is sequentially complete, then the space $(X, \mathcal{D})$ is sequentially complete too.

Proof. Fix $\alpha \in A$. Consider the projection $\pi_\alpha : X \longrightarrow X/\rho_\alpha$ of (X, ρ_α) onto the metric space $(X/\rho_\alpha, \bar{\rho}_\alpha)$ and the mapping $f_\alpha : X/\rho_\alpha \longrightarrow X/\rho_\alpha$, where $f(\pi_\alpha(x)) = \pi_\alpha(f_\alpha(x))$ and $\rho_\alpha(x, y) = \bar{\rho}_\alpha(\pi_\alpha(x), \pi_\alpha(y))$ for all $x, y \in X$. Denote by (X_α, r_α) the Hausdorff completion of the metric space $(X/\rho_\alpha, \bar{\rho}_\alpha)$. Since f is a contraction, there exist a positive number $k_\alpha < 1$ and a continuous extension $g_\alpha : X_\alpha \longrightarrow X_\alpha$ of the mapping f_α such that $r_\alpha(g_\alpha(x), g_\alpha(y)) \leq k_\alpha \cdot r_\alpha(x, y)$ for all $x, y \in X_\alpha$. By virtue of P.R.Meyers'

theorem [19], there exists a complete metric h_α on X_α such that $\mathcal{T}(r_\alpha) = \mathcal{T}(h_\alpha)$ and $h_\alpha(g_\alpha(x), g_\alpha(y)) \leq q \cdot h_\alpha(x, y)$ for all $x, y \in X_\alpha$. Now we put $d_\alpha(x, y) = h_\alpha(\pi_\alpha(x), \pi_\alpha(y))$ for all $x, y \in X$. The proof is complete.

Theorem 5.4. *Let $f : X \longrightarrow X$ be a mapping of a pseudocompact space X into itself. The following assertions are equivalent:*

1. For any positive number $q < 1$ there exists a family of pseudo-metrics $\mathcal{D}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{D})$ is the topology of the space X and $d(f(x), f(y)) < q \cdot d(x, y)$ for all $\rho \in \mathcal{D}$ and $x, y \in X$.

2. There exists a family of pseudo-metrics $\mathcal{P}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{P})$ is the topology of the space X and $\rho(f(x), f(y)) < \rho(x, y)$ for all $\rho \in \mathcal{P}$ and $x, y \in X$.

Proof. The implication $1 \to 2$ is obvious.

Assume that there exists a family of pseudo-metrics $\mathcal{P}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{P})$ is the topology of the space X and $\rho(f(x), f(y)) < \rho(x, y)$ for all $\rho \in \mathcal{P}$ and $x, y \in X$. Fix $\rho \in \mathcal{P}$. Consider the projection $\pi_\rho : X \longrightarrow X/\rho$ of (X, ρ) onto the metric space $(X/\rho, \bar{\rho})$ and the mapping $f_\rho : X/\rho \longrightarrow X/\rho$, where $f(\pi_\rho(x)) = \pi_\rho(f_\alpha(x))$ and $\rho(x, y) = \bar{\rho}(f(x), f(y))$ for all $x, y \in X$. By construction, X/ρ is a compact metric space and $\bar{\rho}(f(x), f(y)) < \bar{\rho}(x, y)$ for all $x, y \in X/\rho$. Suppose that for any $n \in \omega$ there exists $x_n, y_n \in X$ such that $\rho(f(x_n), f(y_n) \geq (1 - 2^{-n})\rho(x.y)$. Since X/ρ is a metrizable compact space, there exist an infinite subset $L \subseteq \omega$, an infinite subset $M \subseteq L$ and two points $x, y \in X$ such that:

- $\pi_\rho(x)$ is the limit of the subsequence $\{\pi_\rho(x_n) : n \in L\}$ and $\rho(x, x_n) < 2^{-n}$ for any $n \in L$;

- $\pi_\rho(y)$ is the limit of the subsequence $\{\pi_\rho(y_n) : n \in M\}$ and $\rho(y, y_n) < 2^{-n}$ for any $n \in M$.

There exists $m \in \omega$ such that $2^{-m+2} < \rho(x, y) - \rho(f(x), f(y))$. By construction, $\rho(f(x), f(y)) = lim\rho(f(x_n), f(y_n)) = lim\rho(x_n, y_n) = \rho(x, y)$, a contradiction. Thus there exists a number $k_\rho < 1$ such that $\rho(f(x), f(y)) \leq \rho(x, y)$ for all $x, y \in X$. Theorem 5.3 completes the proof.

The next assertion for compact metric spaces was proved by V.Niemytzki (see [22, 17]) in 1936.

Corollary 5.5. *Let $(X, \mathcal{P})$ be a countably compact multi-metric space, $f : X \longrightarrow X$ be a mapping and $\rho(f(x), f(y)) < \rho(x, y)$ for all $\rho \in \mathcal{P}$ and $x, y \in X$. Then the mapping f admits one and only one fixed point $b \in X$.*

The next assertion for compact metric spaces was proved by L.Janos (see [18, 17]).

Theorem 5.6. *Let $f : X \longrightarrow X$ be a continuous mapping of a compact space X into itself. The following assertions are equivalent:*

1. If $X_0 = X$ and $X_{n+1} = f(X_n)$ for any $n \in \omega$, then $\cap\{X_n : n \in \omega\}$ is a singleton.

2. For any positive number $q < 1$ there exists a family of pseudo-metrics $\mathcal{D}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{D})$ is the topology of the space X and $d(f(x), f(y)) < q \cdot d(x, y)$ for all $\rho \in \mathcal{D}$ and $x, y \in X$.

3. There exists a family of pseudo-metrics $\mathcal{P}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{P})$ is the topology of the space X and $\rho(f(x), f(y)) < \rho(x, y)$ for all $\rho \in \mathcal{P}$ and $x, y \in X$.

Proof. The implications $2 \to 3$ and $2 \to 1$ are obvious. The implication $3 \to 2$ follows from Theorem 5.4.

Claim 1. Let d be a continuous pseudo-metric on X, $Y_n = \pi_d(X_n)$ and $d(f(x),\ f(y)) = 0$ provided $d(x, y) = 0$. Then on X/d there exists a metric h such that:

- the metrics h and $\bar{d}$ are equivalent on X/d;

- $h(\pi_d(f(x)), \pi_d(f(y))) \leq q \cdot h(\pi_d(x), \pi_d(y))$ for all $x, y \in X$.

There exists a continuous mapping $f_d : X/d \longrightarrow X/d$ such that $f_d(\pi_d(x)) = \pi_d(f(x))$ for any $x \in X$. By construction, $f_d(Y_n) = Y_{n+1}$ and $Y_0 = X/d$. Assume that $x, y \in X$, $\pi_d(x) \neq \pi_d(y)$ and $\pi_d(x), \pi_d(y) \in \cap\{Y_n : n \in \omega\}$. Since $\cap\{Y_n : n \in \omega\} = \{b\}$ is a singleton, we can consider that $\pi_d(b) \neq \pi_d(y)$. From $\pi_d(y) \in \cap\{Y_n : n \in \omega\}$ it follows that $\pi_d^{-1}(\pi_d(y)) \cap (\cap\{Y_n : n \in \omega\}) \neq \varnothing$, a contradiction. Thus $\cap\{Y_n : n \in \omega\}$ is a singleton and X/d is a metrizable compact space. The L.Janos' theorem (see [18, 17]) completes the proof of the Claim 1.

Claim 2. There exists a family of pseudo-metrics $\mathcal{P}$ such that $\mathcal{T} = \mathcal{T}(\mathcal{P})$ is the topology of the space X and $\rho(f(x), f(y)) = 0$ provided $\rho \in \mathcal{P}$, $x, y \in X$ and $\rho(x, y) = 0$.

Fix a continuous function $g : X \to I = [0, 1]$. We put $g_0 = g$ and $g_{n+1}(x) = g_n(f(x))$ for all $n \in \omega$ and $x \in X$.

Consider the mapping $\pi_g : X \longrightarrow Y_g \subseteq I^\omega$, where $\pi_g(x) = (g_n(x) : n \in \omega)$ and $Y_g = \pi_g(X)$. On the metrizable compact space Y_g fix some metric d_g. Let $\rho_g(x, y) = d_g(\pi_g(x), \pi_g(y))$. By construction, if $\rho_g(x, y) = 0$, then $g_n(x) = g_n(y)$ for any $n \in \omega$.

Since $g_{n+1}(z) = g_n(f(z))$, we have that $g_n(f(x)) = g_n(f(y))$ for any $n \in \omega$. Thus $\rho_g(f(x), f(y)) = 0$ too.

If $\mathcal{F}$ is a family of continuous functions which separates the points of X, then $\mathcal{P} = \{\rho_g : g \in \mathcal{F}\}$ is the desired family of pseudo-metrics. Claim 2 is proved. The implication $1 \to 2$ is proved too. The proof is complete.

The using of this construction has some obstacles.

In the first, the space $(X, \mathcal{P})$ may be not sequentially complete. In the second, the space $(X, \mathcal{P})$ may be complete and each space $(X/\rho, \hat{\rho})$ may be non-complete.

Example 5.7. Let $I_n = [0, 1]$ for any $n \in \mathbb{N}$. We put $X = \{x = (x_n \in I_n : n \in \mathbb{N})$ *the set* $\{n \in \mathbb{N} : x_n \neq 0\}$ *is finite*$\}$. Assume that X is a subspace of the metrizable compact space $I^{\mathbb{N}} = \Pi\{I_n : n \in \mathbb{N}\}$. If $x = (x_n : n \in \mathbb{N})$ and $y = (y_n : n \in \mathbb{N})$, then we put $\rho_n(x, y) =\mid x_n - y_n \mid$. Then the pseudo-metrics $\mathcal{P} = \{\rho_n : n \in \mathbb{N}\}$ generates the topology of the space X. Thus $(X, \mathcal{P})$ is a metrizable multi-metric space. For any $x = (x_n : n \in \mathbb{N})$ we put $f(x) = ((1-2^{-n})x_n : n \in \mathbb{N})$. Then $f : X \longrightarrow X$ is a contraction of the multi-metric space $(X, \mathcal{P})$ with the coefficient $k = (k_n = 1 - 2^{-n} : n \in \mathbb{N}) \ll 1$. If $0 = (0 \in I_n : n \in \mathbb{N})$ the $f(0) = 0$ and 0 is the unique fixed point of the mapping f. The space $(X, \mathcal{P})$ is not sequentially complete and the space $X/\rho = I_n$ is compact for any $n \in \mathbb{N}$.

Remark 5.8. A space X is called a sequential space if a set $F \subseteq X$ is closed if and only if it contains the limit of any sequence $\{x_n \in F : n \in \omega\}$. In a sequential space X any countably compact subset F is closed. Therefore, the assertions 1 - 3 from Theorem 5.6 are equivalent for any countably compact sequential space X.

Remark 5.9. In a first countable space X any pseudocompact subset F is closed. Therefore, the assertions 1 - 3 from Theorem 5.6 are equivalent for any pseudocompact first countable space X and a mapping f for which $Fix(f) \neq \varnothing$.

Example 5.10. Let $\mathcal{Q}$ be the space of all rational numbers. Denote by $\mathcal{P}_1$ the family of all continuous pseudo-metrics on $\mathcal{Q}$. Then $(\mathcal{Q}, \mathcal{P}_1)$ is a complete multi-metric space and $\mathcal{T}(\mathcal{P}_1)$ coincides with the topology of the space $\mathcal{Q}$. Let $X = \mathbb{R} \times \mathcal{Q}$. For any $\rho \in \mathcal{P}_1$ we put $d_\rho((x_1, t_1), (x_2, t_2)) = |x_1 - x_2| + |t_1 - t_2| + \rho(t_1, t_2)$. Then d_ρ is a continuous metric on X. If $\mathcal{P} = \{d_\rho : \rho \in \mathcal{P}_1\}$, then $(X, \mathcal{P})$ is a complete multi-metric space and $\mathcal{T}(\mathcal{P})$ coincides with the topology of the space X. The metric d_ρ generates on X the topology of X. Thus $X = X/d_\rho$ for any $\rho \in \mathcal{P}_1$. Hence the space

$(X/d, \hat{d})$ is not complete for any $d \in \mathcal{P}$ and $(X, \mathcal{P})$ is a complete multi-metric space. If $f(x,t) = (2^{-1}x, 0)$, then $f : X \longrightarrow X$ is a contraction with the coefficient $k = (k_d = 2^{-1} : d \in \mathcal{P}) = 2^{-1}$. Moreover, any contraction of the multi-metric space $(X, \mathcal{P})$ and $(\mathcal{Q}, \mathcal{P}_1)$ admits one and only one fixed point.

Remark 5.11. Let $(X, \mathcal{P})$ be a sequentially complete multi-metric space. If the set $\mathcal{P}$ is finite, then we put $d(x,y) = sup\{\rho(x,y) : \rho \in \mathcal{P}\}$ for any $x, y \in X$. The metric d generate the topology of X and (X,d) is a complete metric space. If $f : X \longrightarrow X$ is a contraction of the multi-metric space $(X, \mathcal{P})$ with the coefficient $(k_\rho : \rho \in \mathcal{P}) < 1$, then f is a contraction of the metric space (X,d) with the coefficient $k = sup\{k_\rho : \rho \in \mathcal{P}\} < 1$. If the set $\mathcal{P} = \{\rho_n : n \in \mathbb{N}\}$ is countable and the space $(X, \mathcal{P})$ is sequentially complete, then the space X is complete metrizable.

Question 5.12. *Let $(X, \mathcal{P})$ be a sequentially complete multi-metric space, the set $\mathcal{P}$ be countable and $f : X \longrightarrow X$ be a contraction. Under which conditions on X there exists a complete metric d of X such that f is a contraction of X?*

If $(X, \mathcal{P})$ is a multi-metric space, $f : X \longrightarrow X$ is a mapping and $x_0 \in X$ is a unique fixed point of the mapping f, then, by virtue of the C.Besaga's theorem [7], on X there exists a complete metric d such that $d(f(x), f(y)) < r^{-1}d(x,y)$ for all $x, y \in X$. But, the topologies $\mathcal{T}(d)$ and $\mathcal{T}(\mathcal{P})$ may be distinct.

6 Dilations of multi-metric spaces

Let $(X, \mathcal{P})$ be a non-empty multi-metric space.

A mapping $f : X \longrightarrow X$ is called *a dilation* if there exists a $\mathcal{P}$-number $k = (k_\rho : \rho \in \mathcal{P}) > 0$ such that $k_\rho \neq 1$ and $\rho(f(x), f(y)) = k_\rho \rho(x,y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$.

Any dilation is a uniform homeomorphism. Moreover, the inverse mapping $f^{-1} : X \longrightarrow X$ of the dilation with the coefficient $k = (k_\rho : \rho \in \mathcal{P})$ is a dilation with the coefficient $k^{-1} = (k_\rho^{-1} : \rho \in \mathcal{P})$.

From Theorem 5.1 it follows

Corollary 6.1. *Let $f : X \longrightarrow X$ be a mapping, $k = (k_\rho : \rho \in \mathcal{P})$ be a $\mathcal{P}$-number and $k_\rho > 0$, $\rho(f(x), f(y)) = k_\rho f(x,y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$. Then for any $\rho \in \mathcal{P}$ there exists a mapping $f_\rho : X/\rho \longrightarrow X/\rho$ such that $\pi_\rho(f(x)) = f_\rho(\pi_\rho(x))$ and $\bar{\rho}(f_\rho(u), f_\rho(v)) = k_\rho \bar{\rho}(u,v)$ for all $x \in X$ and $u, v \in X/\rho$.*

Corollary 6.1 permits to study the construction of the general dilation of the multi-metric space. Let $(X, \mathcal{P})$ be a non-empty multi-metric space and $k = (k_\rho : \rho \in \mathcal{P}) > 0$ be a positive $\mathcal{P}$-number. Suppose that $f : X \longrightarrow X$ is a mapping and $\rho(f(x), f(y)) = k_\rho f(x, y)$ for all $x, y \in X$ and $\rho \in \mathcal{P}$. We put $\mathcal{P}' = \{\rho \in \mathcal{P} : k_\rho < 1\}$, $\mathcal{P}'' = \{\rho \in \mathcal{P} : k_\rho > 1\}$ and $\mathcal{P}^0 = \{\rho \in \mathcal{P} : k_\rho = 1\}$. If $\mathcal{P}^0 \neq 0$, then we say that f is *a general dilation*. For any $\rho \in \mathcal{P}$ consider the mapping $f_\rho : X/\rho \longrightarrow X/\rho$, where $\pi_\rho(f(x)) = f_\rho(\pi_\rho(x))$ and $\bar{\rho}(f_\rho(\pi_\rho(x))), f_\rho(\pi_\rho(y)) = k_\rho \cdot \rho(x, y)$ for all $x, y \in X$. If $k_\rho = 1$, then f_ρ is an isometry of the metric space $(X/\rho, \bar{\rho})$. If $k_\rho \neq 1$, then f_ρ is a dilation with the coefficient k_ρ.

Fix $\rho \in \mathcal{P}'$. We put $N_{(\rho,f)}(u) = f_\rho(u)$, $S_{(\rho,f)(u)} = q_{(\rho,f)}(u) = u$ for any $u \in X/\rho$.

Fix $\rho \in \mathcal{P}^0$. We put $N_{(\rho,f}(u) = Q_{(\rho,f}(u) = u$. $S_{(\rho,f)} = f_\rho(u)$ for any $u \in X/\rho$.

Now we consider the embedding $e_{\mathcal{P}} : X \longrightarrow \Pi\{X/\rho : \rho \in \mathcal{P}\}$ and the mappings $N_f(x) = e_{\mathcal{P}}^{-1}(N_{(\rho,f)}(\pi_\rho(x)) : \rho \in \mathcal{P})$, $S_f(x) = e_{\mathcal{P}}^{-1}(S_{(\rho,f)}(\pi_\rho(x)) : \rho \in \mathcal{P})$, $Q_f(x) = e_{\mathcal{P}}^{-1}(Q_{(\rho,f)}(\pi_\rho(x)) : \rho \in \mathcal{P})$.

The mappings N_f, S_f, Q_f are homeomorphisms of the space $(X, \mathcal{P})$ associated with the mapping f. By construction, we have the following four properties.

Property 1. $f = S_f \cdot N_f \cdot Q_f = N_f \cdot S_f \cdot Q_f = Q_f \cdot N_f \cdot S_f = S_f \cdot Q_f \cdot N_f$.

Property 2. $\rho(S_f(x), S_f(y)) = \rho(x, y)$ for any $x, y \in X$ and $\rho \in \mathcal{P}$.

Property 3. $\rho(N_f(x), N_f(y)) = k_\rho \rho(x, y)$ for any $\rho \in \mathcal{P}'$ and $\rho(N_f(x), N_f(y)) = \rho(x, y)$ for any $\rho \in \mathcal{P} \setminus \mathcal{P}'$ and all $x, y \in X$.

Property 4. $\rho(Q_f(x), Q_f(y)) = k_\rho(x, y)$ for any $\rho \in \mathcal{P}''$ and $\rho(N_f(x), N_f(y)) = \rho(x, y)$ for any $\rho \in \mathcal{P} \setminus \mathcal{P}''$ and all $x, y \in X$.

We say that S_f is the isometrically component of the mapping f, N_f is the negative component of the mapping f and Q_f is the positive component of the mapping f.

Assume that $\mathcal{P}^0 = \varnothing$, i.e. f is a dilation. In this case $D_f = N_f \circ Q_f^{-1}$ is a contraction mapping. By construction, $Fix(f) = Fix(D_f)$. Thus from Corollary 6.1 it follows the next assertion.

Corollary 6.2. *Let $(X, \mathcal{P})$ be a sequentially complete multi-metric space and $f : X \longrightarrow X$ be a dilation. Then the mapping f admits one and only one fixed point.*

7 Dilations of E-metric spaces

Fix an m-scale E.

Let (X, d, E) be an E-metric space.

A mapping $f : X \longrightarrow X$ is *a dilation mapping* if there exist an element $q \in E^{(+,1)}$ and an element $k \geq 0$ such that $q \cdot k = 1$ and $d(f(x), f(y)) = k \cdot d(x, y)$ for all $x, y \in X$. Every dilation mapping f is a uniformly continuous homeomorphism and f^{-1} is a contraction with the Lipschitz constant $q \in E^{(+,1)}$.

From Corollary 3.2 it follows

Corollary 7.1. *Let (X, d, E) be a sequentially complete E-metric space and $f : X \longrightarrow X$ be a dilation mapping. Then the mapping f admits one and only one fixed point $b \in X$. Moreover, $b = lim_{n\to\infty} n(x, f^{-1})$ for any point $x \in X$.*

8 On M-complete multi-metric spaces

We say that the pseudo-metric d on a space X is *M-complete* if it is continuous, the metric space $(X/d, \bar{d})$ is complete and the sequence $\{x_n : n \in \mathbb{N}\}$ has an accumulation point provided $lim_{n\to\infty} d(x, x_n) = 0$ for some $x \in X$.

Remark 8.1. The pseudo-metric d on a space X is M-complete if and only if the metric space $(X/d, \bar{d})$ is complete and the projection $\pi_d : X \longrightarrow X/d$ is a continuous closed mapping with the countably compact fibers $\pi_d^{-1}(u)$. Thus a space with a complete pseudo-metric is an M-space [20]. The spaces related to M-spaces were studied in [4, 5, 9, 10].

A multi-metric space $(X, \mathcal{P})$ is said to be *an M-complete* multi-metric space if there exists some M-complete pseudo-metric $d \in \mathcal{P}$. Denote by $s(\mathcal{P})$ the set of all M-complete pseudo-metrics $d \in \mathcal{P}$.

Theorem 8.2. *Any M-complete multi-metric space $(X, \mathcal{P})$ is sequentially complete.*

Proof. Fix $d \in s(\mathcal{P})$. Let $L = \{x_n \in X : n \in \mathbb{N}\}$ be a fundamental sequence. Since the metric space $(X/d, \bar{d})$ is complete, there exists a point $b \in X$ such that $lim_{n\to\infty} d(b, x_n) = 0$. In this case $F = \{x \in X : d(b, x) = 0\}$ and $H = F \cup L$ are countably compact closed subsets of the space X. Thus the sequence $\{x_n : n \in \mathbb{N}\}$ is convergent in the subspace H. The proof is complete.

Theorem 8.3. *Let $(X, \mathcal{P})$ be an M-complete multi-metric space, $d \in s(\mathcal{P})$ and $f : X \longrightarrow X$ be a mapping with the properties:*

1. $\rho(f(x), f(y)) < \rho(x, y)$ for all $\rho \in \mathcal{P}$ and $x, y \in X$;

2. There exists a positive number $k < 1$ such that $d(f(x), f(y)) < k \cdot d(x, y)$ for all $x, y \in X$.

Then the mapping f admits one and only one fixed point $b \in X$. Moreover, $b = lim_{n\to\infty} n(x, f)$ for any point $x \in X$.

Proof. Fix a point $x \in X$. Then $\{x_n = n(x, f) : n \in \mathbb{N}\}$ is a d-fundamental sequence and there exists a point $b \in X$ such that $lim_{n\to\infty} d(b, x_n) = 0$. In this case $F = \{x \in X : d(b, x) = 0\}$ and $H = F \cup L$ are countably compact closed subsets of the space X. By construction, $f(F) \subseteq F$ and $f(H) \subseteq H$. Thus the mapping $g = f|H : H \longrightarrow H$ satisfies the conditions of Corollary 5.5. The proof is complete.

9 Locally convex linear spaces

Let L be a locally convex linear topological space. Then the topology of the space L is generated by a family $\mathcal{N} = \{n_\alpha : \alpha \in A\}$ of pseudo-norms. Any pseudo-norm n_α generates the invariant pseudo-metric $\rho_\alpha(x, y) = n_\alpha(x - y)$. Thus any locally convex space can be studied as a multi-metric space.

We say that $(L, \mathcal{N})$ is a multi-normed space. For any pseudo-norm $\nu \in \mathcal{N}$ we have the linear mapping $\pi_\nu : L \longrightarrow L/\nu$ onto a normed space $(L/\nu, \bar{\nu})$ and $\bar{\nu}(f(x)) = \nu(x)$ for any $x \in L$.

Consider the embedding $e_{\mathcal{N}} : L \longrightarrow \Pi\{L/\nu : \nu \in \mathcal{N}\}$.

For any $\nu \in \mathcal{N}$ let $g_\nu : L/\nu \longrightarrow L/\nu$ be a dilation with the positive coefficient $k_\nu \neq 1$. Then the mapping $g : L \longrightarrow L$, where $g(x) = e^{-1}_{\mathcal{N}}(g_\nu(\pi_\nu(x)) : \nu \in \mathcal{N})$, is a dilation with the $\mathcal{N}$-coefficient $k = (k_\nu : \nu \in \mathcal{N})$. If the space $(L, \mathcal{N})$ is sequentially complete, then f admits one and only one fixed point.

If $\mathcal{N}$ is a finite set then L is a normed space. In this case the mapping $f : L \longrightarrow L$ may be a dilation for a multi-normed space $(L, \mathcal{N})$ and not a dilation for a normed space L.

Example 9.1. Let $E^2 = \{(x, y) : x, y \in \mathbb{R}\}$ be the Euclidean plane with the norm $\| (x, y) \| = (x^2 + y^2)^{1/2}$. On E^2 consider the pseudo-norms $\mathcal{N} = \{\nu_1, \nu_2\}$, where $\nu_1(x, y) = \mid x \mid$ and $\nu_2(x, y) = \mid y \mid$. Now we consider the homeomorphism $f : E^2 \longrightarrow E^2$, where $f(x, y) = (k_1 x, k_2 y)$, $0 < k_1 < 1$ and $1 < k_2$. By construction, f is a dilation of the multi-normed space $(E^2, \mathcal{N})$. For the normed space $(E^2, \| \cdot \|)$ the mapping f may do not be a dilation.

Let $k_1 = 2^{-1}$ and $k_2 = 2$. Then:

1) $f(2,1) = (1,2)$ and $\| f(0,0) - f(2,1) \| = \| (0,0) - (1,2) \| = \| (0,0) - (2,1) \|$;

2) $f(2,2) = (1,4)$ and $\| f(0,0) - f(2,2) \|^2 = \| (0,0) - (1,4) \|^2 = 17$ and $\| (0,0) - (2,2) \|^2 = 8$;

3) $f(1,2^{-2}) = (2^{-1},2^{-1})$, $\| f(0,0) - f(1,2^{-2}) \|^2 = \| (2^{-1},2^{-1}) \|^2 = 2^{-1}$ and $\| (0,0) - (1,2^{-4}) \|^2 = \| (1,2^{-4}) \|^2 = 17/16$.

10 Spaces with chainable pseudo-metrics

The method of construction of representations of a mapping $f : X \longrightarrow X$ proposed in Section 5 permits using effectively the methods of the fixed point theory in metric spaces.

Fix an m-scale E. Assume that E as the lattice is *lower reticulated*, i.e. every non-empty lower bounded subset $B \subseteq E$ has the infimum $\wedge B$ in E.

A mapping $d : X \times X \longrightarrow E$ is called a pseudo-metric over m-scale E or an E-pseudo-metric if it is satisfying the following properties: $d(x,x) = 0$, $d(x,y) = d(y,x)$, $d(x,y) \leq d(x,z) + d(yz,y)$ for all $x,y,z \in X$. Every E-pseudo-metric is non-negative.

Let d be an E-pseudo-metric on a space X. If $a \in E$ and $x \in X$, then $B(x,d,a) = \{y \in X : d(x,y) < a\}$

The pseudo-metric d is continuous on X if the set $B(x,d,U) = \{y \in X : d(x,y) \in U\}$ is open for each $U \in N(0,E)$ and any $x \in X$. The pseudo-metric d generates on X the canonical equivalence relation: $x \sim y$ iff $d(x,y) = 0$. Let X/d be the quotient set with the canonical projection $\pi_d : X \longrightarrow X/d$ and metric $\bar{d}(u,v) = \rho(\pi_d^{-1}(u), \pi_d^{-1}(v))$. Then $(X/d, d, E)$ is an E-metric space and the pseudo-metric d is continuous on X if and only if the mapping π_d is continuous.

Let $r \in E$. We say that the pseudo-metric d is *r-chainable* if for any two distinct points $x,y \in X$ there exist $n \in \mathbb{N}$, $U = U(x) \in N(0,E)$ and a chain $x_0, x_1, ..., x_n \in X$ such that $x = x_0 \in B(x,d,U) \subseteq B(x,d,r)$, $y = x_n$ and $d(x_{i-1}, x_i) \leq r$ for each $i \leq n$. We say that $x_0, x_1, ..., x_n \in X$ is an r-chain between x, y.

If the space X is connected, the pseudo-metric d is continuous, $r \in E$ and for any point $x \in X$ there exists $U = U(x) \in N(0,E)$ such that $B(x,d,U) \subseteq B(x,d,r)$, then the pseudo-metric d is r-chainable.

Assume that d is a continuous r-chainable E-pseudo-metric on a space X. For any two points $x, y \in X$ it is determined the element $d_r(x,y) = \wedge\{d(x_0,x_1) + d(x_1,x_2) + ... + d(x_{n-1},x_n) : x_0, x_1, ..., x_n \in X$ *is an* r*-chain between* $x, y\}$.

Property 1. $d(x,y) \leq d_r(x,y)$ for all $x, y \in X$. Moreover, if $d(x,y) \leq r$, then $d(x,y) = d_r(x,y)$.

Proof. We have $d(x,y) \leq d(x_0,x_1) + d(x_1,x_2) + ... + d(x_{n-1},x_n)$ for any r-cain $x_0, x_1, ..., x_n \in X$ between x, y. If $d(x,y) \leq r$, then $x_0 = x, x_1 = y$ is an r-chain between x, y.

Property 2. $X/d = X/d_r$ and the E-metrics $\bar{d}$ and $\bar{d}_r$ on X/d are equivalent, i. e. $\mathcal{T}(\bar{d}) = \mathcal{T}(\bar{d}_r)$.

Property 3. If the E-metric space $(X/d, \bar{d}, E)$ is complete (sequentially complete, respectively), then the E-metric space $(X/d_r, \bar{d}_r, E)$ is complete (sequentially complete, respectively) too.

Property 4. Let $f : X \longrightarrow X$ be a mapping $k \in E$ and $d(f(x), f(y)) \leq k \cdot d(x,y)$ provided $d(x,y) \leq r$. Then $d_r(f(x), f(y)) \leq k \cdot d_r(x,y)$ for all $x, y \in X$.

From Properties 1 - 4 and Theorems 3.1 and 4.1 it follows.

Corollary 10.1. *Let* $f : X \longrightarrow X$ *be a mapping,* $k \in E^{(+,1)}$, $r \in E$, (X,d,E) *be an* r*-chainable* E*-metric space and* $d(f(x), f(y)) \leq k \cdot d(x,y)$ *provided* $d(x,y) \leq r$. *Then for each point* $x \in X$ *the Picard orbit* $T(f,x) = \{n(x,f) : n \in \omega\}$ *is a fundamental sequence of the metric space* (X,d,E). *Moreover, if the space* (X,d,E) *is sequentially complete, then the mapping* f *admits one and only one fixed point* $b \in X$.

Corollary 10.2. *Let* $(X, \mathcal{P})$ *be a sequentially complete multi-metric space,* $f : X \longrightarrow X$ *be a mapping and for any* $\rho \in \mathcal{P}$ *there exist two positive numbers* $r(\rho)$ *and* $k(\rho) < 1$ *such that the pseudo-metric* ρ *is* $r(\rho)$*-chainable and* $\rho(f(x), f(y)) \leq k(\rho)$ *provided* $\rho(x,y) \leq r(\rho)$. *Then the mapping* f *admits one and only one fixed point* $b \in X$.

The Corollaries 10.1 and 10.2 for metric spaces were proved by M. Edelstein [11, 13, 17].

Example 10.3. Let A be an uncountable set. We put $I_\alpha = I = [0,1]$ for any $\alpha \in A$. Let $Y = I^A = \Pi\{I_\alpha : \alpha \in A\}$, $(\lambda)_A = (\lambda_\alpha = \lambda : \alpha \in A) \in Y$ and $X_\lambda = \{x = (x_\alpha : \alpha \in A) \in Y :$ *the set* $\{\alpha : x_\alpha \neq \lambda\}$ *is countable*$\}$ for any $\lambda \in I$. The subspace X_λ is countably compact and dense in Y for any $\lambda \in I$. Thus $X = X_0 \cup X_1$ is a countably compact space and X_0, X_1 are dense countably compact subspaces of the spaces X and Y. The space $E = \mathbb{R}^A$ is

a topological semifield and an m-scale. By construction, $X \subseteq I^A \subseteq E$. We put $d(x,y) = (|x_\alpha - y_\alpha| : \alpha \in A)$ for any pair of points $x = (x_\alpha : \alpha \in A) \in Y$ and $y = (y_\alpha : \alpha \in A) \in Y$. Thus (X, d, E) is an E-metric space.

For any $x = (x_\alpha : \alpha \in A) \in Y)$ we put $S_Y(x) = (1 - x_\alpha : \alpha \in A)$ and $S_X = S_Y|X$.

Obviously $S_Y : Y \longrightarrow Y$ and $S_X : X \longrightarrow X$ are continuous involutions and the space Y is compact. We have $S_Y(Y) = Y$ and $S_X(X) = X$. The mapping S_Y has a unique fixed point $(2^{-1})_A$. The mapping S_X is without fixed points.

11 Cauchy sequences of sets and mappings

A subset $L \subseteq X$ is a bounded set if any continuous function $f : X \longrightarrow \mathbb{R}$ is bounded on L. A space X is b-complete if the closure of any bounded subset is compact. For any space X there exists the maximal b-complete extension $\mu X = \cup\{cl_{\beta X} L : L \textit{ is a bounded subset of } X\}$.

For any sequence $\{L_n : n \in \mathbb{N}\}$ of subsets of a space X we put $Lim_X\{L_n : n \in \mathbb{N}\} = \cap\{cl_X(\cup\{L_i : i \geq n\}) : n \in \mathbb{N}\}$.

A sequence $\{L_n : n \in \mathbb{N}\}$ of subsets of a space X is called a Cauchy sequence if for any two sets $A = \{x_i \in A_{n_i} : i \in \mathbb{N}, n_i < n_{i+1}\}$ and $B = \{y_j \in A_{n_j} : j \in \mathbb{N}, n_j < n_{j+1}\}$ do not exists a continuous function $f : X \longrightarrow \mathbb{R}$ such that $A \subseteq f^{-1}(0)$ and $B \subseteq f^{-1}(1)$.

For a mapping $g : X \longrightarrow X$ and $n \in \mathbb{N}$ we put $g^{(1)} = g$ and $g^{(n+1)} = g \circ g^{(n)}$. Then $n(x, g) = g^{(n)}(x)$.

Theorem 11.1. *Let X be a b-complete space, $g : X \longrightarrow X$ be a continuous mapping and $\{g^{(n)}(x) : n \in \mathbb{N}\}$ is a Cauchy sequence for any point $x \in X$. Then for any point $x \in X$ there exists a unique fixed point $b = b(x) \in Fix(g)$ such that $lim g^{(n)}(x) = b$.*

Proof. If $\{F_n : n \in \mathbb{N}\}$ is a Cauchy sequence of the space X, then:

- $Lim_X\{F_n : n \in \mathbb{N}\}$ is a singleton;

- any sequence $\{x_i \in A_{n_i} : i \in \mathbb{N}, n_i < n_{i+1}\}$ is convergent.

This complete the proof.

For any continuous mapping $g : X \longrightarrow X$ there exists a unique continuous extension $\beta g : \beta X \longrightarrow \beta X$ on the Stone-Čech compactification βX of the space X. Let $\mu g = g|\mu X$.

Theorem 11.2. *Let X be a space, $g : X \longrightarrow X$ be a continuous maping and $\{g^{(n)}(F) : n \in \mathbb{N}\}$ is a Cauchy sequence for any non-empty finite set $F \subseteq X$. Then:*

1. There exists a unique points $b \in \mu X$ such that $Fix(\mu g) = \{b\}$.

2. $lim_{n\to\infty} g^{(n)}(x) = b$ for any $x \in \mu X$.

3. $\{g^{(n)}(F) : n \in \mathbb{N}\}$ is a Cauchy sequence for any non-empty finite set $F \subseteq \mu X$.

Proof. Any continuous function $f : X \longrightarrow \mathbb{R}$ admits a continuous extension $\mu f : \mu X \longrightarrow \mathbb{R}$.

Let $x \in X$. Since $\{g^{(n)}(x) : n \in \mathbb{N}\}$ is a Cauchy sequence, there exists a unique point $b = b(x, g)$ such that $lim_{n\to\infty} g^{(n)}(x) = b$. Then $b(x, g) \in Fix(\mu g)$.

Suppose that $b, c \in Fix(\mu g)$ and $b \neq c$. There exist two open subsets V and W of μX and two bounded subsets H and L of X such that $V \cap W = \varnothing$, $b \in cl_X H \subseteq V$ and $c \in cl_X L \subseteq W$. There exist two sequences of open subsets $\{V_n : n \in \mathbb{N}\}$ and $W_n : n \in \mathbb{N}\}$ of μX such that $b \in cl_{\mu X} V_{n+1} \subseteq V_n \subseteq V$, $\mu g(V_{n+1} \subseteq V_n$, $c \in cl_{\mu X} W_{n+1} \subseteq W_n \subseteq W$ and $\mu g(V_{W+1} \subseteq W_n$. By construction, $F = \cap\{V_n : n \in \mathbb{N}\}$ and $K = \cap\{W_n : n \in \mathbb{N}\}$ are closed G_δ-subsets of μX, $\mu g(F) \subseteq F$ and $\mu g(K) \subseteq K$. Since the sets H and L are bounded, $b \in F \cap cl_{\mu X} H \subseteq V$ and $c \in K \cap cl_{\mu X} L \subseteq W$, there exist two point $x_1 \in F \cap H$ and $y_1 \in K \cap L$. We put $x_{n+1} = g^{(n)}(x_1)$, $y_{n+1} = g^{(n)}(y_1)$ and $\Phi_n = \{x_n, y_n\}$. By conditions, $\{\Phi_n : n \in \mathbb{N}\}$ is a Cauchy sequence. By constructions, $b(x_1, g) \in F$ and $b(y_1, g) \in K$. Then there exists a continuous function $f : \mu X \longrightarrow \mathbb{R}$ such that $f(x_n) = 0$ and $f(y_n) = 1$ for any $n \in \mathbb{N}$, a contradiction. The assertion 1 is proved and the assertion 2 is proved for any $x \in X$.

For every point $x \in \mu X$ and any open subset U of μX we put $i(x, U) = sup\{n \in \mathbb{N} : \mu g^{(n)}(x) \notin cl_{\mu X} U\}$.

Let U be an open subset of μX and $b \in U$.

Then $i(x, U) < \infty$ for any $x \in X$. We affirm that $i(x, U) < \infty$ for any $x \in \mu X$. Let $x \in \mu X \mu$ and $i(x, U) = \infty$. There exist an infinite subset $M \subseteq \mathbb{N}$ and a sequence $\{W_n : n \in M\}$ of open subsets of μX such that $x \in cl_X W_m \subseteq W_n$ and $U \cap \mu g^{(m)}(W_m) = \varnothing$ for all $n, m \in M$ and $m < n$. Let $F = \cap\{W_n : n \in M\}$. Then $x \in F$, $F \cap X \neq \varnothing$ and $U \cap \mu g^{(n)}(F) = \varnothing$ for each $n \in M$, a contradiction. The assertion 2 is proved and the assertion 3 is proved for any non-empty finite subset $F \subseteq \mu X$. The proof is complete.

One of the Meyers' theorem [19, 17] one can formulated in the following way

Theorem 11.3. *Let $g :: X \longrightarrow X$ be a continuous mapping of a metrizable space X, $b \in X$ and $g(b) = b$. The next assertions are equivalent:*

1. For each positive number $k < 1$ there exists a metric d on X such that $\mathcal{T}(d)$ is the topology of the space X and $d(g(x), g(y)) \leq kd(x, y)$ for all $x, y \in X$. Moreover, if the space X is complete metrizable, then (X, d) is a completete metric space.

2. There exists an open subset U of X such that $b \in U$ and $\{g^{(n)}(U); n \in \mathbb{N}\}$ is a Cauchy sequence.

From Theorem 11.3 it follows

Corollary 11.4. *Let $g : X \longrightarrow X$ be a continuous mapping of a space X, $b \in X$ and $g(b) = b$. The next assertions are equivalent:*

1. For each positive number $k < 1$ there exists a family of pseudo-metrics $\mathcal{P}$ on X such that $\mathcal{T}(\mathcal{P})$ is the topology of the space X and $d(g(x), g(y)) \leq kd(x, y)$ for all $x, y \in X$ and $d \in \mathcal{P}$. Moreover, if the space X is complete multi-metrizable, then $(X, \mathcal{P})$ is a complete multi-metric space.

2. There exists a family of pseudo-metrics $\mathcal{D}$ such that:

- $\mathcal{T}(\mathcal{D})$ is the topology of the space X;

- if $x, y \in X$, $d \in \mathcal{D}$ and $d(x, y) = 0$, then $d(g(x), g(y)) = 0$;

- for any $d \in \mathcal{D}$ there exists an open subset $U(d)$ of X for which $B(b, d, r) \subseteq U(d)$ for some $r > 0$ and for each $\epsilon > 0$ there exists $m \in \mathbb{N}$ such that $\cup\{g^{(n)}(U(d)) : n \in \mathbb{N}, n \geq m\} \subset B(b, d, \epsilon)$.

Example 11.5. Let $X = \mathbb{R}^{\mathbb{N}}$, $0 < k < 1$ and for any point $x = (x_1, x_2, ...) \in X$ we put $g(x) = (2^{-1}kx_1, 2^{-2}kx_2, ...)$. Then:

- X is a complete metrizable space, a topological ring and a linear locally covex topological space;

- $\{g^{(n)}(F) : n \in \mathbb{N}\}$ is a Cauchy sequence for any non-empty compact set $F \subseteq X$;

- for each positive number $q < 1$ there does not exist a metric d on X such that $\mathcal{T}(d)$ is the topology of the space X and $d(g(x), g(y)) \leq qd(x, y)$ for all $x, y \in X$;

- on X there exists a complete invariant metric d such that $\mathcal{T}(d)$ is the topology of the space X and $d(g(x), g(y)) \leq d(x, y)$ for all $x, y \in X$;

- the mapping g is continuous;

- if $d_n(x, y) = |x_n - y_n|$ for all $x = (x_1, x_2, ...) \in X$ and $y = (y_1, y_2, ...) \in X$, then $\mathcal{P} = \{d_n : n \in \mathbb{N}\}$ is a complete family of pseudo-metrics on X, $\mathcal{T}(\mathcal{P})$ is the topology of the space X and $d_n(g(x), g(y)) \leq 2^{-n}kd_n(x, y)$ for all $x, y \in X$ and $n \in \mathbb{N}$.

References

[1] M. Ya. Antonovskii, V. G. Boltyanskii and T. A. Sarymsakov, *Topological semifields*, Tashkent, 1960.

[2] M. Ya. Antonovskii, V. G. Boltyanskii and T. A. Sarymsakov, *An outline of the theory of topological semifields*, Russian Math. Surveys 21:4 (1966), 163-192.

[3] M. Ya. Antonovskii, V. G. Boltyanskii and T. A. Sarymsakov, *Topological semifields and their applications to general topology*, American Math. Society. Translations, Seria 2, 106, 1977.

[4] A. V. Arhangel'skii, *A class of spaces which contains all metric and all locally compact spaces*, Amer. Math. Soc. Transl. 92 (1970), 1-39.

[5] A. V. Arhangel'skii, *Mappings and spaces*, Russian Math. Surveys 21:4 (1966), 115-162.

[6] V. Berinde and M. Choban, *Remarks on some completeness conditions involved in several common fixed point theorems*, Creative Mathematics and Informatics 19:1 (2010), 3-12.

[7] C. Bessaga, *On the converse of the Banach fixed point principle*, Colloq. Math. 7:1 (1959), 41-43.

[8] L.Calmuţchi and M. Choban, *On a mappings with fixed points*, Buletin Ştiinţific. Universitatea din Piteşti, Matematica şi Informatica, 3 (1999), 91-96.

[9] J. Chaber, M.M.Čoban and K. Nagami, *On monotonic generalizations of Moore spaces, Čech complete spaces and p-spaces,* Fund. Math. 84 (1974) 107-119.

[10] M. Choban, *The open mappings and spaces*, Supl. Rend. Circolo Matem. di Palermo. Serie II, numero 29 (1992) 51-104.

[11] M. Edelstein, *An Extension of Banach's contraction principle*, Proceed. Amer. Math. Society 12 (1961), 7-10.

[12] R.Engelking, *General Topology*, PWN. Warszawa, 1977.

[13] A. Granas and J. Dugundji, *Fixed point theory*, Springer, New York, 2003.

[14] G. Gruenhage, *Generalized Metric spaces*, In. Handbook of Set-Theoretic Topology, K. Kunen and J. E. Vaughan, eds., Elsevier, Amsterdam, 1984.

[15] G. Gruenhage, *Metrizable spaces and generalizations*, In: Recent Progress in General Topology II, M. Hušek and J. van Mill, eds., Elsevier, Amsterdam, 2002, 201-225.

[16] K. Iseki, *On Banach theorem on contractive mappings*, Proceed. Japan Academy, 41 :2 (1965), 145-146.

[17] A. A. Ivanov, *Fixed points of mappings of metric spaces*, Journa of Soviet Mathematics 12:1 (1979), 1-64.

[18] L. Janos, *A convers of Banach's contraction theorem*, Proceed. Amer. Math. Soc. 18:2 (1967), 287-289.

[19] P. R. Meyers, *A converse to Banach's contraction theorem*, J. Res. Nat. Bur. Standards 71B (1967), 73-76.

[20] K. Morita, *A survey of the theory of M-spaces*, General Topology and Appl. 1 (1971), 49-55.

[21] S.I.Nedev and M.M.Choban, *General conception of the metrizability of topological spaces*, Annales of the Sophia University, Mathematics, 65 (1973), 111-165.

[22] V.Niemytzki, *The method of fixed points in analysis*, Uspekhi Matem. Nauk 1 (1936), 141-174.

[23] B. N. Sadovskii, *Limit-compact and condensing operators*, Russian Math. Surveys 27:1 (1972), 85-155.

[24] J.Stepfans, S.Watson and W.Just, *A topological fixed point theorem for compact Hausdorff spaces*, York University, Preprint, 1991.

[25] I.V.Yashchenko, *Fixed points of mappings and convergence of the iterations of the dual mappings*, In: Obshchaya Topologia: Otobrajenia, Proizvedenia i Razmernosti Prostranstv, Moskva: Moskov. Gosud. Unit, 1995, 131-142.

[26] T. Zamfirescu, *Fixed point theorem in metric spaces*, Arch. Math. (Basel) 23 (1972), 292-298.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

FOLDED SADDLE-NODES AND THEIR NORMAL FORM REDUCTION IN A NEURONAL RATE MODEL*

Rodica Curtu†

Abstract

The paper investigates the existence of folded singularities in a dynamical system of two fast and two slow equations. The normal form of the system near its fold curve is constructed. Then it is used to determine the analytical conditions satisfied by a folded singularity. In particular, we find that there is a parameter region where folded saddle-nodes of type II exist. In the neighborhood of those points the system possesses a stable folded node and an unstable true equilibrium, and the local dynamics is complex.

MSC 2010: 37G05, 34E13, 92C20

keywords: inhibitory neural networks, folded singularities, canards

Foreword

This work is a tribute to my mentor, Professor Dr. Adelina Georgescu.

I met Dr. Georgescu at the beginning of the year 1997, in Bucharest, Romania. At that time she was the Director of the Institute of Applied Mathematics of the Romanian Academy, and she was very busy reorganizing

*Accepted for publication on December 12, 2010.

†rodica-curtu@uiowa.edu Department of Mathematics, and Program in Applied Mathematical and Computational Sciences, University of Iowa, Iowa City, IA 52242, USA

its activity. Nevertheless, she still found time to talk with me and agreed to supervise my doctoral thesis in mathematics.

Our encounter was of incalculable value to my professional development: she enlarged my horizon by pointing out that mathematics can be successfully used to study biological systems; she introduced me to the exciting field of applied dynamical systems and bifurcation theory; she even taught me with patience and critical view how to write a scientific paper. Moreover, when I continued my studies in the United States, she has been supportive; that allowed me to write and finish in parallel two doctoral theses.

I have always admired and respected Professor Adelina Georgescu: she was extremely energetic; passionate about mathematics; dedicated to her work, her family, and her country; a wonderful mentor and collaborator. But most of all, she was an excellent researcher and an example of human and scientific integrity. She passed away at the beginning of May 2010 after a long battle with cancer that she fought with courage and dignity. It was a sad day! Romania lost an important scientist, but we, her disciples, lost much more; we lost a very good friend.

I thank the editors of this special issue for giving me the opportunity to express my deep respect and admiration to my mentor. This paper is written In The Memory of Adelina Georgescu!

1 Introduction

This article is the second in a series of three papers investigating the formation of mixed mode oscillations in a neuronal competition model of two reciprocally inhibitory populations.

Previous studies [4], [8], [9] showed that the system can exhibit a large range of dynamics such as approaching a steady state (equal level of activity) for both populations (the *fusion*), anti-phase oscillations with the period of oscillations decreasing with strength of the external stimulus (*escape*), anti-phase oscillations with their period increasing with stimulus strength (*release*), or a bistability regime of two distinct equilibria assimilated to a *winner-take-all* situation.

In a more recent paper [2] we reported another possible behavior. This is a more complex pattern of activity called the *mixed mode oscillations* (MMOs). MMOs consist of two distinct amplitudes in a cycle; some are small amplitude oscillations but they are followed by large exchanges of relaxation-

type. While the formation of small amplitude oscillations can be partially explained through the presence of a singular Hopf bifurcation point [2], the complete mechanism of MMOs is still unclear.

We continue our work from [2] by showing here that there exists a parameter regime where the neuronal rate model possesses folded saddle-node singularities of type II. Note that the model is a slow-fast dynamical system, and its layer problem (or fast sub-system) has a fold curve (see Section 2). A folded singularity is a point on the fold curve which is an equilibrium of an associated *desingularized flow* [10] (see also Section 4). Obviously, it is not an equilibrium of the full (original) system and therefore it is not easily detected. However its presence is important because it may lead to the formation of canards, and consequently to the formation of MMOs. The canards are solutions with the peculiarity that they cross the fold curve from the attractive slow manifold of the slow-fast system into the repelling branch of the slow manifold, and they stay there for finite time before following a relaxation oscillator trajectory. In the case of folded nodes the canards have rotational properties due to the folded node funnel [10]. Therefore the rotations of the trajectories in the funnel together with the fast relaxation-type part of the trajectory form an MMO solution.

As already mentioned, we find in this paper that folded saddle-nodes singularities of type II exist. These are even more interesting points than the commonly seen folded nodes: in their neighborhood the system has a stable folded node and an unstable true equilibrium. Therefore the local dynamics becomes much more complex; the canard trajectories passing through the folded node funnel into the repelling side of the slow manifold are then influenced by the local stable and unstable manifolds of the true equilibrium. A geometrical approach explaining this interaction and thus completing the proof of how MMOs form in the model is the topic of a next paper [3]. In the present manuscript we focus on preparing the ground necessary to the geometrical approach. We construct the normal form of the system near the fold curve and show that indeed, folded saddle-nodes of type II exist.

2 Slow-fast dynamics and its characteristics in a neuronal rate model

The system we investigate in this paper results from an inhibitory network of two populations of neurons. The activity (spike frequency rate) level of

each population is monitored by variables u_j, $j = 1, 2$ which, if taken in isolated environment, would reach a steady state with exponential decay. However since the populations are coupled through inhibitory connections and are subject to an intrinsic slow negative feedback process (the neuronal adaptation), their dynamics is much complex. Moreover, a constant external input is applied, and it modulates the behavior as well. In summary, the system is defined by two pairs of fast-slow equations of the form $du_j/dt = -u_j + S(I - \beta u_k - g a_j)$, $\tau da_j/dt = -a_j + u_j$ with $j, k = 1, 2$, $k \neq j$. Inhibition has a negative impact on the population-target and is assumed to have strength β; the input is quantified by parameter I; The adaptation variables are a_j and they evolve slowly in time, as opposed to u_j, according to a timescale $\tau \gg 1$; the adaptation strength is g; the system's nonlinearities are defined by function S of typical sigmoid shape such as $S(x) = 1/\left(1 + e^{-r(x-\theta)}\right)$ (the parameters r and θ are said to control the slope of the gain and the activation threshold). All parameters I, β, g, τ, and r are considered to be positive.

From the point of view of the analysis it is important to mention that τ is assumed to be large enough such that $\varepsilon = 1/\tau$, $0 < \varepsilon \ll 1$ is true. Moreover, we need to summarize some important properties of the function S. For consistency let us assume that S is invertible with inverse $F = S^{-1}$, and that S and F are differentiable and monotonically increasing with $S(\theta) = u_0 \in (0, 1)$, $\lim_{x\to-\infty} S(x) = 0$, $\lim_{x\to\infty} S(x) = 1$ and so $\lim_{u\to 0} F(u) = -\infty$, $\lim_{u\to 1} F(u) = \infty$; then $\lim_{u\to 0} F'(u) = \lim_{u\to 1} F'(u) = \infty$. Moreover we assume F' has a local (positive) minimum at u_0 , so $F''(u) < 0$ for $u \in (0, u_0)$, $F''(u) > 0$ for $u \in (u_0, 1)$ and $F''(u_0) = 0$. Note that, in general, these properties are satisfied by the sigmoid functions used in neuronal applications such as the example above.

The system under analysis is thus

$$\begin{aligned} du_1/dt &= -u_1 + S(I - \beta u_2 - g a_1), \\ du_2/dt &= -u_2 + S(I - \beta u_1 - g a_2), \\ da_1/dt &= \varepsilon(-a_1 + u_1), \\ da_2/dt &= \varepsilon(-a_2 + u_2). \end{aligned} \tag{1}$$

In the singular limit case $\varepsilon = 0$, variables a_1 and a_2 are constant, say $a_1 = \bar{a}_1$, $a_2 = \bar{a}_2$ and play the simple role of parameters in the u_j-equations. This is called *the layer problem* or *the fast sub-system*. The set of equilibrium points for the layer problem is a manifold called *the critical manifold*; it is

defined by $-u_1 + S(I - \beta u_2 - g a_1) = 0$, $-u_2 + S(I - \beta u_1 - g a_2) = 0$ and, as in most examples of slow-fast dynamical systems, it has a cubic shape [4]. In an equivalent form, the critical manifold, say Σ can be described as follows

$$\begin{aligned} \Sigma = \{\ (u_1, u_2, a_1, a_2) \quad &: \quad u_1, u_2 \in (0,1),\ a_1, a_2 \in \mathbb{R} \text{ and} \\ \mathscr{F}(u_1, a_1, a_2) \quad &= \quad I - F(u_1) - \beta S(I - \beta u_1 - g a_2) - g a_1 = 0, \\ u_2 \quad &= \quad S(I - \beta u_1 - g a_2)\ \} \end{aligned} \tag{2}$$

where $F = S^{-1}$. Importantly, it can be shown that the layer problem can have either three, two or one equilibrium points depending on the values of a_1 and a_2 [4]. The transition from three to one equilibrium occurs at a double-equilibrium point, that is a saddle-node (fold) bifurcation point. A short calculation in (2) shows this happening at $-F'(u_1) + \beta^2 S'(I - \beta u_1 - g a_2) = 0$ for any constant values a_1, a_2. Due to the invertibility of S and since at the equilibrium point $u_2 = S(I - \beta u_1 - g a_2)$ is true, we get $-F'(u_1) + \beta^2 S'(F(u_2)) = -F'(u_1) + \beta^2 S'(S^{-1}(u_2)) = -F'(u_1) + \beta^2/(S^{-1})'(u_2) = -F'(u_1) + \beta^2/F'(u_2) = 0$. So, the *fold curve* (or, the *curve of saddle-nodes*) is defined by

$$\mathscr{L}^{\pm}: \quad F'(u_1)F'(u_2) = \beta^2 \tag{3}$$

together with (2).

Obviously, the fold condition can be also verified by looking into the eigenvalues of the linearized problem. The partial derivatives of the u_j-equations with respect to u_1 and u_2 are evaluated at a critical point of the layer problem $(u_1^*, u_2^*, a_1^*, a_2^*) \in \Sigma$ and the linearization matrix becomes

$$\mathscr{A} = \begin{bmatrix} -1 & -\beta/F'(u_1^*) \\ -\beta/F'(u_2^*) & -1 \end{bmatrix}. \tag{4}$$

Clearly, $\mathscr{A}$ has a zero eigenvalue if and only if condition (3) is true.

The cubic shape of Σ has the following significance: its outer branches $\Sigma_a^{\pm}$ consist of stable nodes for the layer problem while the middle branch Σ_r is a set of saddles points. That is obtained by testing the sign of the determinant in (4), or equivalent, the sign of $\mathscr{F}_{u_1}$. It results indeed that $\mathscr{F}_{u_1}(u_1, a_1, a_2) < 0$ on Σ_a^- and Σ_a^+ as opposed to $\mathscr{F}_{u_1}(u_1, a_1, a_2) > 0$ on Σ_r [4]. In the perturbed system (1) the dynamics is attracted to either of $\Sigma_a^{\pm}$ and repelled away from Σ_r. For this reason, $\Sigma_a^{\pm}$ are called attractive manifolds and Σ_r is called the repelling (critical) manifold. Thus we can decompose Σ into several significant components like $\Sigma = \Sigma_a^- \cup \Sigma_a^+ \cup \Sigma_r \cup \mathscr{L}^{\pm}$.

From the point of view of the fast-slow analysis of (1), the critical manifold has an additional role. Assume in (1) that we change the time according to $\tilde{t} = \varepsilon t$ ($' = d/d\tilde{t}$). System (1) becomes $\varepsilon u_j' = -u_j + S(I - \beta u_k - g a_j)$, $a_j' = -a_j + u_j$. Setting now $\varepsilon = 0$ we see that Σ is in fact the manifold where the solution of the so-called *reduced system* (or *slow sub-system*) lays. The reduced system evolves according to equations $a_1' = -a_1 + u_1(a_1, a_2)$, $a_2' = -a_2 + u_2(a_1, a_2)$ where u_1, u_2 are implicit functions defined by (2). But note that the formula of $u_1(a_1, a_2)$ and $u_2(a_1, a_2)$ on Σ_a^- (Σ_a^+) changes when curve $\mathscr{L}^{\pm}$ is reached because at $\mathscr{L}^-$ ($\mathscr{L}^+$) a node and a saddle of the layer problem collide and annihilate each other. However, another (stable) node exists on the opposite branch Σ_a^+ (Σ_a^-); the trajectory of the full system will be attracted to it and the equations of the reduced system will change accordingly. We say that a 'jump' takes places from Σ_a^- (Σ_a^+) to Σ_a^+ (Σ_a^-). The trajectory of the full system is thus a relaxation oscillator [11].

For the perturbed system ($\varepsilon > 0$), the dynamics have similar properties away from the fold curve. For ε sufficiently small Fenichel theory [5] proves the existence of a smooth locally invariant normally hyperbolic manifold Σ_ε; this is an $\mathscr{O}(\varepsilon)$ perturbation of Σ and the slow dynamics of (1) takes place close to it. Consequently, to fully describe system (1)'s dynamics one only needs to analyze its trajectories close to the fold curve $\mathscr{L}^{\pm}$. This is especially important if system (1) has complex trajectories such as mixed-mode oscillations (MMOs). Indeed, MMOs were observed and reported in [2]; they are trajectories that combine small amplitude oscillations with large excursions of relaxation type. While the relaxation oscillator can be explain through classical Fenichel theory and slow-fast analysis (see also [11]), the small amplitude oscillations cannot. MMOs exist in an interval of parameter I close to a Hopf bifurcation point but the Hopf is subcritical and MMOs exist on the side of it where the equilibrium is unstable. Therefore there is a need to explain how it is possible for the trajectory to stay close to the unstable equilibrium (situated on $\Sigma_{r,\varepsilon}$) for a finite time and then jump to the opposite attractive branch of $\Sigma_{a,\varepsilon}$, instead of directly jumping to it. The answer is found in the theory of canards [10]. The canards are solutions that pass from the attractive manifold Σ_a into the repelling branch Σ_r through a particular type of point on the fold curve. Such a point, say $p_s \in \mathscr{L}^{\pm}$, is called a *folded singularity*. As we will show in Section 4 folded singularities do exist in system (1) suggesting that canards may be possible in (1). We note that the existence of canards per se is not proven here and it is the

topic of a future paper [3]. Instead we focus now only on the preliminary (but necessary) step of showing the existence of folded singularities. For this, a normal form reduction of (1) near the fold curve $\mathscr{L}^{\pm}$ is necessary. We take this approach in the next section.

3 Normal form reduction of the rate model near the fold curve

Let us consider an arbitrary point on the fold curve $p \in \mathscr{L}^{\pm}$ of coordinates $p = (u_1^*, u_2^*, a_1^*, a_2^*)$.

We translate the point $p \in \mathscr{L}^{\pm}$ into the origin with $U_j := u_j - u_j^*$, $y_j := a_j - a_j^*$ $(j = 1, 2)$ and consider the expansion of the U_j-equations in power series. The equation for U_1 (and similar for U_2) becomes $dU_1/dt = -U_1 - u_1^* + S(I - \beta U_2 - g y_1 - \beta u_2^* - g a_1^*) = -U_1 - u_1^* + S(F(u_1^*) - [\beta U_2 + g y_1]) = -U_1 - S'(F(u_1^*))[\beta U_2 + g y_1] + \frac{1}{2} S''(F(u_1^*))[\beta U_2 + g u_1]^2 + \ldots = -U_1 - \frac{1}{F(u_1^*)}[\beta U_2 + g y_1] - \frac{F''(u_1^*)}{2F'(u_1^*)^3}[\beta U_2 + g u_1]^2 + \ldots$ (Here the lower dots stand for the higher order terms.) Then system (1) can be written as

$$\begin{aligned} dU/dt &= \mathscr{V}(\mathbf{y}) + \mathscr{A} U + \mathscr{A}_0(\mathbf{y}) U + \frac{1}{2}\mathscr{B}(U, U) + \ldots, \\ dy_1/dt &= \varepsilon(u_1^* - a_1^* - y_1 + U_1), \\ dy_2/dt &= \varepsilon(u_2^* - a_2^* - y_2 + U_2) \end{aligned} \tag{5}$$

where $U = (U_1, U_2)^T$, $\mathbf{y} = (y_1, y_2)^T$, $\mathscr{A}$ is defined by (4) and

$$\mathscr{B}(U, U) = \begin{pmatrix} -\frac{\beta^2 F''(u_1^*)}{F'(u_1^*)^3} U_2^2 \\ -\frac{\beta^2 F''(u_2^*)}{F'(u_2^*)^3} U_1^2 \end{pmatrix}, \quad \mathscr{V}(\mathbf{y}) = \begin{pmatrix} -\frac{g}{F'(u_1^*)} y_1 - \frac{g^2 F''(u_1^*)}{2F'(u_1^*)^3} y_1^2 + \mathscr{O}(y_1^3) \\ -\frac{g}{F'(u_2^*)} y_2 - \frac{g^2 F''(u_2^*)}{2F'(u_2^*)^3} y_2^2 + \mathscr{O}(y_2^3) \end{pmatrix},$$

$$\mathscr{A}_0(\mathbf{y}) = \begin{bmatrix} 0 & -\frac{\beta g F''(u_1^*)}{F'(u_1^*)^3} y_1 + \mathscr{O}(y_1^2) \\ -\frac{\beta g F''(u_2^*)}{F'(u_2^*)^3} y_2 + \mathscr{O}(y_2^2) & 0 \end{bmatrix}.$$

Here T stands for the transpose.

As mentioned in the previous section, $\mathscr{L}^{\pm}$ is the set of points that correspond to a saddle-node (fold) bifurcation in the layer problem. Since the fold has a one-dimensional normal form we should be able to reduce (1), or its equivalent form (5), to a system of only three variables, two of which being

slow and only one fast. This can be achieved by projection on the center manifold associated with the zero eigenvalue of $\mathscr{A}$.

The point $U = (0, 0)$ is an equilibrium of the layer problem for $\varepsilon = 0$ and $y_1 = y_2 = 0$. Its associated Jacobian matrix is $\mathscr{A}$ which has a zero ($\lambda_1 = 0$) and a negative ($\lambda_2 = -2$) eigenvalue. The corresponding eigenvectors are $q = (-\sqrt{F'(u_2^*)}, \sqrt{F'(u_1^*)})^T$ such that $\mathscr{A}q = 0$, and $\tilde{q} = (\sqrt{F'(u_2^*)}, \sqrt{F'(u_1^*)})^T$ with $\mathscr{A}\tilde{q} = -2\tilde{q}$. We will use the adjoint vector n of the matrix $\mathscr{A}$ ($\mathscr{A}^T n = 0$ with scalar product $n \cdot q = n_1 q_1 + n_2 q_2 = 1$) to construct the projection on the center manifold. (Note that n is defined by $n = (-\frac{\sqrt{F'(u_1^*)}}{2\beta}, \frac{\sqrt{F'(u_2^*)}}{2\beta})^T$.)

The solution of the layer problem $U = (U_1, U_2)^T$ is decomposed into its projection on the center manifold (σq) and a complementary component V orthogonal to n, that is: $U = \sigma q + V$ [7]. Then the coordinate σ is the variable on the center manifold that replaces u_1 and u_2 in system (1) according to the relationship $\sigma = U \cdot n$. This is $\sigma = -\frac{\sqrt{F'(u_1^*)}}{2\beta}(u_1 - u_1^*) + \frac{\sqrt{F'(u_2^*)}}{2\beta}(u_2 - u_2^*)$. The component V depends on y_1, y_2, σy_1, σy_2, and ε, $\varepsilon\sigma$, εy_1, εy_2 but includes only σ-terms starting with quadratic order (σ^2, σ^3, ...); it is defines by

$$\begin{aligned} V = {} & (y_1 q_{10} + y_2 q_{01} + y_1^2 q_{20} + y_1 y_2 q_{11} + y_2^2 q_{02} + \ldots) + (\sigma^2 h_2 + \sigma^3 h_3 + \ldots) \\ & + (\sigma y_1 h_{10} + \sigma y_2 h_{01} + \ldots) + (\varepsilon h_{000} + \varepsilon\sigma h_{001} + \varepsilon y_1 h_{100} + \varepsilon y_2 h_{010}) \\ & + \mathscr{O}(\varepsilon y_1^2, \varepsilon y_2^2, \varepsilon y_1 y_2, \varepsilon\sigma^2, \varepsilon^2\sigma, \varepsilon^2 y_1, \varepsilon^2 y_2, \varepsilon^i \sigma^j y_1^k y_2^l),\ 4 - i = j = k + l. \quad (6) \end{aligned}$$

The differential equation that σ satisfies on the center manifold is a direct consequence of (5). However its coefficients depend in equal measure on the coefficients of (5) and the admissible values of the vectors h_j, q_{ij}, h_{ijk} ... (all orthogonal on n) from the definition of V.

The projection of system (5) on the center manifold is given below.

Theorem 1. *Let ε be a sufficiently small positive number ($0 < \varepsilon \ll 1$), and parameters I, β, g such that system (1) has a fold curve $\mathscr{L}^{\pm}$.*

Then, in the neighborhood of any point $p \in \mathscr{L}^{\pm}$, $p = (u_1^*, u_2^*, a_1^*, a_2^*)$, system (1) is topologically equivalent to

$$\begin{aligned} d\sigma/dt = {} & c_{10} y_1 + c_{01} y_2 + c_{20} y_1^2 + c_{11} y_1 y_2 + c_{02} y_2^2 + b_{00}\sigma^2 + b_{10}\sigma y_1 \\ & + b_{01}\sigma y_2 + \mathscr{O}(\varepsilon(\sigma + y_1 + y_2), \varepsilon^2, (\sigma + y_1 + y_2)^3), \\ dy_1/dt = {} & \varepsilon \left[(u_1^* - a_1^*) + \left(-1 - \frac{g}{4F'(u_1^*)} \right) y_1 + \left(-\frac{g}{4\beta} \right) y_2 - \sqrt{F'(u_2^*)}\sigma \right. \end{aligned}$$

$$+\mathscr{O}(\varepsilon(\sigma+y_1+y_2),\varepsilon,(\sigma+y_1+y_2)^2)\Big],$$

$$dy_2/dt=\varepsilon\left[(u_2^*-a_2^*)+\left(-\frac{g}{4\beta}\right)y_1+\left(-1-\frac{g}{4F'(u_2^*)}\right)y_2+\sqrt{F'(u_1^*)}\sigma\right.$$
$$\left.+\mathscr{O}(\varepsilon(\sigma+y_1+y_2),\varepsilon,(\sigma+y_1+y_2)^2)\right] \tag{7}$$

with coefficients c_{ij}, b_{ij} defined by

$$b_{00}=\frac{1}{4\beta^2}\left(F'(u_2^*)^{3/2}F''(u_1^*)-F'(u_1^*)^{3/2}F''(u_2^*)\right) \tag{8}$$

and

$$c_{10}=\frac{g}{2\beta\sqrt{F'(u_1^*)}},\quad c_{01}=-\frac{g}{2\beta\sqrt{F'(u_2^*)}},\quad c_{11}=-\frac{3g^2}{8\beta^3}b_{00},$$
$$c_{20}=\frac{g^2F''(u_1^*)}{8\beta F'(u_1^*)^{\frac{5}{2}}}+\frac{g^2}{16\beta^2F'(u_1^*)}b_{00},\ c_{02}=-\frac{g^2F''(u_2^*)}{8\beta F'(u_2^*)^{\frac{5}{2}}}+\frac{g^2}{16\beta^2F'(u_2^*)}b_{00},$$
$$b_{10}=\frac{gF''(u_1^*)}{4F'(u_1^*)^2}+\frac{g}{2\beta\sqrt{F'(u_1^*)}}b_{00},\quad b_{01}=\frac{gF''(u_2^*)}{4F'(u_2^*)^2}-\frac{g}{2\beta\sqrt{F'(u_2^*)}}b_{00}. \tag{9}$$

Proof. Since (5) is a translation of the original system (1), it is obviously topological equivalent to it. Therefore we will focus here only on the proof of the topological equivalence between (5) and (7).

In order to simplify our calculation we will work with vector equations; for this we consider beneficial to introduce the following notation: $\mathbf{e}_1=(1,0)^T$, $\mathbf{e}_2=(0,1)^T$ and

$$A_{10}=\begin{bmatrix}0&1\\0&0\end{bmatrix},\qquad A_{01}=\begin{bmatrix}0&0\\1&0\end{bmatrix}.$$

Then we express the y_j-equations from (5) in terms of $U=\sigma q+V$ with V defined by (6). It results

$$\begin{aligned}dy_1/dt &= \varepsilon(u_1^*-a_1^*)+\varepsilon y_1(\mathbf{e}_1\cdot q_{10}-1)+\varepsilon y_2(\mathbf{e}_1\cdot q_{01})-\varepsilon\sigma\sqrt{F'(u_2^*)}\\ &\quad +\mathscr{O}(\varepsilon(\sigma+y_1+y_2)^2,\varepsilon^2(\sigma+y_1+y_2),\varepsilon^2), \qquad (10)\\ dy_2/dt &= \varepsilon(u_2^*-a_2^*)+\varepsilon y_1(\mathbf{e}_2\cdot q_{10})+\varepsilon y_2(\mathbf{e}_2\cdot q_{01}-1)+\varepsilon\sigma\sqrt{F'(u_1^*)}\\ &\quad +\mathscr{O}(\varepsilon(\sigma+y_1+y_2)^2,\varepsilon^2(\sigma+y_1+y_2),\varepsilon^2). \qquad (11)\end{aligned}$$

A similar calculation apply to the first equation in (5) and implies

$$\begin{aligned}
dU/dt = {} & y_1[\mathscr{A}q_{10} - \frac{g}{F'(u_1^*)}\mathbf{e}_1] + y_2[\mathscr{A}q_{01} - \frac{g}{F'(u_2^*)}\mathbf{e}_2] \\
& + y_1^2[\mathscr{A}q_{20} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^3}A_{10}q_{10} - \frac{g^2F''(u_1^*)}{2F'(u_1^*)^3}\mathbf{e}_1 + \frac{1}{2}\mathscr{B}(q_{10},q_{10})] \\
& + y_2^2[\mathscr{A}q_{02} - \frac{\beta g F''(u_2^*)}{F'(u_2^*)^3}A_{01}q_{01} - \frac{g^2F''(u_2^*)}{2F'(u_2^*)^3}\mathbf{e}_2 + \frac{1}{2}\mathscr{B}(q_{01},q_{01})] \\
& + y_1y_2[\mathscr{A}q_{11} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^3}A_{10}q_{01} - \frac{\beta g F''(u_2^*)}{F'(u_2^*)^3}A_{01}q_{10} + \mathscr{B}(q_{10},q_{01})] \\
& + \sigma^2[\mathscr{A}h_2 - \frac{\beta^2F''(u_1^*)}{2F'(u_1^*)^2}\mathbf{e}_1 - \frac{\beta^2F''(u_2^*)}{2F'(u_2^*)^2}\mathbf{e}_2] \\
& + \sigma y_1[\mathscr{A}h_{10} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^{5/2}}\mathbf{e}_1 + \mathscr{B}(q,q_{10})] + \varepsilon(\mathscr{A}h_{000}) \\
& + \sigma y_2[\mathscr{A}h_{01} + \frac{\beta g F''(u_2^*)}{F'(u_2^*)^{5/2}}\mathbf{e}_2 + \mathscr{B}(q,q_{01})] + \varepsilon\sigma[\mathscr{A}h_{001} + \mathscr{B}(q,h_{000})] \\
& + \varepsilon y_1[\mathscr{A}h_{100} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^3}A_{10}h_{000} + \frac{1}{2}\mathscr{B}(q_{10},h_{000})] \\
& + \varepsilon y_2[\mathscr{A}h_{010} - \frac{\beta g F''(u_2^*)}{F'(u_2^*)^3}A_{01}h_{000} + \frac{1}{2}\mathscr{B}(q_{01},h_{000})] + \dots \qquad (12)
\end{aligned}$$

The equation of σ on the center manifold due to a fold should be at least quadratic in order (with respect to σ) so it should take the form

$$\begin{aligned}
d\sigma/dt = {} & (c_{10}y_1 + c_{01}y_2 + c_{20}y_1^2 + c_{11}y_1y_2 + c_{02}y_2^2) + b_{00}\sigma^2 + b_{10}\sigma y_1 + b_{01}\sigma y_2 + \\
& \varepsilon(d_0\sigma + d_1y_1 + d_2y_2) + \mathscr{O}(y_1^iy_2^j\sigma^k, \varepsilon y_1^iy_2^j, \varepsilon\sigma^iy_1^j, \varepsilon\sigma^iy_2^j, \varepsilon^2). \qquad (13)
\end{aligned}$$

Of course its coefficients c_{ij}, b_{ij}, ... are unknown but they are specific to the system that is projected on the center manifold. We will compute them from equation (12).

For this, let us note that $U = \sigma q + V$ implies $dU/dt = (d\sigma/dt)q + (dV/dt)$, so $dU/dt = (d\sigma/dt)q + [(dy_1/dt)q_{10} + (dy_2/dt)q_{01} + 2y_1(dy_1/dt)q_{20} + (dy_1/dt)y_2q_{11} + y_1(dy_2/dt)q_{11} + 2y_2(dy_2/dt)q_{02} + \dots] + [2\sigma(d\sigma/dt)h_2 + \dots] + (\,y_1h_{10} + y_2h_{01}\,)(d\sigma/dt) + \sigma[\,(dy_1/dt)h_{10} + (dy_2/dt)h_{01}\,] + \varepsilon(d\sigma/dt)\,h_{001} + \varepsilon(dy_1/dt)h_{100} + \varepsilon(dy_2/dt)h_{010} + \dots$.

We then replace $d\sigma/dt$, dy_1/dt, dy_2/dt according to (13), (10) and (11) and obtain

$$\begin{aligned} dU/dt &= y_1(c_{10}q) + y_2(c_{01}q) + y_1^2(c_{20}q + c_{10}h_{10}) + y_2^2(c_{02}q + c_{01}h_{01}) \\ &\quad + y_1y_2(c_{11}q + c_{01}h_{10} + c_{10}h_{01}) + \sigma^2(b_{00}q) + \sigma y_1(b_{10}q + 2c_{10}h_2) \\ &\quad + \sigma y_2(b_{01}q + 2c_{01}h_2) + \varepsilon[(u_1^* - a_1^*)q_{10} + (u_2^* - a_2^*)q_{01}] \\ &\quad + \varepsilon\sigma[d_0q + (u_1^* - a_1^*)h_{10} + (u_2^* - a_2^*)h_{01} + \sqrt{F'(u_1^*)}q_{01} \\ &\quad - \sqrt{F'(u_2^*)}q_{10}] + \varepsilon y_1[c_{10}h_{001} + d_1q + (\mathbf{e}_1 \cdot q_{10} - 1)q_{10} + (\mathbf{e}_2 \cdot q_{10})q_{01} \\ &\quad + 2(u_1^* - a_1^*)q_{20} + (u_2^* - a_2^*)q_{11}] + \varepsilon y_2[c_{01}h_{001} + d_2q + (\mathbf{e}_1 \cdot q_{01})q_{10} \\ &\quad + (\mathbf{e}_2 \cdot q_{01} - 1)q_{01} + 2(u_2^* - a_2^*)q_{02} + (u_1^* - a_1^*)q_{11}] + \dots \end{aligned} \tag{14}$$

The compatibility condition between (12) and (14) allows us to determine the vectors q_{ij}, h_j, h_{ij}, ... in (6) together with coefficients c_{ij}, b_{ij}, ... in (13). This implies

$$\begin{aligned} c_{10}q &= \mathscr{A}q_{10} - \frac{g}{F'(u_1^*)}\mathbf{e}_1 \quad \text{and} \quad c_{01}q = \mathscr{A}q_{01} - \frac{g}{F'(u_2^*)}\mathbf{e}_2, \\ c_{20}q + c_{10}h_{10} &= \mathscr{A}q_{20} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^3}A_{10}q_{10} - \frac{g^2F''(u_1^*)}{2F'(u_1^*)^3}\mathbf{e}_1 + \frac{1}{2}\mathscr{B}(q_{10}, q_{10}), \\ c_{02}q + c_{01}h_{01} &= \mathscr{A}q_{02} - \frac{\beta g F''(u_2^*)}{F'(u_2^*)^3}A_{01}q_{01} - \frac{g^2F''(u_2^*)}{2F'(u_2^*)^3}\mathbf{e}_2 + \frac{1}{2}\mathscr{B}(q_{01}, q_{01}), \\ c_{11}q + c_{10}h_{01} + c_{01}h_{10} &= \mathscr{A}q_{11} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^3}A_{10}q_{01} - \frac{\beta g F''(u_2^*)}{F'(u_2^*)^3}A_{01}q_{10} \\ &\qquad + \mathscr{B}(q_{10}, q_{01}), \\ b_{00}q &= \mathscr{A}h_2 - \frac{\beta^2F''(u_1^*)}{2F'(u_1^*)^2}\mathbf{e}_1 - \frac{\beta^2F''(u_2^*)}{2F'(u_2^*)^2}\mathbf{e}_2, \\ b_{10}q + 2c_{10}h_2 &= \mathscr{A}h_{10} - \frac{\beta g F''(u_1^*)}{F'(u_1^*)^{5/2}}\mathbf{e}_1 + \mathscr{B}(q, q_{10}), \\ b_{01}q + 2c_{01}h_2 &= \mathscr{A}h_{01} + \frac{\beta g F''(u_2^*)}{F'(u_2^*)^{5/2}}\mathbf{e}_2 + \mathscr{B}(q, q_{01}), \quad \text{and so forth.} \end{aligned}$$

The orthogonality principle $n \cdot V = 0$ (equivalent to $n \cdot q_{ij} = 0$, $n \cdot h_j = 0$, $n \cdot h_{ij} = 0$, ...) together with the property $n \cdot q = 1$ implies $q_{10} = (-\frac{g}{4F'(u_1^*)}, -\frac{g}{4\beta})^T$, $q_{01} = (-\frac{g}{4\beta}, -\frac{g}{4F'(u_2^*)})^T$ and $h_2 = -\frac{1}{8\beta^2}(F'(u_2^*)^{3/2}F''(u_1^*) + F'(u_1^*)^{3/2}F''(u_2^*))\tilde{q}$ and determines the coefficients from (8) and (9).

Therefore the projection on the center manifold is successful and it satisfies (7). The topological equivalence between (5) and (7) is then a direct consequence of the center manifold theorem [1], [7]. □

Remark 1. *In this paper we took the Lyapunov-Schmidt projection approach to construct (7) from (5). However, a similar result is obtained if Carr's center manifold reduction method is used [1]. If the latter approach is considered, the system under analysis will be system (5) together with an additional equation for* ε *(*$d\varepsilon/dt = 0$*). All* U_1, U_2, y_1, y_2, ε *are treated as variables and the reduction is made around the point* $(0, 0, 0, 0; 0)$*. It can be verified that at* $(0, 0, 0, 0; 0)$ *the linearization matrix of the 5-dimensional dynamical system has four zero eigenvalues and one negative eigenvalue* (-2)*, and that the theory developed by Carr applies.*

System (1) can now be reduced to its fold normal form in the neighborhood of a point on the fold curve. The goal is to use the 3-dimensional system (7) and apply transformations that change the fast equation of σ into the fold normal form $dz/dt = x + z^2$ plus higher order terms.

Before we proceed let us mention that the coefficient b_{00} of σ^2 in (7) can take either sign. If $u_1^* < u_0 < u_2^*$ we have $F''(u_1^*) < 0 < F''(u_2^*)$; so b_{00} is negative. On the other hand if $u_1^* > u_0 > u_2^*$ then $F''(u_1^*) > 0 > F''(u_2^*)$ and b_{00} is positive. (u_0 is the local minimum point of F'; see page 72.) For example, let us take parameter values $\beta = 1.1$, $g = 0.5$, $I = 1.315$ and function $S(x) = 1/(1 + e^{-r(x-\theta)})$ with $r = 10$, $\theta = 0.2$. Then $p_- = (0.2980253, 0.9587985, 0.2919871, 0.944903,) \in \mathscr{L}^-$ and by symmetry $p_+ = (0.9587985, 0.2980253, 0.944903, 0.2919871) \in \mathscr{L}^+$. Since $u_0 = S(\theta) = 0.5$ we have $b_{00}(p_-) < 0$ and $b_{00}(p_+) > 0$. In fact, in this example almost all points of $\mathscr{L}^-$ have $u_1^* < u_2^*$ and $b_{00} < 0$ while almost all points of $\mathscr{L}^+$ have $u_1^* > u_2^*$ and $b_{00} > 0$. Only at the intersection of $\mathscr{L}^- \cap \mathscr{L}^+$ there are two points with $b_{00} = 0$; they satisfy $u_1^* = u_2^*$ such that $F'(u_1^*) = F'(u_2^*) = \beta$, and they correspond to a cusp bifurcation in the layer problem (not discussed here).

There are three main steps of the reduction to the normal form of the 3-dimensional fast-slow system (7): i) a timescale proportional to b_{00} followed by ii) a linear transformation depending on all variables (σ, y_1, y_2), then iii) a close to linear change of variables depending only on y_1 and y_2. They are explained in detail in the proof of the next theorem.

Theorem 2. *Let ε be a sufficiently small positive number ($0 < \varepsilon \ll 1$), and parameters I, β, g such that system (1) has a fold curve $\mathscr{L}^{\pm}$.*

Let $p \in \mathscr{L}^{\pm}$, $p = (u_1^*, u_2^*, a_1^*, a_2^*)$ be a point on the fold curve such that $b_{00} \neq 0$ where b_{00} is defined by (8).

Then in the neighborhood of p, system (1) is topologically equivalent to

$$\begin{aligned} x' &= \alpha_0 + \alpha_1 y - \alpha_2 z + \mathscr{O}(\varepsilon, x, (y+z)^2), \\ y' &= \alpha_3 + \eta_2 y + \eta_3 z + \mathscr{O}(\varepsilon, x, (y+z)^2), \\ \varepsilon z' &= x + z^2 + \mathscr{O}(\varepsilon, \varepsilon(x+y+z), (x+y+z)^3) \end{aligned} \tag{15}$$

with coefficients

$$\alpha_0 = \frac{g}{2\beta b_{00}|b_{00}|}\left(\frac{u_1^* - a_1^*}{\sqrt{F'(u_1^*)}} - \frac{u_2^* - a_2^*}{\sqrt{F'(u_2^*)}}\right), \tag{16}$$

and

$$\begin{aligned} \alpha_1 &= \frac{g^2}{8\beta^3\sqrt{F'(u_1^*)}|b_{00}|^3}\left[F''(u_1^*)\sqrt{F'(u_2^*)} + F''(u_2^*)\sqrt{F'(u_1^*)}\right. \\ &\qquad \left. - F''(u_1^*)F''(u_2^*)\left(\frac{u_1^* - a_1^*}{2\sqrt{F'(u_1^*)}} + \frac{u_2^* - a_2^*}{2\sqrt{F'(u_2^*)}}\right)\right], \\ \alpha_2 &= \frac{g}{2\beta^2 b_{00}^2}[F'(u_1^*) + F'(u_2^*)], \quad \alpha_3 = \frac{u_1^* - a_1^*}{|b_{00}|}, \\ \eta_2 &= \frac{1}{|b_{00}|}\left(\frac{gF''(u_2^*)}{4\beta b_{00}\sqrt{F'(u_2^*)}} - 1\right), \quad \eta_3 = -\frac{\sqrt{F'(u_2^*)}}{b_{00}}. \end{aligned} \tag{17}$$

Proof. Based on theorem 1 it is sufficient to show that system (7) is topologically equivalent to (15).

Scaling the time with b_{00} allows us to reduce the coefficient of σ^2 to the unity in the fast equation. In order to maintain the initial orientation along the trajectories, we make the transformation independent of the sign of b_{00}. That is achieved with the time change $t \mapsto \tilde{t} = |b_{00}|t$ and the equation for σ in (7) becomes $d\sigma/d\tilde{t} = \frac{1}{|b_{00}|}(c_{10}y_1 + c_{01}y_2 + c_{20}y_1^2 + c_{11}y_1y_2 + c_{02}y_2^2 + b_{00}\sigma^2 + b_{10}\sigma y_1 + b_{01}\sigma y_2 + \cdots)$

The next step is to group all second-order terms involving σ into a unique term. We would like to have the coefficient of the quadratic term in the normal form equal to 1. For this reason, if $b_{00} < 0$ we need to consider a

reflection of the variable σ according to $\sigma \mapsto (-\sigma)$. However this issue can be easily resolved if the coefficient of σ in the new transformation is simply $b_{00}/|b_{00}|$; this will take care of the eventual sign change in the case of $b_{00} < 0$.

We define the linear change of variables $z = \frac{b_{00}}{|b_{00}|}\sigma + \frac{b_{10}}{2|b_{00}|}y_1 + \frac{b_{01}}{2|b_{00}|}y_2$ and use it to modify the fast equation. The new fast variable is z and it satisfies the equation $dz/d\tilde{t} = \frac{c_{10}}{b_{00}}y_1 + \frac{c_{01}}{b_{00}}y_2 + \left(\frac{c_{20}}{b_{00}} - \frac{b_{10}^2}{4b_{00}^2}\right)y_1^2 + \left(\frac{c_{11}}{b_{00}} - \frac{b_{10}b_{01}}{2b_{00}^2}\right)y_1y_2 + \left(\frac{c_{02}}{b_{00}} - \frac{b_{01}^2}{4b_{00}^2}\right)y_2^2 + z^2 + \mathscr{O}(\varepsilon, \varepsilon(z+y_1+y_2), (z+y_1+y_2)^3)$.

The slow equations of y_1, y_2 change as well and they become:

$\varepsilon^{-1}dy_1/d\tilde{t} = \frac{1}{|b_{00}|}(u_1^* - a_1^*) + y_1\frac{1}{|b_{00}|}\left(-1 - \frac{g}{4F'(u_1^*)} + \frac{b_{10}\sqrt{F'(u_2^*)}}{2b_{00}}\right) - z\frac{\sqrt{F'(u_2^*)}}{b_{00}} + y_2\frac{1}{|b_{00}|}\left(-\frac{g}{4\beta} + \frac{b_{01}\sqrt{F'(u_2^*)}}{2b_{00}}\right) + \mathscr{O}(\varepsilon, \varepsilon(z+y_1+y_2), (z+y_1+y_2)^2)$, and

$\varepsilon^{-1}dy_2/d\tilde{t} = \frac{1}{|b_{00}|}(u_2^* - a_2^*) + y_1\frac{1}{|b_{00}|}\left(-\frac{g}{4\beta} - \frac{b_{10}\sqrt{F'(u_1^*)}}{2b_{00}}\right) + z\frac{\sqrt{F'(u_1^*)}}{b_{00}} + y_2\frac{1}{|b_{00}|}\left(-1 - \frac{g}{4F'(u_2^*)} - \frac{b_{01}\sqrt{F'(u_1^*)}}{2b_{00}}\right) + \mathscr{O}(\varepsilon, \varepsilon(z+y_1+y_2), (z+y_1+y_2)^2)$.

At last, we use an almost linear transformation to reduce the fast equation to the normal form of a fold bifurcation.

The change of variables $(y_1, y_2) \mapsto (x, y)$ defined by $x = \frac{c_{10}}{b_{00}}y_1 + \frac{c_{01}}{b_{00}}y_2 + \left(\frac{c_{20}}{b_{00}} - \frac{b_{10}^2}{4b_{00}^2}\right)y_1^2 + \left(\frac{c_{11}}{b_{00}} - \frac{b_{10}b_{01}}{2b_{00}^2}\right)y_1y_2 + \left(\frac{c_{02}}{b_{00}} - \frac{b_{01}^2}{4b_{00}^2}\right)y_2^2 + \ldots$ and $y = y_1$, and the change to slow time $\tilde{t} \mapsto \hat{t} = \varepsilon\tilde{t}$ ($' = d/d\hat{t}$) leads directly to system (15). Systems (15) and (1) are indeed topologically equivalent. □

Remark 2. *The theory of canards associated with folded singularities is developed from the normal form of fast-slow systems with one fast and two slow equations [6]. Therefore this theory can be used as a tool in the study of the system (15); the latter is now in the required normal form. Previous studies on folded singularities and canards [6], [10] show that the first order x-terms in the x- and y- equations play no essential role in the analysis. For this reason we did not specifically include them here. But their coefficients can be calculated in a similar way to those in (16) and (17). For example, the coefficient η_1 of x in the y- equation is: $\eta_1 = \frac{1}{|b_{00}|}\left(b_{00}\sqrt{F'(u_2^*)} - \frac{\beta F''(u_2^*)}{4F'(u_2^*)}\right)$. Similarly, the coefficients of ε-terms in all equations can be determined and they are: $\hat{\varepsilon}_x = \frac{g^2}{16\beta^3 b_{00}|b_{00}|}\left(\sqrt{F'(u_2^*)} - \sqrt{F'(u_1^*)}\right)\left(\frac{u_1^* - a_1^*}{\sqrt{F'(u_1^*)}} + \frac{u_2^* - a_2^*}{\sqrt{F'(u_2^*)}}\right)$, $\hat{\varepsilon}_y =$*

$\frac{g}{8|b_{00}|\sqrt{F'(u_1^*)}}\left(\frac{u_1^*-a_1^*}{\sqrt{F'(u_1^*)}}+\frac{u_2^*-a_2^*}{\sqrt{F'(u_2^*)}}\right)$ *and* $\hat{\varepsilon}_z=\frac{1}{2b_{00}^2}[b_{10}(u_1^*-a_1^*)+b_{01}(u_2^*-a_2^*)]$ *with* b_{00} *and* b_{10}, b_{01} *defined by (8) and (9).*

4 An example of folded saddle-node singularity of type II in the neuronal model (1)

Through reduction to the normal form, the critical manifold Σ of (1) has been transformed (up to the quadratic terms in (15) into the paraboloid $\tilde{\Sigma}$: $\mathscr{G}(x,y,z)=x+z^2=0$. The fold curve has been projected locally into a straight line, the y-axis. This is because the condition for fold is $\mathscr{G}=\mathscr{G}_z=0$ that implies $x=z=0$; so the projection of the fold curve $\mathscr{L}^{\pm}$ is $\{(0,y,0) : |y|<\delta\}$. The attractive branch $\tilde{\Sigma}_a^{\pm}$ is defined by $\mathscr{G}_z<0$, i.e. $z<0$ while the repelling branch $\tilde{\Sigma}_r$ is defined by $\mathscr{G}_z>0$, or $z>0$. The origin $(0,0,0)$ is the point on the resulting fold that corresponds to $p\in\mathscr{L}^{\pm}$.

The analysis of the trajectories along the paraboloid (critical manifold) $\tilde{\Sigma}$ in the neighborhood of $(0,0,0)$ can be done though a blow-up approach [6], [10]. Thus, starting from $x=-z^2$ we get $x'=-2zz'$ and so (15) implies $-2zz'=\alpha_0+\alpha_1 y-\alpha_2 z+\mathscr{O}((y+z)^2)$ and $y'=\alpha_3+\eta_2 y+\eta_3 z+\mathscr{O}((y+z)^2)$.

Last step clarifies why we did not particularly take into account the linear x-terms in the first two equations in (15); simply because they turn into higher order (quadratic) terms when computed on the critical manifold. So they do not change the system's dynamical characteristics in the neighborhood of $(0,0,0)$ (see below).

The two-dimensional system in z and y is however singular at $z=0$; but the blow-up technique deals with it by time-rescaling $\hat{t}\mapsto s=\hat{t}/(-2z)$ (notation: $\cdot=d/ds$). Therefore we obtain the so-called *desingularized flow*

$$\begin{aligned}\dot{z} &= \alpha_0+\alpha_1 y-\alpha_2 z+\mathscr{O}((y+z)^2),\\ \dot{y} &= -2\alpha_3 z+\mathscr{O}((y+z)^2).\end{aligned}\tag{18}$$

A point of the fold curve that is an equilibrium of the desingularized system *without* being an equilibrium of the original (full) system is called *a folded singularity* [6]. Therefore (assuming we work in the singular case $\varepsilon=0$) let us check when $(0,0,0)$ satisfies this property for (15) and (18) respectively; we conclude that $(0,0,0)$ is a folded singularity if and only if $\alpha_0=0$ and $\alpha_3\neq 0$. Based on equations (16) and (17) that is equivalent to $\frac{u_1^*-a_1^*}{\sqrt{F'(u_1^*)}}=\frac{u_2^*-a_2^*}{\sqrt{F'(u_2^*)}}$ with $u_1^*\neq a_1^*$, $u_2^*\neq a_2^*$.

For example, at $\beta = 1.1$, $g = 0.5$, $I = 1.343$ and function $S(x) = 1/(1 + e^{-r(x-\theta)})$ with $r = 10$, $\theta = 0.2$ we found at least one folded singularity: $p_f = (0.3307008, 0.9611521, 0.3124687, 0.9167623)$. Interesting, in the neighborhood of p_f there exist an equilibrium of the full system (1) with coordinates $e = (0.32903, 0.95431, 0.32903, 0.95431)$. Obviously, e satisfies the conditions $u_1 = a_1$ and $u_2 = a_2$ which are not true for p_f.

Assume in the following that $\alpha_0 = 0$ (and that $\varepsilon \approx 0$).

In the normal form (15), the equilibrium e corresponds to the following point: $x = -z^2$, $z = \frac{\alpha_1}{\alpha_2} y$ and $\alpha_3 + \eta_2 y + \eta_3 z = 0$, that is e maps into $\left(-\frac{\alpha_1^2\alpha_3^2}{(\eta_2\alpha_2+\eta_3\alpha_1)^2}, -\frac{\alpha_2\alpha_3}{\eta_2\alpha_2+\eta_3\alpha_1}, -\frac{\alpha_1\alpha_3}{\eta_2\alpha_2+\eta_3\alpha_1}\right)$. However if $\alpha_3 \to 0$ then $e \to p_f$ (the regular singularity collides with the folded singularity p_f). This is the general case of the *folded saddle-node singularity of type II* analyzed in detail by Krupa and Wechselberger [6].

We can identify now what conditions system (1)'s parameters need to satisfy in order to have a folded saddle-node singularity of type II. They are $\alpha_3 = 0$ (and of course $\alpha_0 = 0$) together with the critical manifold and fold curve constraints. In terms of original system, these conditions become $u_1 = a_1$, $u_2 = a_2$ with $F(u_1) = I - \beta u_2 - g a_1$, $F(u_2) = I - \beta u_1 - g a_2$, and $F'(u_1)F'(u_2) = \beta^2$. Consequently, we get $I = F(u_1) + \beta u_2 + g u_1$ with $F'(u_1)F'(u_2) = \beta^2$ and $F(u_1) - F(u_2) = (g - \beta)(u_1 - u_2)$. A more detailed study of the system's dynamics in the neighborhood of a folded saddle-node singularity of type II can be found in [3]. Here we only show that this particular type of points exists in (1).

Theorem 3. *There exist values of parameters β, g, I and gain functions S such that system (1) has folded saddle-node singularities of type II.*

Proof. It is enough to provide an example. As above, we consider $\beta = 1.1$, $g = 0.5$ and function $S(x) = 1/(1 + e^{-r(x-\theta)})$ with $r = 10$, $\theta = 0.2$. The value of I results after solving for appropriate u_1 and u_2 solutions of the algebraic system $F'(u_1)F'(u_2) = \beta^2$ and $F(u_1) - F(u_2) = (g - \beta)(u_1 - u_2)$. That happens at about $u_1 = 0.2841539$ and $u_2 = 0.9575702$ and implies $I = 1.303009$. Therefore (independent of the value of parameter ε), at $\beta = 1.1$, $g = 0.5$, $I = 1.303009$ and $r = 10$, $\theta = 0.2$ in function $S(x) = 1/(1 + e^{-r(x-\theta)})$, system (1) has a type II folded saddle-node singularity. □

5 Discussion

We have investigated the existence of folded singularities in a neuronal rate model of reciprocally inhibitory populations. In particular, we found that folded saddle-nodes of type II exist and we constructed the normal form reduction of the system in their neighborhood. The importance of the folded saddle-node of type II stays in its property to have near it (through perturbation of the system's parameters) of both a stable folded node and an unstable true equilibrium. The former generates a funnel through which canard solutions can pass while the latter modulates the canard trajectory through its stable/unstable manifolds (not shown). Therefore the presence of folded saddle-nodes of type II in this model offers a hint on where to search (in the parameter space) for more complex behaviors. Indeed, based on the results from this paper, a detailed geometrical description of the system in the neighborhood of a folded saddle-node of type II can be obtained. This will be presented in a future manuscript [3].

Acknowledgment. This work was partially supported by The University of Iowa Presidential Faculty Fellowship 2010, and by the Romanian grant PNCDI-2 11-039.

References

[1] J. Carr. *Applications of Centre Manifold Theory.* Springer, New York, 1981.

[2] R. Curtu. Singular Hopf bifurcations and mixed-mode oscillations in a two-cell inhibitory neural network. *Physica D* 239:504-514, 2010.

[3] R. Curtu, J. Rubin – Interaction of canard and singular Hopf mechanisms in a neural model, submitted 2011.

[4] R. Curtu, A. Shpiro, N. Rubin, J. Rinzel. Mechanisms for frequency control in neuronal competition models. *SIAM J. Appl. Dyn. Syst.* 7(2):609-649, 2008.

[5] N. Fenichel. Geometric singular perturbation theory. *J. Diff. Eq.* 31:53-98, 1979.

[6] M. Krupa, M. Wechselberger. Local analysis near a folded saddle-node singularity. *J. Diff. Eq.* 248:2841-2888, 2010.

[7] Y. Kuznetsov. *Elements of applied bifurcation theory.* 2-nd ed, Springer, New York, 1998.

[8] C.R. Laing, C.C. Chow. A spiking neuron model for binocular rivalry. *J. Comput. Neurosci.* 12:39-53, 2002.

[9] A. Shpiro, R. Curtu, J. Rinzel, N. Rubin. Dynamical characteristics common to neuronal competition models. *J. Neurophysiol.* 97:462-473, 2007.

[10] P. Szmolyan, M. Wechselberger. Canards in $\mathbb{R}^3$. *J. Diff. Eq.* 177:419-453, 2001.

[11] F. Verhulst. *Nonlinear differential equations and dynamical systems.* 2-nd ed, Springer, New York, 2000.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

H_2 OPTIMAL CONTROLLERS FOR A LARGE CLASS OF LINEAR STOCHASTIC SYSTEMS WITH PERIODIC COEFFICIENTS*

Vasile Dragan† Toader Morozan† Adrian–Mihail Stoica‡

Abstract

In this paper the H_2 type optimization problem for a class of time varying linear stochastic systems modeled by Ito differential equations and Markovian jumping with periodic coefficients is considered. The main goal of such an optimization problem is to minimize the effect of additive white noise perturbations on a suitable output of the controlled system. It is assumed that only an output is available for measurements. The solution of the considered optimization problem is constructed via the stabilizing solutions of some suitable systems of generalized Riccati differential equations with periodic coefficients.

MSC: 93E20, 93E15, 93E03.

keywords: H_2 norms; linear stochastic systems; periodic coefficients; output based controllers; Riccati differential equations.

*Accepted for publication on December 15, 2010.

†Institute of Mathematics "Simion Stoilow" of the Romanian Academy, P.O.Box. 1-764, RO-014700, Bucharest, Romania, vasile.dragan@imar.ro, toader.morozan@imar.ro

‡University "Politehnica" of Bucharest, Faculty of Aerospace Engineering, Bucharest, Romania, RO-011061, email:adrian.stoica@upb.ro

1 Introduction

The H_2 and the linear quadratic control problems for linear stochastic systems have been widely studied in the current literature. A particular attention was paid to two classes of stochastic systems, namely Markov jump linear systems and systems subject to multiplicative white noise.When an important and unpredictible variation causes a discrete change in the plant characterization at isolated points in time, a Markov chain with a finite state space is a natural model for the plant parameter processes.

Some illustrative applications of these systems can be found for example in [2, 13, 16, 17] and their references, where stochastic stability properties and useful results concerning controllability, observability and optimal control are presented.

More recently, the H_2 control problem for Markov jump linear systems has been studied in [3] for the state feedback case and [11] for the output feedback case. The stochastic systems with multiplicative white noise naturally arise in control problems of linear uncertain systems with stochastic uncertainty (see [12, 15, 19] and the references therein). Results concerning the H_2 control problem for this type of systems are derived for instance in [4, 6]. In [8] the H_2 optimal state feedback control problem is addressed for time varying periodic linear stochastic systems subject to both Markov jumps and multiplicative white noise. The afore mentioned paper extends to the time varying periodic case a part of the results from [7].

In the present paper we extend the results of [8] to the case when only an output is available for measurements. Lately, there is an increasing interest in the consideration of control problems for systems modeled by differential equations with periodic coefficients. For the reader's convenience we refer to [1].

The outline of the paper is: Section 2 contains the description of the mathematical model of the considered controlled systems. Also the H_2 optimization problems are stated. Section 3 collects several auxiliary results which are required for the proof of the main result. Formulae for the computation of H_2-norms of a linear stochastic system with periodic coefficients are provided. The main result of the paper is given in Section 4.

2 The problem formulation

Consider the controlled system (**G**) modeled by a system of the Ito differential equations perturbed by a Markov process of the form:

$$\begin{aligned} dx(t) &= (A_0(t,\eta_t)x(t) + B_0(t,\eta_t)u(t))dt \\ &\quad + \sum_{k=1}^{r}(A_k(t,\eta_t)x(t) + B_k(t,\eta_t)u(t))dw_k(t) + B_v(t,\eta_t)dv(t) \\ dy(t) &= C_0(t,\eta_t)x(t)dt + \sum_{k=1}^{r} C_k(t,\eta_t)x(t)dw_k(t) + D_v(t,\eta_t)dv(t) \quad (2.1) \\ z(t) &= C_z(t,\eta_t)x(t) + D_z(t,\eta_t)u(t) \end{aligned}$$

where $x(t) \in \mathbf{R}^n$ is the state vector, $u(t) \in \mathbf{R}^m$ are the control parameters, $y(t) \in \mathbf{R}^{n_y}$ are the measurements, while $z(t) \in \mathbf{R}^{n_z}$ is the controlled output. In (2.1) $\{\eta_t\}_{t\geq 0}$ is an homogeneous right continuous Markov process on a given probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with the set of the states $\mathfrak{S} = \{1, 2, ..., N\}$ and the transition probability matrix $P(t) = e^{Qt}, t \geq 0$, where $Q \in \mathbf{R}^{N\times N}$ is a matrix whose elements have the properties: $q_{ij} \geq 0$, if $i \neq j$ and $\sum_{j=1}^{N} q_{ij} = 0$ for all $1 \leq i \leq N$. Also, the existence of $\lim_{t\to\infty} P(t)$ is valid. For details see for example [5]. Here, $(w^T(t), v^T(t))^T$ is an $(r + m_v)$-dimensional standard Wiener process. $w(t) = (w_1(t), ..., w_r(t))^T$, $v(t) = (v_1(t), ..., v_{m_v}(t))^T$ (see [14, 18]).

Throughout this paper, we make the following assumptions:

H$_1$: $\{w(t)\}_{t\geq 0}$, $\{v(t)\}_{t\geq 0}$, $\{\eta_t\}_{t\geq 0}$ are independent stochastic processes and $\mathcal{P}\{\eta_0 = i\} > 0$, $1 \leq i \leq N$.

H$_2$: $A_k(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n\times n}$, $B_k(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n\times m}$, $C_k(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n_y\times n}$, $0 \leq k \leq r$, $B_v(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n\times m_v}$, $D_v(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n_y\times m_v}$, $C_z(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n_z\times n}$, $D_z(\cdot, i) : \mathbf{R} \to \mathbf{R}^{n_z\times m}$, $1 \leq i \leq N$, are continuous matrix valued functions which are periodic with the period $\theta > 0$.

To control the system (2.1) we consider the following admissible controllers $(\mathbf{G}_c)$ having the following state space representation:

$$\begin{aligned} dx_c(t) &= A_{c_0}(t,\eta_t)x_c(t)dt + \sum_{k=1}^{r} A_{ck}(t,\eta_t)x_c(t)dw_k(t) + B_c(t,\eta_t)du_c(t) \\ y_c(t) &= C_c(t,\eta_t)x_c(t) \end{aligned} \tag{2.2}$$

where $x_c(t) \in \mathbf{R}^{n_c}$ is the vector of the state parameters of the controller, $u_c(t) \in \mathbf{R}^{n_y}$ is the vector of the inputs of the controller and $y_c \in \mathbf{R}^m$ is the output of the controller, $A_{ck}(\cdot,i)$, $0 \le k \le r$, $B_c(\cdot,i)$, $C_c(\cdot,i)$ are continuous matrix valued functions which are periodic with period θ. As in the time invariant case, (see [7],[10]), the order n_c of an admissible controller is not prefixed. It will be determined in the process of the construction of the solution of the optimization problems. The closed-loop system obtained when coupling an admissible controller (2.2) to the system (2.1) by taking $u_c(t) = y(t)$ and $u(t) = y_c(t)$ has the state space representation given by:

$$(\mathbf{G}_{cl}) : \begin{cases} dx_{cl}(t) = A_{0cl}(t,\eta_t)x_{cl}(t)dt + \sum\limits_{k=1}^{r} A_{kcl}(t,\eta_t)x_{cl}(t)dw_k(t) + \\ \qquad\qquad + B_{vcl}(t,\eta_t)dv(t), \\ z_{cl}(t) = C_{cl}(t,\eta_t)x_{cl}(t). \end{cases} \tag{2.3}$$

where,

$$x_{cl}(t) = \begin{pmatrix} x^T(t) & x_c^T(t) \end{pmatrix}^T, A_{kcl}(t,i) = \begin{pmatrix} A_k(t,i) & B_k(t,i)C_c(t,i) \\ B_c(t,i)C_k(t,i) & A_{ck}(t,i) \end{pmatrix},$$

$$\begin{aligned} 0 \le k \le r, \;\; B_{vcl}(t,i) &= \begin{pmatrix} B_v(t,i) \\ B_c(t,i)D_v(t,i) \end{pmatrix}, \\ C_{cl}(t,i) &= \begin{pmatrix} C_z(t,i) & D_z(t,i)C_c(t,i) \end{pmatrix} \end{aligned} \tag{2.4}$$

for all $t \in \mathbf{R}$, $1 \le i \le N$.

Definition 2.1. An admissible controller $(\mathbf{G}_c)$ of the form (2.2) is a stabilizing controller for the systems $(\mathbf{G})$ if the zero state equilibrium of the linear system

$$dx_{cl}(t) = A_{0cl}(t,\eta_t)x_{cl}(t)dt + \sum_{k=1}^{r} A_{kcl}(t,\eta_t)x_{cl}dw_k(t) \tag{2.5}$$

is exponentially stable in mean square (ESMS).

We denote $\mathcal{K}_s(\mathbf{G})$ the set of all stabilizing controllers of type (2.2).

Now, we construct the following two cost functionals associated to the system $(\mathbf{G})$:
$J_l : \mathcal{K}_s(\mathbf{G}) \to \mathbf{R}^+, \ \ l \in \{1, 2\}$ by

$$J_1(\mathbf{G}_c) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t_0}^{t_0+\tau} E[|z_{cl}(t)|^2]dt \tag{2.6}$$

and

$$J_2(\mathbf{G}_c) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \sum_{i=1}^{N} E[|z_{cl}(t)|^2/\eta_{t_0} = i]dt \tag{2.7}$$

In this paper we shall solve the following optimization problems, which will be called stochastic H_2 optimal control problems:

OP$_1$: Construct a stabilizing controller $\mathbf{G}_c^1 \in \mathcal{K}_s(\mathbf{G})$ with the property that

$$J_1(\mathbf{G}_c^1) = min\{J_1(\mathbf{G}_c)|\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})\} \tag{2.8}$$

OP$_2$: Construct an admissible controller $(\mathbf{G}_c^2) \in \mathcal{K}_s(\mathbf{G})$ with the property that

$$J_2(\mathbf{G}_c^2) = min\{J_2(\mathbf{G}_c)|\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})\}. \tag{2.9}$$

Remark 2.1. a) In the next section we shall see that both $J_1(\mathbf{G}_c)$ and $J_2(\mathbf{G}_c)$ do not depend upon the initial time t_0 and the initial state $x_{cl}(t_0)$. The values of these cost functionals are expressed in terms of bounded solutions of some suitable affine differential equations which extend to this framework the differential equations of the controllability Gramian and observability Gramian.

b) Also in Section 3 we shall see that the value of the cost functional $J_1(\mathbf{G}_c)$ depends upon the initial distribution $\pi_0 = (\pi_0(1), ..., \pi_0(N))$, $(\pi_0(i) = \mathcal{P}\{\eta_0 = i\})$ of the Markov process, while in the case of the second optimization problem, the value of the cost functional $J_2(\mathbf{G}_c)$ does not depend upon the initial distribution of the Markov process.

3 Several preliminary results

Let $\mathcal{S}_n \subset \mathbf{R}^{n\times n}$ be the linear subspace of the real symmetric matrices. Define $\mathcal{S}_n^N$ by $\mathcal{S}_n^N = \mathcal{S}_n \oplus \mathcal{S}_n \oplus ... \oplus \mathcal{S}_n$. We recall that $\mathcal{S}_n^N$ is a real Hilbert space with respect to the inner product

$$< \mathbf{X}, \mathbf{Y} >= \sum_{i=1}^{N} Tr[X(i)Y(i)] \tag{3.1}$$

for all $\mathbf{X} = (X(1), ..., X(N))$, $\mathbf{Y} = (Y(1), ..., Y(N)) \in \mathcal{S}_n^N$.

Additionally, $\mathcal{S}_n^N$ is the ordered linear space, via the ordering induced by the cone

$$\mathcal{S}_n^{N+} = \{\mathbf{X} \in \mathcal{S}_n^N | \mathbf{X} = (X(1), X(2), ..., X(N)),\ X(i) \geq 0, 1 \leq i \leq N\}. \tag{3.2}$$

Here $X(i) \geq 0$ means that $X(i)$ is positive semidefinite. For more details concerning the properties of the cone $\mathcal{S}_n^{N+}$ we refer to [9].

Based on the coefficients of the linear system (2.5) we construct the following operator valued function $t \to \mathcal{L}_{cl}(t)$ by $\mathcal{L}_{cl}(t)\mathbf{X} = ((\mathcal{L}_{cl}(t)X)(1), (\mathcal{L}_{cl}(t)X)(2),$ $..., (\mathcal{L}_{cl}(t)X)(N))$ where

$$\begin{aligned}(\mathcal{L}_{cl}(t)\mathbf{X})(i) = A_{0cl}(t,i)X(i) + X(i)A_{0cl}^T(t,i) + \\ + \sum_{k=1}^{r} A_{kcl}(t,i)X(i)A_{kcl}^T(t,i) + \sum_{j=1}^{N} q_{ji}X(j)\end{aligned} \tag{3.3}$$

for all $\mathbf{X} = (X(1), ..., X(N)) \in \mathcal{S}_{n+n_c}^N$. By direct calculation one obtains that the adjoint operator of $\mathcal{L}_{cl}(t)$ with respect to the inner product (3.1) is given by
$\mathcal{L}_{cl}^*(t)\mathbf{X} = ((\mathcal{L}_{cl}^*(t)\mathbf{X})(1), ..., (\mathcal{L}_{cl}^*(t)\mathbf{X})(N)),$

$$\begin{aligned}(\mathcal{L}_{cl}^*\mathbf{X})(i) = A_{0cl}^T(t,i)X(i) + X(i)A_{0cl}(t,i) + \\ + \sum_{k=1}^{r} A_{kcl}^T(t,i)X(i)A_{kcl}(t,i) + \sum_{j=1}^{N} q_{ij}X(j)\end{aligned} \tag{3.4}$$

for all $\mathbf{X} \in \mathcal{S}_{n+n_c}^N$.

In our developments an important role is played by the following affine differential equations on $\mathcal{S}_{n+n_c}^N$:

$$\frac{d}{dt}Y(t) = \mathcal{L}_{cl}(t)Y(t) + \mathcal{B}^{\epsilon}(t) \tag{3.5}$$

$$\frac{d}{dt}X(t) + \mathcal{L}_{cl}^*(t)X(t) + \mathcal{C}(t) = 0 \tag{3.6}$$

where $\mathcal{B}^{\epsilon}(t) = (\mathcal{B}^{\epsilon}(t,1), \mathcal{B}^{\epsilon}(t,2), ..., \mathcal{B}^{\epsilon}(t,N))$,

$$\mathcal{B}^{\epsilon}(t,i) = \varepsilon(i)B_{vcl}(t,i)B_{vcl}^T(t,i) \tag{3.7}$$

and $\mathcal{C}(t) = (\mathcal{C}(t,1), \mathcal{C}(t,2), ..., \mathcal{C}(t,N))$,

$$\mathcal{C}(t,i) = C_{cl}^T(t,i)C_{cl}(t,i) \tag{3.8}$$

In (3.7) $\varepsilon(i)$ are given nonnegative scalars. Applying Theorem 4.9 and Theorem 4.7 in [9] in the case of the equations (3.5) and (3.6), respectively, we obtain:

Corollary 3.1. *Under the considered assumptions, if the zero state equilibrium of the linear system (2.5) is (ESMS), each of the affine differential equations (3.5) and (3.6), has a unique bounded on* $\mathbf{R}$ *solution* $\mathcal{Y}_{cl}^{\epsilon}(t)$ *and* $\mathcal{X}_{cl}(t)$*, respectively. Additionally, these solutions have the properties:*

(i) $\mathcal{Y}_{cl}^{\epsilon}(t) \in \mathcal{S}_{n+n_c}^{N+}$, $\mathcal{X}_{cl}(t) \in \mathcal{S}_{n+n_c}^{N+}$ for all $t \in \mathbf{R}$.

(ii) $\mathcal{Y}_{cl}^{\epsilon}(t+\theta) = \mathcal{Y}_{cl}^{\epsilon}(t)$, $\mathcal{X}_{cl}(t+\theta) = \mathcal{X}_{cl}(t), \forall t \in \mathbf{R}$.

Remark 3.1. In the special case of $N = 1$, $A_k(t,1) = 0$, $1 \leq k \leq r$ the differential equations (3.5) and (3.6) reduce to the well known differential equations of the controllability Gramian and observability Gramian, respectively from the deterministic case.

The following result provides values of the cost functionals (2.6), (2.7) respectively, in terms of the bounded solutions of the affine differential equations (3.5) and (3.6).

Theorem 3.2. *Under the assumptions* $\mathbf{H_1}$ *and* $\mathbf{H_2}$ *for each stabilizing controller* $\mathbf{G}_c$*) the following equalities hold:*

(i)

$$J_1(\mathbf{G}_c) = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^N \pi_{j\infty} Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)]ds \tag{3.9}$$

$$= \frac{1}{\theta}\int_0^\theta \sum_{j=1}^N Tr[C_{cl}(s,j)\mathcal{Y}_{cl}^{\pi\infty}(s,j)C_{cl}^T(s,j)]ds$$

(ii)

$$J_2(\mathbf{G}_c) = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^N \delta(j) Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)]ds \tag{3.10}$$

$$= \frac{1}{\theta}\int_0^\theta \sum_{j=1}^N Tr[C_{cl}(s,j)\mathcal{Y}_{cl}^{\delta}(s,j)C_{cl}^T(s,j)]ds$$

where $\mathcal{X}_{cl}(t) = (\mathcal{X}_{cl}(t,1),...,\mathcal{X}_{cl}(t,N))$ *is the unique bounded solution of the affine differential equation (3.6), (3.8), while* $\mathcal{Y}_{cl}^{\pi\infty}(t) = (\mathcal{Y}_{cl}^{\pi\infty}(t,1),...,\mathcal{Y}_{cl}^{\pi\infty}(t,N))$ *is the unique bounded on* $\mathbf{R}$ *solution of affine differential equation (3.5), (3.7) with* $\varepsilon(j) = \pi_{j\infty}$ *and* $\mathcal{Y}_{cl}^{\delta}(t) = (\mathcal{Y}_{cl}^{\delta}(t,1),...,\mathcal{Y}_{cl}^{\delta}(t,N))$ *is the unique bounded on* $\mathbf{R}$ *solution of the affine differential equation (3.5), (3.7) for* $\varepsilon(j) = \delta(j)$. *The scalars* $\pi_{j\infty}$ *and* $\delta(j)$ *are defined by*

$$\pi_{j\infty} = \sum_{i=1}^N \tilde{p}_{ij}\mathcal{P}\{\eta_0 = i\}, \quad \delta(j) = \sum_{i=1}^N \tilde{p}_{ij} \tag{3.11}$$

where $\tilde{p}_{ij}$ *are the elements of the matrix* $\tilde{P} = \lim_{t\to\infty} P(t)$.

Proof. The first equalities of (3.9) and (3.10), respectively, are obtained directly applying Theorem 4.2 and Theorem 4.3, respectively in [8].

We rewrite the second part of (3.9) and (3.10) in an unified manner, as follows:

$$\frac{1}{\theta}\int_0^\theta \sum_{j=1}^N \varepsilon(j) Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)]ds \tag{3.12}$$

$$= \frac{1}{\theta}\int_0^\theta \sum_{j=1}^N Tr[C_{cl}(s,j)\mathcal{Y}_{cl}^{\epsilon}(s,j)C_{cl}^T(s,j)]ds$$

So, to complete the proof of the theorem it is sufficient to show that (3.12) is true for some given nonnegative scalars $\varepsilon(j)$.

Using (3.7) together with (3.1) we obtain

$$\sum_{j=1}^{N} \varepsilon(j) Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)] =< \mathcal{X}_{cl}(s), \mathcal{B}^{\epsilon}(s) > .$$

Using successively equations (3.5) and (3.6) we deduce

$$\sum_{j=1}^{N} \varepsilon(j) Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)] = \frac{d}{ds} < \mathcal{X}_{cl}(s), \mathcal{Y}_{cl}^{\epsilon}(s) > + < \mathcal{C}(s), \mathcal{Y}_{cl}^{\epsilon}(s) > .$$

Integrating the last equality from $s = 0$ to $s = \theta$ we obtain via Corollary 3.1. (ii) that

$$\int_0^{\theta} \sum_{j=1}^{N} \varepsilon(j) Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)]ds = \int_0^{\theta} < \mathcal{C}(s), \mathcal{Y}_{cl}^{\epsilon}(s) > ds =$$

$$= \int_0^{\theta} \sum_{j=1}^{N} Tr[C_{cl}(s,j)\mathcal{Y}_{cl}^{\epsilon}(s,j)C_{cl}^T(s,j)].$$

For the last equality we used (3.1) together with (3.8). Thus the proof is complete.

Remark 3.2. From (3.9) and (3.10) one sees that the values of the cost functionals (2.6) and (2.7), respectively, do not depend upon the initial conditions $(t_0, x_{cl}(t_0))$ of the trajectories of the closed loop system $(\mathbf{G}_{cl})$. These values may be seen as measures of the effect of the additional white noise on an output of the closed-loop system. So, the optimization problems we want to solve in this work minimize the effect of the additive white noise perturbations on a suitable output of the closed-loop system.

To construct the optimal controllers of the two optimization problems stated before, we need the stabilizing solution of the following systems of Riccati equations:

a) System of generalized Riccati differential equations of control SGRDE-C

$$\begin{aligned}\frac{d}{dt}X(t,i)+A_0^T(t,i)X(t,i)+X(t,i)A_0(t,i)+\sum_{k=1}^{r}A_k^T(t,i)X(t,i)A_k(t,i)+\\ +\sum_{j=1}^{N}q_{ij}X(t,j)-(X(t,i)B_0(t,i)+\sum_{k=1}^{r}A_k^T(t,i)X(t,i)B_k(t,i)+\\ +C_z^T(t,i)D_z(t,i))(D_z^T(t,i)D_z(t,i)+\\ +\sum_{k=1}^{r}B_k^T(t,i)X(t,i)B_k(t,i))^{-1}(B_0^T(t,i)X(t,i)+\\ +\sum_{k=1}^{r}B_k^T(t,i)X(t,i)A_k(t,i)+D_z^T(t,i)C_z(t,i))+\\ +C_z^T(t,i)C_z(t,i)=0,\quad 1\le i\le N.\end{aligned}\tag{3.13}$$

b) System of generalized Riccati differential equations of filtering SGRDE-F

$$\begin{aligned}\frac{d}{dt}Y(t,i)=A_0(t,i)Y(t,i)+Y(t,i)A_0^T(t,i)+\sum_{k=1}^{r}A_k(t,i)Y(t,i)A_k^T(t,i)+\\ +\sum_{j=1}^{N}q_{ji}Y(t,j)-(Y(t,i)C_0^T(t,i)+\sum_{k=1}^{r}A_k(t,i)Y(t,i)C_k^T(t,i)+\\ \varepsilon(i)B_v(t,i)D_v^T(t,i))(\varepsilon(i)D_v(t,i)D_v^T(t,i)\\ +\sum_{k=1}^{r}C_k(t,i)Y(t,i)C_k^T(t,i))^{-1}(C_0(t,i)Y(t,i)\\ +\sum_{k=1}^{r}C_k(t,i)Y(t,i)A_k^T(t,i)+\varepsilon(i)D_v(t,i)B_v^T(t,i))+\\ +\varepsilon(i)B_v(t,i)B_v^T(t,i),\quad 1\le i\le N.\end{aligned}\tag{3.14}$$

We recall that a global solution $\mathbf{X}_s:\mathbf{R}\to\mathcal{S}_n^N$ of SGRDE-C (3.13) is called stabilizing solution if the zero state equilibrium of the following closed-loop system

$$\begin{aligned}dx(t)=(A_0(t,\eta_t)+B_0(t,\eta_t)F_s(t,\eta_t))x(t)dt\\ +\sum_{k=1}^{r}(A_k(t,\eta_t)+B_k(t,\eta_t)F_s(t,\eta_t))x(t)dw_k(t)\end{aligned}\tag{3.15}$$

is ESMS, where

$$F_s(t,i) = -(D_z^T(t,i)D_z(t,i)+\sum_{k=1}^{r} B_k^T(t,i)X_s(t,i)B_k(t,i))^{-1}(B_0^T(t,i)X_s(t,i) + \sum_{k=1}^{r} B_k^T(t,i)X_s(t,i)A_k(t,i)+D_z^T(t,i)C_z(t,i)), 1\leq i\leq N. \quad (3.16)$$

Also, a global solution $\mathbf{Y}_s : \mathbf{R} \to \mathcal{S}_n^N$ of SGRDE-F (3.14) is called stabilizing solution if the zero state equilibrium of the closed-loop system

$$dx(t) = (A_0(t,\eta_t) + K_s(t,\eta_t)C_0(t,\eta_t))x(t)dt + \sum_{k=1}^{r}(A_k(t,\eta_t) + K_s(t,\eta_t)C_k(t,\eta_t))x(t)dw_k(t) \quad (3.17)$$

is ESMS, where

$$K_s(t,i) = -(Y_s(t,i)C_0^T(t,i)+\sum_{k=1}^{r} A_k(t,i)Y_s(t,i)C_k^T(t,i) + \varepsilon(i)B_v(t,i)D_v^T(t,i))(\sum_{k=1}^{r} C_k(t,i)Y_s(t,i)C_k^T(t,i)+\varepsilon(i)D_v(t,i)D_v^T(t,i))^{-1}, \quad (3.18)$$
$$1\leq i\leq N.$$

It must be remarked that in (3.14) and (3.18), $\varepsilon(i)$ is replaced by $\pi_{i\infty}$ in the case of $\mathbf{OP_1}$ and by $\delta(i)$, in the case of $\mathbf{OP_2}$, respectively.

Applying Theorem 7 from Chapter 4 in [10] one obtains a set of necessary and sufficient conditions for the existence of the bounded on $\mathbf{R}$ and stabilizing solution
$\mathbf{X}_s(t) = (X_s(t,1),...,X_s(t,N))$ of SGRDE-C (3.13) which satisfies the condition

$$D_z^T(t,i)D_z(t,i) + \sum_{k=1}^{r} B_k^T(t,i)X_s(t,i)B_k(t,i) > 0,\ 0\leq t\leq\theta,\ 1\leq i\leq N. \quad (3.19)$$

Moreover, $X_s(t+\theta,i) = X_s(t,i)\ \ \forall t\in\mathbf{R},\ \ 1\leq i\leq N.$

Also, applying Theorem 7 from Chapter 4 in [10] to a suitable dual equation one obtains a set of necessary and sufficient conditions for the existence of a

bounded on $\mathbf{R}$ and stabilizing solution $\mathbf{Y}_s(t) = (Y_s(t,1), ..., Y_s(t,N))$ of the SGRDE-F (3.14) which satisfies the following sign condition:

$$\varepsilon(i)D_v(t,i)D_v^T(t,i)+\sum_{k=1}^{r}C_k(t,i)Y_s(t,i)C_k^T(t,i)>0,\ 0\leq t\leq\theta,\ 1\leq i\leq N. \quad (3.20)$$

Additionally, we have $Y_s(t+\theta,i) = Y_s(t,i) \quad (\forall)t \in \mathbf{R}, \quad i \in \mathfrak{S}$.

Several aspects concerning the numerical computation of the stabilizing solutions of (3.13) and (3.14), respectively, via some Lyapunov iterations can be found in [8] or [10] Chapter 4.

4 The main result

Let us introduce the following performance index $W_\epsilon : \mathcal{K}_s(\mathbf{G}) \to \mathbf{R}^+$ defined by:

$$W_\epsilon(\mathbf{G}_c) = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N}\varepsilon(j)Tr[B_{vcl}^T(s,j)\mathcal{X}_{cl}(s,j)B_{vcl}(s,j)]ds \quad (4.1)$$

where $\varepsilon(i) \geq 0$ are given and $\mathcal{X}_{cl}(t) = (\mathcal{X}_{cl}(t,1), ..., \mathcal{X}_{cl}(t,N))$ is the unique bounded on $\mathbf{R}$ solution of the affine differential equation on $\mathcal{S}_{n+n_c}^N$ (3.6), (3.8).

From Theorem 3.2 we deduce that if $\varepsilon(i) = \pi_{i\infty}$ then $W_\epsilon(\mathbf{G}_c)$ coincides with $J_1(\mathbf{G}_c)$ while if $\varepsilon(i) = \delta(i)$ then $W_\epsilon(\mathbf{G}_c)$ recovers $J_2(\mathbf{G}_c)$. Therefore, the finding of a controller which minimizes (4.1) allows us to obtain in an unified manner the solutions of the two optimization problems stated in Section 2.

Theorem 4.1. *Assume: a) the assumptions* $\mathbf{H_1}$*) and* $\mathbf{H_2}$*) are fulfilled.*

b) The SGRDE-C (3.13) has a θ *periodic and stabilizing solution* $\mathbf{X}_s(\cdot)$ *which verifies condition (3.19).*

c) The SGRDE-F (3.14) has a θ*-periodic and stabilizing solution* $\mathbf{Y}_s(\cdot)$ *which verifies condition (3.20).*

Consider the controller $\tilde{\mathbf{G}}_c^\epsilon$ *having the state space representation*

$$\begin{aligned} d\tilde{x}_c(t) &= \tilde{A}_{c0}(t,\eta_t)\tilde{x}_c(t)dt + \sum_{k=1}^{r} \tilde{A}_{ck}(t,\eta_t)\tilde{x}_c(t)dw_k(t) + \tilde{B}_c(t,\eta_t)du_c(t) \\ dy_c(t) &= \tilde{C}_c(t,\eta_t)\tilde{x}_c(t) \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} \tilde{A}_{ck}(t,i) &= A_k(t,i) + B_k(t,i)F_s(t,i) + K_s(t,i)C_k(t,i), \;\; 0 \le k \le r \\ \tilde{B}_c(t,i) &= -K_s(t,i), \qquad \tilde{C}_c(t,i) = F_s(t,i). \end{aligned} \tag{4.3}$$

$F_s(t,i)$ *and* $K_s(t,i)$ *being constructed as in (3.16), (3.18) respectively. Under the considered assumptions* $\tilde{\mathbf{G}}_c^\epsilon \in \mathcal{K}_s(\mathbf{G})$ *and* $W_\epsilon(\tilde{\mathbf{G}}_c^\epsilon) \le W_\epsilon(\mathbf{G}_c)$, *for all* $\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})$.

The minimal value achieved by the cost (4.1) is

$$\begin{aligned} W_\epsilon(\tilde{\mathbf{G}}_c^\epsilon) = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N} \{\varepsilon(j) Tr[B_v^T(s,j)X_s(s,j)B_v(s,j)] \\ + Tr[V(s,j)F_s(s,j)Y_s(s,j)F_s^T(s,j)V(s,j)]\}ds \end{aligned} \tag{4.4}$$

where

$$V(s,j) = (D_z^T(s,j)D_z(s,j) + \sum_{k=1}^{r} B_k^T(s,j)X_s(s,j)B_k(s,j))^{\frac{1}{2}}. \tag{4.5}$$

Proof. From (4.3) one sees that $\tilde{\mathbf{G}}_c^\epsilon$ depends upon ϵ, via $K_s(t,i)$. In the sequel we do not write explicitly the dependence of $\tilde{\mathbf{G}}_c$ upon the parameter ϵ.

To show that $\tilde{\mathbf{G}}_c \in \mathcal{K}_s(\mathbf{G})$ we consider the linear system of type (2.5) obtained when coupling (4.2), (4.3) to (2.1), taking $u_c(t) = y(t)$ and $u(t) = y_c(t)$.

If $\tilde{x}_{cl}(t) = (\tilde{x}(t), \tilde{x}_c^T(t))^T$ is the state vector of this system, we perform the change of the state variables as:

$$\tilde{e}(t) = \tilde{x}(t) - \tilde{x}_c(t).$$

Thus, we obtain the system of stochastic differential equations:

$$\begin{aligned} d\tilde{x}(t) &= [(A_0(t,\eta_t)+B_0(t,\eta_t)F_s(t,\eta_t))\tilde{x}(t)-B_0(t,\eta_t)F_s(t,\eta_t)\tilde{e}(t)]dt \\ &\quad +\sum_{k=1}^{r}[(A_k(t,\eta_t)+B_k(t,\eta_t)F_s(t,\eta_t))\tilde{x}(t) \\ &\quad -B_k(t,\eta_t)F_s(t,\eta_t)\tilde{e}(t)]dw_k(t) \qquad (4.6) \\ d\tilde{e}(t) &= [A_0(t,\eta_t)+K_s(t,\eta_t)C_0(t,\eta_t)]\tilde{e}(t)dt \\ &\quad +\sum_{k=1}^{r}[A_k(t,\eta_t)+K_s(t,\eta_t)C_k(t,\eta_t)]\tilde{e}(t)dw_k(t) \end{aligned}$$

The exponential stability in mean square of the closed loop systems (3.15) and (3.17), respectively, together with Theorem 32, (iii) Chapter 2 in [10] allows us to deduce that the trajectories of (4.6) satisfy:

$$\lim_{t\to\infty} E[|\tilde{x}(t)|^2+|\tilde{e}(t)|^2|\eta_0=i]=0$$

or equivalently,

$$\lim_{t\to\infty} E[|\tilde{x}_{cl}(t)|^2|\eta_0=i]=0 \qquad (4.7)$$

for all $1\le i\le N$ and all $\tilde{x}_{cl}(0)\in\mathbf{R}^{2n}$.

Further (4.7) together with Theorem 23 in Chapter 2 in [10] lead to

$$E[|\tilde{x}_{cl}(t)|^2|\eta_{t_0}=i]\le\beta e^{-\alpha(t-t_0)}|\tilde{x}_{cl}(t_0)|^2$$

for all $t\ge t_0\ge 0$, $\tilde{x}_{cl}(t_0)\in\mathbf{R}^{2n}$, for some $\beta>0$ and $\alpha>0$.

This shows that the controller (4.2), (4.3) is stabilizing. It remains to prove that the controller $\tilde{\mathbf{G}}_c$ minimizes the cost (4.1). First we rewrite (4.1) in the form

$$\begin{aligned} W_\epsilon(\mathbf{G}_c) &= \frac{1}{\theta}\int_0^\theta\sum_{j=1}^{N}\varepsilon(j)Tr[B_v^T(s,j)X_s(s,j)B_v(s,j)]ds \qquad (4.8) \\ &\quad +\frac{1}{\theta}\int_0^\theta\sum_{j=1}^{N}\varepsilon(j)Tr[B_{vcl}^T(s,j)\hat{X}(s,j)B_{vcl}(s,j)]ds \end{aligned}$$

for all $\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})$, where $\hat{X}(t,j) = \mathcal{X}_{cl}(t,j) - diag(X_s(t,j),0)$. By direct calculations one obtains that $t \to \hat{X}(t) = (\hat{X}(t,\cdot),...,\hat{X}(t,N))$ verifies the affine differential equation

$$\frac{d}{dt}\hat{\mathbf{X}}(t) + \mathcal{L}^*_{cl}(t)\hat{\mathbf{X}}(t) + \mathbf{\Theta}(t) = 0 \tag{4.9}$$

where $\mathbf{\Theta}(t) = (\mathbf{\Theta}(t,1),...,\mathbf{\Theta}(t,N))$ with

$$\begin{aligned} \mathbf{\Theta}(t,i) &= (\Theta^T(t,i)\Theta(t,i)) \\ \Theta(t,i) &= V(t,i)\begin{pmatrix} F_s(t,i) & -C_c(t,i) \end{pmatrix}. \end{aligned} \tag{4.10}$$

So $\hat{X}(t,i) \geq 0, \quad (\forall)\ t \in \mathbf{R}, 1 \leq i \leq N$. Reasoning as in the proof of the equality (3.12), one gets

$$\sum_{j=1}^{N} \varepsilon(j) Tr[B^T_{vcl}(s,j)\hat{X}(s,j)B_{vcl}(s,j)] = \sum_{j=1}^{N} Tr[\Theta(s,j)\mathcal{Y}^{\varepsilon}_{cl}(s,j)\Theta^T(s,j)].$$

This allows us to transform (4.8) as follows:

$$\begin{aligned} W_\epsilon(\mathbf{G}_c) = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N} \varepsilon(j) Tr[B_v^T(s,j)X_s(s,j)B_v(s,j)]ds \\ + \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N} Tr[\Theta(s,j)\mathcal{Y}^{\epsilon}_{cl}(s,j)\Theta^T(s,j)]ds. \end{aligned}$$

Further, we write

$$W_\epsilon(\mathbf{G}_c) = \tilde{\mu} + \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N} Tr[\Theta(s,j)\hat{\mathcal{Y}}(s,j)\Theta^T(s,j)]ds \tag{4.11}$$

where we denote by

$$\begin{aligned} \tilde{\mu} = \frac{1}{\theta}\int_0^\theta \sum_{j=1}^{N} \{\varepsilon(j) Tr[B_v^T(s,j)X_s(s,j)B_v(s,j)] \\ + Tr[V(s,j)F_s(s,j)Y_s(s,j)F_s^T(s,j)V(s,j)]\}ds \end{aligned} \tag{4.12}$$

and $\hat{Y}(s,j) = \mathcal{Y}^{\epsilon}_{cl}(s,j) - diag(Y_s(s,j),0)$.

By direct calculations one obtains that $t \to \hat{\mathbf{Y}}(t) = (\hat{Y}(t,1), ..., \hat{Y}(t,N))$ is a bounded solution of the affine differential equation on $\mathcal{S}^N_{n+n_c}$:

$$\frac{d}{dt}\hat{\mathbf{Y}}(t) = \mathcal{L}_{cl}(t)\hat{\mathbf{Y}}(t) + \mathbf{\Psi}(t) \tag{4.13}$$

where $\mathbf{\Psi}(t) = (\mathbf{\Psi}(t,1), ..., \mathbf{\Psi}(t,N))$ with

$$\begin{aligned} \mathbf{\Psi}(t,i) &= \Psi(t,i)\Psi^T(t,i) \\ \Psi(t,i) &= \begin{pmatrix} K_s(t,i) \\ -B_c(t,i) \end{pmatrix} \hat{V}(t,i) \\ \hat{V}(t,i) &= (\varepsilon(i)D_v(t,i)D_v^T(t,i) + \sum_{k=1}^{r} C_k(t,i)Y_s(t,i)C_k^T(t,i))^{\frac{1}{2}}. \end{aligned} \tag{4.14}$$

Applying Theorem 4.9 in [9] to the equation (4.13), (4.14) we deduce that

$$\hat{Y}(t,i) \geq 0 \tag{4.15}$$

for all $t \in \mathbf{R}$, $1 \leq i \leq N$ and for all $\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})$.

From (4.12) one sees that $\tilde{\mu}$ does not depend upon the controller $\mathbf{G}_c$. Moreover, from (4.11) and (4.15) we deduce that

$$W_\epsilon(\mathbf{G}_c) \geq \tilde{\mu} \tag{4.16}$$

for all $\mathbf{G}_c \in \mathcal{K}_s(\mathbf{G})$. To complete the proof we have to show that in (4.16) the equality takes place if $\mathbf{G}_c = \tilde{\mathbf{G}}_c$. To this end let us remark that in the case of the controller $\tilde{\mathbf{G}}_c$ we have

$$\begin{aligned} \Theta(s,j)\hat{Y}(s,j)\Theta^T(s,j) = V(s,j)F_s(s,j)\mathcal{J}\hat{Y}(s,j)\mathcal{J}^T F_s^T(s,j)V(s,j) = \\ = V(s,j)F_s(s,j)Z_{11}(s,j)F_s^T(s,j)V(s,j) \end{aligned} \tag{4.17}$$

where $\mathcal{J} = \begin{pmatrix} I_n & -I_n \end{pmatrix}$ and $Z_{11}(s,j)$ is the 11-block of the matrix $Z(t,j) = \mathcal{T}\hat{\mathcal{Y}}(s,j)\mathcal{T}^T$, with $\mathcal{T} = \begin{pmatrix} I_n & -I_n \\ 0 & I_n \end{pmatrix}$. One obtains the equation

$$\begin{aligned} \frac{d}{dt}Z(t,i) = \hat{A}_0(t,i)Z(t,i) + Z(t,i)\hat{A}_0^T(t,i) + \sum_{k=1}^{r} \hat{A}_k(t,i)Z(t,i)\hat{A}_k^T(t,i) \\ + \sum_{j=1}^{N} q_{ji}Z(t,j) + \mathcal{T}\Psi(t,i)\Psi^T(t,i)\mathcal{T}^T \end{aligned} \tag{4.18}$$

where $\hat{A}_k(t,i) \in \mathbf{R}^{2n\times 2n}$, $\hat{A}_k(t,i) = \mathcal{T}\tilde{A}_{kcl}(t,i)\mathcal{T}^{-1}$, $\tilde{A}_{kcl}(t,i)$ being constructed via (2.4) using (4.3). Taking the (1,1)-block of (4.18) one obtains that $t \to (Z_{11}(t,1),...,Z_{11}(t,N))$ is a bounded solution of the differential equation on $\mathcal{S}_n^N$:

$$\begin{aligned}\frac{d}{dt}Z_{11}(t,i) &= (A_0(t,i)+K_s(t,i)C_0(t,i))Z_{11}(t,i)\\ &\quad +Z_{11}(t,i)(A_0(t,i)+K_s(t,i)C_0(t,i))^T\\ &+\sum_{k=1}^{r}(A_k(t,i)+K_s(t,i)C_k(t,i))Z_{11}(t,i)(A_k(t,i)\\ &+K_s(t,i)C_k(t,i))^T+\sum_{j=1}^{N}q_{ji}Z_{11}(t,j),\quad 1\le i\le N\end{aligned} \tag{4.19}$$

Having in mind the fact that $K_s(t,i)$ is the stabilizing injection associated to the stabilizing solution $\mathbf{Y}_s(\cdot)$ of SGRDE-F (3.14) we conclude that (4.19) admits a unique bounded on $\mathbf{R}$ solution. Hence $Z_{11}(t,i) = 0$ for all $t \in \mathbf{R}$, $1 \le i \le N$. So, we deduce that in (4.16) we have equality if $\mathbf{G}_c = \tilde{\mathbf{G}}_c$. This completes the proof.

Remark 4.1. In the special case $N = 1$, $A_k(t,1) = 0$, $B_k(t,1) = 0$, $C_k(t,1) = 0$, $1 \le k \le r$, $t \in \mathbf{R}$ the optimal controller (4.2), (4.3) reduces to the well known Kalman filter (see e.g. [20]).

References

[1] S. Bitanti, P. Colaneri, *Periodic Systems, Filtering and Control*, Springer Verlag, London, 2009.

[2] W.P.Blair jr, D.D. Sworder, Continuous time regulation of a class of econometric models, *IEEE Trans Systems Man Cybernet*, 1975, 5, 341-346.

[3] O.L.V. Costa, J.B.R. do Val, J.C. Geromel, Continuous time state feedback H_2 control of Markovian jump linear systems via convex analysis, *Automatica*, 35, pp.259-268, 1999.

[4] G. DaPrato, A. Ichikawa, Quadratic control for linear time varying systems, *SIAM J. Control Optimiz.*, **28**, pp 359-381, 1990.

[5] J.L.Doob, *Stochastic processes*,John Weley, New York, 1967.

[6] V. Dragan, A. Halanay, T. Morozan, Optimal stabilizing compensator for linear systems with state-dependent noise, *Stochastic Analysis and Appl.*, (1992), vol. 10, nr. 5, 557-573.

[7] V. Dragan, T. Morozan, A. Stoica, H2 optimal control for linear stochastic systems, *Automatica*, (2004), vol. 40, 7, 1103- 1113.

[8] V. Dragan, T. Morozan, Stochastic H2 Optimal Control for a Class of Linear Systems with Periodic Coeścients, *European Journal of Control*, (2005), 6, 11, 619-631.

[9] V. Dragan, T. Damm, G. Freiling and T. Morozan, Differential equations with positive evolutions and some applications, *Result. Math.*, (2005), 48, 206-236.

[10] V. Dragan, T. Morozan and A.M. Stoica, *Mathematical methods in robust control of linear stochastic systems*, Springer, New York, 2006.

[11] D.P.De Farias, J.C. Geromel, J.B.R.do Val, O.L.V. Costa, Output feedback control of Markov jump linear systems in continuous -time, *IEEE Trans. on Automatic Control*, 45, pp. 944-949, 2000.

[12] El- Ghaoui, State feedback control of systems with multiplicative noise via linear matrix inequalities , *Syst. Control Letters*, 1995, 24, 223-228.

[13] X. Feng, K.A.Loparo, Y.Ji, H.J. Chizeck, Stochastic stability properties of jump linear systems, *IEEE Trans Autom Control*, 1992, 37, 58-53.

[14] A. Friedman, *Stochastic differential equations and applications*, vol.1, Academic Press, New York, 1971.

[15] D. Hinrichsen, A.J. Pritchard, Stability radii of systems with stochastic uncertainty and their optimization by output feedback,*SIAM J. Control and Optimization*, 1996, 34, 1972-1998.

[16] Y. Ji, H.J. Chizeck, Controllability, stabilizability and continuous time linear quadratic control, *IEEE Trans Autom control*, 1990, 35, 777-788.

[17] M.Mariton, P. Bertrand, Output feedback for a class of linear systems with stochastic jump parameters, *IEEE Trans Autom Control*, 1985, 30, 898-900.

[18] B. Oksendal, *Stochastic differential equations*, Springer Verlag, Berlin, Heidelberg, 1998.

[19] V.A. Ugrinovski, Robust H_∞ in the presence of stochastic uncertainty, *Int. J. Control*, 1998,71, 219-237.

[20] W.M. Wonham, Random differential equations in control theory, *Probabilistic Methods in Applied Mathematics, 2, Academic Press, New York*, 1970, pp.131-212.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

ON THE NONLINEAR STABILITY OF A BINARY MIXTURE WITH CHEMICAL SURFACE REACTIONS*

Adelina Georgescu† Lidia Palese ‡

Abstract

In this work we consider the non linear stability of a chemical equilibrium of a thermally conducting two component reactive viscous mixture which is situated in a horizontal layer heated from below and experiencing a catalyzed chemical reaction at the bottom plate. The evolution equation for the perturbation energy is deduced with an approach which generalizes the Joseph's parametric differentiation method. Moreover, the nonlinear stability bound for the chemical equilibrium of the fluid mixture is derived in terms of thermal and concentrational non dimensional numbers.

MSC:76E15 - 76E30

keywords: Nonlinear Stability - Horizontal Thermal Convection - Energy Method

1 Introduction

The convective instability and the nonlinear stability of a chemically inert fluid in a gravitational field heated from below (the classical Bénard problem)

*Accepted for publication on December 18, 2010.

†Academy of Romanian Scientists, Splaiul Independentei, No. 54, 050094 Bucharest, Romania

‡Dept. of Math., University of Bari Via E. Orabona, 4-70125 Bari, palese1@dm.uniba.it

have been studied and present a well known interesting problem in several fields of fluid mechanics. More recently, [1]-[4] have considered reactive fluids of technological interest for whose chemical reactions can give temperature and concentration gradients which influence the transport process, for example, the dissociation of nitrogen, oxygen or hydrogen gas near the gas-solid interface of a space vehicle when returning to the earth's atmosphere, (see Bdzil and Frisch [1], [2] and Loper and Roberts [5],) and can alter hydrodynamic stabilities.

In the present paper, begun in 2009 when the first Author was still alive and then finished by the second author also developing some A. Georgescu's ideas and suggestions, we consider a fluid mixture composed of the dimer A_2 and the monomer A in a horizontal layer heated from below, the bottom plate being catalytic. We evaluate the effects of heterogeneous surface catalyzed reactions on the hidrodynamic stability of the chemical equilibrium.

The model adopted is that of Bdzil and Frisch.

We consider a Newtonian fluid model and derive the evolution equation for the perturbation energy with the approach from [6], [7],[8], which generalizes the Joseph's parametric differentiation method reported in [9], [10].

A non linear stability bound is derived in terms of all involved physical parameters.

2 The initial/boundary value problem for perturbation

We consider the mixture (A_2, A) described by a Newtonian model to which we apply the Boussinesq approximation in the layer bounded by the surfaces $z = 0$ and $z = 1$ with the lower surface being catalytic, that is, the interconversion $(A_2 \rightleftharpoons A)$ occurs via the surface $z = 0$. However, the conditions that must be satisfied at the catalyzed boundary $z = 0$, are [3]:

$$\vec{J} \cdot \vec{k} = 0 \qquad \vec{Q} \cdot \vec{k} = 0$$

where $\vec{J}$ is the mass flux, $\vec{k}$ is the unit vector in the vertical upward direction, and Q is the heat flux. The chemical equilibrium S_0 is characterized by the following temperature $(\overline{T})$ and degree of dissociation (fraction of pure monomers present) $(\overline{C})$ fields [2], [3]:

$$\overline{T}(z) = T_1 + \beta(1 - z), \qquad \overline{C}(z) = C_1 + \gamma(1 - z), \tag{1}$$

where C_1 and T_1 are the values of C and T at $z = 1$ and the constants β and γ are given in [1], [3].

Let us now perturb S_0 up to a cellular motion (convection-diffusion) characterized by a velocity $\vec{u} = \vec{0} + \vec{u}$, a pressure $p = \bar{P} + p'$ a temperature $T = \bar{T} + \theta$ and a concentration $C = \bar{C} + \gamma$ fields, where $\vec{u}, p', \theta, \gamma$ are the corresponding perturbation fields and $\vec{0}, \bar{P}, \bar{T}, \bar{C}$ represent the basic state S_0 (the expression of $\bar{P}$ follows from the momentum equation for S_0). The perturbation fields satisfy the following equations which express the conservation of the momentum, energy and concentration, written in nondimensional coordinates [4], [11]

$$\frac{\partial}{\partial t}\vec{u} + (\vec{u}\cdot\nabla)\vec{u} = -\nabla p' + \Delta\vec{u} + (\mathcal{R}\theta + \mathcal{C}\gamma)\vec{k},, \tag{2}$$

$$P_r(\frac{\partial}{\partial t}\theta + \vec{u}\cdot\nabla\theta) = \Delta\theta - \mathcal{R}w, \qquad (t,\vec{x}) \in (0,\infty)\times V \tag{3}$$

$$S_c(\frac{\partial}{\partial t}\gamma + \vec{u}\cdot\nabla\gamma) = \Delta\gamma + \mathcal{C}w, \tag{4}$$

in a subset of L_2, namely ,

$$\mathcal{N} = \{(\vec{u}, p, \theta, \gamma) \in L^2 \mid div\vec{u} = 0; \qquad \frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = w = 0 \,\mathrm{on}\, \partial \mathrm{V}_2, \tag{5}$$

$$\vec{u} = 0 \,\mathrm{on}\, \partial \mathrm{V}_1 \qquad \theta = \gamma = 0 \,\mathrm{on}\, \partial \mathrm{V}_2 \qquad \frac{\partial\theta}{\partial \mathrm{z}} = -\mathrm{s}\gamma,\ \frac{\partial\gamma}{\partial \mathrm{z}} = \mathrm{r}\gamma \,\mathrm{on}\, \partial \mathrm{V}_1\}.$$

where $\vec{u} = (u, v, w)$, V is a periodicity cell in the x, y-directions, ∂V is the boundary of V, $\partial V_1 = \partial V \cap \{z = 0\}, \partial V_2 = \partial V \cap \{z = 1\}$. The perturbation fields depend on the time t and space $\vec{x} = (x, y, z)$ and $\mathcal{R}^2, \mathcal{C}^2, P_r$ and S_c are the thermal and concentrational numbers of Rayleigh, Prandtl and Schmidt, respectively. In addition, $r, s > 0$ are dimensionless surface reactions numbers.

The basic state S_0 corresponds to the zero solution of the initial-boundary value problem for (2)-(4) in the class $\mathcal{N}$. This state is called non linearly stable if a Liapunov function $E(t)$, called energy, remains bounded when $t \to \infty$ in the sense of $lim_{t\to\infty}\int_0^t E(t')dt' < \infty$ [9], [10]. It is asymptotically nonlinearly stable if $E(t) \to 0$ when $t \to \infty$. The stability or instability of S_0 depends on six physical parameters occurring in (2)-(5): $P_r, S_c = \tau P_r, \mathcal{R}, \mathcal{C}, r$ and s.

3 The evolution equation for the perturbation energy

Integrating over V the sum of the equation (4) multiplied by $P_r^{-1}\gamma$ and the equation (3) multiplied by $S_c^{-1}\theta$ we obtained:

$$\frac{d}{dt}<\theta\gamma>=-\tau\mathcal{R}S_c^{-1}<\gamma w>+\mathcal{C}S_c^{-1}<\theta w>-(1+\tau)S_c^{-1}<\nabla\theta\cdot\nabla\gamma>+ \quad (6)$$

$$S_c^{-1}\int_V\nabla\cdot(\theta\nabla\gamma)dV+P_r^{-1}\int_V\nabla\cdot(\gamma\nabla\theta)dV.$$

Multyplying (2) by $\mathbf{u}$, (3) by θ, (4) by γ and integrating the resulted equations over V and taking into account the boundary conditions from (5) we have respectively

$$\frac{1}{2}\frac{d}{dt}<|\mathbf{u}|^2>=-<|\nabla\mathbf{u}|^2>+\mathcal{R}<\theta w>+\mathcal{C}<\gamma w>, \quad (7)$$

$$\frac{1}{2}\frac{d}{dt}<P_r\theta^2>=-\mathcal{R}<\theta w>-<|\nabla\theta|^2>+\int_V\nabla\cdot(\theta\nabla\theta)dV, \quad (8)$$

$$\frac{1}{2}\frac{d}{dt}<S_c\gamma^2>=\mathcal{C}<\gamma w>-<|\nabla\gamma|^2>+\int_V\nabla\cdot(\gamma\nabla\gamma)dV. \quad (9)$$

We perform the sum of (7) to (8) multiplied by $a>0$, (9) multiplied by $b>0$ and (6) multiplied by $c>0$, and introducing the functions

$$E_2(t)=<|\mathbf{u}|^2+d_1\phi_1^2+d_2\phi_1\phi_2>/2, \qquad \Psi(t)=<d_3\phi_2^2>/2, \quad (10)$$

we obtain

$$\frac{dE_2}{dt}+\frac{d\Psi}{dt}=-<|\nabla\mathbf{u}|^2+(a_1^2d_4+b_1^2d_5+a_1b_1d_6)|\nabla\phi_1|^2+$$

$$(a_2^2d_4+b_2^2d_5+a_2b_2d_6)|\nabla\phi_2|^2+\Big[2a_1a_2d_4+2b_1b_2d_5+(a_1b_2+a_2b_1)d_6\Big]|\nabla\phi_1\cdot\nabla\phi_2|>$$

$$+\mathcal{R}<(a_1d_7+b_1d_8)\phi_1w>+\mathcal{R}<(a_2d_7+b_2d_8)\phi_2w>+$$

$$(aa_1^2+bb_1^2+cS_c^{-1}a_1b_1+cP_r^{-1}a_1b_1)\int_V\nabla\cdot(\phi_1\nabla\phi_1)dV+$$

$$(aa_2^2+bb_2^2+cS_c^{-1}a_2b_2+cP_r^{-1}a_2b_2)\int_V\nabla\cdot(\phi_2\nabla\phi_2)dV+$$

$$(aa_1a_2 + bb_1b_2 + cS_c^{-1}a_1b_2 + cP_r^{-1}a_2b_1)\int_V \nabla\cdot(\phi_1\nabla\phi_2)dV +$$

$$(aa_1a_2 + bb_1b_2 + cS_c^{-1}a_2b_1 + cP_r^{-1}a_1b_2)\int_V \nabla\cdot(\phi_2\nabla\phi_1)dV, \qquad (11)$$

where

$$\theta = a_1\phi_1 + a_2\phi_2, \quad \gamma = b_1\phi_1 + b_2\phi_2. \qquad (12)$$

Here a_1, a_2, b_1 and b_2 are unknown parameters and d_i, $i = 1\cdots 8$ are functions of a, b, c and the physical parameters, defined by

$$\begin{array}{ll} d_1 = aP_ra_1^2 + bS_cb_1^2 + 2ca_1b_1; & d_2 = aP_ra_1a_2 + bS_cb_1b_2 + c(a_1b_2 + a_2b_1), \\ d_3 = aP_ra_2^2 + bS_cb_2^2 + 2ca_2b_2; & d_4 = a, \\ d_5 = b; & d_6 = cS_c^{-1}(1+\tau), \\ d_7 = 1 - a + c\alpha S_c^{-1}; & d_8 = b\alpha + \alpha - c\tau S_c^{-1}, \end{array}$$

where $\alpha = \frac{\mathcal{C}}{\mathcal{R}}$.

The seven constants a, b, c, a_1, b_1, a_2 and b_2 shall be determined from the requirement that (11) assumes the form [6], [7], [8]

$$\frac{dE_2}{dt} + \frac{d\Psi}{dt} = -<|\nabla\mathbf{u}|^2 + |\nabla\phi_1|^2> + \mathcal{R}<(a_1d_7 + b_1d_8)\phi_1 w>, \qquad (13)$$

where the energy E_2 has the form

$$E_2(t) = <|\mathbf{u}|^2 + d_1|\phi_1|^2>/2. \qquad (14)$$

In the case $\tau = 1$ the right-hand side of (11) assumes the form from (13) and instead of (10) the energy E_2 assumes the form (14) if

$$\begin{cases} d_2 = 0, \\ a_1^2d_4 + b_1^2d_5 + a_1b_1d_6 = 1, \\ a_2^2d_4 + b_2^2d_5 + a_2b_2d_6 = 0, \\ 2a_1a_2d_4 + 2b_1b_2d_5 + (a_1b_2 + a_2b_1)d_6 = 0, \\ a_2d_7 + b_2d_8 = 0, \\ sb_2 + ra_2 = 0. \end{cases} \qquad (15)$$

ϕ_1 and ϕ_2 as linear combinations of θ and γ are given by

$$\phi_1 = a_1'\theta + a_2'\gamma, \qquad \phi_2 = b_1'\theta + b_2'\gamma, \qquad (16)$$

where

$$a_1' = b_2/M, \quad a_2' = a_2/M, \quad b_1' = -b_1/M, \quad b_2' = a_1/M,$$
$$a_1 = b_2'/M', \quad a_2 = -a_2'/M', \quad b_1 = -b_1'/M', \quad b_2 = a_1'/M', \tag{17}$$

and $M = a_1b_2 - a_2b_1$ and $M' = a_1'b_2' - a_2'b_1'$.

It follows

$$\phi_1 = (b_2\theta - a_2\gamma)/M \qquad \text{and} \qquad \phi_2 = (-\mathrm{b_1}\theta + \mathrm{a_1}\gamma)/\mathrm{M}. \tag{18}$$

The system (15) can be considered as yielding $a_1, b_1, a_2/b_2, b, c$ as functions of a, $(15)_3$ follows from $(15)_1$ and $(15)_2$, so, another relationship between these parameters is necessary.

In order to find it we followed the Joseph's generalized method of parametric differentiation [6], [7], [8].

Denoting

$$2A = \mathcal{R}|a_1d_7 + b_1d_8|, \tag{19}$$

relation (13) implies

$$\frac{dE_2}{dt} \leq -\xi^2\Big(1 - A/\sqrt{R_{a*}}\Big)E_2(t), \tag{20}$$

where [12], [13], [14], [15]

$$\xi^2 = min_{\mathbf{u},\phi_1}\frac{2 < |\nabla\mathbf{u}|^2 + |\nabla\phi_1|^2 >}{< |\mathbf{u}|^2 + |\phi_1|^2 >}, \frac{1}{\sqrt{R}_{a*}} = max_{\mathbf{u},\phi_1}\frac{2 < \phi_1 w >}{< |\nabla\mathbf{u}|^2 + |\nabla\phi_1|^2 >}. \tag{21}$$

Therefore, the stability criterion reads

$$\mathcal{R} < 2\sqrt{R}_{a*}/|a_1d_7 + b_1d_8|. \tag{22}$$

As a consequence $\mathcal{R}$ will be maximal if $|a_1d_7 + b_1d_8|$ will be minimal. Since $|a_1d_7 + b_1d_8|$ is a function of the parameter a this requirement will be fullfilled iff

$$\frac{d(a_1d_7 + b_1d_8)}{da} = 0. \tag{23}$$

This equation represents the equation determining a.

In this way, the stability bound

$$\mathcal{R}_E = 2\sqrt{R}_{a*}/|a_1d_7 + b_1d_8| \tag{24}$$

will be obtained once the system (15), (23), admits real solutions and it can be solved explicitly in terms of the physical parameters.

Of course, all values of the physical parameters ensuring the negativity of $a_1 d_7 + b_1 d_8$ are in the stability domain.

In this section we applied a Joseph's generalized method [6], [7], [8] to derive the evolution equation for E_2.

The Joseph's idea of using (6) was generalized by us in the following way [6], [7], [8]:

We used from the beginning an integral relation, i.e. the equation (6) (already followed by suitable multiplications, addition and integration over V of the balance equations for temperature and concentration (2)-(4)), in [7] we proved that (6) is nothing else but the projection of a system, equivalent to (2)-(4), including the equations which generate (6) and with a symmetrizable linear part, for a suitable choice of the constants [7].

As a consequence, the initial equations (2)-(4) were replaced by some others in which the equations which generated (6) were present. In this way drastically changed the linear part of the initial equations allowing a much more advantageous symmetrization. By contrast, the symmetric operator for (2)-(4) does not contain the effect of terms in $\mathbf{u}$ from (2) and those of terms in θ from (3) because they are opposite.

4 Nonlinear stability bound

From (15) we deduce the relations

$$d_6^2 = 4 d_4 d_5, \tag{25}$$

$$d_8^2 d_4 = d_7^2 d_5, \tag{26}$$

$$\frac{s}{r} = \alpha, \tag{27}$$

Then we determine explicit expression for a_1, b_1 and a_2/b_2in terms of $d_8/d_7, d_6/d_4$ and d_4, or, taking into account (25) and (26), in terms of $\sqrt{d_5/d_4}$ and d_4.This implies

$$a_1 d_7 + b_1 d_8 = d_7/\sqrt{d_4}. \tag{28}$$

On the other hand, (23) was written as an equation of the form

$$\frac{d}{da}\frac{1 + a(\frac{s}{r}\alpha - 1)}{\sqrt{a}} = 0. \tag{29}$$

If $\frac{s}{r}\alpha > 1$, the solution

$$a = \frac{1}{\frac{s}{r}\alpha - 1} \tag{30}$$

of (29) gives, in terms of the physical quantities, the non linear stability bound (24)

$$\mathcal{R}_E = \sqrt{R_{a*}}\Big(\sqrt{(\frac{s}{r})^2 - 1}\Big)^{-1} \tag{31}$$

Theorem. For physical parameters $\mathcal{R}$, $\mathcal{C} = \alpha\mathcal{R}$, $\frac{s}{r} = \alpha$, $(\frac{s}{r})^2 > 1$, *the zero solution of (2)-(4), corresponding to the basic conduction state, is non linearly asymptotically stable if* $\mathcal{R} < \mathcal{R}_E$, *where* $\mathcal{R}_E$ *is given by (31). Or, equivalently, if*

$$\mathcal{C}^2 - \mathcal{R}^2 < R_{a*}$$

where R_{a*} *is given by (21).*

5 Conclusions

We treated the problem of the non linear stability of an equilibrium for a binary mixture in a horizontal layer heated from below and experiencing a catalyzed chemical reaction at bottom plate, using the energy method, improved as in [6] by taking into account an idea from [9] [10] . The given problem governing the perturbation evolution was changed in order to obtain an optimum energy inequation. Then the non linear stability bound was obtained with the aid of some appropriately chosen multiplication constants.

The presence of derivatives in the boundary conditions heavily influences the possibility to relate linear and non linear bounds because of the lack of corresponding maximum principle for the Laplace equation. However, the generalized method, as in [6], [7], [8] gives us the possibility to drastically change the linear problem derived by the evolution equations, so that allows us an easier handling of the linear problem to determine a generalized linearization principle (in the sense of the coincidence of linear and nonlinear stability bounds) [16].

References

[1] J. Bdzil, H.L. Frisch ,*Chemical Instabilities. II. Chemical Surface Reactions and Hydrodynamic Instability* Phys. Fluids, **14** (3), 1971, 475-481.

[2] J. Bdzil, H.L. Frisch,*Chemical Instabilities. IV. Nonisothermal Chemical Surface Reactions and Hydrodynamic Instability*, Phys. Fluids **14** (6), 1971, 1077-1086.

[3] C. L. Mc Taggart, B. Straughan, *Chemical Surface Reactions and Nonlinear Stability by the Method of Energy*, SIAM J. Math. Anal., **17** (2), 1986, 342-351.

[4] B. Straughan, *Nonlinear Convective Stability by the Method of Energy: Recents Results*, Atti delle Giornate di Lavoro su Onde e Stabilità nei mezzi continui, Cosenza (Italy), 1983, 323-338.

[5] D.E. Loper, P.H. Roberts, *On The Motion of an Airon-Alloy Core Containing a Slurry,* Geophys. Astrophys. Fluid Dynamics, **9**, 1978, 289-331.

[6] A. Georgescu, L. Palese, *Extension of a Joseph's criterion to the non linear stability of mechanical equilibria in the presence of thermodiffusive conductivity,* Theoretical and Computational Fluid Mechanics, **8**, 1996, 403-413.

[7] A. Georgescu, L. Palese, A. Redaelli, *On a New Method inHydrodynamic Stability Theory,* Mathematical Sciences Research. Hot Line, **4** (7), 2000, 1-16.

[8] A. Georgescu, L. Palese, A. Redaelli, *The Complete Form for the Joseph Extended Criterion,* Annali Università di Ferrara, Sez. VII, Sc. Mat. **48**, 2001, 9-22.

[9] D.D. Joseph, *Global Stability of the Conduction-Diffusion Solution,* Arch. Rational Mech. Anal., **36** (4), 1970, 285-292.

[10] D.D. Joseph, *Stability of fluid motions* I-II, Springer, Berlin, 1976.

[11] S.Chandrasekhar,*Hydrodynamic and Hydromagnetic stability*(Oxford, Clarendon Press), 1968.

[12] A. Georgescu, *Hydrodynamic stability theory*, Kluwer, Dordrecht, 1985.

[13] A. Georgescu, L. Palese, *Stability Criteria for Fluid Flows,* Advances in Math. for Appl. Sc., **81**, World Scient. Singapore, 2010.

[14] O. A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flow*, Gordon and Breach, New York, 1969.

[15] B. Straughan, *The Energy Method, Stability and Nonlinear Convection*, Springer, New York, 2003.

[16] A. Georgescu, L. Palese, *A Linearization Principle for the Stability of the Chemical Equilibrium of a Binary Mixture,* ROMAI J., **2**, 2010, 131-138.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

AN APPLICATION OF DOUBLE-SCALE METHOD TO THE STUDY OF NON-LINEAR DISSIPATIVE WAVES IN JEFFREYS MEDIA*

Adelina Georgescu† Liliana Restuccia‡

Abstract

In previous papers we sketched out the general use of the double-scale method to nonlinear hyperbolic partial differential equations (PDEs) in order to study the asymptotic waves and as an example the model governing the motion of a rheological medium (Maxwell medium) with one mechanical internal variable was studied. In this paper the double scale method is applied to investigate non-linear dissipative waves in viscoanelastic media without memory of order one (Jeffreys media), that were studied by one of the authors (L. R.) in more classical way. For these media the equations of motion include second order derivative terms multiplied by a very small parameter. We give a physical interpretation of the new (fast) variable, related to the surfaces across which the solutions or/and some of their derivatives vary steeply. The paper concludes with one-dimensional application containing original results.

MSC: 34E05, 34E10, 34E13, 73B20, 73B99.

*Accepted for publication on January 18, 2011.

† Academy of Romanian Scientists, Splaiul Independentei, nr. 54, sector 5, 050094 Bucharest, Romania.

‡lrest@dipmat.unime.it, Department of Mathematics, University of Messina, Viale F. Stagno D'Alcontres 31, 98166 Messina, Italy.

keywords: Partial differential equations, double-scale method, non-linear dissipative waves, asymptotic methods, rheological media.

Introduction

The mathematical aspects involved into the study of asymptotic waves belong to singular perturbation theory, namely the double-scale method ([1, 10, 14, 16, 24, 30, 31, 35, 36, 40, 42, 43]. The multiple-scale method, and, in particular, the double-scale approach, is appropriate to phenomena which possess qualitatively distinct aspects at various scales. For instance, at some well-determined times or space coordinates, the characteristics of the motion vary steeply, while at larger scale the characteristics are slow and describe another type of motion. In addition, the scales are defined by some small parameters.

The theoretical interest in nonlinear waves was manifest as early as the years '50 and '60 of the last century and a lot of applications to various branches of physics were worked out [2, 3, 4, 12, 13, 19, 20, 21, 22, 23, 32, 33].

In the context of rheological media studies on non-linear waves were carried out in [7]-[9]. In previous papers (see [17] and [39]) we sketched out the general use of the double-scale method to nonlinear hyperbolic partial differential equations (PDEs) in order to study the asymptotic waves and as an application the model governing the motion of anelastic media without shape and bulk memory (Maxwell media) was studied.

In this paper the double scale method (see [16]) is applied to investigate non-linear dissipative waves in isotropic viscoanelastic media without memory of order one in which a viscous flow phenomenon occurs (Jeffreys media), that were studied in [8] by one of the authors (L. R.) in more classical way, following the methodologies developed in [3] and generalized in [15]. Only shear phenomena are taken into consideration and the hydrostatic pressure is assumed constant and uniform. Furthermore, the isothermal case is considered. For these media the equations of motion include second order derivative terms multiplied by a very small parameter, that play a very important role because they usually have a balancing effect on the non-linear steepening of waves. In Section 1 the various steps in applying the double scale method are introduced and the asymptotic approximations of first and second order are obtained. In Section 2 the propagation into an uniform unperturbed state is discussed and in Section 3 the first approximation of wavefront and of $\mathbf{U}$ are

derived. In Section 4 the equations governing the motion of Jeffreys media are treated and the mechanical relaxation equation valid for these media is described in the framework of classical irreversible thermodynamics (TIP) with internal variables [5, 11, 25, 26, 27, 28, 29, 34, 37, 38]. In Section 5 an one-dimensional application is carried out containing original results.

1 Asymptotic dissipative waves from the point of view of double-scale method

Let E^{3+1} be an Euclidean space and let $P \in E^{3+1}$ be a current point. Let $\mathbf{U} = \mathbf{U}(P)$ be the unknown vector function, solution of a system of PDEs written in the following matrix form

$$\mathbf{A}^{\alpha}(\mathbf{U})\mathbf{U}_{\alpha} + \omega^{-1}\left[\mathbf{H}^{k}\frac{\partial^{2}\mathbf{U}}{\partial t\partial x^{k}} + \mathbf{H}^{ik}\frac{\partial^{2}\mathbf{U}}{\partial x^{i}\partial x^{k}}\right] = \mathbf{B}(\mathbf{U}),$$

$$(\alpha = 0, 1, 2, 3);\ (i, k = 1, 2, 3), \tag{1}$$

where $x^0 = t$ (time), x^1, x^2, x^3 are the space coordinates, $\mathbf{U}$ depends on x^{α}, $\mathbf{U}_{\alpha} = \frac{\partial \mathbf{U}}{\partial x^{\alpha}}$, $\mathbf{A}^{\alpha}$, $\mathbf{H}^{k}$, $\mathbf{H}^{ik}$ are appropriate matrices 9×9 and

$$\mathbf{A}^{\alpha}(\mathbf{U})\mathbf{U}_{\alpha} = \mathbf{B}(\mathbf{U}) \tag{2}$$

is the associated system of nonlinear hyperbolic PDEs.

In [8] it was shown that the motion of viscoanelastic media without memory, in the isothermal case, where only shear phenomena are taken into consideration and the hydrostatic pressure is constant and uniform, is described by a system of nonlinear PDEs having the form (1). The system of PDEs (1) includes terms containing second order derivatives multiplied by a very small parameter. These terms play a very important role because they usually have a balancing effect on the non-linear steepening of the waves. In [41], using (1), the propagation of linear acoustic waves was considered and the velocity and attenuation of the waves were investigated. In [8] the non-linear dissipative waves were worked out (see [2, 3, 4, 12, 13, 19, 20, 21, 22, 23, 32, 33]) and, in particular, a method, developed by G. Boillat [3] and generalized by D. Fusco [15], was applied to construct asymptotic approximations of order 1 of solutions of the system of equations (1).

In this Section we study these non-linear dissipative waves from the point of view of double scale-method.

Following A. Jeffrey in [23], let us introduce for systems of PDEs of type (1) (or type (2)) the concepts of waves (called *dissipative waves* only for the system (1)) and associated wavefronts. Precisely, the solution hypersurfaces of systems of type (1) (or type (2)) are referred to as waves, because they may be interpreted as representing propagating wavefronts. When physical problems are associated with such interpretation the solution on the side of the wavefront towards which the propagation takes place may then be regarded as being the *undisturbed solution* ahead of the wavefront, whilst the solution on the other side may be regarded as a propagating *disturbance wave* which is entering a region occupied by the undisturbed solution. This is because the solution at a point in the undisturbed region characterises the state of the physical system at that time and place before the advancing wave has reached it.

The smooth solutions of systems of type (1) (or type(2)) that present a steep variation in the normal direction to the associated wavefront are called *asymptotic waves*. Then, there exists a family of hypersurfaces S (defined by the equation $\varphi(x^\alpha) = 0$) moving in the Euclidean space E^{3+1} (consisting of points of coordinates x^α, $\alpha = 0, 1, 2, 3$, or, equivalently of the time $t = x^0$ and the space coordinates x^i, i=1, 2, 3, having equation

$$\varphi(t, x^i) = \bar{\xi} = const, \tag{3}$$

such that the solutions $\mathbf{U}$ or/and some of their derivatives vary steeply across S while along S their variation is slow [1]. From the double scale method point of view this means that around S there exist *asymptotic internal layers* (see [16]) such that *the order of magnitude (i.e. the scale)* of the solutions or/and of some of their derivatives inside these layers and far away from them differs very much. In systems of equations of type (1) the coefficient ω^{-1} is the small parameter, that is associated with the order of magnitude of the interior layer. Therefore, it is natural to introduce a new independent variable ξ, related to the hypersurfaces S,

$$\xi = \omega\bar{\xi} = \omega\varphi(t, x^i), \tag{4}$$

with $\xi = \frac{\varphi(t,x^i)}{\omega^{-1}}$ asymptotically fixed, i.e. $\xi = Ord(1)$ as $\omega^{-1} \to 0$,
and $\omega \gg 1$ a very large parameter, to assume that the solution depends on the old as well as the new variable, i.e. $\mathbf{U} = \mathbf{U}(x^\alpha, \xi)$, and to consider that x^α and ξ are independent.

Taking into account that $\mathbf{U}$ is sufficiently smooth, hence it has sufficiently many bounded derivatives, it follows that, except for the terms containing ω, all the other terms are asymptotically fixed and the computation can proceed formally. In this way, if $x^\alpha = x^\alpha(s)$ are the parametric equations of a curve C in E^{3+1}, we have

$$\frac{d\mathbf{U}}{ds} = \omega \frac{\partial \mathbf{U}}{\partial \xi}\frac{\partial \varphi}{\partial s} + \frac{\partial U^\alpha}{\partial x^\alpha}\frac{dx^\alpha}{ds}$$

(where the dummy index convention is understood). This relation shows that, indeed, along C, $\mathbf{U}$ does not vary too much if C belongs to the hypersurface S (in this case $\frac{d\varphi}{ds} = 0$) but has a large variation if C is not situated on S. For these reasons ξ is referred to as the fast variable.

Let us sketch the various steps in applying the double-scale method.

First, we look for the solution of the equations as an asymptotic series of powers of the small parameter, say ϵ, namely with respect to the asymptotic sequence $\left\{1, \varepsilon^{a+1}, \varepsilon^{a+2}, ...,\right\}$ or $\left\{1, \varepsilon^{\frac{1}{p}}, \varepsilon^{\frac{2}{p}}, ...,\right\}$, as $\varepsilon \to 0$. In [7] - [9] it is considered $p = 1$ and $\varepsilon = \omega^{-1}$, such that $\mathbf{U}(x^\alpha, \xi)$ is written as an asymptotic power series of the small parameter ω^{-1}, i.e. with respect to the asymptotic sequence $1,\ \omega^{-1},\ \omega^{-2}, ..., \text{as} \quad \omega^{-1} \to 0$, the $\mathbf{U}^i$ $(i = 1, 2, ...)$ being functions of x^α and ξ,

$$\mathbf{U}(x^\alpha, \xi) \sim \mathbf{U}^0(x^\alpha, \xi) + \omega^{-1}\mathbf{U}^1(x^\alpha, \xi) + O(\omega^{-2}), \text{ as } \quad \omega^{-1} \to 0. \tag{5}$$

In (5) $\mathbf{U}^0(x^\alpha, \xi)$ is a known solution [15] of

$$\mathbf{A}^\alpha(\mathbf{U}^0)\mathbf{U}_\alpha(\mathbf{U}^0) = \mathbf{B}(\mathbf{U}^0), \tag{6}$$

where $\mathbf{U}^0$ is taken as the initial unperturbed state.

The next step of the double-scale method consists in expressing the derivatives with respect to x^α, $\frac{\partial}{\partial x^\alpha}$, in terms of the derivatives with respect to x^α and ξ, i.e. $\frac{\partial}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha} + \frac{\partial}{\partial \xi}\frac{\partial \xi}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha} + \omega\frac{\partial}{\partial \xi}\frac{\partial \varphi}{\partial x^\alpha}$, so that the derivative $\mathbf{U}_\alpha = \frac{\partial \mathbf{U}}{\partial x^\alpha}$ has the form

$$\frac{\partial \mathbf{U}}{\partial x^\alpha} \sim \omega^{-1}\left(\frac{\partial \mathbf{U}^1}{\partial x^\alpha} + \omega\frac{\partial \mathbf{U}^1}{\partial \xi}\frac{\partial \varphi}{\partial x^\alpha}\right) + \omega^{-1}\frac{\partial \mathbf{U}^2}{\partial \xi}\frac{\partial \varphi}{\partial x^\alpha} + O(\omega^{-2}), \text{ as } \omega^{-1} \to 0, \tag{7}$$

where we have assumed that the first approximation $\mathbf{U}^0$ is constant.

Then, taking into account the form of $\mathbf{A}^\alpha$, $\mathbf{H}^k$, $\mathbf{H}^{ik}$ and $\mathbf{B}$, the following asymptotic expansions are deduced:

$$\mathbf{A}^\alpha(\mathbf{U}) \sim \mathbf{A}^\alpha(\mathbf{U}^0) + \frac{1}{\omega}\nabla\mathbf{A}^\alpha(\mathbf{U}^0)\mathbf{U}^1 + O\left(\frac{1}{\omega^2}\right), \quad \text{as} \quad \omega^{-1} \to 0, \tag{8}$$

$$\mathbf{H}^k(\mathbf{U}) \sim \mathbf{H}^k(\mathbf{U}^0) + \frac{1}{\omega}\nabla\mathbf{H}^k(\mathbf{U}^0)\mathbf{U}^1 + O\left(\frac{1}{\omega^2}\right), \text{ as } \omega^{-1} \to 0, \ (k = 1, 2, 3), \tag{9}$$

$$\mathbf{H}^{ik}(\mathbf{U}) \sim \mathbf{H}^{ik}(\mathbf{U}^0) + \frac{1}{\omega}\nabla\mathbf{H}^{ik}(\mathbf{U}^0)\mathbf{U}^1 + O\left(\frac{1}{\omega^2}\right), \text{ as } \omega^{-1} \to 0, \ (i, k = 1, 2, 3), \tag{10}$$

$$\mathbf{B}(\mathbf{U}) \sim \mathbf{B}(\mathbf{U}^0) + \frac{1}{\omega}\nabla\mathbf{B}(\mathbf{U}^0)\mathbf{U}^1 + O\left(\frac{1}{\omega^2}\right), \quad \text{as} \ \ \omega^{-1} \to 0, \tag{11}$$

where $\nabla = \frac{\partial}{\partial \mathbf{U}}$.

The last point of the method consists in introducing the asymptotic expansions (7) - (11) into (1) and matching the obtained series.

It follows

$$(\mathbf{A}^\alpha)_0\Phi_\alpha\frac{\partial \mathbf{U}^1}{\partial \xi} = \mathbf{0} \quad (\alpha = 0, 1, 2, 3), \tag{12}$$

$$(\mathbf{A}^\alpha)_0\left(\Phi_\alpha\frac{\partial \mathbf{U}^2}{\partial \xi}\right) = -\left[(\mathbf{A}^\alpha)_0\frac{\partial \mathbf{U}^1}{\partial x^\alpha} + (\nabla\mathbf{A}^\alpha)_0\mathbf{U}^1\left(\Phi_\alpha\frac{\partial \mathbf{U}^1}{\partial \xi}\right)\right.$$

$$\left. +(\mathbf{H}^k)_0\Phi_0\Phi_k\frac{\partial^2 \mathbf{U}^1}{\partial \xi^2} + (\mathbf{H}^{ik})_0\Phi_i\Phi_k\frac{\partial^2 \mathbf{U}^1}{\partial \xi^2} - (\nabla\mathbf{B})_0\mathbf{U}^1\right], \tag{13}$$

where $\Phi_\alpha = \frac{\partial\varphi}{\partial x^\alpha}$ ($\Phi_k = \frac{\partial\varphi}{\partial x^k}$, $k = 1, 2, 3$) and the symbol $(...)_0$ indicates that the quantities are calculated in $\mathbf{U}^0$. Equation (12) is linear in $\mathbf{U}^1$, while (13) is affine in $\mathbf{U}^2$.

Remind that the wavefront φ is still an unknown function. In order to determine it, we recall its equation is $\varphi(t, x^1, x^2, x^3) = 0$). This implies that along the wavefront we have $\frac{d\varphi}{dt} = 0$, implying $\frac{\partial\varphi}{\partial t} + \mathbf{v} \cdot grad\varphi = 0$, or equivalently, $\frac{\frac{\partial\varphi}{\partial t}}{|grad\varphi|} + \mathbf{v} \cdot \frac{grad\varphi}{|grad\varphi|} = 0$. Obviously,

$$\frac{grad\varphi}{|grad\varphi|} = \mathbf{n}, \tag{14}$$

such that the previous equality reads

$$\frac{\frac{\partial\varphi}{\partial t}}{|grad\varphi|} + \mathbf{v}\cdot\mathbf{n} = 0. \tag{15}$$

Introduce the notation

$$\lambda = -\frac{\frac{\partial\varphi}{\partial t}}{|grad\varphi|}, \tag{16}$$

so that

$$\lambda(\mathbf{U},\mathbf{n}) = \mathbf{v}\cdot\mathbf{n}, \tag{17}$$

where λ is called the *velocity normal to the progressive wave*, being $\mathbf{n}$ the unit vector normal to the wave front.

Following the general theory [3] we introduce the quantity

$$\Psi(\mathbf{U},\Phi_\alpha) = \varphi_t + |grad\varphi|\lambda(\mathbf{U},\mathbf{n}). \tag{18}$$

The characteristic equations for (18) are

$$\frac{dx^\alpha}{d\sigma} = \frac{\partial\Psi}{\partial\Phi_\alpha}, \qquad \frac{d\Phi_\alpha}{d\sigma} = -\frac{\partial\Psi}{\partial x^\alpha} \quad (\alpha = 0,1,2,3), \tag{19}$$

where σ is the time along the rays.
The i-th component of the radial velocity $\mathbf{\Lambda}$ is defined by

$$\Lambda_i(\mathbf{U},\mathbf{n}) \equiv \frac{dx^i}{d\sigma} = \frac{\partial\Psi}{\partial\phi_i} = \lambda n_i + \frac{\partial\lambda}{\partial n_i} - \left(\mathbf{n}\cdot\frac{\partial\lambda}{\partial\mathbf{n}}\right) n_i = \lambda n_i + v_i - (n_k v_k) n_i, \tag{20}$$

$$(i = 1,2,3).$$

Hence,

$$\mathbf{\Lambda}(\mathbf{U},\mathbf{n}) = \mathbf{v} - (v_n - \lambda)\mathbf{n}. \tag{21}$$

The theory in [3] enables us to deduce the equation for φ by using (15). Of course, equations of asymptotic approximations of higher order can be written and they are affine, but their solutions are very difficult. Just to solve the linear equation (12), a method was developed by G. Boillat [3] and generalized by D. Fusco [15] (see Sections 2, 3).

2 Propagation into an uniform unperturbed state

Consider an uniform unperturbed state $\mathbf{U}^0$, solution of (6). If the quantities (14) and (16) are introduced in the expression (12) we obtain

$$(A_{0n} - \lambda I)\frac{\partial \mathbf{U}^1}{\partial \xi} = 0, \tag{22}$$

where $(\mathbf{A}_n)_0 = A_{0n}$ and $\mathbf{A}_n(\mathbf{U}) = \mathbf{A}^i n_i$. In the case where the eigenvalues are real and the eigenvectors of the matrix $\mathbf{A}_n$ are linearly independent, the system of PDEs (2) is hyperbolic (see [17] for the definition of hyperbolicity). Furthermore, in [3] it was shown that only for the waves propagating with a velocity λ such that $\nabla\lambda \cdot \mathbf{r} \neq 0$ (with $\mathbf{r}$ the right eigenvector of $(\mathbf{A}_n)_0$ corresponding to the eigenvalue λ), i.e. with a velocity that does not satisfy the Lax - Boillat exceptionality condition $\nabla\lambda \cdot \mathbf{r} = 0$, our results are valid. Eq. (22) shows that $\mathbf{U}^1(x^\alpha, \xi)$, by integration, has the form

$$\mathbf{U}^1(x^\alpha, \xi) = u(x^\alpha, \xi)\mathbf{r}(\mathbf{U}^0, \mathbf{n}) + \mathbf{v}^1(x^\alpha), \tag{23}$$

where u is a scalar function to be determined and $\mathbf{v}^1$ is an arbitrary function of integration which can be taken as zero, without loss of generality. It may be observed that in (23) u gives rise to the phenomenon of the distortion of the signals and this term governs the first-order perturbation obeying a non-linear partial differential equation (see Section 3). We conclude this section by showing how the wave front $\varphi(t, x^1, x^2, x^3) = 0$ can be determined (see [8]). Since we are considering the propagation into an uniform unperturbed state, it is known [3] that the wave front φ satisfies the partial differential equation

$$\Psi(\mathbf{U}^0, \Phi_\alpha) = \varphi_t + |grad\varphi|\lambda(\mathbf{U}^0, \mathbf{n}^0) = \Psi^0 = 0, \tag{24}$$

where $\mathbf{n}^0$ is a constant value of $\mathbf{n}$, and so

$$\Lambda_i(\mathbf{U}^0, \mathbf{n}^0) = \frac{\partial \Psi^0}{\partial \Phi_i} \quad (i = 1, 2, 3). \tag{25}$$

The characteristic equations for (24) are

$$\frac{dx^\alpha}{d\sigma} = \frac{\partial \Psi^0}{\partial \Phi_\alpha}, \qquad \frac{d\Phi_\alpha}{d\sigma} = -\frac{\partial \Psi^0}{\partial x^\alpha} \quad (\alpha = 0, 1, 2, 3), \tag{26}$$

where σ is the time along the rays.

By integration of (26) one obtains

$$x^0 = t = \sigma,\ x^i = (x^i)^0 + \Lambda_i^0(\mathbf{U}^0, \mathbf{n}^0)t,\ \text{with}\ (x^i)^0 = (x^i)_{t=0}\ (i=1,2,3). \quad (27)$$

If we denote by φ^0 the given initial surface, we have $(\varphi)_{t=0} = \varphi^0\left[(x^i)^0\right]$ and $\mathbf{n}^0$ represents the normal vector at the point $(x^i)^0$ defined by $\mathbf{n}^0 = \left(\frac{grad\varphi}{|grad\varphi|}\right)_{t=0} = \frac{grad^0\varphi^0}{|grad^0\varphi^0|}$, where $(grad^0)_i \equiv \frac{\partial}{\partial(x^i)^0}$ $(i=1,2,3)$. Then, $\mathbf{x} = \mathbf{x}|_{t=0} + \mathbf{\Lambda}^0 t$ and since the Jacobian J of the transformation $\mathbf{x} \to \mathbf{x}|_{t=0}$ is nonvanishing, i.e. $J = det|\delta_{ik} + \frac{\partial\Lambda_k^0}{\partial(x^i)^0}t| \neq 0$ $(i,k=1,2,3)$, $(x^i)^0$ can be deduced from equations $(27)_2$ and φ in the first approximation takes the following form

$$\varphi(t, x^i) = \varphi^0(x^i - \Lambda_i^0 t). \quad (28)$$

3 First approximation of the wavefront and of U

In [8] it is shown that, by utilizing (13) and (23) (see [3] and [15]), the following equation for $u(x^\alpha, \xi)$ can be obtained:

$$\frac{\partial u}{\partial\sigma} + (\nabla\Psi\cdot\mathbf{r})_0\, u\frac{\partial u}{\partial\xi} + \frac{1}{\sqrt{J}}\frac{\partial\sqrt{J}}{\partial\sigma}u + \mu^0\frac{\partial^2 u}{\partial\xi^2} = \nu^0 u, \quad (29)$$

where

$$(\nabla\Psi\cdot\mathbf{r})_0 = (|grad\varphi|)_0\,(\nabla\lambda\cdot\mathbf{r})_0\,, \quad (30)$$

$$\mu^0 = \frac{\left[\mathbf{l}\cdot\left(\mathbf{H}^k\frac{\partial\varphi}{\partial t}\frac{\partial\varphi}{\partial x^k} + \mathbf{H}^{ik}\frac{\partial\varphi}{\partial x^i}\frac{\partial\varphi}{\partial x^k}\right)\mathbf{r}\right]_0}{(\mathbf{l}\cdot\mathbf{r})_0}, \quad (31)$$

$$\nu^0 = \frac{(\mathbf{l}\cdot\nabla\mathbf{B}\,\mathbf{r})_0}{(\mathbf{l}\cdot\mathbf{r})_0}, \quad (32)$$

with $\mathbf{l}$ the left eigenvector and $\mathbf{r}$ the right eigenvector corresponding to the eigenvalue λ, that does not satisfy the Lax- Boillat condition.

By using the transformation of variables (see [15])

$$u = \frac{v}{\sqrt{J}}e^w, \qquad \kappa = \int_0^\sigma \frac{e^w}{\sqrt{J}}(\nabla\Psi\cdot\mathbf{r})_0\, d\sigma, \quad \text{with} \quad w = \int_0^\sigma \nu^0 d\sigma, \quad (33)$$

equation (29) can be reduced to an equation of the type

$$\frac{\partial v}{\partial \kappa} + v\frac{\partial v}{\partial \xi} + \hat{\mu}^0 \frac{\partial^2 v}{\partial \xi^2} = 0, \quad \text{with} \quad \hat{\mu}^0 = \frac{\mu^0 \sqrt{J} e^{-w}}{(\nabla \Psi \cdot \mathbf{r})_0}, \tag{34}$$

which is similar to Burger's equation and is valid along the characteristic rays. Equation $(34)_1$ can be reduced to the semilinear heat equation [18]

$$\frac{\partial h}{\partial \kappa} = \hat{\mu}^0 \frac{\partial^2 h}{\partial \xi^2} - h log \frac{h}{\hat{\mu}^0} \frac{d\hat{\mu}^0}{d\kappa}, \tag{35}$$

for which the solution is known, using the following Hopf transformation

$$v(\xi, \kappa) = \hat{\mu}_0 \frac{\partial}{\partial \xi} log h(\xi, \kappa). \tag{36}$$

4 Equations governing the motion of Jeffreys media and their matrix form

In [27] a theory for mechanical relaxation phenomena, based on thermodynamics of irreversible processes [11, 29, 34, 37] with internal variables [29], was developed by G. A. Kluitenberg. It was assumed that several microscopic phenomena occur, which give rise to inelastic deformation, such that the tensor of the total strain $\varepsilon_{\alpha\beta}$ can be split in two parts: $\varepsilon_{\alpha\beta} = \varepsilon^{el}_{\alpha\beta} + \varepsilon^{in}_{\alpha\beta}$, where the tensors $\varepsilon^{el}_{\alpha\beta}$ and $\varepsilon^{in}_{\alpha\beta}$ describe the elastic and inelastic strains, respectively. Contrary to the elastic strain, the inelastic deformation is due to the effects of lattice defects (slip, dislocations,...) and to the influence of microscopic stress fields, surrounding imperfections in the medium, that can give rise to memory effects on the mechanical and thermodynamic behavior of rheological media. Experiments show that there exist several types of such independent and simultaneous contributions to the inelastic strain, so that, assuming that they are of n different types, then $\varepsilon^{in}_{\alpha\beta}$ can be split in n contributions $\varepsilon^{(k)}_{\alpha\beta} (k = 1, 2, ..., n)$: $\varepsilon^{in}_{\alpha\beta} = \sum_{k=1}^{n} \varepsilon^{(k)}_{\alpha\beta}$ (with n arbitrary), that are introduced as *internal variables* in the thermodynamical state vector.

In the theory of Kluitenberg Eulerian formalism is used and it is assumed that the gradient of the displacement field is small. This implies that the deformations are supposed to be small from a kinematical (or geometrical) point of view. However the translations and the velocity of the medium

may be large [29]. Then, the strain tensor ε_{ik} is assumed to be small, i.e. $\varepsilon_{ik} = \frac{1}{2}\left(\frac{\partial}{\partial x^k}u_i + \frac{\partial}{\partial x^i}u_k\right)(i,k = 1,2,...,n)$, where u_i is the i-th component of the displacement field $\mathbf{u}$ and x^i is the i-th component of the position vector $\mathbf{x}$ in Eulerian coordinates in a Cartesian reference frame. It should be emphasized, however, that the same physical ideas which are developed in this theory can also be reformulated for the case where the deformations are large from a kinematical point of view [29].

In [27], eliminating the internal tensorial variables, for shear phenomena in the isotropic case, the following mechanical relaxation equation between the deviators $\tilde{\tau}_{ik}$ of the mechanical stress tensor (which occurs in the equation of motion and in the first law of thermodynamics) and $\tilde{\epsilon}_{ik}$ of the strain tensor was derived

$$R^{(\tau)}_{(d)0}\tilde{\tau}_{ik} + \sum_{m=1}^{n-1} R^{(\tau)}_{(d)m}\frac{d^m}{dt^m}\tilde{\tau}_{ik} + \frac{d^n}{dt^n}\tilde{\tau}_{ik} = R^{(\epsilon)}_{(d)0}\tilde{\epsilon}_{ik} + \sum_{m=1}^{n+1} R^{(\epsilon)}_{(d)m}\frac{d^m}{dt^m}\tilde{\epsilon}_{ik}$$

$$(i,k = 1,2,3). \tag{37}$$

In the above equations $\frac{d}{dt}$ is the material derivative with respect to time [29] and $R^{(\tau)}_{(d)m}$ $(m = 0,1,...,n-1)$ and $R^{(\epsilon)}_{(d)m}$ $(m = 0,1,...,n+1)$ are algebraic functions of the coefficients occurring in the phenomenological equations and in the equations of state. The rheological relations for ordinary viscous fluids, for thermoelastic media and for Maxwell, Kelvin (Voigt), Jeffreys, Burgers, Poynting-Thomson, Prandtl-Reuss, Bingham, Saint Venant and Hooke media are special cases of this more general mentioned above relation (see also [5, 11, 25, 26, 27, 28, 29, 34, 37, 38]). Assuming that only one microscopic phenomenon gives rise to inelastic strain (n = 1), in the isothermal and isotropic case, for shear phenomena, when the hydrostatic pressure is assumed constant and uniform, the mechanical relaxation equation (37) describing the behaviour of viscoanelastic media without memory (Jeffreys media) can be written in the following form [8]

$$R^{(\tau)}_{(d)0}\tilde{P}_{ik} + \frac{d}{dt}\tilde{P}_{ik} + R^{(\epsilon)}_{(d)1}\frac{d}{dt}\tilde{\epsilon}_{ik} + R^{(\epsilon)}_{(d)2}\frac{d^2}{dt^2}\tilde{\epsilon}_{ik} = 0, \tag{38}$$

where $\tilde{P}_{ik}$ and $\tilde{\epsilon}_{ik}$ are the deviators of the mechanical pressure tensor P_{ik} and of the strain tensor ϵ_{ik}, respectively, and $\frac{d\varepsilon_{ik}}{dt} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x^k} + \frac{\partial v_k}{\partial x^i}\right)$. We define P_{ik} in terms of the symmetric Cauchy stress tensor $P_{ik} = -\tau_{ik}$ $(i,k = 1,2,3)$

and the following quantities

$$\tilde{P}_{ik} = P_{ik} - \frac{1}{3}P_{ss}\delta_{ik}, \quad P = \frac{1}{3}P_{ss}, \quad P_{ss} = trP,$$

$$P_{ik} = \tilde{P}_{ik} + P\delta_{ik}, \quad \tilde{P}_{ss} = 0,$$

where the hydrostatic pressure P is the scalar part of the tensor P_{ik}. Analogous definitions are valid for the deviator $\tilde{\epsilon}_{ik}$ and the scalar part ϵ of the strain tensor. In eq. (38) the coefficients satisfy the relations

$$R^{(\tau)}_{(d)0} = a^{(0,0)}\eta^{(1,1)}_s \geq 0, \tag{39}$$

$$R^{(\epsilon)}_{(d)1} = a^{(0,0)}\left[\left(1+\eta^{(0,1)}_s\right)^2 + \eta^{(0,0)}_s\eta^{(1,1)}_s\right] \geq 0, \tag{40}$$

$$R^{(\epsilon)}_{(d)2} = \eta^{(0,0)}_s \geq 0, \tag{41}$$

where $a^{(0,0)}$ is a scalar constant which occurs in the equations of state, while the coefficients $\eta^{(0,0)}_s$, $\eta^{(0,1)}_s$ and $\eta^{(1,1)}_s$ are called *phenomenological coefficients* and represent fluidities.

The balance equations for the mass density and momentum in the case of Jeffreys media read

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho v_i) = 0, \qquad (i = 1,2,3) \tag{42}$$

$$\rho\left(\frac{\partial}{\partial t}v_i + v_k\frac{\partial}{\partial x^k}v_i\right) + \frac{\partial}{\partial x^k}\tilde{P}_{ik} = 0, \tag{43}$$

where $v_i = \frac{du_i}{dt}$ is the i-th component of the velocity field and the force per unit mass is neglected.

5 One-dimensional case

In this Section the one-dimensional case is studied, containing original results. Consider the system of equations (38), (42) and (43). Assume that $v_2 = v_3 = 0$, $x_2 = x_3 = 0$ and that the involved physical quantities depend only on x_1, denoted by x. Denote $v_1(x,t)$ by v and the components of the

deviator of the mechanical pressure tensor P_{ik}, $\tilde{P}_{ik}$, by D_{ik}. Then, the system of equations (38), (42) and (43) read

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial x} + \rho\frac{\partial v}{\partial x} = 0, \tag{44}$$

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + \frac{1}{\rho}\frac{\partial D_{11}}{\partial x} = 0, \tag{45}$$

$$\frac{\partial D_{21}}{\partial x} = 0, \tag{46}$$

$$\frac{\partial D_{31}}{\partial x} = 0, \tag{47}$$

$$\frac{\partial D_{11}}{\partial t} + v\frac{\partial D_{11}}{\partial x} + \frac{2}{3}R^{(\epsilon)}_{(d)1}\frac{\partial v}{\partial x} + \frac{2}{3}R^{(\epsilon)}_{(d)2}\frac{\partial^2 v}{\partial t \partial x} + \frac{2}{3}R^{(\epsilon)}_{(d)2}\frac{\partial^2 v}{\partial x^2}v + R^{(\tau)}_{(d)0}D_{11} = 0, \tag{48}$$

$$\frac{\partial D_{12}}{\partial t} + v\frac{\partial D_{12}}{\partial x} + R^{(\tau)}_{(d)0}D_{12} = 0, \tag{49}$$

$$\frac{\partial D_{13}}{\partial t} + v\frac{\partial D_{13}}{\partial x} + R^{(\tau)}_{(d)0}D_{13} = 0, \tag{50}$$

$$\frac{\partial D_{22}}{\partial t} + v\frac{\partial D_{22}}{\partial x} - \frac{1}{3}R^{(\epsilon)}_{(d)1}\frac{\partial v}{\partial x} - \frac{1}{3}R^{(\epsilon)}_{(d)2}\frac{\partial^2 v}{\partial t \partial x} - \frac{1}{3}R^{(\epsilon)}_{(d)2}\frac{\partial^2 v}{\partial x^2}v + R^{(\tau)}_{(d)0}D_{22} = 0, \tag{51}$$

$$\frac{\partial D_{23}}{\partial t} + v\frac{\partial D_{23}}{\partial x} + R^{(\tau)}_{(d)0}D_{23} = 0, \tag{52}$$

where $D_{ik} = D_{ki}$.
Thus, equs. (46) and (47) show that $D_{21} = f(t)$, $D_{31} = f_1(t)$, where f and f_1 are functions of t. Therefore, from eqs. (49) and (50) we have

$$D_{12} = e^{-R^{(\tau)}_{(d)0}t} + D^0_{12}, \qquad D_{13} = e^{-R^{(\tau)}_{(d)0}t} + D^0_{13}.$$

Remark that, due to the presence of a tensorial internal variable, there is a response time of the medium possessing mechanical relaxation properties.

Then, the remained system of equations (44), (45), (48), (51) and (52) takes the matrix form (1)

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x + \omega^{-1}\left[\mathbf{H}^1\frac{\partial^2 \mathbf{U}}{\partial t \partial x} + \mathbf{H}^{11}\frac{\partial^2 \mathbf{U}}{\partial x^2}\right] = \mathbf{B}(\mathbf{U}), \tag{53}$$

having the following associated system of nonlinear hyperbolic PDEs

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{B}(\mathbf{U}), \tag{54}$$

where $\mathbf{A}^0(\mathbf{U}) = \mathbf{I}$ is the identity matrix,

$$\mathbf{U} = (\rho, v, D_{11}, D_{22}, D_{23})^T, \quad \mathbf{B} = (0, 0, -R^{(\tau)}_{(d)0}D_{11}, -R^{(\tau)}_{(d)0}D_{22}, -R^{(\tau)}_{(d)0}D_{23})^T,$$

$$\mathbf{A} = \begin{pmatrix} v & \rho & 0 & 0 & 0 \\ 0 & v & \frac{1}{\rho} & 0 & 0 \\ 0 & \frac{2}{3}R^{(\epsilon)}_{(d)1} & v & 0 & 0 \\ 0 & -\frac{1}{3}R^{(\epsilon)}_{(d)1} & 0 & v & 0 \\ 0 & 0 & 0 & 0 & v \end{pmatrix}, \tag{55}$$

$$\mathbf{H}^1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3}R'^{(\epsilon)}_{(d)2} & 0 & 0 & 0 \\ 0 & -\frac{1}{3}R'^{(\epsilon)}_{(d)2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{H}^{11} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3}R'^{(\epsilon)}_{(d)2}v & 0 & 0 & 0 \\ 0 & -\frac{1}{3}R'^{(\epsilon)}_{(d)2}v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{56}$$

with $R^{(\epsilon)}_{(d)2} = \omega^{-1}R'^{(\epsilon)}_{(d)2}$. The symbol $(...)^T$ means that $\mathbf{U}$ and $\mathbf{B}$ are column vectors of 5 components.

The eigenvalues of the matrix $\mathbf{A}$ are:

- $\lambda_1 = v$ (of multiplicity equal to 3);
- the simple eigenvalues $\lambda_2^{(\pm)} = v \pm \gamma$, with $\gamma = \sqrt{\frac{2R^{(\epsilon)}_{(d)1}}{3\rho}}$.

The right eigenvectors corresponding to $\lambda_2^{(\pm)}$ can be taken as

$$\mathbf{r}_2^{(\pm)} = \left(\rho, -(v - \lambda_2^{(\pm)}),\ \frac{2}{3}R^{(\epsilon)}_{(d)1}, -\frac{R^{(\epsilon)}_{(d)1}}{3},\ 0\right)^T. \tag{57}$$

The left eigenvectors are taken as

$$\mathbf{l}_2^{(\pm)} = \left(0, -\left(v - \lambda_2^{(\pm)}\right),\ \frac{1}{\rho},\ 0,\ 0\right). \tag{58}$$

Only, the eigenvalues $\lambda_2^{(\pm)}$ do not satisfy the Lax - Boillat exceptionality condition because $\nabla\lambda_2^{(\pm)} \cdot \mathbf{r}_\mathbf{2}^{(\pm)} \neq 0$. Thus, our results are valid for $\lambda_2^{(\pm)}$. Now,

let us consider only the longitudinal wave traveling in the right direction and the case where the propagation is in a constant state $\mathbf{U}^0$, i. e.

$$\lambda_2^{(+)} = v + \gamma \quad \text{and} \quad \mathbf{U}^0 = (\rho^0,\ 0,\ 0,\ 0,\ 0),$$

with ρ^0 constant. The characteristic rays are

$$x^0 = \sigma = t, \quad x = (x)^0 + \lambda_2^{(+)}(\mathbf{U}^0)\sigma = (x)^0 + \gamma^0 t, \tag{59}$$

whence the wave front is

$$\varphi(t,x) = \varphi^0\left(x(t) - \gamma^0 t\right), \quad \text{where} \quad \gamma^0 = \gamma(\mathbf{U}^0), \quad \gamma^0 = \sqrt{\frac{2R^{(\epsilon)}_{(d)1}}{3\rho^0}}, \tag{60}$$

implying $\varphi_x = 1$.
In order to compute the terms in (29) we start with

$$\nabla\Psi \cdot r_2^{(+)} = \varphi_x(\nabla\lambda_2^{(+)} \cdot r_2^{(+)}), \text{ with } \nabla \equiv \left(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial v}, \frac{\partial}{\partial D_{11}}, \frac{\partial}{\partial D_{22}}, \frac{\partial}{\partial D_{23}}\right). \tag{61}$$

$$\text{Hence,} \quad \left(\nabla\lambda_2^{(+)} \cdot \mathbf{r}_2^{(+)}\right)_0 = \frac{1}{2}\gamma^0, \quad \text{being} \quad (\nabla\lambda_2^{(+)})_0 = \left(-\frac{\gamma^0}{2\rho^0}, 1, 0, 0, 0, 0\right). \tag{62}$$

Furthermore, a direct easy computation gives

$$\left(\mathbf{l}_2^{(+)} \cdot \nabla\mathbf{B}\ \mathbf{r}_2^{(+)}\right)_0 = -\frac{2R^{(\tau)}_{(d)0}R^{(\epsilon)}_{(d)1}}{3\rho^0}, \quad \left(\mathbf{l}_2^{(+)} \cdot \mathbf{r}_2^{(+)}\right)_0 = 2(\gamma_0)^2 = \frac{4R^{(\epsilon)}_{(d)1}}{3\rho^0}, \tag{63}$$

and so from (32) we have

$$\nu^0 = -\frac{R^{(\tau)}_{(d)0}}{2}. \tag{64}$$

Finally, we have

$$\mu^0 = \frac{\left[\mathbf{l}_2^{(+)} \cdot \left(\mathbf{H}^1 \frac{\partial\varphi}{\partial t}\frac{\partial\varphi}{\partial x} + \mathbf{H}^{11}\frac{\partial^2\varphi}{\partial x^2}\right)\left(\mathbf{r}_2^{(+)}\right)\right]_0}{\left(\mathbf{l}_2^{(+)} \cdot \mathbf{r}_2^{(+)}\right)_0} = \frac{\left(\frac{\partial\varphi}{\partial t}\right)_0 R'^{(\epsilon)}_{(d)2}}{\sqrt{6\rho^0 R^{(\epsilon)}_{(d)1}}}. \tag{65}$$

This example, in spite of its simplicity, shows the influence of a tensorial internal variable on the motion of Jeffreys media.

Note of L. R. *The present paper is one of a series of works started and planned in 2004 during a visit of Adelina Georgescu at the Department of Mathematics of the University of Messina in occasion of study days on "Asymptotic Methods with Applications to Waves and Shocks". These papers contain a systematic formulation of previous studies on nonlinear dissipative waves on rheological media, performed in a classical way by the second author L. R., following the modern point of view of double scale method as in the book [16] on asymptotic treatments of A.G. These works were continued during the scientific collaboration of the two authors, in particular at Messina in 2005 and 2007, during meetings dedicated to series of lectures of A. G. on "Applied Mathematics" and next visits of L. R. at Bucharest in 2007 and 2009. They were finished in 2009 and written in final version in 2010. These studies come from enlightening discussions on some mathematical tools, together with their physical interpretations. The author L. R. is very grateful to A. G. for her precious encouragement to this joint study regarding a revision of previous studies and the derivation of original results on the same subject.*

Acknowledgement. This work was supported by Bonino - Pulejo Foundation of Messina and by University of Studies of Messina.

References

[1] N. N. Bogoliubov and Yu. A. Mitropolskü. *Asymptotic methods in the theory of nonlinear oscillations.* Gordon and Breach, New York, 1961.

[2] G. Boillat, *La propagation des ondes.* Gauthier-Villars, Paris, 1965.

[3] G. Boillat. Ondes asymptotiques nonlinéaires, *Annali di Matematica Pura ed Applicata.* 91, IV: 31-44, 1976.

[4] Y. Choquet-Bruhat, Ondes asymptotiques et approchées pour systèmes d'équations aux dériveés partielles nonlinéaires, *J. Math. Pures et Appl.* 1968.

[5] V. Ciancio and G. A. Kluitenberg. On linear dynamical equations of state for isotropic media - II - Some cases of special interest, *Physica* 99 A: 592-600, 1979.

[6] V. Ciancio and E. Turrisi. Sulle onde di discontinuità nei mezzi anelastici di ordine uno, *Atti Accademia Peloritana dei Pericolanti*, LVII: 1-11, 1979.

[7] V. Ciancio and L. Restuccia, Asymptotic waves in anelastic media without memory (Maxwell media). *Physica* 131 A: 251-262, 1985.

[8] V. Ciancio and L. Restuccia. Nonlinear dissipative waves in viscoanelastic media. *Physica* 132 A: 606-616, 1985.

[9] V. Ciancio and L. Restuccia. The generalized Burgers equation in viscoanelastic media with memory. *Physica* 142 A: 309-320, 1987.

[10] J. D. Cole. *Perturbation methods in applied mathematics.* Blaisdell, Waltham, MA, 1968.

[11] S. R. De Groot and P. Mazur. *Non-equilibrium Thermodynamics.* North-Holland Publishing Company, Amsterdam, 1962.

[12] A. Donato. Lecture notes of the course on "Nonlinear wave propagation" held by A. Donato during academic year 1979-1980, attended by one of us (L.R.).

[13] A. Donato and A. M. Greco. *Metodi qualitativi per onde non lineari*, Quaderni del C. N. R., Gruppo Nazionale di Fisica Matematica, 11th Scuola Estiva di Fisica Matematica, Ravello, 1986, 8-20 September (1986).

[14] W. Eckhaus. *Asymptotic analysis of singular perturbation*, North-Holland, Amsterdam, 1979.

[15] D. Fusco. Onde non lineari dispersive e dissipative, *Bollettino U.M.I* 5, 16-A: 450-458, 1979.

[16] A. Georgescu. *Asymptotic treatment of differential equations.* Chapman and Hall, London, 1995.

[17] A. Georgescu and L. Restuccia. Asymptotic waves from the point of view of double-scale method. *Atti Accademia Peloritana dei Pericolanti.* LXXXIV DOI:10.1478/C1A0601005, 2006.

[18] E. Hopf. The partial differential equation $u_t + uu_x = \mu u_{xx}$, *Comm. Pure Appl. Math.* 3: 201-230, 1950.

[19] J. K. Hunter and J. B. Keller. Weakly nonlinear high frequency waves, *Comm. Pure and Appl. Math.*, 36: 1983.

[20] A. Jeffrey. The development of jump discontinuities in nonlinear hyperbolic systems of equations in two independent variables, *Arch. Rational Mech. Anal.* 14: 27-37,1963.

[21] A. Jeffrey. The propagation of weak discontinuities in quasilinear symmetric hyperbolic systems. *ZAMP* 14: 301-314, 1963.

[22] A. Jeffrey and T. Taniuti. *Nonlinear wave propagation.* Academic, New York, 1964.

[23] A. Jeffrey. *Quasilinear hyperbolic systems and waves.* Pitman, London, 1976.

[24] J. Kevorkian, Méthodes des échelles multiple, Séminaire de l'École Nationale Supérieure de Techniques Avancées, 1972.

[25] G. A. Kluitenberg. On the thermodynamics of viscosity and plasticity. *Physica* 29: 633-652, 1963.

[26] G. A. Kluitenberg. On heat dissipation due to irreversible mechanical phenomena in continuous media. *Physica.* 35: 117-192, 1967.

[27] G. A. Kluitenberg. A thermodynamic derivation of the stress-strain relations for Burgers media and related substances. *Physica* 38: 513-548, 1968.

[28] G. A.Kluitenberg and V. Ciancio. On linear dynamical equations of state for isotropic media - I - General formalism. *Physica* 93 A: 273-286, 1978.

[29] G. A. Kluitenberg. *Plasticity and Non-equilibrium Thermodynamics*, CISM Lecture Notes. Springer, Wien, New York, 1984.

[30] N. M. Krylov and N. N. Bogoliubov. *Introduction to nonlinear mechanics*, Izd. AN USS, 1937. (Russian)

[31] P. A. Lagerstrom and R. G. Casten. Basic concepts underlying singular perturbation technique. *SIAM Rev.* 14, 1: 63-120, 1972.

[32] P. D. Lax. *Contributions to the theory of partial differential equations.* Princeton University Press, 1954.

[33] P. D. Lax. Nonlinear hyperbolic equations. *Comm. Pure Appl. Math.* 6: 231-258, 1983.

[34] J. Meixner and H. G. Reik. *Thermodynamik der Irreversiblen Prozesse*, Handbuch der Physik, Band III/2. Springer, Berlin, 1959.

[35] Yu. A. Mitropolskü. *Problèmes de la théorie asymptotique des oscillations nonstationnaires*, Gauthier-Villars, Paris, 1966.

[36] R. E. Jr. O'Malley. Topics in singular perturbations, *Adv. Math.* 2, 4: 365-470, 1968.

[37] I. Prigogine. *Introduction to Thermodynamics of Irreversible Processes.* Interscience Publishers-John Wiley & Sons, New York-London, 1961.

[38] L. Restuccia and G. A. Kluitenberg. On the heat dissipation function for irreversible mechanical phenomena in anisotropic media, *Rendiconti del Seminario Matematico di Messina.* 7, II, Tomo XXII: 169-187, 2000.

[39] L. Restuccia and A. Georgescu. Determination of asymptotic waves in Maxwell media by double-scale method. *Technische Mechanik* 28, 2: 140-151, 2008.

[40] D. R. Smith. The multivariable method in singular perturbation analysis, *SIAM Rev.* 17: 221-273, 1975.

[41] E. Turrisi, V. Ciancio and G.A. Kluitenberg. On the propagation of linear transverse acoustic waves in isotropic media with mechanical relaxation phenomena due to viscosity and a tensorial internal variable II. Some cases of special interest (Poynting-Thomson, Jeffreys, Maxwell, Kelvin-Voigt, Hooke and Newton media), *Physica* 116 A: 594, 1982.

[42] G. Veronis. A note on the method of multiple scales. *A. Appl. Math.* 38: 363-368, 1980.

[43] D. J. Wolkind. Singular perturbation techniques. *SIAM Rev.* 19, 3: 502-516, 1977.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

FINITE SINGULARITIES OF TOTAL MULTIPLICITY FOUR FOR A PARTICULAR SYSTEM WITH TWO PARAMETERS*

Raluca Mihaela Georgescu† Simona Cristina Nartea ‡

Abstract

A particular Lotka-Volterra system with two parameters describing the dynamics of two competing species is analyzed from the algebraic viewpoint. This study involves the invariants and the comitants of the system determinated by the application of the affine transformations group. First, the conditions for the existence of four (different or equal) finite singularities for the general system are proofed, then is studied the particular case.

MSC: 37H25, 37B05

keywords: dynamical system, affine transformations group, invariant, comitant

1 Introduction

In this paper we study a particular family of planar vector fields with two parameters modeling the dynamics of two competing populations.

*Accepted for publication on January 10, 2011.

†gemiral@yahoo.com Department of Mathematics, University of Pitesti

‡Department of Mathematics and Informatics, Technical University of Civil Engineering, Bucharest

We consider the general form of a Lotka -Volterra system as [4], [8]

$$\begin{cases} \dot{x} = x(c + gx + hy), \\ \dot{y} = y(f + mx + ny), \end{cases} \tag{1}$$

where x, y represent the number of the populations of the two species, c, f represent the growth rates of the species, and g, h, m, n represent the competitive impacts of one specie to another. The equilibrium points of (1) are: $M_1(0,0)$, $M_2(-c/g, 0)$, $M_3(0, -f/n)$ and $M_4((fh - cn)/(gn - hm), (cm - fg)/(gn - hm))$. All these points are in the finite part of the phase plane if and only if $gn(gn - hm) \neq 0$. On the other hand, for the system (1), we have $\mu_0 = gn(gn - hm)$, where μ_0 is defined in the Appendix.

Therefore, for $\mu_0 \neq 0$ the system (1) has four different or equal equilibrium points.

The following two theorems holds, and their proofs can be found in [3].

Theorem 1. *[3]. For $\mu_0 \neq 0$ the number of the four finite singularities of the system (1) are determinated by the following conditions:*

$$\begin{array}{lcl} 4 \textit{ simple} & \Leftrightarrow & \mathbf{D} \neq 0; \\ 2 \textit{ simple},\ 1 \textit{ double} & \Leftrightarrow & \mathbf{D} = 0,\ \mathbf{S} \neq 0; \\ 2 \textit{ double} & \Leftrightarrow & \mathbf{D} = \mathbf{S} = 0,\ \mathbf{P} \neq 0; \\ 1 \textit{ of multiplicity } 4 & \Leftrightarrow & \mathbf{D} = \mathbf{S} = \mathbf{P} = 0, \end{array}$$

where $\mathbf{D}, \mathbf{S}, \mathbf{P}$ are defined in the Appendix.

Since $\mu_0 \neq 0$, due to the transformation $(x, y) \mapsto (x/g, y/n)$, we can consider $g = n = 1$. Therefore, the system (1) becomes

$$\begin{cases} \dot{x} = x(c + x + hy), \\ \dot{y} = y(f + mx + y), \end{cases} \tag{2}$$

for which $\mu_0 = 1 - hm$, $\mathbf{D} = -c^2 f^2 (c - fh)^2 (f - cm)^2$ and

$$\mathbf{S} = 3c^4 m^2 (x + hy)^2 [3m^2 x^2 - 2m(hm - 4)xy + (3h^2 m^2 - 8hm + 8)y^2],$$
$$\mathbf{P} = c^4 y^2 (mx + y)^2 (\text{ if } cf = 0),$$

or

$$\mathbf{S} = 3c^4 m^2 (hm - 1)^4 x^2 (3m^2 x^2 + 8mxy + 8y^2),$$
$$\mathbf{P} = c^4 (hm - 1)^2 y^2 (mx + y)^2 (\text{ if } (c - fh)(f - cm) = 0).$$

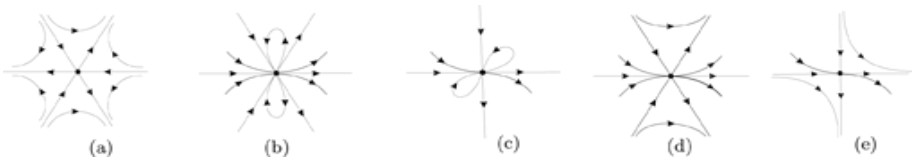
We use the following abbreviations: S=saddle, N=node, F=focus, C=center, SN=saddle-node.
In addition, K, W_3, W_4 are defined in the Appendix.

Theorem 2. *[3]. Let us consider the system (1) with $\mu_0 \neq 0$. Then the type of the finite singularities of this system is determinated by the following affine-invariant conditions:*

1) $S,\ S,\ S,\ N \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K < 0,\ W_4 \geq 0;$
2) $S,\ S,\ S,\ F \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K < 0,\ W_4 < 0,\ B_3 \neq 0;$
3) $S,\ S,\ S,\ C \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K < 0,\ W_4 < 0,\ B_3 = 0;$
4) $S,\ N,\ N,\ N \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K > 0$ and $\begin{cases} W_4 > 0 \text{ or} \\ W_4 = 0,\ W_3 \geq 0; \end{cases}$
5) $S,\ N,\ N,\ F \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K > 0$ and $\begin{cases} W_4 < 0,\ B_3 \neq 0 \text{ or} \\ W_4 = 0,\ W_3 < 0 \end{cases}$;
6) $S,\ N,\ N,\ C \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 < 0,\ K > 0$ and $W_4 < 0, B_3 = 0;$
7) $S,\ S,\ N,\ N \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 > 0$ and $\begin{cases} W_4 > 0 \text{ or} \\ W_4 = 0,\ W_3 \geq 0; \end{cases}$
8) $S,\ S,\ N,\ F \quad \Leftrightarrow \quad \mathbf{D} \neq 0,\ \mu_0 > 0$ and $\begin{cases} W_4 < 0 \text{ or} \\ W_4 = 0,\ W_3 < 0; \end{cases}$
9) $SN,\ S,\ S \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0, \mu_0 < 0,\ K < 0;$
10) $SN,\ N,\ N \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0,\ \mu_0 < 0,\ K > 0$
and $\begin{cases} W_4 > 0 \text{ or} \\ W_4 = 0,\ W_3 \geq 0; \end{cases}$

11) $SN,\ N,\ F \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0,\ \mu_0 < 0,\ K > 0,\ W_4 < 0;$
12) $SN,\ N,\ C \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0,\ \mu_0 < 0,\ K > 0,\ W_4 = 0,\ W_3 < 0;$
13) $SN,\ S,\ N \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0,\ \mu_0 > 0,\ W_4 \geq 0;$
14) $SN,\ S,\ F \quad \Leftrightarrow \quad \mathbf{D} = 0,\ \mathbf{S} \neq 0,\ \mu_0 > 0,\ W_4 < 0;$
15) $SN,\ SN \quad \Leftrightarrow \quad \mathbf{D} = \mathbf{S} = 0,\ \mathbf{P} \neq 0;$
16) a degenerated nonhyperbolic point of the multiplicity 4
(a) $\Leftrightarrow \quad \mathbf{D} = \mathbf{S} = \mathbf{P} = 0,\ \mu_0 < 0,\ \eta > 0,\ \chi > 0;$
(b) $\Leftrightarrow \quad \mathbf{D} = \mathbf{S} = \mathbf{P} = 0,\ \mu_0 < 0,\ \eta > 0,\ \chi < 0;$
(c) $\Leftrightarrow \quad \mathbf{D} = \mathbf{S} = \mathbf{P} = 0,\ \mu_0 < 0,\ \eta = 0;$
(d) $\Leftrightarrow \quad \mathbf{D} = \mathbf{S} = \mathbf{P} = 0,\ \mu_0 > 0,\ \eta > 0;$
(e) $\Leftrightarrow \quad \mathbf{D} = \mathbf{S} = \mathbf{P} = 0, \mu_0 > 0,\ \eta = 0,$

where (a)-(e) have the representations:



2 The particular competing species model

The model we study in this paper is proposed as an application by M.W. Hirsch, S. Smale and R. L. Devaney in [5] and has the form

$$\begin{cases} \dot{x} &= x(a-x-ay), \\ \dot{y} &= y(b-bx-y), \end{cases} \tag{3}$$

where $x,\ y$ represent the number of the populations of the two species, a and b are positive parameters.

In order to apply the Theorems 1 and 2, we transform the system (3) into the system (1).

The system (3) is equivalent with

$$\begin{cases} \dot{x} &= -x(-a+x+ay), \\ \dot{y} &= -y(-b+bx+y), \end{cases}$$

and, by the change of the sense of the time $t \mapsto -t$ we obtain the system

$$\begin{cases} \dot{x} &= x(-a+x+ay), \\ \dot{y} &= y(-b+bx+y), \end{cases} \tag{4}$$

which is the system we are concerned herein.

Due to physical reasons, the phase space must be the first quadrant (without axes of coordinates). However, for mathematical (namely bifurcation) reasons we consider, in addition, the origin and the half-axes.

Remark 1. *The system (4) has the same equilibrium points as (3), but the attractive properties of the equilibria of the system (4) are opposite of those of the system (3).*

3 The equilibrium points

By convention, we say that an equilibrium exists if its coordinates are finite and positive. Therefore, this is a biological, not a mathematical existence.

The equilibrium points of the system (4) are $M_1(0,0)$, $M_2(a,0)$, $M_3(0,b)$, $M_4(a(1-b))/(1-ab), b(1-a)/(1-ab)$. For these points we compute Δ_i,

ρ_i, δ_i, $(i = 1, 2, 3, 4)$ given in the Appendix.

$$\begin{gathered}\Delta_1 = ab,\ \rho_1 = -a - b,\ \delta_1 = (a-b)^2,\\ \Delta_2 = ab(a-1),\ \rho_2 = a - b + ab,\ \delta_2 = (a + b - ab)^2,\\ \Delta_3 = ab(b-1),\ \rho_3 = -a + b + ab,\ \delta_3 = (a + b - ab)^2,\\ \Delta_4 = ab(a-1)(b-1)/(1-ab),\ \rho_4 = (a + b - 2ab)/(1-ab),\\ \delta_4 = [(a-b)^2 + 4a^2b^2(a-1)(b-1)]/(1-ab)^2\end{gathered} \tag{5}$$

For the system (4) we have

$$\begin{gathered}\mu_0 = 1 - ab,\ \mu_1 = (-2b + ab + ab^2)x + (2a - ab - a^2b)y,\\ \mathbf{D} = -a^4b^4(a-1)^2(b-1)^2,\ K = 2(bx^2 + 2xy + ay^2),\\ W_4 = (a-b)^2(ab - a - b)^2[(a-b)^2 + 4a^2b^2(a-1)(b-1)].\end{gathered} \tag{6}$$

In the following, we study the nature of the finite singularities of the system (4) for the case $\mu_0 \neq 0$ (i.e. $ab \neq 1$).

Case D $\neq 0$. This case is equivalent with $1 - ab \neq 0$, $a \notin \{0, 1\}$, $b \notin \{0, 1\}$. From Theorem 1, it follows that the system (4) has four simple equilibrium points.

• If $\mu_0 < 0$ then $1 - ab < 0$. Since a and b are positive parameters, it follows that $K > 0$. If $W_4 > 0$, then we have $a > 1,\ b > 1$ and we are in the case 4 from the Theorem 2 (i.e. the system (4) has three nodes and a saddle) or $(a > 1,\ b < 1)$, $(a < 1,\ b > 1)$, where the point M_4 is not in the first quadrant, therefore it does not exist from biological viewpoint. In this case there are only three points from biological viewpoint (two nodes and a saddle). On the other hand, W_4 can not be negative. Indeed, if $W_4 < 0$ then $(a-1)(b-1) < 0$, therefore $a > 1,\ b < 1$ or $a < 1,\ b > 1$. It follows that M_4 is not in the first quadrant, therefore it does not exist from biological viewpoint. Again there are only three points from biological viewpoint (two nodes and a saddle).

Thus, the finite singularities of total multiplicity four of the system (3) which exist from biological viewpoint are as follows: if $a > 1,\ b > 1$, then M_1 is a repulsive node, $M_2,\ M_3$ are attractive nodes and M_4 is a saddle; if $a > 1,\ b < 1$, then M_1 is a repulsive node, M_2 is an attractive node and M_3 is a saddle; if $a < 1,\ b > 1$, then M_1 is a repulsive node, M_2 is a saddle and M_3 is an attractive node.

• If $\mu_0 > 0$ then $1 - ab > 0$. If $W_4 > 0$, then we have $a < 1,\ b < 1$ and we are in the case 7 from the Theorem 2 (i.e. the system (4) has two nodes and

two saddles), or $(a > 1,\ b < 1)$, $(a < 1,\ b > 1)$, when M_4 is not in the first quadrant, therefore it does not exist from biological viewpoint. In this case there are only three points (two nodes and a saddle). On the other hand, W_4 can not be negative. Indeed, if $W_4 < 0$ then $(a-1)(b-1) < 0$, equivalently with $a > 1,\ b < 1$ or $a < 1,\ b > 1$ therefore, the point M_4 is not in the first quadrant, so it does not exist from biological viewpoint. Again there are only three points from biological viewpoint (two nodes and a saddle).

Thus, the finite singularities of total multiplicity four of the system (3) that exist from biological viewpoint are as follows: if $a < 1,\ b < 1$, then M_1 is a repulsive node, $M_2,\ M_3$ are saddles and M_4 is an attractive node; if $a > 1,\ b < 1$, then M_1 is a repulsive node, M_2 is an attractive node and M_3 is a saddle; if $a < 1,\ b > 1$, then M_1 is a repulsive node, M_2 is a saddle and M_3 is an attractive node.

Case $\mathbf{D} = 0$. We have two subcases: $ab = 0$ or $(a-1)(b-1) = 0$.

• For $ab = 0$, without loss of generality, due to the change $x \leftrightarrow y$, $a \leftrightarrow b$, which keeps the system (4) unchanged, we can consider only $a = 0$. In this case $\mathbf{S} = 0$ and $\mathbf{P} = b^4x^4$.

If $b \neq 0$, then $\mathbf{P} \neq 0$ and we are in the case 15 from the Theorem 2 (i.e. the system (4) has two saddle-nodes).

If $b = 0$, then $\mathbf{P} = 0$, $\mu_0 = 1 > 0$ and $\eta = 1 > 0$, therefore we are in the case 16 (d) from the Theorem 2 (i.e. the system (4) has a point of multiplicity 4).

Thus, in the plane, the type of the finite singularities for the system (3) are as follows: if $a = 0,\ b \neq 0$ $(a \neq 0,\ b = 0)$, then $M_1 = M_2$, $M_3 = M_4$ $(M_1 = M_3,\ M_2 = M_4)$ are saddle-nodes ; if $a = 0,\ b = 0$, then $M_1 = M_2 = M_3 = M_4$, i.e. we have a nonhyperbolic point of multiplicity 4.

• For $(a-1)(b-1) = 0$, without loss of generality, due to the change $x \leftrightarrow y$, $a \leftrightarrow b$, which keeps the system (4) unchanged, we can consider only $a = 1$. In this case $\mathbf{S} = 3b^2(b-1)^4x^2(3b^2x^2 + 8bxy + 8y^2)$.

If $\mathbf{S} \neq 0$, then $b \notin \{0, 1\}$ and $\mu_0 = 1 - b$, $\eta = 0$, $W_4 = (b-1)^2$.

For $\mu_0 < 0$ (i.e. $b > 1$), we have $K > 0$, $W_4 > 0$, therefore we are in the case 10 from the Theorem 2 (i.e. the system (4) two nodes and a saddle-node).

For $\mu_0 > 0$ (i.e. $b < 1$), we have $K > 0$, $W_4 > 0$ therefore we are in the case 13 from the Theorem 2 (i.e. the system (4) has a node, a saddle and a saddle-node).

Thus, the finite singularities of total multiplicity four of the system (3) are as follows: if $a = 1,\ b < 1$ $(b > 1)$, then M_1 is a repulsive node, M_3 is a saddle, $M_2 = M_4$ is a saddle-node (M_1 is a repulsive node, M_3 is an attractive node, $M_2 = M_4$ is a saddle-node); if $b = 1,\ a < 1$ $(a > 1)$ then M_1 is a repulsive node, M_2 is a saddle, $M_3 = M_4$ is a saddle-node (M_1 is a repulsive node, M_2 is an attractive node, $M_3 = M_4$ is a saddle-node).

If $\mathbf{S} = 0$, then $b = 0$ (if $b = 1$ we obtain a contradiction: $\mu_0 = 0$). For $b = 0$ we have two saddle-nodes $M_1 = M_3$ and $M_2 = M_4$.

4 The phase portraits

In [2] the system (3) was studied by the topological methods and the dynamic bifurcation diagram was representing. Here we represent only the phase portraits that have a biological significance (i.e the equilibria are in the first quadrant) and where the equilibrium points have total multiplicity four (fig.2). The parametric portrait (fig.1) is representing by the strata 0-10 without the curve T (corresponding to $ab = 1$).

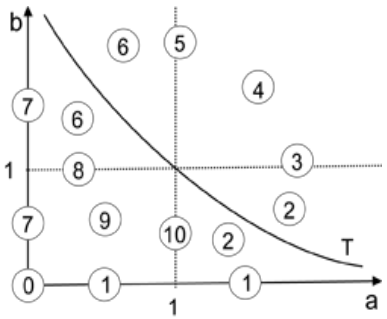


Fig. 1. The parametric portrait.

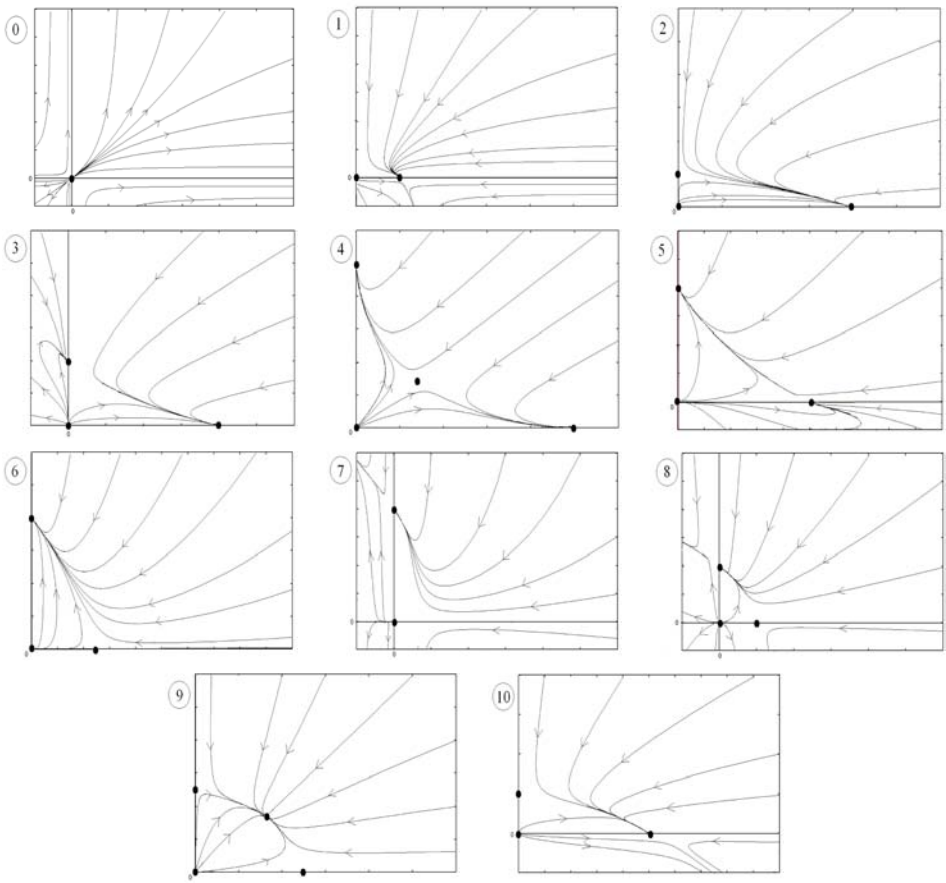


Fig. 2. The phase portraits for the various strata from Fig. 1

5 Appendix

Consider the two-dimensional nonlinear system of ordinary differential equations

$$\begin{cases} \dot{x} = p_0(x,y) + p_1(x,y) + p_2(x,y) \equiv p(x,y), \\ \dot{y} = q_0(x,y) + q_1(x,y) + q_2(x,y) \equiv q(x,y), \end{cases} \tag{7}$$

where p_i and q_i, i=0,1,2 homogenous polynomials of i degree.

For a singular point $M_i(x_i, y_i)$ we use the notations:

$$\rho_i = (p'_x(x,y) + q'_y(x,y))|_{(x_i,y_i)} = \mathrm{tr}\mathbf{A}_i,$$

$$\Delta_i = \left| \begin{array}{cc} p'_x(x,y) & p'_y(x,y) \\ q'_y(x,y) & q'_y(x,y) \end{array} \right|_{(x_i,y_i)} = \det \mathbf{A}_i,$$

$$\delta_i = \rho_i^2 - 4\Delta_i = \mathrm{tr}^2\ \mathbf{A}_i - 4 \det \mathbf{A}_i,$$

where $\mathbf{A}_i$ is the matrix of the linear terms from the linearized system around the point (x_i, y_i).

The following polynomials are the GL-comitants and T-comitants of the system (7) [1], [6], [7]:

$$\begin{array}{l} C_i(\mathbf{a},x,y) = yp_i(\mathbf{a},x,y) - xq_i(\mathbf{a},x,y), \quad i = 0,1,2; \\ \eta(\mathbf{a}) = Discrim(C_2(\mathbf{a},x,y)); \\ K(\mathbf{a},x,y) = Jacob(p_2(\mathbf{a},x,y), q_2(\mathbf{a},x,y)); \\ \mu_0(\mathbf{a}) = Res_x(p_2,q_2)/y^4 = Discrim(K(\mathbf{a},x,y))/16; \\ \mathbf{D}(\mathbf{a}) = -\Big(\big((D,D)^{(2)},D\big)^{(1)},D\Big)^{(3)}/576 \equiv -Discrim(D); \\ \mathbf{P}(\mathbf{a},x,y) = \mu_2^2 - 3\mu_1\mu_3 + 12\mu_0\mu_4; \\ \mathbf{S}(\mathbf{a},x,y) = \left[3\mu_1^2 - 8\mu_0\mu_2\right]^2 - 16\mu_0^2\mathbf{P}; \\ B_3(\mathbf{a},x,y) = (C_2,D)^{(1)} = Jacob(C_2,D), \\ W_3 = \mu_0^2 \sum\limits_{1\le i<j<l\le 4} \delta_i\delta_j\delta_l, \\ W_4 = \mu_0^2\delta_1\delta_2\delta_3\delta_4. \end{array}$$

References

[1] J.C. Artes, J. Llibre, N. Vulpe. Singular points of quadratic systems: a complete classification in the coefficient space $\mathbf{R}^{12}$, Preprint no 39/2005, Dept. of Math., Univ. Autonòma de Barcelona, 1-84.

[2] R.M. Georgescu. Bifurcation in a two competing species model, *ICAMCS, Automation Computers Applied Mathematics*, Cluj-Napoca, 15(1): 147-155, 2006.

[3] R.M. Georgescu. *Bifurcatie in dinamica biologica cu metode de teoria grupurilor*, Ed. Univ. Pitesti, Pitesti, 2009.

[4] X.Z. He. The Lyapunov functionals for delay Lotka-Volterra-type models, *SIAM J. Appl. Math.*, 58(4): 1222-1236, 1998.

[5] M.W. Hirsch, S. Smale, R.L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos*, Academic Press-Elsevier, New York-Amsderdam, 2004.

[6] D. Schlomiuk, N. Vulpe. Geometry of quadratic differential systems in the neighborhood of infinity, *J. Diff. Eq.*, 215:357-400, 2005.

[7] D. Schlomiuk, N. Vulpe. Planar quadratic differential systems with invariant straight lines of total multiplicity 4, URL: www.crm.umontreal.ca/pub/Rapports/2900-2999/2940.pdf

[8] P. Yan. *Limit cycles for generalized Liénard-type and Lotka-Volterra systems*, Painosalama Oy, Turku, 2005.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

A FINITE VOLUME METHOD FOR SOLVING GENERALIZED NAVIER-STOKES EQUATIONS*

Stelian Ion† Anca Veronica Ion‡

Abstract

In this paper we set up a numerical algorithm for computing the flow of a class of pseudo-plastic fluids. The method uses the finite volume technique for space discretization and a semi-implicit two steps backward differentiation formula for time integration. As primitive variables the algorithm uses the velocity field and the pressure field. In this scheme quadrilateral structured primal-dual meshes are used. The velocity and the pressure fields are discretized on the primal mesh and the dual mesh respectively. A certain advantage of the method is that the velocity and pressure can be computed without any artificial boundary conditions and initial data for the pressure. Based on the numerical algorithm we have written a numerical code. We have also performed a series of numerical simulations.

MSC: 35Q30, 65M08, 76A99, 76D05, 76M12.

keywords: generalized Navier-Stokes equations, pseudo-plastic fluids, finite volume methods, admissible primal-dual mesh, discrete derivatives operators, discrete Hodge formula.

*Accepted for publication on December 30, 2010.

†ro_diff@yahoo.com "Gheorghe Mihoc-Caius Iacob" Institute of Statistical Mathematics and Applied Mathematics, Bucharest, Romania; Paper written with financial support of ANCS Grant 2-CEX06-11-12/2006.

‡"Gheorghe Mihoc-Caius Iacob" Institute of Statistical Mathematics and Applied Mathematics, Bucharest, Romania.

1 Introduction

In this paper we are interested in the numerical approximation of a class of pseudo-plastic fluid flow. The motion of the fluid is described by the generalized incompressible Navier-Stokes equations

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla p + \nabla \cdot \sigma(\mathbf{u}) + \mathbf{f}, \\ \nabla \cdot \mathbf{u} &= 0, \end{cases} \tag{1}$$

where $\mathbf{u}$ is the velocity vector field, p is the hydrodynamic pressure field, σ the extra stress tensor field and $\mathbf{f}$ is the body force. The extra stress tensor $\sigma(\mathbf{u})$ obeys a constitutive equation of the type

$$\sigma_{ab}(\mathbf{u}) = 2\nu(|\widetilde{\partial \mathbf{u}}|)\widetilde{\partial u}_{ab} \tag{2}$$

where $\widetilde{\partial u}$ is the strain rate tensor given by

$$\widetilde{\partial u}_{ab} = \frac{1}{2}\left(\partial_a u_b + \partial_b u_a\right),$$

∂_a standing for the partial derivative with respect to the space coordinate x_a, and for any square matrix $\mathbf{e}$, $|\mathbf{e}|$ being defined as

$$|\mathbf{e}| = \left(\sum_{i,j} e_{ij}^2\right)^{1/2}.$$

Concerning the viscosity function $\nu(s)$, we assume that it is a continuous differentiable, decreasing function, with bounded range

$$\begin{cases} 0 < \nu_\infty \leq \nu(s) \leq \nu_0 < \infty, \forall s > 0, \\ \left(\nu(s_1) - \nu(s_2)\right)\left(s_1 - s_2\right) < 0, \forall s_1, s_2 > 0, \end{cases} \tag{3}$$

and it satisfies the constraint

$$\nu(s) + s\dot{\nu}(s) > c > 0. \tag{4}$$

The model of the Newtonian fluid corresponds to $\nu = constant$.

We consider the case when the flow takes place inside a fixed and bounded domain $\Omega \subset \mathbf{R}^2$ and we assume that the fluid adheres to its boundary $\partial\Omega$, hence we impose a Dirichlet type boundary condition for the velocity field,

$$\mathbf{u} = \mathbf{u}_D(x),\ x \in \partial\Omega, t > 0. \tag{5}$$

To the equations (1) we append the initial condition for the velocity

$$\mathbf{u}(x, 0) = \mathbf{u}_0(x),\ x \in \Omega. \tag{6}$$

The initial boundary value problem (IBV), which we intend to solve numerically, consists in finding the velocity field $\mathbf{u}(x, t)$ and the pressure field $p(x, t)$ that satisfy the partial differential equations (1), boundary condition (5) and the initial condition (6).

A constitutive function as (2) is used , for example, to describe the behavior of polymeric fluids, [5], [8], [11], and the flow of the blood through the vessels, [19], [9], [6], [18].

In writing down a numerical algorithm for the non-stationary incompressible generalized Navier-Stokes equations three main difficulties occur, namely: (i) the velocity field and the pressure field are coupled by the incompressibility constraint [12], (ii) the presence of the nonlinear convection term and (iii) the nonlinear dependence of the viscosity on the share rate.

The first two problems are common to the Navier-Stokes equations and in the last fifty years several methods were developed to overcome them: the projection method, [12], [13], [7],[14], [3], and gauge method, [20]- to mention the most significant methods for our case.

When one deals with a non-Newtonian fluid, the nonlinearity of the viscosity rises a new problem in obtaining a discrete form for the generalized Navier-Stokes equations. The new issue is the development of an appropriate discrete form of the action of the stress tensor on the boundary of the volume-control. A similar dificulty is raised by the discretization of the p-laplacean, see [2] for that.

The outline of the paper is as follows. In Section 2 we define the weak solution of IBV (1), (5) and (6) and we present an existence theorem of the weak solution for a class of pseudo-plastic fluids that satisfy (4). In Section 3 we establish the semi-discrete, space discrete coordinates and continuum time variable form of the equation (1) and we present some general concepts concerning the space discretization and related notions like admissible mesh, primal and dual mesh, the discretization of the derivative operators etc. In Section 4 we present an algorithm for solving a 2D model. In the last section we present the results of some numerical simulations of the lid driven cavity flow.

2 The Existence of the Weak Solution

To define the weak solution we need the following functional frame, [16], [17].

By $L^p(\Omega)$ and $W^{m,p}(\Omega)$, $m = 0, 1, \cdots$, we denote the usual Lebesgue and Sobolev spaces, respectively. The scalar product in L^2 is indicated by $(\cdot, \cdot)$. For $\mathbf{u}$, $\mathbf{v}$ vector functions defined on Ω we put

$$(\mathbf{u}, \mathbf{v}) = \int\limits_{\Omega} u^a v_a \mathrm{d}x,$$

$$(\nabla \mathbf{u}, \nabla \mathbf{v}) = \sum_{a,b=1} \int\limits_{\Omega} \partial_a u^b \partial_a v^b \mathrm{d}x.$$

We denote by $||\cdot||$ the norm in L^2 associate to $(\cdot, \cdot)$. The norm in $W^{m,p}$ is denoted by $||\cdot||_{m,p}$. Consider the space

$$\mathcal{V} = \{\psi \in C_0^{\infty}(\Omega), \mathrm{div}\psi = 0\}.$$

We define $\mathbf{H}(\Omega)$ the completion of $\mathcal{V}$ in the space $\mathbf{L}^2(\Omega)$. We denote by $\mathbf{H}^1(\Omega)$ the completion of $\mathcal{V}$ in the space $\mathbf{W}^{1,2}$.

For $T \in (0, \infty]$ we set $Q_T = \Omega \times [0, T)$ and define

$$\mathcal{V}_T = \{\phi \in C_0^{\infty}(Q_T);\ \mathrm{div}\phi(x,t) = 0 \text{ in } Q_T\}.$$

The weak solution of IBV is defined as follow.

Definition 1. *Let* $\mathbf{f} \in \mathbf{L}^2(\Omega)$. *Let* $\mathbf{u}_0(x) \in \mathbf{L}^2(\Omega)$ *and* u_D *be such that*

$$\begin{cases} \mathrm{div}\, \mathbf{u}_0 = 0, \\ \mathbf{u}_D \cdot \mathbf{n} = 0, x \in \partial\Omega, \\ \mathbf{u}_0 = \mathbf{u}_D,\ x \in \partial\Omega, \end{cases} \tag{7}$$

and there exists $\mathbf{v} \in \mathbf{W}^{1,2}(\Omega) \cap \mathbf{L}^4(\Omega)$ *a vector function that satisfies*

$$\begin{cases} \mathrm{div}\, \mathbf{v} = 0, \\ \mathbf{v} = \mathbf{u}_D,\ x \in \partial\Omega. \end{cases} \tag{8}$$

Then $\mathbf{u}$ *is a weak solution of* IBV (1,5,6) *if*

$$\mathbf{u} - \mathbf{v} \in L^2\left((0,T); \mathbf{H}^1(\Omega)\right) \cap L^{\infty}\left((0,T); \mathbf{H}(\Omega)\right) \tag{9}$$

and $\mathbf{u}$ *verifies*

$$\begin{aligned} -\int_0^\infty \left(\mathbf{u}, \frac{\partial \phi}{\partial t}\right) \mathrm{d}t - \int_0^\infty (\mathbf{u} \otimes \mathbf{u}, \nabla \phi)\, \mathrm{d}t + \int_0^\infty \left(\sigma(\mathbf{u}), \widetilde{\partial \phi}\right) \mathrm{d}t = \\ = \int_0^\infty (\mathbf{f}, \phi)\, \mathrm{d}t + (u_0, \phi) \end{aligned} \tag{10}$$

for any test function $\phi \in \mathcal{V}_T$.

Concerning the existence of the weak solution of the IBV we proved the following result, [15]:

Theorem 1. *If the constitutive function* $\nu(\cdot)$ *satisfies the relations* (3) *and* (4) *then there exists a weak solution of the* IBV (1), (5) *and* (6).

3 Semi-discrete Finite Volume Method

The finite volume method (FVM) is a method for approximating the solution of a partial differential equation (PDE). It basically consists in partitioning the domain Ω, on which the PDE is formulated, into small polygonal domains ω_i (control volumes) on which the unknown is approximated by constant values, [10].

We consider a class of finite-volume schemes that includes two types of meshes: the *primal mesh*, $\mathcal{T} = \{\omega_\mathcal{I}, \mathbf{r}_\mathcal{I}\}$ and the *dual mesh*, $\widetilde{\mathcal{T}} = \{\widetilde{\omega}_\mathcal{J}, \widetilde{\mathbf{r}}_\mathcal{J}\}$. The space discrete form of the GNS equations are obtained from the integral form of the balance of momentum equation and mass balance equation on the primal mesh and the dual mesh respectively.

For any ω_i of the primal mesh $\mathcal{T}$ the integral form of the balance of momentum equation reads as,

$$\partial_t \int_{\omega_i} \mathbf{u}(\mathbf{x}, t) \mathrm{d}x + \int_{\partial \omega_i} \mathbf{u}\mathbf{u} \cdot \mathbf{n} \mathrm{d}s + \int_{\omega_i} \nabla p \mathrm{d}x = \int_{\partial \omega_i} \sigma \cdot \mathbf{n} \mathrm{d}s, \tag{11}$$

and for any $\widetilde{\omega}_\alpha$ of the dual mesh $\widetilde{\mathcal{T}}$ the integral form of mass balance equation is given by

$$\int_{\partial \widetilde{\omega}_\alpha} \mathbf{u} \cdot \mathbf{n} \mathrm{d}s = 0. \tag{12}$$

The velocity field $\mathbf{u}(\mathbf{x},t)$ and the pressure field $p(\mathbf{x},t)$ are approximated by the piecewise constant functions on the primal mesh and the dual mesh respectively,

$$\mathbf{u}(\mathbf{x},t) \approx \mathbf{u}_i(t),\ \forall \mathbf{x} \in \omega_i,\ p(\mathbf{x},t) \approx p_\alpha(t),\ \forall \mathbf{x} \in \widetilde{\omega}_\alpha.$$

By using certain approximation schemes of the integrals as functions of the discrete variables $\{u_i(t)\}_{i\in\mathcal{I}}$, $\{p_\alpha(t)\}_{\alpha\in\mathcal{J}}$ one can define:

$$\begin{gathered}\mathcal{F}_i(\mathbf{u}) \approx \int_{\partial\omega_i} \mathbf{u}\mathbf{u}\cdot\mathbf{n}\mathrm{d}s,\ \mathcal{S}_i(\mathbf{u}) \approx \int_{\partial\omega_i} \sigma\cdot\mathbf{n}\mathrm{d}s,\\ \mathbf{Grad}_i(p) \approx \int_{\omega_i} \nabla p\mathrm{d}x,\ ,\mathrm{Div}_\alpha(\mathbf{u}) \approx \int_{\partial\widetilde{\omega}_\alpha} \mathbf{u}\cdot\mathbf{n}\mathrm{d}s.\end{gathered} \tag{13}$$

The semi-discrete form of GNS equations, continuous with respect to time variable and discrete with respect to space variable, can be written as:

$$\begin{aligned} m_i\frac{\mathrm{d}\mathbf{u}_i}{\mathrm{d}t} + \mathcal{F}_i(\{\mathbf{u}\}) + \mathbf{Grad}_i(\{p\}) - \mathcal{S}_i(\{\mathbf{u}\}) = 0, \quad i\in\mathcal{I}\\ \mathrm{Div}_\alpha(\{\mathbf{u}\}) = 0, \quad \alpha\in\mathcal{J}\end{aligned} \tag{14}$$

where m_i stands for the volume of the ω_i.

Now the problem is to find the functions $\{\mathbf{u}_i(t)\}_{i\in\mathcal{I}}, \{p_\alpha(t)\}_{\alpha\in\mathcal{J}}$ that satisfy the differential algebraic system of equations (DAE) (14) and the initial condition

$$\mathbf{u}_i(t)|_{t=t_0} = \mathbf{u}_i^0,\ \forall i\in\mathcal{I}. \tag{15}$$

In solving the Cauchy problem (14) and (15), an essential step is to define a primal-dual mesh $\left(\mathcal{T},\widetilde{\mathcal{T}}\right)$ that allows one to calculate the velocity field independent of the pressure field.

In the next subsections we define a pair of quadrilateral admissible primal-dual (QAPD) meshes $\left(\mathcal{T},\widetilde{\mathcal{T}}\right)$, and we define the discrete gradient of the scalar functions and the discrete divergence of the vector functions such that the discrete space of the vector fields admits an orthogonal decomposition into two subspaces: one of discrete divergences free vectors fields, and other consisting of vectors that are the discrete gradient of some scalar fields.

3.1 Quadrilateral primal-dual meshes

Let Ω be a polygonal domain in $\mathbf{R}^2$. Let $\mathcal{T} = \{\omega_\mathcal{I}, \mathbf{r}_\mathcal{I}\}$ be a quadrilateral mesh defined as follows:

(1) ω_i is a quadrilateral, $\overline{\cup_{i\in I}\omega_i} = \overline{\Omega}$,
(2) $\forall i \neq j \in I$ and $\overline{\omega}_i \cap \overline{\omega}_j \neq \Phi$, either $\mathcal{H}_1(\overline{\omega}_i \cap \overline{\omega}_j) = 0$, or
$\sigma_{ij} := \overline{\omega}_i \cap \overline{\omega}_j$ is a common $(n-1)-$face of ω_i and ω_j,
(3) $\mathbf{r}_i \in \omega_i$, if $\omega_i = [ABCD]$, then $\mathbf{r}_i = [M_{AB}M_{DC}] \cap [M_{AD}M_{BC}]$,
(4) for any vertex P $\in \Omega$ there exists only four quadrilateral ω with the common vertex P,

where $\mathcal{H}_1$ is the one-dimensional Hausdorff measure, and M_{AB} denotes the midpoint of the line segment $[AB]$.

Let $\widetilde{\mathcal{T}} = \{\widetilde{\omega}_{\mathcal{J}}, \widetilde{\mathbf{r}}_{\mathcal{J}}\}$ be another mesh defined as follows:

(1) $\forall \alpha \in \mathcal{J}$, $\widetilde{\mathbf{r}}_\alpha$ is a vertex of $\mathcal{T}$,
(2) $\widetilde{\mathbf{r}}_\alpha \in \widetilde{\omega}_\alpha, \forall \alpha \in \mathcal{J}$,
(3) $\forall \widetilde{\mathbf{r}}_\alpha \in \overline{\Omega}$, the poligon $\widetilde{\omega}_\alpha$ has the vertexes :
the centers of the quadrilaterals with the common vertex $\widetilde{\mathbf{r}}_\alpha$
and the midpoints of the sides emerging from $\widetilde{\mathbf{r}}_\alpha$,

where by "center" of the quadrilateral we understand the intersection of the two segments determined by the midpoints of two opposed sides.

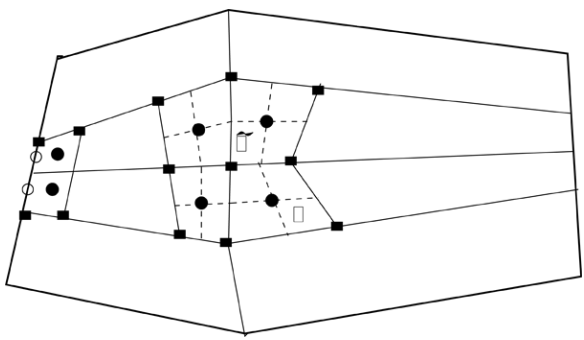


Figure 1: Quadrilateral mesh.

We call $(\mathcal{T}, \widetilde{\mathcal{T}})$ - a pair of QAPD meshes.

We denote by $H_{\widetilde{\mathcal{T}}}(\Omega)$ the space of piecewise constant scalar functions that are constant on each volume $\widetilde{\omega}_\alpha \in \widetilde{\omega}_{\mathcal{J}}$, by $\mathbf{H}_{\mathcal{T}}(\Omega)$ the space of piecewise constant vectorial functions that are constant on each volume $\omega_i \in \omega_{\mathcal{I}}$ and by $\mathbf{H} \otimes \mathbf{H}_{\widetilde{\mathcal{T}}}(\Omega)$ the space of piecewise constant tensorial functions of order two that are constant on each volume $\widetilde{\omega}_a \in \widetilde{\omega}_{\mathcal{J}}$.

For any quantity ψ that is piecewise constant on $\widetilde{\omega}_{\mathcal{J}}$ we denote by ψ_α the constant value of ψ on $\widetilde{\omega}_a$, analogously ψ_i stands for the constant value of a piecewise constant quantity ψ on $\omega_{\mathcal{I}}$.

We define the discrete derivative operators:
$\mathrm{Div}_{(\mathcal{T},\widetilde{\mathcal{T}})} : \mathbf{H}_{\mathcal{T}}(\Omega) \to H_{\widetilde{\mathcal{T}}}(\Omega)$, by

$$\mathrm{Div}_\alpha(\mathbf{u}) := \int\limits_{\partial\widetilde{\omega}_\alpha} \mathbf{u}\cdot\mathbf{n}\mathrm{d}s = \sum_i u_i^a \int\limits_{\partial\widetilde{\omega}_\alpha\cap\omega_i} n_a \mathrm{d}s, \tag{16}$$

$\partial_{(\mathcal{T},\widetilde{\mathcal{T}})} : \mathbf{H}_{\mathcal{T}}(\Omega) \to \mathbf{H}\otimes\mathbf{H}_{\widetilde{\mathcal{T}}}(\Omega)$ by

$$\partial_b\, u^a|_\alpha =: \frac{1}{\mathrm{m}(\widetilde{\omega}_\alpha)} \int\limits_{\partial\widetilde{\omega}_\alpha} u^b n_a \mathrm{d}s = \frac{1}{\mathrm{m}(\widetilde{\omega}_a)} \sum_i u_i^b \int\limits_{\partial\widetilde{\omega}_\alpha\cap\omega_i} n_a \mathrm{d}s, \tag{17}$$

$\mathbf{Grad}_{(\mathcal{T},\widetilde{\mathcal{T}})} : H_{\widetilde{\mathcal{T}}}(\Omega) \to \mathbf{H}_{\mathcal{T}}(\Omega)$ by

$$\mathbf{Grad}_i(\phi) \int\limits_{\partial\omega_i} \phi\,\mathbf{n}\mathrm{d}s = \sum_\alpha \phi_\alpha \int\limits_{\partial\omega_i\cap\widetilde{\omega}_\alpha} \mathbf{n}\mathrm{d}s, \tag{18}$$

$\mathbf{rot}_{(\mathcal{T},\widetilde{\mathcal{T}})} : H_{\widetilde{\mathcal{T}}}(\Omega) \longrightarrow \mathbf{H}_{\mathcal{T}}(\Omega)$ by

$$\mathbf{rot}_i(\phi) := \frac{1}{\mathrm{m}(\omega_i)} \int\limits_{\partial\omega_i} \phi\mathrm{d}\mathbf{r} = \frac{1}{\mathrm{m}(\omega_i)} \sum_\alpha \phi_\alpha \int\limits_{\partial\omega_i\cap\widetilde{\omega}_\alpha} \mathrm{d}\mathbf{r}. \tag{19}$$

On the space $\mathbf{H}_{\mathcal{T}}(\Omega)$ we define the scalar product $\langle\langle\cdot,\cdot\rangle\rangle$ by

$$\langle\langle\mathbf{u},\mathbf{v}\rangle\rangle = \sum_{i\in\mathcal{I}} \mathbf{u}_i\cdot\mathbf{v}_i, \tag{20}$$

and on the space $H_{\widetilde{\mathcal{T}}}(\Omega)$ we define the scalar product $\langle\cdot,\cdot\rangle$ by

$$\langle\phi,\psi\rangle = \sum_{\alpha\in\mathcal{J}} \phi_\alpha\psi_\alpha. \tag{21}$$

In the next lemma we prove certain properties of the discrete derivative operators.

Lemma 1. *Let* $\left(\mathcal{T},\widetilde{\mathcal{T}}\right)$ *be a pair of QAPD meshes and the discrete divergence, the discrete gradient and the discrete rotation be defined, respectively, by* (16), (18), *and* (19). *Then:*

(a1) *Discrete Stokes formula. For any* $\mathbf{u} \in \mathbf{H}_{\mathcal{T}}(\Omega)$ *and any* $\phi \in H_{\widetilde{\mathcal{T}}}(\Omega)$*, a discrete integration by parts formula holds, that is*

$$\left\langle \mathrm{Div}_{(\mathcal{T},\widetilde{\mathcal{T}})}(\mathbf{u}), \phi \right\rangle + \left\langle \left\langle \mathbf{u}, \mathbf{Grad}_{(\mathcal{T},\widetilde{\mathcal{T}})}(\phi) \right\rangle \right\rangle = 0. \tag{22}$$

(a2) *For any* $\psi \in H_{\widetilde{\mathcal{T}}}(\Omega)$*,* $\psi|_{\partial\Omega} = 0$ *one has*

$$\mathrm{Div}_{(\mathcal{T},\widetilde{\mathcal{T}})} \mathbf{rot}_{(\mathcal{T},\widetilde{\mathcal{T}})}\, \psi = 0. \tag{23}$$

Proof. To prove (a1) we use the fact that for any domain ω

$$\int_{\partial\omega} \mathbf{n} \mathrm{d}s = 0$$

and the definitions of the two operators.

To prove (a2), we note firstly that

$$\mathrm{Div}_\alpha(\mathbf{rot}_{(\mathcal{T},\widetilde{\mathcal{T}})}\psi) = \sum_i \mathbf{rot}_i(\psi) \cdot \int_{\partial\widetilde{\omega}_\alpha \cap \omega_i} \mathbf{n} \mathrm{d}s =$$

$$= \sum_i \frac{1}{\mathrm{m}(\omega_i)} \sum_\beta \psi_\beta \int_{\widetilde{\omega}_\beta \cap \partial\omega_i} \mathrm{d}\mathbf{r} \cdot \int_{\partial\widetilde{\omega}_\alpha \cap \omega_i} \mathbf{n} \mathrm{d}s.$$

Then, let $\omega_{i_a^\alpha}$, $a = \overline{1,4}$ be the primal volumes with the common vertex P_α and numbered such that $\omega_{i_a^\alpha}$ and $\omega_{i_{a+1}^\alpha}$ have a common side. For each i_a^α let $P_{\alpha_b^{i_a^\alpha}}$, $b = \overline{1,4}$ be the vertexes of the quadrilateral $\omega_{i_a^\alpha}$ anticlockwise numbered and $P_{\alpha_1^{i_a^\alpha}} = P_\alpha$. We have

$$\frac{1}{\mathrm{m}(\omega_{i_a})} \sum_b \psi_{\alpha_b^{i_a^\alpha}} \int_{\widetilde{\omega}_{\alpha_b^{i_a^\alpha}} \cap \partial\omega_{i_a^\alpha}} \mathrm{d}\mathbf{r} \cdot \int_{\partial\widetilde{\omega}_\alpha \cap \omega_{i_a}} \mathbf{n} \mathrm{d}s = \psi_{\alpha_2^{i_a^\alpha}} - \psi_{\alpha_4^{i_a^\alpha}}.$$

Finally, by summing up for $a = \overline{1,4}$, we have

$$\mathrm{Div}_\alpha(\mathbf{rot}_{(\mathcal{T},\widetilde{\mathcal{T}})}\psi) = \sum_a (\psi_{\alpha_2^{i_a^\alpha}} - \psi_{\alpha_4^{i_a^\alpha}}) = 0,$$

for any α such that $P_\alpha \in \Omega$. If for some α, $P_\alpha \in \partial\Omega$, we use the fact that $\psi_\beta = 0$ on any boundary dual-volumes $\widetilde{\omega}_\beta$.

Now we prove an orthogonal decomposition of the space $\mathbf{H}_{\mathcal{T}}(\Omega)$ that resembles the one for the non-discrete case. Let $\{\Psi^{\alpha}\}_{\alpha\in\mathcal{J}}, \Psi^{\alpha}\in H_{\widetilde{\mathcal{T}}}(\Omega)$ be a basis of the space $H_{\widetilde{\mathcal{T}}}(\Omega)$ given by

$$\Psi^{\alpha}(x)=\left\{\begin{array}{ll} 1, & \text{if } x\in\widetilde{w}^{\alpha},\\ 0, & \text{if } ,x\notin\widetilde{w}^{\alpha}.\end{array}\right. \tag{24}$$

Define the discrete vector field $\mathcal{U}^{\alpha}\in\mathbf{H}_{\mathcal{T}}(\Omega)$ by

$$\mathcal{U}^{\alpha}=\mathbf{rot}(\Psi^{\alpha}). \tag{25}$$

Let $\mathbf{W}_{\mathcal{T}}(\Omega)$ be the linear closure of the set $\{\mathcal{U}^{\alpha};\alpha\in\mathrm{Int}(\mathcal{J})\}$ in the space $\mathbf{H}_{\mathcal{T}}(\Omega)$ and let $\mathbf{G}_{\mathcal{T}}(\Omega)$ be the subspace orthogonal to it, so that

$$\mathbf{H}_{\mathcal{T}}(\Omega)=\mathbf{W}_{\mathcal{T}}(\Omega)\oplus\mathbf{G}_{\mathcal{T}}(\Omega). \tag{26}$$

We state and prove the following proposition.

Proposition 1. $\mathbf{G}_{\mathcal{T}}(\Omega)$ *consists of elements* $\mathbf{Grad}_{(\mathcal{T},\widetilde{\mathcal{T}})}$ *with* $\phi\in H_{\widetilde{\mathcal{T}}}(\Omega)$.

Proof. Let $\mathbf{u}\in\mathbf{G}_{\mathcal{T}}(\Omega)$, i.e.

$$\langle\langle\mathbf{u},\mathcal{U}^{\alpha}\rangle\rangle=0,\ \forall\alpha\in\mathrm{Int}(\mathcal{J}). \tag{27}$$

We construct a function $\phi\in H_{\widetilde{\mathcal{T}}}(\Omega)$ such that

$$\mathbf{Grad}_i(\phi)=\mathbf{u}_i,\ \forall i\in\mathcal{I}.$$

For a given ω_i we denote by $P_{\alpha_b^i}$, $b=\overline{1,4}$ its vertexes counterclockwise numbered. The gradient of a scalar field ϕ can be written as

$$\mathbf{Grad}_i(\phi)=\overrightarrow{\tau}_{1,3}(\phi_{\alpha_3^i}-\phi_{\alpha_1^i})+\overrightarrow{\tau}_{2,4}(\phi_{\alpha_4^i}-\phi_{\alpha_2^i}),$$

where $\overrightarrow{\tau}_{1,3}$ is a vector orthogonal to $\overrightarrow{P_{\alpha_2^i}P_{\alpha_4^i}}$ oriented from $P_{\alpha_1^i}$ to $P_{\alpha_3^i}$ and $|\overrightarrow{\tau}_{1,3}|=\left|\overrightarrow{P_{\alpha_2^i}P_{\alpha_4^i}}\right|/2$ and $\overrightarrow{\tau}_{2,4}$ is a vector orthogonal to $\overrightarrow{P_{\alpha_1^i}P_{\alpha_3^i}}$ oriented from $P_{\alpha_2^i}$ to $P_{\alpha_4^i}$ and $|\overrightarrow{\tau}_{2,4}|=\left|\overrightarrow{P_{\alpha_1^i}P_{\alpha_3^i}}\right|/2$. Hence, we have

$$\begin{array}{l}\dfrac{\mathbf{u}_i\cdot\overrightarrow{P_{\alpha_2^i}P_{\alpha_4^i}}}{\mathrm{m}(\omega_i)}=\phi_{\alpha_4^i}-\phi_{\alpha_2^i},\\ \dfrac{\mathbf{u}_i\cdot\overrightarrow{P_{\alpha_1^i}P_{\alpha_3^i}}}{\mathrm{m}(\omega_i)}=\phi_{\alpha_3^i}-\phi_{\alpha_1^i}.\end{array} \tag{28}$$

The point is that if $\mathbf{u}$ satisfies the ortogonalty conditions (27) then one can solve the equations (28) inductively, i.e. starting from two adjacent values and following some path of continuation. For a general discrete field $\mathbf{u}$ the different paths lead to different values!

Corollary 1 (Discrete Hodge formula)**.** *Let* $(\mathcal{T}, \widetilde{\mathcal{T}})$ *be a pair of QAPD meshes. Then for any* $\mathbf{w} \in \mathbf{H}_{\mathcal{T}}(\Omega)$ *there exists an element* $\mathbf{u} \in \mathbf{H}_{\mathcal{T}}(\Omega)$ *and a scalar function* $\phi \in H_{\widetilde{\mathcal{T}}}(\Omega)$ *such that*

$$\mathbf{w} = \mathbf{u} + \mathbf{Grad}(\phi) \quad \textit{with } \mathrm{Div}_{(\mathcal{T},\widetilde{\mathcal{T}})}(\mathbf{u}) = 0. \tag{29}$$

Proof. We search for a divergence free vector $\mathbf{u}$ of the form

$$\mathbf{u} = \sum_{a \in \mathcal{J}} \alpha_a \mathcal{U}^a.$$

By inserting this form into (29), one obtains a linear algebraic system of equation for the determination of the unknowns $\{\alpha_a\}_{a\in\mathcal{J}}$,

$$\left\langle\left\langle \mathbf{w}, \mathcal{U}^b \right\rangle\right\rangle = \sum_{a \in \mathcal{J}} \alpha_a \left\langle\left\langle \mathcal{U}^a, \mathcal{U}^b \right\rangle\right\rangle.$$

The matrix of the system is the Gram matrix of a linear independent family, hence there exists an unique solution $\mathbf{u}$.

Since $\langle\langle \mathbf{w} - \mathbf{u}, \mathcal{U}^a \rangle\rangle = 0$ for any function in the basis, it follows that $\mathbf{w} - \mathbf{u}$ is orthogonal to $\mathbf{G}^{\perp}$, thus $\mathbf{w} - \mathbf{u} \in \mathbf{G}$. Hence there exists $\phi \in H_{\widetilde{\mathcal{T}}}(\Omega)$ such that

$$\mathbf{w} - \mathbf{u} = \mathbf{Grad}(\phi).$$

3.2 Discrete convective flux and discrete stress flux

To cope with the boundary value problems one defines a partition $\{\partial_k \omega\}_{k\in\mathcal{K}}$ of the boundary $\partial\Omega$ mesh induced by the primal mesh i.e

$$\partial_k \omega = \partial\Omega \cup \partial\omega_{i_k}, \partial\Omega = \cup_{k\in\mathcal{K}} \partial_k \omega.$$

On each $\partial_k \omega$ the boundary data $\mathbf{u}_D$ are approximated by constant values $\mathbf{u}_{D\,k}$.

Several formulas to calculate the numerical convective flux (NCF) are available, most of them derived from the theory of hyperbolic equations.

In the case of a hyperbolic equation, the numerical convective flux must satisfy, besides the accuracy of the approximation requirements, a number of conditions in order that the implied solution be physically relevant. In the case of Navier-Stokes equation at high Reynolds number, the way in which NCF is evaluated is also very important. We propose to define the NCF as follow. Consider the tensorial product $\mathbf{u} \oplus \mathbf{u}$ constant on the dual mesh and, for any control volume ω_i that does not lie on the boundary $\mathcal{F}$, set for the NCF:

$$\mathcal{F}_i^a = \sum_\alpha (u^a u^b)_\alpha \int_{\widetilde{\omega}_\alpha \cap \partial\omega_i} n_b \mathrm{d}s. \tag{30}$$

The tensorial product $\mathbf{u} \oplus \mathbf{u}$ is approximated by

$$(u^a u^b)_\alpha = \frac{1}{\mathrm{m}(\widetilde{\omega}_\alpha)} \int_{\widetilde{\omega}_\alpha} u^a \mathrm{d}x \, \frac{1}{\mathrm{m}(\widetilde{\omega}_\alpha)} \int_{\widetilde{\omega}_\alpha} u^b \mathrm{d}x. \tag{31}$$

The numerical stress flux is set up by considering that the gradient of the velocity is piecewise constant on the dual mesh. This fact implies that the stress tensor is also piecewise constant on the dual mesh. So we can write for the numerical stress flux

$$\mathcal{S}_i(\mathbf{u}) = \sum_\alpha \sigma_\alpha(\mathbf{u}) \cdot \int\limits_{\partial\omega_i \cap \widetilde{\omega}_\alpha} \mathbf{n} \mathrm{d}s. \tag{32}$$

The values of $\sigma_\alpha(\mathbf{u})$ are evaluated as

$$\sigma_\alpha(\mathbf{u}) = 2\nu(|\mathbf{D}_\alpha(\mathbf{u})|)\mathbf{D}_\alpha(\mathbf{u}), \tag{33}$$

where the discrete strain rate tensor $\mathbf{D}_\alpha$ is given by

$$D_{ab}(\mathbf{u})|_\alpha = \frac{1}{2} \left(\partial_a u^b + \partial_b u^a \right) \Big|_\alpha . \tag{34}$$

Dirichlet Boundary conditions. The boundary conditions for the velocity are taken into account through the numerical convective flux and numerical stress flux. If for some α the dual volume $\widetilde{\omega}_\alpha$ intersects the boundary $\partial\Omega$ the gradient of the velocity is given by:

$$\partial_a u^b \Big|_\alpha = \frac{1}{\mathrm{m}(\widetilde{\omega}_\alpha)} \int\limits_{\partial\widetilde{\omega}_\alpha} u^b n_a \mathrm{d}s =$$

$$= \frac{1}{\mathrm{m}(\widetilde{\omega}_\alpha)} \left(\int\limits_{\partial_{\mathrm{ext}}\widetilde{\omega}_\alpha} u_D^b n_a \mathrm{d}s + \sum_i u_i^b \int\limits_{\partial_{\mathrm{int}}\widetilde{\omega}_\alpha \cap \omega_i} n_a \mathrm{d}s \right). \tag{35}$$

For a primal volume ω_i adjacent to the boundary $\partial\Omega$ the NCF (30) is given by

$$\mathcal{F}_i^a = u_D^a u_D^b \int\limits_{\partial\Omega\cap\partial\omega_i} n_b \mathrm{d}s + \sum_\alpha (u^a u^b)_\alpha \int\limits_{\Omega\cap\partial\omega_i} u n_b \mathrm{d}s. \tag{36}$$

4 Fully-Discrete Finite Volume Method

We set up a time integration scheme of the Cauchy problem (14) and (15) that determines the velocity field independently on the pressure field. The pressure field results from the discrete balance momentum equation (14-1). The scheme resembles the Galerkin method and it makes use of the orthogonal decomposition (26) of the space $\mathbf{H}_{\mathcal{T}}(\Omega)$ and of the set of divergence free vectorial fields $\{\mathcal{U}^\alpha\}_{\alpha\in\mathcal{J}^0}$.

We write the unknown velocity field $\mathbf{u}(t)$ as linear combination of $\{\mathcal{U}^\alpha\}_{\alpha\in\mathcal{J}^0}$

$$\mathbf{u} = \sum_\alpha \xi_\alpha(t)\mathcal{U}^\alpha \tag{37}$$

where the coefficients $\xi_\alpha(t)$ are required to satisfy the ordinary differential equations

$$\sum_\alpha \frac{\mathrm{d}\xi_\alpha}{\mathrm{d}t} \left\langle\left\langle m\mathcal{U}^\alpha, \mathcal{U}^\beta \right\rangle\right\rangle + \left\langle\left\langle \mathcal{F}(\xi), \mathcal{U}^\beta \right\rangle\right\rangle - \left\langle\left\langle \mathcal{S}(\xi), \mathcal{U}^\beta \right\rangle\right\rangle = 0,\ \forall\beta \in \mathcal{J}^0, \tag{38}$$

with the initial conditions

$$\sum_\alpha \xi_\alpha(0) \left\langle\left\langle \mathcal{U}^\alpha, \mathcal{U}^\beta \right\rangle\right\rangle = \left\langle\left\langle \mathbf{u}^0, \mathcal{U}^\beta \right\rangle\right\rangle,\ \forall\beta \in \mathcal{J}^0. \tag{39}$$

If the functions ξ_α satisfy (38) and (39) then $m\dfrac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + \mathcal{F}(\{\mathbf{u}\}) - \mathcal{S}(\{\mathbf{u}\})$ belongs to the space $\mathbf{G}_{\mathcal{T}}(\Omega)$ which implies that there exists a scalar field $p(t)$ such that

$$-\mathbf{Grad}_{(\mathcal{T},\widetilde{\mathcal{T}})} p = m\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + \mathcal{F}(\{\mathbf{u}\}) - \mathcal{S}(\{\mathbf{u}\}). \tag{40}$$

Concerning the initial conditions (15), we note that for $t = 0$ the solution (37) equals not u^0 but its projection on the space $\mathbf{W}_{\mathcal{T}}(\Omega)$.

Now we develop a time integration scheme for the equation (38) derived from two steps implicit backward differentiation formulae (BDF).

Let $\{t^n\}$ be an increasing sequence of moments of time. We make the notations: $\xi_\alpha^n = \xi_\alpha(t^n)$, $\mathbf{u}^n = \sum_\alpha \xi_\alpha^n \mathcal{U}^\alpha$. Supposing that one knows the values $\{\xi^{n-1}, \xi^n\}$ one calculates the values ξ^{n+1} at the next moment of time t_{n+1} as follows. Define the second degree polynomial $P(t)$ which interpolates the unknown ξ^{n+1} and known quantities $\{\xi^{n-1}, \xi^n\}$ at the moments of time t^{n+1}, t^n, t^{n-1}, respectively. The unknowns ξ^{n+1} are determined by imposing to the polynomial $P(t)$ to satisfy the equations (38).

For a constant time step $\triangle t$ one has

$$\frac{\mathrm{d}P_\alpha(t^{n+1})}{\mathrm{d}t} = \left(\frac{3}{2}\xi_\alpha^{n+1} - 2\xi_\alpha^n + \frac{1}{2}\xi_\alpha^{n-1}\right)/\triangle t$$

that leads to the following nonlinear equations for ξ^{n+1}

$$\begin{aligned}\sum_I \frac{3}{2}\xi_\alpha^{n+1}\left\langle\left\langle m\mathcal{U}^\alpha, \mathcal{U}^\beta\right\rangle\right\rangle + \quad & \triangle t\left\langle\left\langle \mathcal{F}(\xi^{n+1}), \mathcal{U}^\beta\right\rangle\right\rangle - \triangle t\left\langle\left\langle \mathcal{S}(\xi^{n+1}), \mathcal{U}^\beta\right\rangle\right\rangle = \\ & = \left\langle\left\langle 2\mathbf{u}^n - 0.5\mathbf{u}^{n-1}, m\mathcal{U}^\beta\right\rangle\right\rangle.\end{aligned} \tag{41}$$

To overcome the difficulties implied by the nonlinearity, we consider a linear algorithm:

$$\begin{aligned}\sum_\alpha \frac{3}{2}\lambda_\alpha^{n+1}\left\langle\left\langle m\mathcal{U}^\alpha, \mathcal{U}^\beta\right\rangle\right\rangle - \triangle t\left\langle\left\langle \mathcal{S}(\mathbf{u}; \lambda^{n+1}), \mathcal{U}^\beta\right\rangle\right\rangle = \\ 0.5\left\langle\left\langle \mathbf{u}^n - \mathbf{u}^{n-1}, m\mathcal{U}^\beta\right\rangle\right\rangle - \\ -\triangle t\left\langle\left\langle \frac{3}{2}\mathcal{F}(\mathbf{u}^n) - \frac{1}{2}\mathcal{F}(\mathbf{u}^{n-1}), \mathcal{U}^\beta\right\rangle\right\rangle + \triangle t\left\langle\left\langle \mathcal{S}(\mathbf{u}^n), \mathcal{U}^\beta\right\rangle\right\rangle,\end{aligned} \tag{42}$$

where

$$\lambda^{n+1} := \xi^{n+1} - \xi^n.$$

For the first step one can use a Euler step

$$\begin{aligned}\sum_\alpha \lambda_\alpha^{n+1}\left\langle\left\langle m\mathcal{U}^\alpha, \mathcal{U}^\beta\right\rangle\right\rangle - \triangle t\left\langle\left\langle \mathcal{S}(\mathbf{u}^n; \lambda^{n+1}), \mathcal{U}^\beta\right\rangle\right\rangle = \\ -\triangle t\left\langle\left\langle \mathcal{F}(\mathbf{u}^n), \mathcal{U}^\beta\right\rangle\right\rangle + \triangle t\left\langle\left\langle \mathcal{S}(\mathbf{u}^n), \mathcal{U}^\beta\right\rangle\right\rangle.\end{aligned} \tag{43}$$

In both (42) and (43) schemes we use the notation

$$\mathcal{S}(\mathbf{u}^n;\lambda^{n+1}) = 2\nu(|\mathbf{D}(\mathbf{u}^n)|)\sum_{\alpha}\lambda_{\alpha}^{n+1}\mathbf{D}(\mathcal{U}^{\alpha}).$$

5 Numerical Simulations

We present the results of some numerical experiments designed to test the numerical method presented in the previous sections.

We consider the pseudo-plastic fluid modeled by the Carreau-Yasuda law,

$$\nu(\dot{\gamma}) = \nu_{\infty} + (\nu_0 - \nu_{\infty})\left(1 + (\Lambda\dot{\gamma})^a\right)^{(n-1)/a}.$$

In the current study the problem was solved for a series of rectangular regular or non-regular meshes. The code incorporates the time integration scheme (42) and (43); the numerical convective flux $\mathcal{F}$ defined by the formulae (30), (31), (36) and the numerical stress flux $\mathcal{S}$ defined by formulae (32), (33), (34), (17), (35). In all the numerical simulations we consider that at the initial time the fluid is at rest.

Lid Driven Cavity Flow

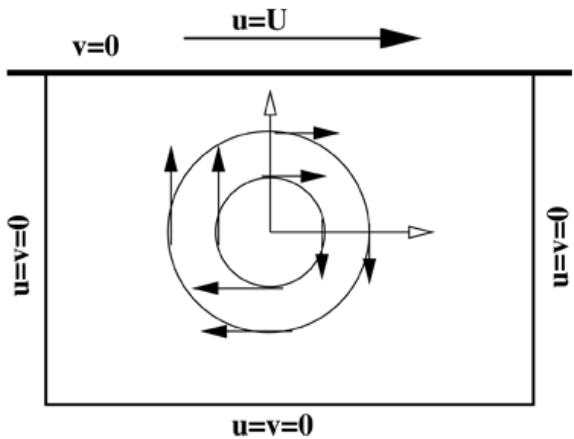


Figure 2: Lid Driven 2D Cavity Flow.

The fluid is moving in a rectangular box, the side and bottom walls are static while the top wall is moving across the cavity with a constant velocity

$u = U, v = 0$ as in Fig. 2. We assume the non-slip boundary conditions on the walls.

In the first set of computations we test the capabilty of the method to catch the behavior of the pseudo-plastic fluid. To be more precise, we chose a pseudo-plastic fluid and two Navier Stokes fluids having the viscosities equal to ν_0, and ν_∞, respectively.

Figure 3 shows the contours plots of the steady solutions for the three type of fluids. Each flow consists of a core of fluid undergoing solid body rotation and small regions in the bottom corners of counter-rotating vortex. The intensity of the counter-rotating vortex is decreasing with respect to viscosity. The velocity profile along the vertical centerline is shown in Figure 4. We observe that, in the lower part of the cavity, the fluid is moving in contrary sense to the sense of the motion of the top wall. The maximum of the negative velocity depends decreasingly on the viscosity of the fluid.

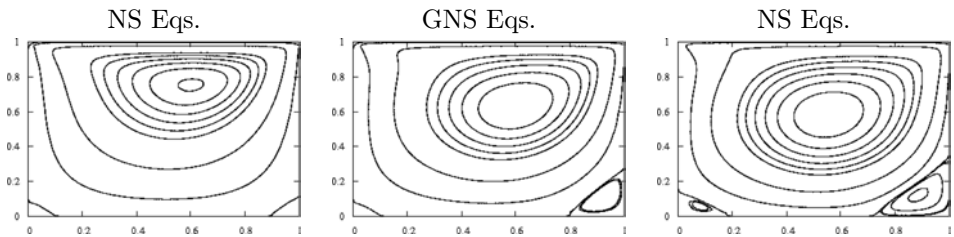


Figure 3: $U = 0.01\mathrm{ms}^{-1}$, $a = 0.144$. Contour plot of stream functions, steady solutions. Regular grid, 51×51 grid points.

The second set of computations analyzes the response of the numerical method to the variation of the parameters of the fluid. The results are shown in Figure 5.

Final Remarks

A certain advantage of our method is that there is no need to introduce artificial boundary conditions for the pressure field or supplementary boundary conditions for additional velocity field as in the projection methods or gauge methods. The preliminary numerical results prove a good agreement with

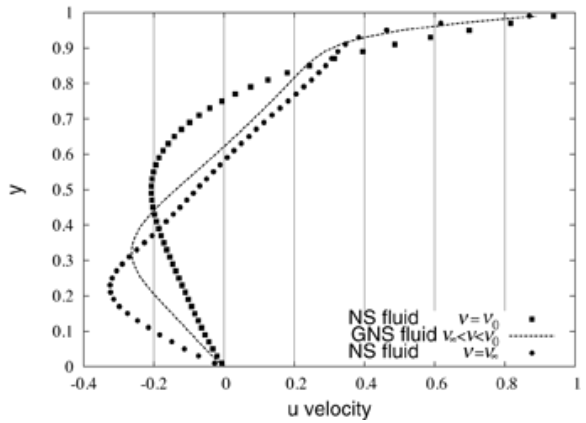
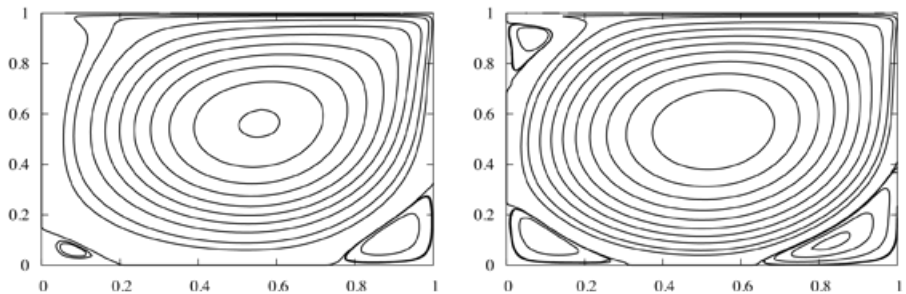


Figure 4: $U = 0.01\text{ms}^{-1}$, $a = 0.144$. Distribution of u-velocity along at vertical centre line of the cavity. Regular grid, 51×51 grid points.



$a = 0.644$
steady solution

$a = 0.144$
$t = 200$

Figure 5: GNS Eqs. $U = 0.1\text{ms}^{-1}$ $\text{Re}_\infty = 10^4/1.57$, $\text{Re}_0 = 10^3/1.57$. Stretched grid, 51×51 grid points.

the results obtained by other methods. At the present moment we do not know if it is possible to extended the method to the 3D case and this is a drawback of the method. The study of this extension might be a task for our future work.

References

[1] P.D. Anderson, O.S. Galaktionov, G.W.M. Peters, F.N. van de Vose, H.E.H. Meijer. Mixing of non-Newtonian fluids in time-periodic cavity flow. *J. Non-Newtonian Fluid Mech.* 93:265-286, 2000.

[2] B. Andreianov, F. Boyer and F. Hubert. Finite volume scheme for the p-laplacian on cartezian meshes. M^2AN. 38(6):931–960, 2004.

[3] J. Bell, P. Colella, and H. Glaz. A second-order projection method for incompressible Navier-Stokes equations. *J. Comp. Phys.* 85:257-283, 1989.

[4] K.L. Brenan, S.L. Campbell, L.R. Petzold. *Numerical Solution of Initial Values Problems in Differential-Algebraic Equations.* Classics in Applied Mathematics, SIAM, 1996.

[5] G.F. Carey, W. Barth, J. A. Woods, B. S. Kork, M. L. Anderson, S. Chow, and W. Bangerth. Modelling error and constitutive relations in simulattion of flow and transport. *Int. J. Num. Meth. Fluids.* 46:1211-1236, 2004.

[6] P. Carreau. Rheological equations from molecular network theories. *Transactions of Society of Rheology.* 16:99-127, 1972.

[7] A.J. Chorin. Numerical solution of the Navier-Stokes equations. *Math. Comput.* 22:745-762, 1968.

[8] S.D. Cramer, J.M. Marchello. Numerical evolution of models describing non-Newtonian behavior. *AIChE Journal.* 4(6):980-983, 1968.

[9] M.M. Cross. Rheology of non newtonian fluids: A new flow equation for pseudoplastic systems. *Journal of Colloid and Interface Science.* 20:417-437, 1965.

[10] R. Eymard, Th. Gallouet and R. Herbin. Finit Volume Method. In P.G. Ciarlet and J.L. Lions, (eds.) *Handbook of Numerical Analysis.* North Holland, 2000.

[11] H. Eyring. Viscosity, plasticity, and diffusion as examples of absolute reaction rates. *Journal of Chemical Physics.* 4:283-287, 1936.

[12] J.L. Guermond, P. Minev, and Jie Shen. An overview of projection methods for incompressible flows. *Comput. Methods Appl. Mech. Engrg.* 195:6011-6045, 2006.

[13] F.H. Harlow and J.E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids.* 8:2182-2189, 1965.

[14] J. Kim and P. Moin. Application of a fractional-step method to incompressible Navier-Stokes equations. *J. Comp. Phys.* 59:308-323, 1985.

[15] S. Ion. Solving Generalized 2D Navier-Stokes Equations. Technical report, Institute of Mathematical Statistics and Applied Mathematics, Department of Applied Mathematics, 2009, http://www.ima.ro/publications/reports/rth_030605.pdf.

[16] O. A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow.* Gordon and Breach, 1969.

[17] R. Temam. *Navier–Stokes Equations.* Elsevier Science, 1984.

[18] K. Yasuda, R.C. Armstrong, and R.E. Cohen. Shear flow properties of concentrated solutions of linear and star branched polystyrenes. *Rheologica Acta.* 20:163-178, 1981.

[19] K.K. Yeleswarapu. *Evaluation of continuum models for characterizing the constitutive behavior of blood.* Ph. D. thesis, University of Pittsburgh, 1996.

[20] Weinan E and J.-G. Liu. Gauge method for viscous incompressible flows. *Comm. Math. Sci.* 1:317-332, 2003.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

APPROXIMATION FORMULAE GENERATED BY EXPONENTIAL FITTING*

Liviu Ixaru †

Abstract

We present the main elements of the exponential fitting technique for building up linear approximation formulae. We cover the two main components of this technique, that is the error analysis and the way in which the coefficients of the new formulae can be determined. We present briefly the recently developed error analysis of Coleman and Ixaru, whose main result is that the error of the formulae based on the exponential fitting (ef, for short) is a sum of *two* Lagrange-like terms, in contrast to the case of the classical formulae where it consists of a *single* term. For application we consider the case of two quadrature formulae (extended Newton-Cotes and Gauss), which are indistinguishable in the frame of the traditional error analysis, to find out that the Gauss rule is more accurate. As for the determination of the coefficients, we show how the *ef* procedure can be applied for deriving formulae of *classical* type. We re-obtain wellknown formulae and also derive some new ones.
MSC: 65D30, 65D32, 65D20, 65L70

keywords: Error formula, exponential fitting, quadrature rules

*Accepted for publication on January 20, 2011.

†`ixaru@theory.nipne.ro` Address:(1) "Horia Hulubei" National Institute of Physics and Nuclear Engineering, Department of Theoretical Physics, P.O. Box MG–6, Platforma Magurele, Bucharest, Romania, and (2) Academy of Romanian Scientists, 54 Splaiul Independentei, 050094, Bucharest, Romania. Paper written with the partial financial support by research project PN 09370102 of Romanian Ministry of Education and Research.

1 Introduction

The exponential fitting (ef for short) is a powerful technique for the construction of approximation formulae for operations on functions with special behaviour, in particular when these are oscillatory functions. The following simple examples are of help for understanding the object of this technique.

First derivative. The simplest approximation for this operation is the popular central difference formula

$$f'(X) \approx \frac{1}{2h}[f(X+h) - f(X-h)]\,, \tag{1.1}$$

which gives good results when f has a smooth variation on $[X-h, X+h]$. Much less known is the fact that when f is an oscillatory function of form

$$f(x) = f_1(x)\sin(\omega x) + f_2(x)\cos(\omega x) \tag{1.2}$$

with smooth f_1 and f_2, then the slightly modified formula

$$f'(X) \approx \frac{\theta}{2h\sin(\theta)}[f(X+h) - f(X-h)], \text{ where } \theta = \omega h\,, \tag{1.3}$$

becomes appropriate; it tends to the former when $\theta \to 0$.
Second derivative. Three-point approximation

$$f''(X) \approx \frac{1}{h^2}\{a_1[f(X+h) + f(X-h)] + a_2 f(X)\}, \tag{1.4}$$

has the constant coefficients $a_1 = 1$, $a_2 = -2$ for the classical case, but the θ dependent coefficients

$$a_1(\theta) = \frac{\theta}{\sin\theta} \qquad \text{and} \qquad a_2(\theta) = \frac{\theta(\sin\theta - 2\cos\theta)}{\sin\theta}$$

for oscillatory functions of form (1.2).
Quadrature. Trapezium rule

$$\int_{X-h}^{X+h} f(z)dz \approx h[a_1 f(X+h) + a_2 f(X-h)]\,, \tag{1.5}$$

has the classical coefficients $a_1 = a_2 = 1$ but

$$a_1(\theta) = a_2(\theta) = \frac{\sin(\theta)}{\theta\cos(\theta)}\,,$$

for functions of form (1.2).

Interpolation. Let $f(X \pm h)$ be given and we want to interpolate at some $x' \in (X-h,\, X+h)$ with the formula

$$f(x') \approx a_- f(X-h) + a_+ f(X+h). \tag{1.6}$$

In the classical case (usual linear interpolation) the coefficients $a_\pm$ depend only on x'; with $t = (x'-X)/h$ these are $a_\pm(x') = (1 \pm t)/2$. However, for treating oscillatory functions they depend also on θ,

$$a_\pm(x',\, \theta) = \frac{\sin[(1 \pm t)\theta]}{\sin(2\theta)}.$$

For other examples see e.g. [1], [2], [3].
The purpose of the exponential fitting procedure is to produce such new forms for the approximation formulae and to evaluate their error. The expression 'exponential fitting' indicates that the procedure has a larger area: in general it covers the cases where f is a linear combination of exponential functions with different frequencies. The oscillatory function (1.2) represents only one of the possible combinations of such functions (two imaginary frequencies $\pm i\omega$ are actually involved in it) but in practice this case is by far the most popular of all. The reason is related to the existence of a tremendously large variety of phenomena governed by oscillatory functions; think, for example, of phenomena involving oscillations, rotations, vibrations, wave propagation, behavior of quantum particles etc.
The paper is organized in two parts. In the first part (Section 2) we consider the error analysis while in the second part (Sections 3-5) we show how the ef technique is used to build up new formulae. In the first part we present briefly the recently developed error analysis of Coleman and Ixaru [4], whose results might be of interest well beyond the area covered by the ef procedure. The main finding of this analysis is that the error of the ef-based approximation formulae is a sum of two Lagrange-like terms, in contrast to the case of the classical formulae (that is where the coefficients are constants) where it consists of a single term. For application we consider the case of two quadrature formulae (extended Newton-Cotes and Gauss), which are indistinguishable in the frame of the existing error analysis, to find out that the Gauss rule is more accurate.
The unusual feature in the second part is that we apply the ef procedure for deriving formulae of classical type. We re-obtain wellknown formulae and also derive some new ones.

2 A two-term Lagrange-like formula of the error

When the value of a function f at $X+h$ is approximated by a truncated Taylor expansion about X, that is by $f_K(X+h)=\sum_{k=0}^{K} h^k f^{(k)}(X)/k!$, the resulting error may be expressed in the Lagrange form

$$E[f]=f(X+h)-f_K(X+h)=\frac{h^{K+1}}{(K+1)!}f^{(K+1)}(\eta)\,, \tag{2.7}$$

for some $\eta\in(X,\,X+h)$, if $f^{(K+1)}(x)$ is continuous on $(X,\,X+h)$. That error may also be written, less usefully, as the formal expansion

$$E[f]=\sum_{k=K+1}^{\infty}\frac{h^k}{k!}f^{(k)}(X)\,. \tag{2.8}$$

Expressions of Lagrange type are also available for the truncation errors of many other classical approximations. For example, the error of the simplest approximation for the first derivative, eq.(1.1), has the Lagrange-like expression

$$E[f]=-\frac{1}{6}h^2 f^{(3)}(\eta) \tag{2.9}$$

where $\eta\in(X-h,\,X+h)$, but a formal expansion as in eq.(2.8) can also be written, whose leading term is

$$lte=-\frac{1}{6}h^2 f^{(3)}(X)\,. \tag{2.10}$$

Note that in both cases considered above the expression of the leading term is the same as that in the Lagrange form except for the interchange of X and η.

Expressions of the leading term of the error can be easily built up for both classical and new forms of the coefficients. Also, since the new coefficients tend to the classical ones when $\theta\to 0$ the same holds true for the leading term of the error. For example, approximation (1.3) has

$$lte=h^2\frac{\sin(\theta)-\theta}{\theta^2\sin(\theta)}[f^{(3)}(X)+\omega^2 f'(X)]\,, \tag{2.11}$$

see [1]. When $\theta\to 0$ (for fixed h this implies $\omega\to 0$ and viceversa) this lte obviously tends to (2.10) which is the same as the whole $E[f]$ of (2.9) except

for the interchange of X and η. This induces the impression that such a link may be more general, in the sense that for any ef-based approximation formula it is sufficient to build up the expression of the *lte* (which, as said, can be derived without difficulty) and to accept simply that this expression represents also the whole error $E[f]$ if X is replaced by some η.

The problem of whether the suggested link can be sustained has been investigated recently by Coleman and Ixaru [4] for linear ef-based approximation formulae on the basis of a theory developed in the book of Ghizzetti and Ossicini [5]. Coleman and Ixaru have shown that $E[f]$ can be written as a sum of two Lagrange-like terms from which only one survives in the limit $\theta \to 0$. The consequence is that the link is justified in the limit case but it does not hold true for big θ, that is, in the region where the ef-based approximation formulae are actually helpful.

The work [5] is concerned with quadrature formulae of the form

$$\int_a^b g(x)f(x)\,dx \approx \sum_{i=1}^{n}\sum_{k=0}^{m-1} A_{ki} f^{(k)}(x_i), \tag{2.12}$$

whose error

$$E[f] = \int_a^b g(x)f(x)\,dx - \sum_{i=1}^{n}\sum_{k=0}^{m-1} A_{ki} f^{(k)}(x_i) \tag{2.13}$$

is such that $E[f] = 0$ when f is a solution of a linear differential equation $Lf = 0$ of order m. It is assumed that

$$a \le x_1 < x_2 < \cdots < x_n \le b$$

and it is convenient to define $x_0 = a$ and $x_{n+1} = b$, to allow for cases where the end-points of the integration interval are not quadrature abscissas.

The operator L has the form

$$L = \sum_{k=0}^{m} w_k(x) D^{m-k}\,,\ x \in [a,\,b]\,,\ \text{ where } \ D^p = \frac{d^p}{dx^p}, \tag{2.14}$$

with $w_0(x) = 1$. Smoothness conditions on the coefficients w_k are specified in [5].

We place the discussion on the case when the coefficients A_{ki} corresponding to the given L are known, to find the expression of the error $E[f]$. The

presence of $g(x)$ in the integrand allows for considerable flexibility. Not only the quadrature formulae are covered by (2.12) but many others, including any known linear approximation formula which is consistent with L of the given form, for operations such as the numerical differentiation, quadrature, solving differential or integral equations, interpolation etc.

For illustration let us examine the approximation formulae listed above from this perspective.

- *First derivative.* The classical and ef-based formulae, eqs.(1.1) and (1.3), respectively, are of form (2.12) for $g(x) \equiv 0$, $n = m = 3$, $x_1 = X - h$, $x_2 = X$, $x_3 = X + h$, $A_{02} = A_{11} = A_{13} = A_{21} = A_{22} = A_{23} = 0$, and $A_{12} = -1$. The other coefficients are $-A_{01} = A_{03} = 1/(2h)$ for the classical formula and $-A_{01} = A_{03} = \theta/[2h\sin(\theta)]$ for the other. Since the classical formula is exact for $f = 1, x, x^2$ i.e. when f satisfies $f^{(3)}(x) = 0$, it follows that $L = D^3$. Likewise, the ef-based formula is exact when $f = 1, \sin(\omega x), \cos(\omega x)$ and since these are three linear independent solutions of differential equation $f^{(3)} + \omega^2 f' = 0$ it results that $L = D(D^2 + \omega^2)$ in this case.

- *Second derivative*, eq.(1.4). This corresponds to (2.12) if we take $g(x) \equiv 0$, $n = 3$, $m = 4$, $x_1 = X - h$, $x_2 = X$, $x_3 = X + h$, $A_{22} = -1$, $A_{21} = 0 = A_{23}$ and $A_{1k} = A_{3k} = 0$ for $k = 1, 2, 3$. The other coefficients are $A_{01} = A_{03} = 1/h^2$, $A_{02} = -2/h^2$ for the classical case, and $A_{01} = A_{03} = a_1(\theta)/h^2$, $A_{02} = a_2(\theta)/h^2$ for the ef-based case. The expressions of the operator are $L = D^4$ and $L = (D^2 + \omega^2)^2$, respectively.

- *Trapezium rule for the quadrature*, eq.(1.5): $g(x) \equiv 1$, $n = m = 2$, $a = x_0 = x_1 = X - h$, $b = x_2 = x_3 = X + h$, $A_{11} = 0 = A_{12}$. The other coefficients depend on the version. They are $A_{01} = A_{02} = h$ for the classical version and $A_{01} = A_{02} = h\sin(\theta)/[\theta\cos(\theta)]$ for the ef-based version. As for the expression of the operator, this is $L = D^2$ and $L = D^2 + \omega^2$, respectively.

- *Two point interpolation*, eq.(1.6): $g(x) \equiv \delta(x - x')$, $n = m = 2$, $a = x_0 = x_1 = X - h$, $b = x_2 = x_3 = X + h$, $A_{11} = A_{12} = 0$. For the classical version we have $A_{01} = a_-(x')$, $A_{02} = a_+(x')$ and $L = D^2$ while $A_{01} = a_-(x', \theta)$, $A_{02} = a_+(x', \theta)$ and $L = D^2 + \omega^2$ for the ef-based version.

The theory of Ghizzetti and Ossicini allows writing $E[f]$ in integral form,

$$E[f] = \int_a^b \Phi(x) L f(x)\, dx, \tag{2.15}$$

where function $\Phi(x)$ is determined piecewise in terms of some other functions, namely

$$\Phi(x) = \phi_i(x) \qquad \text{for} \quad x_i < x < x_{i+1}, \quad i = 0, \ldots, n.$$

The functions $\phi_i(x)$ are constructed as follows. Let K be the resolvent kernel corresponding to the operator L, i.e., $K(x, z)$ is the solution of $Lu(x) = 0$ such that

$$\left[\frac{\partial^k}{\partial x^k} K(x, z)\right]_{x=z} = \delta_{k,m-1}, \tag{2.16}$$

for $k = 0, 1, \ldots, m-1$. This is used to build up function $\phi_0(x)$ by

$$\phi_0(x) = -\int_a^x K(t, x) g(t)\, dt\,. \tag{2.17}$$

Once $K(t, x)$ and $\phi_0(x)$ are known the other ϕ-functions are generated by recurence,

$$\phi_{i+1}(x) = \phi_i(x) + \sum_{k=0}^{m-1} A_{k,i+1} \left[\frac{\partial^k}{\partial t^k} K(t, x)\right]_{t=x_{i+1}}. \tag{2.18}$$

Let us denote

$$T_0 = \int_a^b \Phi(x) dx.$$

The significance of this T_0 is that it represents the front factor in the expression of the leading term of the error. This is easily seen if we take some reference point X on (a, b), and use the Taylor series for $Lf(x)$ around X,

$$Lf(x) = Lf(X) + \frac{(x-X)}{1!}\frac{d}{dx} Lf(x)|_{x=X} + \frac{(x-X)^2}{2!}\frac{d^2}{dx^2} Lf(x)|_{x=X} + \ldots$$

The leading term of the error is integral (2.15) in which only the first term of this expansion is retained:

$$lte = \int_a^b \Phi(x)\, dx \times Lf(X) = T_0\, Lf(X) \tag{2.19}$$

Indeed, the integrals with the next terms will result in higher order contributions, proportional to h, h^2, ...; to see this use the second mean-value theorem. On the other hand, if $\Phi(x)$ does not change the sign on (a, b), then,

assuming that $f \in C^m(a,b)$, the same second mean-value theorem applied on integral (2.15) gives that

$$E[f] = T_0\, Lf(\eta) \tag{2.20}$$

for some $\eta \in (a,\, b)$, such that only in this case one can say that the expressions of *lte* and of $E[f]$ coincide except for the interchange of X and η. However, $\Phi(x)$ may not be of constant sign. For illustration, function $\Phi(x)$ corresponding to the ef-based approximation of the second derivative (1.4) is

$$\Phi(x) = h\frac{\theta(1-|x^*|)\cos[\theta(1-|x^*|)] - \sin[\theta(1-|x^*|)]}{2\theta^2 \sin\theta}\,,$$

where $x^* = (x-X)/h \in [-1,\, 1]$ is associated to $x \in [X-h,\, X+h]$, see [4]. Experimental evidence, also presented in [4], shows that this $\Phi(x)$ is of constant sign if $\theta \in (0,\, \theta_1 \approx 4.4934)$ but it changes the sign for bigger values of θ.

To treat the case when $\Phi(x)$ changes the sign on $(a,\, b)$ we follow [4] to write $\Phi(x) = \Phi_+(x) + \Phi_-(x)$, where

$$\Phi_+(x) := \begin{cases} \Phi(x) & \text{for all } x \text{ such that } \ \Phi(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Phi_-(x) := \begin{cases} \Phi(x) & \text{for all } x \text{ such that } \ \Phi(x) \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

The integral in (2.15) can be expressed as the sum of two integrals,

$$E[f] = \int_a^b \Phi_+(x) Lf(x)\, dx + \int_a^b \Phi_-(x) Lf(x)\, dx\,. \tag{2.21}$$

and, since functions $\Phi_\pm(x)$ are of constant sign, the mean-value theorem can be applied to both integrals to give

$$E[f] = Lf(\eta_+) \int_a^b \Phi_+(x)\, dx + Lf(\eta_-) \int_a^b \Phi_-(x)\, dx, \tag{2.22}$$

for some $\eta_+\,, \eta_- \in (a\,, b)$. With

$$T_\pm = \int_a^b \Phi_\pm(x)\, dx$$

this reads simply

$$E[f] = T_+ Lf(\eta_+) + T_- Lf(\eta_-)\,, \tag{2.23}$$

which is the announced two-term Lagrange-like expression of the error.

To summarize, the error of approximation formula (2.12) admits a Lagrange-like expression whose number of terms depends on the behavior of $\Phi(x)$ on (a, b): it consists of a single term if $\Phi(x)$ does not change the sign but of two terms otherwise. As a matter of fact, no case in which $\Phi(x)$ changes its sign is known to us if L is of the simple form $L = D^m$ (this covers the familiar formulae with constant coefficients such as Simpson, Newton-Cotes or Gauss). In all these cases the error expressions consist in a single Lagrange-like term.
As for new applications, note that the expression of $\Phi(x)$ can be build up in analytic form but the determination of functions $\Phi_\pm(x)$ needs a numerical approach. A final check for the accuracy of the later determination consists in verifying that $T_0 = T_+ + T_-$.
Note also that formula (2.23), whose derivation uses for start the work of Ghizzetti and Ossicini [5], is more general than needed for linear ef-based approximations since it assumes that the coefficients w_k in the operator L may depend on x, while in the exponential fitting these are simply constants.

Application

We consider two ef-based quadrature rules, see also [4].
• Extended Newton-Cotes rule, [7], [2]:

$$\int_a^b f(x)dx = \int_{X-h}^{X+h} f(x)dx \approx h\sum_{n=1}^{N}[a_n^{(0)} f(X + x_n^* h) + h a_n^{(1)} f'(X + x_n^* h)]\,, \tag{2.24}$$

on evenly-spaced abscissas $x_n^* = 2(n-1)/(N-1) - 1 \quad (n = 1, 2, \ldots, N)$. The rule is called extended because it uses the values of f and its derivative, to underline that its structure contrasts that of the versions in current use, where only the values of f are used. As a matter of fact, the Simpson rule is a particular case of the later ($N = 3$); for an adaptation of the Simpson rule to oscillatory integrals see [9] and [10].

• Gauss rule, [11], [2]:

$$\int_a^b f(x)dx = \int_{X-h}^{X+h} f(x)dx \approx h\sum_{n=1}^{N} a_n^{(0)} f(X + x_n^* h)\,. \qquad (2.25)$$

The $2N$ coefficients, that is $a_n^{(0)}$, $a_n^{(1)}$ for the first rule, and $a_n^{(0)}$, x_n^* for the second ($n = 1, \ldots, N$) are determined from the condition that the rule is exact if f satisfies $Lf = 0$ for

$$L = (D^2 + \omega^2)^N = h^{-2N}(D^{*2} + \theta^2)^N.$$

In the last member we have used the dimensionless $x^* = (x - X)/h$ and $D^{*p} = d^p/dx^{*p} = h^p D^p$. Both rules are exact if the integrand f is of form (1.2) where f_1, f_2 are polynomials of degree $N - 1$ or less. The coefficients of each rule depend on θ only.

The *lte* can be expressed either as in (2.19) or in terms of x^*,

$$lte = T_0(D^2 + \omega^2)^N f(X) \; = hT_0^*(D^{*2} + \theta^2)^N f(X)\,,$$

where $T_0^* = h^{-(2N+1)}T_0$. The advantage of the second representation is that it makes the θ dependence more obvious. Indeed, T_0^* depends on θ only, and its expression is formally the same in both rules,

$$T_0^*(\theta) = \frac{2 - \sum_{n=1}^{N} a_n^{(0)}(\theta)}{\theta^{2N}}.$$

As said, the niche for such quadrature rules is that of highly oscillatory integrands, i.e., when big values of θ are involved. Let then keep h fixed and examine the behaviour of *lte* when ω (or θ) tends to infinity. Factor $T_0^*(\theta)$ decreases as θ^{-2N} in both formulae because in each of these the coefficients $a_n^{(0)}(\theta)$ tend to 0 when $\theta \to \infty$. The last factor, $(D^{*2} + \theta^2)^N f(X)$, which is identical in the two, increases as θ^N, see [2], such that the prediction based on the expression of the leading term is that the error should decrease as θ^{-N} in both formulae.

However, the two-term form of the error, eq.(2.23), leads to a different picture. It is convenient to write this equation under the equivalent form

$$E[f](\theta) = h[T_+^*(\theta)(D^{*2} + \theta^2)^N f(\eta_+) + T_-^*(\theta)(D^{*2} + \theta^2)^N f(\eta_-)]\,, \qquad (2.26)$$

where functions $T_\pm^*(\theta)$ satisfy $T_+^*(\theta) \geq 0$, $T_-^*(\theta) \leq 0$, and $T_0^*(\theta) = T_+^*(\theta) + T_-^*(\theta)$. The picture is different because the asymptotic behaviours of $T_0^*(\theta)$,

on one hand, and those of its components $T^*_\pm(\theta)$, on the other, are not necessarily similar.
Indeed, Coleman and Ixaru have shown that for large θ and $N \geq 2$ the sign conserving functions $T^*_\pm(\theta)$ are well described by the approximation

$$T^*_\pm(\theta) \approx \pm c(\theta)\theta^{-(2N-\bar{N})} + c_\pm(\theta)\theta^{-2N}, \tag{2.27}$$

where $\bar{N} \geq 0$, and the functions $c(\theta)$ and $c_\pm(\theta)$, with $c_+(\theta) \neq -c_-(\theta)$, are oscillating between constant limits; think, for example, of functions of the form $c(\theta) = c_+(\theta) = 1 + \cos\theta$ and $c_-(\theta) = -1 + \cos\theta$. Consequently, the errors will damp out as $\theta^{\bar{N}-N}$, and this is slower than the rule θ^{-N} suggested by the behaviour of the *lte*.
Coleman and Ixaru have also shown that the values of $\bar{N}$ are different in the two rules. They are $\bar{N} = N - 2$ for the extended Newton-Cotes rule but $\bar{N} = \lfloor (N-1)/2 \rfloor$ for the Gauss rule, that is $\bar{N} = 0$ for $N = 2$, $\bar{N} = 1$ for $N = 3$, 4, and $\bar{N} = 2$ for $N = 5$, 6 etc. Thus the error damps out like θ^{-2} for the extended Newton-Cotes rule with any $N \geq 2$ but faster and faster when N is increased for the Gauss rule: θ^{-2} for $N = 2$, 3, θ^{-3} for $N = 4$, 5 etc. All these theoretical predictions are nicely confirmed in practice.
We can then conclude that the two-term Lagrange-like expression of the error [4] allows a solid theoretical understanding of the experimental evidence that the ef-based approximation formulae are so well suited for operations on oscillatory functions. It also warns us that the characterization of the error in terms of the *lte*, as largely used in the literature, is often misleading.
The presented application is however rather special: only functions with *one* frequency were involved and also the two selected quadrature rules (extended Newton-Cotes and Gauss) share the property of being defined for any θ. However, such a property is quite exceptional in the family of the ef-based formulae. The typical situation is when some values of θ exist at which the formulae cannot be defined; these are called critical values, see [1], [2]. For example, $\theta_n = (n + 1/2)\pi,\ n = 0, 1, 2, \ldots$ are the critical values for the trapezium rule (1.5) because the coefficients exhibit a factor $\cos(\theta)$ in the denominator. It would be then interesting to see applications on such cases, and also on cases when linear combinations of functions of form (1.2) with *different* frequencies are involved. Situations of the later type also appear in some applications, as in the computation of the normalization constant (two frequencies) or of the Slater integrals (eight frequencies) in quantum mechanics, see, e.g. [2], [12], [13].

It is also important to notice that the approach which has lead to the two-term error formula is restrictive because in the present form it does not give a direct answer for nonlinear approximations. For example, the two-step hybrid algorithm for differential equations in which the phase-fitting technique is used [14], the conditionally P-stable ef-based method for differential equations of form $y'' = f(x, y)$ [15], the ef-based extensions of Runge-Kutta methods as in [16], [6],[17], [18], and references therein, cannot be approached at this moment, and an adaptation is needed.

3 Exponential fitting technique for the construction of the coefficients of approximation formulae

In the previous section we were concerned with the determination of the expression of the error when the coefficients of the approximation formulae are assumed known. The complementary problem, that is the determination of the coefficients, is of equal importance and this is what we consider in this and the next sections in the frame of the ef approach. To fix the ideas we continue to focus our attention on quadrature formulae and, to make the things even simpler, we restrict our concern on the two and three-point formulae with constant coefficients, that is on the classically allowed extensions (in the sense that the frequencies are simply set to zero) of the familiar trapezium and Simpson rules, respectively.

There is a direct practical motivation for such extensions. When approaching problems in natural sciences (physics, chemistry, biology etc.) a succession of numerical operations has to be carried out, where the output from some step is used as input in the next step. For example, let us assume that at some moment we have to solve a second order differential equation, let this be $y'' = f(x, y)$ on $[a, b]$, and just after that we are interested in the evaluation of the integral of y over this interval. If the differential equation is solved by the Runge-Kutta method, then we get not only the values of the solution y at the mesh points but also of its first and second derivative; the second derivative results directly from the expression of function $f(x, y)$. If, alternatively, the equation is solved by a finite difference scheme, then we get the values of y and y'' but not those of y'. As for the calculation of the integral, plenty of versions are presented in the standard literature, see [19] for example, but, surprisingly enough, these typically use only the values of the integrand. Formulae which use also the values of sets of successive

derivatives appeared only recently while formulae in which some of these are missing do not exist although it is clear that all such extended formulae are potentially more accurate whereas they exploit richier input information than that contained in the integrand alone. Expressed in other words, the new formulae provide an advantageous alternative to the standard formulae which, for comparable accuracy, will need repeating the whole computation on finer partitions, thus increasing the computational effort.
We consider the interval $[-h,\, h]$, a partition of this by the meshpoints $x_0 = x_1 = -h$, $x_2 = 0$, $x_3 = x_4 = h$, and a quadrature rule of the form

$$Q[y] = \int_{-h}^{h} y(z)dz \approx \sum_{k=0}^{2} h^k [a_{k1} y^{(k)}(-h) + a_{k2} y^{(k)}(0) + a_{k3} y^{(k)}(h)]\,, \quad (3.28)$$

that is a rule which potentially allows the computation of the integral in terms of the values at the meshpoints of the integrand and of its first and second derivatives. The error of this rule is

$$E[h, \mathbf{a};\, y] \quad (3.29)$$
$$= \int_{-h}^{h} y(z)dz - \sum_{k=0}^{2} h^k [a_{k1} y^{(k)}(-h) + a_{k2} y^{(k)}(0) + a_{k3} y^{(k)}(h)]$$

where the arguments h and $\mathbf{a}$ (this collects all nine coefficients) are explicitly mentioned. The problem consists in the determination of the coefficients such that the error is minimal in a certain sense.
Various particular forms are of interest in terms of the available data. For example, if only the values of y at the three points are known, then we have to impose that all coefficients of the derivatives equal zero, i.e. only a_{01}, a_{02} and a_{02} have to be determined.
Our investigation follows three steps:
1. Find the expressions of $E[h, \mathbf{a};\, y]$ for $y(x) = x^n$, $n = 0, 1, 2, 3, \cdots$.
2. Evaluate the values of the coefficients such that $E[h, \mathbf{a};\, y] = 0$ for as many successive $y(x) = x^n$ as possible (it is assumed that this is actually the way which leads to coefficients which ensure the minimal error for the considered rule) and determine, on this basis, the expression of the operator L, eq.(2.14).
3. Determine the Lagrange-like expression of the error.
Step 1 regards the general form (3.28) while steps 2-3 will treat each particular case separately. We have the following

Lemma 1. *The expressions of* $E[h, \mathbf{a};\, y]$ *for* $y(x) = x^n$, $n = 0, 1, 2, 3, \cdots$ *are of the form*

$$E[h, \mathbf{a};\, x^n] = h^{n+1}\, E_n(\mathbf{a})\,, \tag{3.30}$$

where $E_n(\mathbf{a})$*, called reduced moments, are*

$$\begin{aligned}
E_0(\mathbf{a}) &= 2 - (a_{01} + a_{02} + a_{03})\,,\\
E_1(\mathbf{a}) &= -(-a_{01} + a_{03} + a_{11} + a_{12} + a_{13})\,,\\
E_2(\mathbf{a}) &= \frac{2}{3} - [a_{01} + a_{03} + 2(-a_{11} + a_{13} + a_{21} + a_{22} + a_{23})]\,,\\
E_n(\mathbf{a}) &= -[-a_{01} + a_{03} + n(a_{11} + a_{13}) + n(n-1)(-a_{21} + a_{23})]\,,\\
&\qquad\qquad \text{for odd } n \geq 3\,,\\
E_n(\mathbf{a}) &= \frac{2}{n+1} - [a_{01} + a_{03} + n(-a_{11} + a_{13}) + n(n-1)(a_{21} + a_{23})]\,,\\
&\qquad\qquad \text{for even } n \geq 4\,.
\end{aligned} \tag{3.31}$$

Proof Elementary evaluations on $y(x) = x^n$ give:

$$\begin{aligned}
&y(h) = (-1)^n y(-h) = h^n,\ y(0) = \delta_{n0},\ \text{for any } n \geq 0,\\
&y'(h) = y'(-h) = y'(0) = 0 \text{ for } n = 0\,,\\
&y'(h) = (-1)^{n-1} y'(-h) = n h^{n-1},\ y'(0) = \delta_{n1} \text{ for } n > 0\,,\\
&y''(h) = y''(-h) = y''(0) = 0 \text{ for } n = 0, 1\,,\\
&y''(h) = (-1)^n y''(-h) = n(n-1) h^{n-2},\ y''(0) = 2\delta_{n2} \text{ for } n > 1\,,
\end{aligned}$$

and

$$\int_{-h}^{h} y(z)dz = \begin{cases} \dfrac{2}{n+1} h^{n+1} & \text{for even } n \\ 0 & \text{for odd } n \end{cases}$$

If these are introduced in (3.30) the expressions under eq.(3.31) result directly.
Q. E. D.
Another element of general interest in the subsequent considerations is the resolvent kernel of operator $L = D^m$. We have

Lemma 2. *The resolvent kernel of* $L = D^m$, $m \geq 1$ *is*

$$K(t, z) = \frac{1}{(m-1)!}(t - z)^{m-1}. \tag{3.32}$$

Proof The general solution of the differential equation $D^m\, u(x) = 0$ is the $(m-1)$-th degree polynomial

$$u(x) = \sum_{i=0}^{m-1} a_i x^i\,.$$

Its successive derivatives are

$$\frac{\partial^k}{\partial x^k} u(x) = \sum_{i=0}^{m-(k-1)} (i+1)(i+2)\cdots(i+k) a_{i+k} x^i,\ k = 1, 2, \cdots, m-1\,.$$

The particular solution which satisfies the initial conditions

$$\frac{\partial^k}{\partial x^k} u(x)|_{x=0} = \delta_{k,m-1}$$

is

$$u_p(x) = \frac{1}{(m-1)!} x^{m-1}\,,$$

and the resolvent kernel is this particular solution with argument $x = t - z$. Q. E. D.

For the construction of functions $\phi_i(x)$, eqs.(2.17)-(2.18), the expressions of the integral and successive partial derivatives of the kernel will often be involved:

$$\begin{aligned} I(X,x) &:= \int_X^x K(t,x)dt = -\frac{1}{m!}(X-x)^m\,, \qquad (3.33)\\ K_k(X,x) &:= \frac{\partial^k}{\partial t^k} K(t,x)|_{t=X} = \frac{1}{(m-k-1)!}(X-x)^{m-k-1},\\ &\qquad\qquad k = 0, 1, \cdots, m-1. \end{aligned}$$

Since $x_0 = x_1 = -h$ and $x_3 = x_4 = h$, the function $\Phi(x)$ will have only two piecewise determinations:

$$\Phi(x) = \begin{cases} \phi_1(x) = -I(-h,x) + \sum_{k=0}^{2} h^k a_{k1} K_k(-h,x) & \text{for } -h < x < 0 \\ \\ \phi_2(x) = \phi_1(x) + \sum_{k=0}^{2} h^k a_{k2} K_k(0,x) & \text{for } 0 < x < h \end{cases} \qquad (3.34)$$

In the following we examine two families of quadrature rules of the form (3.28). These are the two-point rules, denoted Q_s^2, where only data at the meshpoints $\pm h$ are accepted, and three-point rules, denoted Q_s^3, where data at all three meshpoints are accepted. Index $s = 1, 2, 3, 4$ identifies versions in the corresponding family in terms of what are actually the data accepted for input:

- Versions Q_1^2 and Q_1^3. Accepted input data: y. These are the trapezium and Simpson rule, respectively.
- Versions Q_2^2 and Q_2^3. Accepted input data: y and y'.
- Versions Q_3^2 and Q_3^3. Accepted input data: y and y''.
- Versions Q_4^2 and Q_4^3. Accepted input data: y, y' and y''.

4 Two-point rules

Remark: Since for these rules we always have $a_{k2} = 0$, $k = 0, 1, 2$, function $\phi_2(x)$ has the same expression as $\phi_1(x)$ and therefore only one determination is active in eq.(3.34): $\Phi(x) = \phi_1(x)$ for $-h < x < h$.
For the trapezium rule Q_1^2 the following result is wellknown, e.g. [19] :

Theorem 1. *The coefficients and the Lagrange-like expression of the error for version Q_1^2 are*

$$a_{01} = a_{03} = 1 \quad \text{and} \quad E[h, \mathbf{a}; y] = -\frac{2}{3} h^3 y''(\eta)\,,$$

for some $\eta \in (-h,\, h)$.

Proof This result can be proved in various ways but here we reconsider the proof again mainly as a first and simple illustration on how the ef-based procedure works.
Since only the values $y(\pm h)$ are accepted, all coefficients in eq.(3.28) are set to zero except for a_{01} and a_{03} which have to be determined. We cover the above mentioned steps 2-3.
Step 2. Since the number of coefficients to be determined is 2 the same is the number of the involved successive reduced moments. For brevity reasons hereinafter the reduced moments will be called simply moments and the parameter $\mathbf{a}$ will be omitted when they are written.

The first two moments are $E_0 = 2 - (a_{01} + a_{03})$, $E_1 = -(-a_{01} + a_{03})$, and the linear system $E_0 = E_1 = 0$ has the solution $a_{01} = a_{03} = 1$. For these coefficients we have $E_2 = -4/3 \neq 0$ such that the error vanishes when $y(x)$ is a first degree polynomial or, equivalently, when $y(x)$ is any solution of the simple second order differential equation $y'' = 0$, that is $L = D^m$ with $m = 2$. As a matter of fact, after the coefficients have been determined a compulsory practice is to check how many next moments are also vanishing. This is because in some situations it may happen that this holds true for a number of such extra moments and therefore the degree of the polynomial may be higher than the number of coefficients. We will meet such a situation for version Q_3^3.
Step 3. For $m = 2$ we have:

$$I(-h, x) = -\frac{1}{2}(h + x)^2,\ K_0(-h, x) = -(h + x),$$

and then

$$\phi_0(x) = \frac{1}{2}(h + x)^2,\ \phi_1(x) = \phi_0(x) + ha_{01}K_0(-h, x) = \frac{1}{2}(x^2 - h^2)\,.$$

$\phi_1(x)$ does not change the sign on $(-h, h)$ (it is negative) and therefore the error is of one-term Lagrange form (2.20) with

$$T_0 = \int_{-h}^{h} \phi_1(x)\, dx = -\frac{2}{3}h^3\,,$$

and this completes the proof.
The following theorem covers the three extensions of the trapezium rule:

Theorem 2. *The extended trapezium rules and the Lagrange-like expression of their errors are as follows:*
- Version Q_2^2 :

$$Q[y] \approx h[y(-h) + y(h)] + \frac{1}{3}h^2[y'(-h) - y'(h)],$$
$$E[h, \mathbf{a};\, y] = \frac{2}{45}h^5 y^{(4)}(\eta)\,; \tag{4.35}$$

- Version Q_3^2 :

$$Q[y] \approx h[y(-h) + y(h)] - \frac{1}{3}h^3[y''(-h) + y''(h)],$$
$$E[h, \mathbf{a};\, y] = \frac{4}{15}h^5 y^{(4)}(\eta)\,; \tag{4.36}$$

- *Version* Q_4^2 *:*

$$Q[y] \approx h[y(-h)+y(h)] + \frac{2}{5}h^2[y'(-h)-y'(h)] + \frac{1}{15}h^3[y''(-h)+y''(h)],$$
$$E[h,\mathbf{a};\,y] = -\frac{2}{1575}h^7 y^{(6)}(\eta)\,, \tag{4.37}$$

for some $\eta \in (-h,h)$*. The value of* η *may vary from one version to another.*

Remarks:

1. The coefficients of the rules Q_2^2 and Q_4^2 are known, [8], but the expressions of their error are new. The rule Q_3^2 is entirely new.

2. One should not remain with the impression that these rules apply only when the integration limits are $-h$ and h. If these are $X-h$ and $X+h$ the coefficients are the same. For example, Q_2^2 reads:

$$\int_{X-h}^{X+h} y(z)dz \approx h[y(X-h)+y(X+h)] + \frac{1}{3}h^2[y'(X-h)-y'(X+h)]$$

Its error is as in eq.(4.35) but $\eta \in (X-h, X+h)$.

Proof This follows the same pattern as for the previous theorem. However, hereinafter we treat explicitly only the rule Q_3^2 which is really new.

Four parameters have to be determined for this version, namely, a_{01}, a_{03}, a_{21} and a_{23}, and the first four moments are $E_0 = 2-(a_{01}+a_{03})$, $E_1 = -(-a_{01}+a_{03})$, $E_2 = 2/3-[a_{01}+a_{03}+2(a_{21}+a_{23})]$, $E_3 = -[-a_{01}+a_{03}+6(-a_{21}+a_{23})]$, see (3.31).

The algebraic system $E_0 = E_1 = E_2 = E_3 = 0$ has the solution

$$a_{01} = a_{03} = 1,\ a_{21} = a_{23} = -\frac{1}{3}\,.$$

With these we get $E_4 = 32/5 \neq 0$ and therefore $L = D^m$ with $m = 4$. Function $\phi_1(x)$ is

$$\begin{aligned}\phi_1(x) &= -I(-h,x) + ha_{01}K_0(-h,x) + h^3 a_{21}K_2(-h,x)\\ &= \frac{1}{4!}(h+x)^4 - \frac{1}{3!}h(h+x)^3 - \frac{1}{3}h^3(h+x)\,.\end{aligned}$$

Separate investigation shows that this $\phi_1(x)$ is positive on $(-h,h)$ and then the error is of form (2.20) with

$$T_0 = \int_{-h}^{h} \phi_1(x)\,dx = \frac{4}{15}h^5\,.$$

Q. E. D.

5 Three-point rules

The following theorem exists:

Theorem 3. *The set of three-point rules and the Lagrange-like expression of their errors are as follows:*
- Version Q_1^3 (standard Simpson rule):

$$Q[y] \approx h[y(-h) + 4y(0) + y(h)]/3,$$
$$E[h, \mathbf{a};\, y] = -\frac{1}{90} h^5 y^{(4)}(\eta)\,; \tag{5.38}$$

- Version Q_2^3:

$$Q[y] \approx \frac{1}{15} h[7y(-h) + 16y(0) + 7y(h)] + \frac{1}{15} h^2[y'(-h) - y'(h)],$$
$$E[h, \mathbf{a};\, y] = \frac{1}{4725} h^7 y^{(6)}(\eta)\,; \tag{5.39}$$

- Version Q_3^3:

$$Q[y] \approx \frac{1}{21} h[5y(-h) + 32y(0) + 5y(h)] - \frac{1}{315} h^3[y''(-h) - 32y''(0) + y''(h)]\,,$$
$$E[h, \mathbf{a};\, y] = \frac{1}{396900} h^9 y^{(8)}(\eta)\,; \tag{5.40}$$

- Version Q_4^3:

$$Q[y] \approx \frac{1}{105} h[41y(-h) + 128y(0) + 41y(h)] + \frac{2}{35} h^2[y'(-h) - y'(h)] + \frac{1}{315} h^3[y''(-h) + 16y(0) + y''(h)],$$
$$E[h, \mathbf{a};\, y] = -\frac{1}{130977000} h^{11} y^{(10)}(\eta)\,, \tag{5.41}$$

for some $\eta \in (-h, h)$. The value of η may vary from one version to another.

Remark: the coefficients of the Simpson rule Q_1^3 and the expression of its error can be found in any standard textbook, e.g., [19]. The coefficients of versions Q_2^3 and Q_4^3 are also known, [8], but the expressions of their error are new. The rule Q_3^3 is entirely new.

Proof Technically, this follows the same steps as for the previous theorem but the volume of calculations is a bit larger. This is due to the fact that the number of involved moments is bigger, on one hand, and that function $\Phi(x)$ now has two piecewise expressions: $\phi_1(x)$ and $\phi_2(x)$. In the following we give details only on the new version Q_3^3.
- Parameters to be determined and their total number N: a_{k1}, a_{k2}, a_{k3}, $k = 0, 2$, i.e., $N = 6$ parameters.
- Expressions of the first N moments: $E_0 = 2 - (a_{01} + a_{02} + a_{03})$, $E_1 = -(-a_{01} + a_{03})$, $E_2 = 2/3 - [a_{01} + a_{03} + 2(a_{21} + a_{22} + a_{23})]$, $E_3 = -[-a_{01} + a_{03} + 6(-a_{21} + a_{23})]$, $E_4 = 2/5 - [a_{01} + a_{03} + 12(a_{21} + a_{23})]$, $E_5 = -[-a_{01} + a_{03} + 20(-a_{21} + a_{23})]$.
- Solution of the algebraic system $E_n = 0$, $n = 0, 1, \cdots, N-1$: $a_{01} = a_{03} = 5/21$, $a_{02} = 32/21$, $a_{21} = a_{23} = -1/315$, $a_{22} = 32/315$.
- Extra checks and the value of m: $E_6 = E_7 = 0$ but $E_8 = 32/315 \neq 0$, therefore $L = D^m$ with $m = 8$. (Notice that the extra check was crucial for this case. Otherwise we might have been tempted to wrongly assign the value $m = 6$.)
- Components of function $\Phi(x)$:

$$\begin{aligned}
\phi_1(x) &= -I(-h, x) + ha_{01}K_0(-h, x) + h^3 a_{21}K_2(-h, x) \\
&= \frac{1}{8!}(h + x)^8 - \frac{5}{21 \cdot 7!}h(h + x)^7 + \frac{1}{315 \cdot 5!}h^3(h + x)^5 \,, \\
\phi_2(x) &= \phi_1(x) + ha_{02}K_0(0, x) + h^3 a_{22}K_2(0, x) \\
&= \phi_1(x) - \frac{32}{21 \cdot 7!}hx^7 - \frac{32}{315 \cdot 5!}h^3 x^5 \,.
\end{aligned}$$

By separate investigation we find that this $\Phi(x)$ is positive on $(-h, h)$ and then the error is of the form (2.20).
- Value of T_0:

$$T_0 = \int_{-h}^{0} \phi_1(x)\, dx + \int_{0}^{h} \phi_2(x)\, dx = \frac{1}{396900}h^9 \,.$$

Q. E. D.

The results listed above for the quadrature rules Q^2 and Q^3 allow drawing some conclusions. First, in all cases the error has the one-term Lagrange form $Ch^{m+1}y^m(\eta)$ where C is some constant (called the error constant) and m is the order of the differential equation $Ly = 0$. Second, we see that, as

expected, the accuracy increases with the number of input data in the corresponding versions. Thus the three-point versions are more accurate than their two-point counterparts (compare the orders) and within each of these two families the order increases with the number of data at each point (one for Q_1^p versions, two for versions Q_2^p and Q_3^p and three for Q_4^p, $p = 2, 3$). Third, and this is a new issue, the results allow answering a question of a different nature: how does the type of data used in versions with *the same* number of input data/point influence the accuracy? This is the case of versions Q_2^p and Q_3^p where the two data are y and y', and y and y'', respectively. For the two-point versions the order is not modified but the error constant is smaller for Q_2^2 and therefore the values of y' are more helpful in increasing the accuracy than those of y''. This is in contrast with the three-point versions where the use of y'' is more advantageous because the corresponding version, that is Q_3^3, has a bigger order than Q_2^2.

Numerical illustration
We compute the integral

$$Q = \int_0^1 e^{5x} \sin 5x \, dx = \frac{1}{10} e^{5x} [\sin(5x) - \cos(5x)]|_0^1 \tag{5.42}$$

by all versions of two and three-point rules. We use $h = 1/2, 1/4, 1/8, 1/16$, $1/32$ and $1/64$, that is with $N = 1, 2, 4, 8, 16$ and 32 two-step intervals. Once the version and h are fixed the integral is computed numerically by that version on each of the N two-step intervals and the individual results are summed. Let denote the value computed this way as $Q^{comput}(h)$. This and its error, $\Delta Q(h) = Q - Q^{comput}(h)$, depend also on the version, of course.

The error $\Delta Q(h)$ behaves as h^m because it is the sum of the N individual errors and $N \cdot h^{m+1} \sim h^m$. As a consequence the ratio of the errors from the same version at $2h$ and h, $\Delta Q(2h)/\Delta Q(h)$, should be around 2^m. Possible deviations from this value are due to the influence of the variation of factor $y^{(m)}$ over four successive intervals of width h. This variation tends to be less and less important when $h \to 0$ and therefore that ratio will tend to the theoretical value in this limit.
We have written a fortran program in double precision and in Table 1 we give the error $\Delta Q(h)$ for the two-point versions. It is seen that, as expected, the decrease of the error with h becomes faster and faster when the number of acccepted data is increased. It is also confirmed the fact that the error

Table 1: Stepwidth dependence of the absolute errors of the results given by the four versions of rule Q^2 for integral (5.42). Notation $a(b)$ means $a \cdot 10^b$.

h	Q_1^2	Q_2^2	Q_3^2	Q_4^2
1/ 2	0.53(+02)	0.11(+02)	0.14(+03)	-0.16(+02)
1/ 4	0.14(+02)	0.30(+01)	0.20(+02)	-0.29(+00)
1/ 8	0.29(+01)	0.24(+00)	0.14(+01)	-0.35(-02)
1/16	0.67(+00)	0.15(-01)	0.93(-01)	-0.50(-04)
1/32	0.17(+00)	0.97(-03)	0.58(-02)	-0.76(-06)
1/64	0.41(-01)	0.61(-04)	0.36(-03)	-0.12(-07)

Table 2: The same as in Table 1 for the versions of rule Q^3. The error from Q_4^3 for $h = 1/64$ is zero within machine accuracy for double precision computations (of approximately 16 decimal figures).

h	Q_1^3	Q_2^3	Q_3^3	Q_4^3
1/ 2	0.52(+00)	0.25(+01)	0.98(-01)	0.14(-01)
1/ 4	-0.70(+00)	0.50(-01)	-0.14(-02)	0.18(-04)
1/ 8	-0.59(-01)	0.60(-03)	-0.80(-05)	0.13(-07)
1/16	-0.38(-02)	0.83(-05)	-0.33(-07)	0.12(-10)
1/32	-0.24(-03)	0.13(-06)	-0.13(-09)	0.11(-13)
1/64	-0.15(-04)	0.20(-08)	-0.52(-12)	0.00(+00)

decrease is similar for versions Q_2^2 and Q_3^2 and that for each stepwidth h the error for the latter is by a factor 6 larger. Table 2 gives the same data for the three-point versions. The errors decrease faster than for the two-point formulae and also, as predicted but in contrast to the two-point case, the errors from Q_3^3 are massively better than from Q_3^3, especially for small h. Table 3 collects the ratios $\Delta Q(2h)/\Delta Q(h)$. The theoretical predictions that these should tend to $4, 16, 16, 64$ for Q^2 versions, and to $16, 64, 256, 1024$ for Q^3 versions when $h \to 0$ are clearly confirmed.

Table 3: The ratio $\Delta Q(2h)/\Delta Q(h)$ for various values of the stepwith h.

h	Q_1^2	Q_2^2	Q_3^2	Q_4^2	Q_1^3	Q_2^3	Q_3^3	Q_4^3
1/ 4	3.9	3.5	7.2	54.7	-0.7	51.1	-67.6	742.3
1/ 8	4.7	12.9	13.7	81.8	12.0	83.4	180.3	1361.9
1/16	4.3	15.4	15.6	70.8	15.3	71.6	242.4	1159.7
1/32	4.1	15.9	15.9	65.8	15.8	66.1	253.0	1081.7
1/64	4.0	16.0	16.0	64.5	16.0	64.5	254.0	–

References

[1] L. Gr. Ixaru, Operations on oscillatory functions, *Comput. Phys. Commun.* 105:1–19, 1997.

[2] L. Gr. Ixaru and G. Vanden Berghe, *Exponential Fitting*, Kluwer, Dordrecht, 2004.

[3] A. Cardone, L. Gr. Ixaru and B. Paternoster, Exponential fitting direct quadrature methods for Volterra integral equations, *Numerical Algorithms* 55: 467–480, 2010.

[4] J. P. Coleman and L. Gr. Ixaru, Truncation errors in exponential fitting for oscillatory problems, *SIAM J. of Numerical Analysis* 44:1441–1465, 2006.

[5] A. Ghizzetti and A. Ossicini, *Quadrature Formulae*, Birkhaüser, Basel, 1970.

[6] J. P. Coleman and S. C. Duxbury, Mixed collocation methods for $y'' = f(x, y)$, *J. Comput. Appl. Math.* 126:47–75, 2000.

[7] J. K. Kim, R. Cools and L. Gr. Ixaru, Quadrature rules using first derivatives for oscillatory integrands, *J. Comput. Appl. Math.* 140:479–497, 2002.

[8] K. J. Kim, R. Cools and L. Gr. Ixaru, Extended quadrature rules for oscillatory integrands, *Appl. Num. Math.* 46:59-73, 2003.

[9] U. T. Ehrenmark, A three-point formula for numerical quadrature of oscillatory integrals with variable frequency , *J. Comput. Appl. Math.* 21:87–99, 1988.

[10] H. De Meyer, J. Vanthournout, G. Vanden Berghe and A. Vanderbauwhede, On the error estimation for the mixed type of interpolation, *J. Comput. Appl. Math.* 32:407–415, 1990.

[11] L. Gr. Ixaru and B. Paternoster, A Gauss quadrature rule for oscillatory integrands, *Comput. Phys. Commun.* 133:177–188, 2001.

[12] J. K. Kim, Quadrature rules for the integration of the product of two oscillatory functions with different frequencies, *Comput. Phys. Commun.* 153:135–144, 2003.

[13] L. Gr. Ixaru, N. S. Scott and M. P. Scott, Fast computation of the Slater integrals, *SIAM J. Sci. Comput.* 28:1252–1274, 2006.

[14] T. E. Simos, Two-step almost P-stable complete in phase methods for the numerical integration of second-order initial-value problems, *Int. J. Comput. Math.* 46:77–85, 1992.

[15] L. Gr. Ixaru and B. Paternoster, A conditionally P-stable fourth-order exponential-fitting method for $y'' = f(x, y)$, *J. Comput. Appl. Math.* 106:87–98, 1999.

[16] B. Paternoster, Runge-Kutta (Nyström) methods for ODEs with periodic solutions based on trigonometric polynomials, *Appl Num. Math.* 28:401–412, 1998.

[17] G. Vanden Berghe, H. De Meyer, M. Van Daele and T. Van Hecke, Exponentially fitted Runge-Kutta methods, *J. Comput. Appl. Math.* 125:107–115, 2000.

[18] H. Van de Vyver, Frequency evaluation for exponentially-fitted Runge-Kutta methods, *J. Comput. Appl. Math.* 184:442–463, 2005.

[19] P. Davis and P. Rabinowitz, *Methods for numerical integration*, 2-nd ed., Academic Press, 1984.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

MATERIAL ELEMENT MODEL FOR EXTRINSIC SEMICONDUCTORS WITH DEFECTS OF DISLOCATION*

Maria Paola Mazzeo† Liliana Restuccia‡

Abstract

In a previous paper we outlined a geometric model for the thermodynamic description of extrinsic semiconductors with defects of dislocation. Applying a geometrization technique, within the rational extended irreversible thermodynamics with internal variables, the dynamical system for *simple material elements* of these media, the expressions of the entropy function and the entropy 1-form were obtained. In this contribution we deepen the study of this geometric model. We give a detailed description of the defective media under consideration and of the dislocation core tensor, we introduce the transformation induced by the process and, applying the closure conditions for the entropy 1-form, we derive the necessary conditions for the existence of the entropy function. These and other results are new in the paper. The derivation of the relevant entropy 1-form is the starting point to introduce an extended thermodynamical phase space.

MSC: 73B20,73B99.

keywords: Extended and rational irreversible thermodynamics, extrinsic semiconductors, solids with defects, dislocations.

*Accepted for publication on January 20, 2011.

†mazzeo@dipmat.unime.it University of Messina, Department of Mathematics, Viale F. Stagno D'Alcontres 31, 98166 Messina, Italy;

‡lrest@dipmat.unime.it University of Messina, Department of Mathematics, Viale F. Stagno D'Alcontres 31, 98166 Messina, Italy;

Introduction

Since in nature there exist no ideal crystals without defects, the aim of this paper is to study the behavior of deformable extrinsic semiconductors with defects of dislocation. The dislocation lines disturb the periodicity of the crystal lattice (see [10] and [22]) and their structure resembles a network of infinitesimally thin channels. The models for defective extrinsic semiconductors may have relevance in several fundamentals technological sectors as electronic microscopy, nanotechnology and technology for integrated circuits VLSI (Very Large Scale Integration).

Semiconductor crystals, as germanium and silicon, are tetravalent elements [11]. In Fig.1_a we have the representation of a germanium crystal that has a behaviour of an insulator at a temperature of $0°K$. But at room temperature, $300°K$ (see Fig.1_b), electrons of the crystal can gain enough thermal energy to jump to the conduction band.

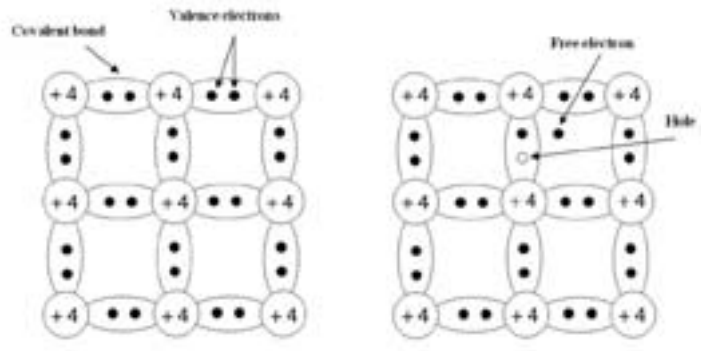


Figure 1: A symbolic representation in 2D of a germanium crystal structure: (a) at $0°K$ and (b) at $300°K$ with a broken covalent bond

To modify the electrical conductivity of an intrinsic semiconductor, impurity atoms adding one electron or one hole are introduced inside the crystal, by means of different techniques of "doping". Using pentavalent impurities, as antimony, a n-type extrinsic semiconductor is obtained, having more free electrons that may flow (see Fig.2_a). By trivalent impurities, as indium, a p-type extrinsic semiconductor crystal is achieved, having more holes that may flow freely (see Fig.2_b).

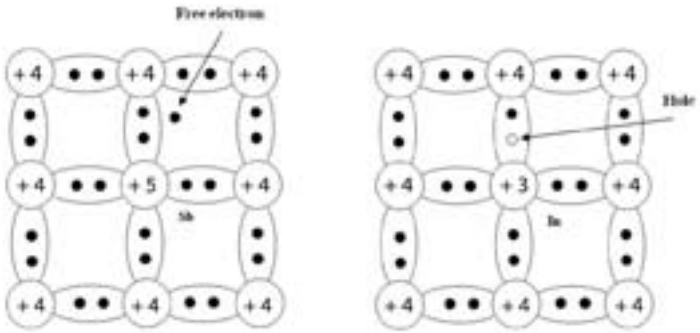


Figure 2: A symbolic representation in 2D of a germanium crystal structure:(a) doped by an atom of a pentavalent impurity (Antimony); (b) doped by an atom of a trivalent impurity (Indium)

In a previous paper [27], in the framework of the rational extended irreversible thermodynamics with internal variables [20], a thermodynamical model for defective extrinsic semiconductors was developed, introducing a dislocation density tensor à la Maruszewski [16] and its flux as internal variables. In [17] taking into account the results obtained in [27], a thermodynamical geometric model was outlined for simple material elements (see [2], [3], [4], [5], [23], [24] and [25]) of these media. The dynamical system and the expressions for the entropy function and the entropy 1-form were obtained. In this paper we deepen the study of this geometric model. In Section 1 we introduce the dislocation core tensor which describes the dislocation lines distribution. In Section 2 we give a detailed thermodynamical description of the defective media under consideration, taking into account the densities and the currents of the free electrons and holes coming from the intrinsic base. Finally, in Section 3 we introduce the transformation induced by the process and, applying the closure conditions for the entropy 1-form, we obtain the necessary conditions for the existence of the entropy function. The derivation of the entropy 1-form is the starting point to introduce a thermodynamical phase space [26]. Furthermore, from the necessary conditions for the existence of the entropy function, constitutive laws can be obtained by a suitable method [7].

In [1], [6], [8], [9], [15] and [18] geometric models for perfect extrinsic

semiconductors, for defective piezoelectric media, for high T_c superconductors of type-II, for porous structures, for polarizable media with internal variables and for deformable dielectrics with a non-Euclidean structure, respectively, were derived in the same geometrized framework.

1 The dislocation core tensor model

In extrinsic semiconductor crystals with defects of dislocation the geometry of the internal structure of these materials can influence the physical fields occurring in the body. These defects, acquired during a process of fabrication, can self propagate, because of changed and favorable surrounding conditions. Thus, they can provoke a premature fracture. The dislocation lines disturb the periodicity of the lattice of the crystal and their structure resembles a network of capillary channels inside the elastic solid (see [11], [16] and [22]). The interatomic distances are not conserved in the direct neighborhood of the dislocation lines in comparison to the distances in the remaining part of the lattice (see Fig.3_a).

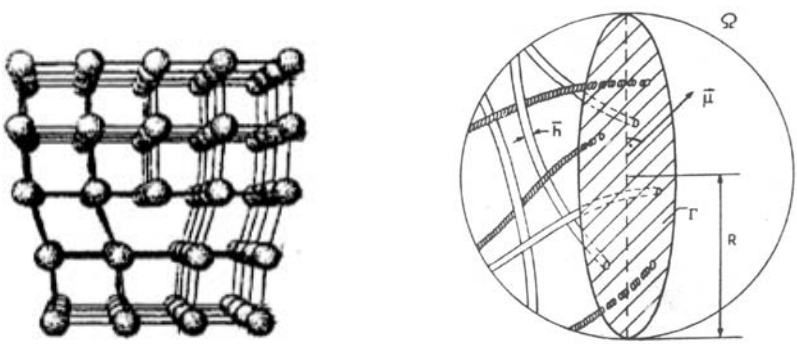


Figure 3: (a) An edge dislocation structure; (b) Characteristics of the pore-core structure ($\bar{h} \ll R$) (after [16])

Moreover, the dislocation lines have their intrinsic orientation, which means, among others, that two dislocations of opposite signs annihilate when lines come close to each other. Their existence should not be omitted in the analysis of such kinetic processes as diffusion of mass or charges, transport of heat, recombination of charge carriers, etc. Thus, we introduce a dislocation

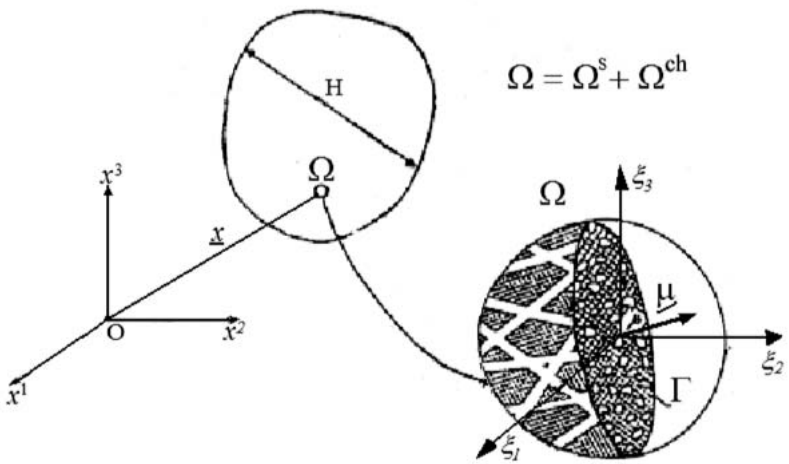


Figure 4: Averaging scheme. Characteristics of a channel structure (see [12])

core tensor à la Maruszewski [16] and its flux in the thermodynamical state space of independent variables for describing these defects. The definition and the introduction of the dislocation core tensor is based (see Fig. 3_b) on a Kubik's geometrical model for porous channels. In [12] Kubik considers a representative elementary sphere volume Ω of a porous structure having capillary tubes, large enough to provide a representation of all the statistical properties of the channel space Ω^{ch} (see Fig. 4). $\Omega = \Omega^s + \Omega^{ch}$, where Ω^s is the solid space. Since all the channels are considered to be interconnected the *effective volume porosity* is completely defined as $f_v = \frac{\Omega^{ch}}{\Omega}$. The analysis is restricted to media which are homogeneous with respect to volume porosity f_v , i.e. f_v remains constant in the medium. To avoid confusion all the microscopic quantities are written with respect to the coordinate system ξ_i , whereas all the macroscopic quantities are written with respect to the coordinate system x_i. Let $\boldsymbol{\alpha}(\xi)$ be any scalar, vectorial or second order tensorial quantity describing a microscopic property of the flux of some physical field flowing through the channel space Ω^{ch} and written with respect to a coordinate system ξ_i. We assume that such quantity is zero in the solid space Ω^s.

The volume averaging procedures give

$$\hat{\boldsymbol{\alpha}}(\mathbf{x}) = \frac{1}{\Omega^{ch}} \int_{\Omega^{ch}} \boldsymbol{\alpha}(\boldsymbol{\xi}) d\Omega, \qquad \bar{\boldsymbol{\alpha}}(\mathbf{x}) = \frac{1}{\Omega} \int_{\Omega^{ch}} \boldsymbol{\alpha}(\boldsymbol{\xi}) d\Omega, \tag{1}$$

where the quantities $\hat{\boldsymbol{\alpha}}(\mathbf{x})$ and $\bar{\boldsymbol{\alpha}}(\mathbf{x})$ (written with respect to the a coordinate system x_i) describe at macroscopic level the same property of the flux of the physical field under consideration. They are averaged quantities on the channel-volume and on the bulk-volume, respectively. Similarly, we define the averaged quantity $\boldsymbol{\alpha}(\boldsymbol{\xi})$ on the channel-area as follows

$$\overset{*}{\boldsymbol{\alpha}}(\mathbf{x}, \boldsymbol{\mu}) = \frac{1}{\Gamma^{ch}} \int_{\Gamma} \boldsymbol{\alpha}(\boldsymbol{\xi}) d\Gamma, \tag{2}$$

where Γ is the central sphere section and Γ^{ch} represents the channel-area of Γ. The orientation of Γ in Ω is given by the normal vector $\boldsymbol{\mu}$. $\Gamma = \Gamma^s + \Gamma^{ch}$, where Γ^s is the solid-area. By definition the quantity $\boldsymbol{\alpha}(\xi)$ is zero on the solid-surface Γ^s. In such a medium, following [12], Maruszewski defines the so called *dislocation tensor*, as follows

$$\bar{\alpha}(\mathbf{x})_i = \mathcal{R}_{ij}(\mathbf{x}, \boldsymbol{\mu}) \overset{*}{\alpha}_j (\mathbf{x}, \boldsymbol{\mu}). \tag{3}$$

Eq. (3) gives a linear mapping between the averaged quantity on the bulk-volume $\bar{\boldsymbol{\alpha}}(\mathbf{x})$ and the average of the same quantity on the channel-area $\overset{*}{\boldsymbol{\alpha}}(\mathbf{x}, \boldsymbol{\mu})$. In [12] Kubik gives an interpretation of $\mathcal{R}_{ij}$ considering the flux of a quantity $\bar{\boldsymbol{\alpha}}(\mathbf{x})$ on a bulk-volume as a superposition of three unidimensional fluxes (along three mutually perpendicular channels) having average values $\overset{*}{\alpha}_i (\mathbf{x}, \boldsymbol{\mu})$ on the orthogonal section areas of these channels. In [16] a new tensor, that refers $\mathcal{R}_{ij}$ to the central sphere section Γ, is defined in the following way

$$\mathcal{R}_{ij}(\mathbf{x}, \boldsymbol{\mu}) = \Gamma a_{ij}(\mathbf{x}, \boldsymbol{\mu}).$$

a_{ij} is called *dislocation core tensor* and its unit is m^{-2}. The components of a_{ij} form a kind of continuous representation of the number of dislocations which cross the surface Γ. Investigations show that a_{ij} is also dependent on time.

2 Governing equations for extrinsic semiconductors with defects of dislocation

Now, we introduce a thermodynamic model for a defective extrinsic semiconductor developed in [27] by one of us (L.R.) in the framework of Thermodynamics of irreversible processes with internal variables. In this paper, we deepen the thermodynamic description of this medium taking into account the densities and currents of the free electrons and holes that come from the intrinsic base of the semiconductor. Furthermore, we derive a set of constitutive relations. We use the standard Cartesian tensor notation in rectangular coordinate systems. We refer the motion of our material system to a current Eulerian configuration $\mathcal{K}_t$. We assume that in defective, extrinsic, thermoelastic semiconductors the following fields interact with each other: *the elastic field* described by the total stress tensor T_{ij} and the small-strain tensor ε_{ij}; *the thermal field* described by the temperature θ, its gradient and the heat flux q_i; *the electromagnetic field* described by the electromotive intensity $\mathcal{E}_i$ (that represents, in the Galilean approximation, the electric field referred to an element of the matter at time t, i.e. to the so called comoving frame $\mathcal{K}_c$) and the magnetic induction B_i per unit volume; *the charge carrier fields* described by the densities of electrons n and holes p, their gradients and their fluxes j_i^n and j_i^p; *the dislocation field* described by the dislocation core tensor a_{ij}, its gradient and the dislocation flux $\mathcal{V}_{ijk}$.

The independent variables are represented by the set

$$C = \{\varepsilon_{ij}, \mathcal{E}_i, B_i, n, p, \theta, a_{ij}, \mathcal{V}_{ijk}, j_i^n, j_i^p, q_i, n_{,i}, p_{,i}, \theta_{,i}, a_{ij,k}\}. \tag{4}$$

All the processes, occurring in the considered body, are governed by the following laws:

Maxwell's equations having the form:

$$\varepsilon_{ijk}E_{k,j} + \frac{\partial B_i}{\partial t} = 0, \quad D_{i,i} - \rho Z = 0, \tag{5}$$

$$\varepsilon_{ijk}H_{k,j} - j_i^Z - \frac{\partial D_i}{\partial t} = 0, \quad B_{i,i} = 0, \tag{6}$$

where $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{H}$ denote the electric field, the magnetic induction, the electric displacement and the magnetic field, respectively. Furthermore,

$$H_i = \frac{1}{\mu_0}B_i, \qquad E_i = \frac{1}{\varepsilon_0}\left(D_i - P_i\right), \tag{7}$$

where ε_0 and μ_0 denote the permittivity and permeability of vacuum and $\mathbf{P}$ is the polarization per unit volume. The magnetization $\mathbf{M}$ is assumed to be zero.

The total charge density Z and the density of the total current $\mathbf{j}^Z$ are defined as follows:

$$Z = n + \bar{n} + p + \bar{p},$$

$$j_i^Z = \rho n v_i^n + \rho \bar{n} v_i^{\bar{n}} + \rho p v_i^p + \rho \bar{p} v_i^{\bar{p}} = \rho Z v_i + j_i^n + j_i^{\bar{n}} + j_i^p + j_i^{\bar{p}},$$

where $n < 0$, $\bar{n} < 0$, $p > 0$, $\bar{p} > 0$, $j_i^n = \rho n(v_i^n - v_i)$, $j_i^{\bar{n}} = 0$ (being $v_i^{\bar{n}} = v_i$), $j_i^p = \rho p(v_i^p - v_i)$, $j_i^{\bar{p}} = 0$ (being $v_i^{\bar{p}} = v_i$), ρ denotes the mass density, v_i are the components of the barycentric velocity of the body, v_i^n, $v_i^{\bar{n}}$, v_i^p, $v_i^{\bar{p}}$, are the velocities of the electric charges n, $\bar{n}$, p, $\bar{p}$, respectively, and j_i^n, $j_i^{\bar{n}}$, j_i^p, $j_i^{\bar{p}}$ their conduction currents, i.e. the electric currents due to the relative motion of the electric charges respect to the barycentric motion of the body. $\rho Z v_i$ is the electric current due to the convection.

In particular, n is a total negative electric charge density coming from: the density of free electrons created doping the semiconductor by pentavalent impurities, denoted by N (see Fig. 2_a), and the density of free electrons coming from the intrinsic base of the semiconductor, denoted by n^* (see Fig. 1_b). $\bar{n}$ is the charge density of the fixed and negative ionized atoms of doping tetravalent impurities, having velocity $\mathbf{v}$ (i.e. they are comoving with the body). Thus, we have the following charge conservation laws

$$\rho \dot{N} + j_{i,i}^N = g^N, \qquad \rho \dot{n}^* + j_{i,i}^{n^*} = g^{n^*}, \qquad \rho \dot{\bar{n}} = \bar{g}^n \qquad \rho \dot{n} + j_{i,i}^n = g^n, \quad (8)$$

where the superimposed dot denotes the material derivative, ρ is the mass density, $n = N + \overset{*}{n}$, $j_{i,i}^n = j_{i,i}^N + j_{i,i}^{n^*}$, $j_{i,i}^{\bar{n}} = 0$ and $g^n = g^N + g^{n^*}$.

Similarly, p are the positive electric charge density coming from: the density of holes created doping the semiconductor by tetravalent impurities, denoted by P (see Fig. 2_b), and the density of holes coming from the intrinsic base of the semiconductor denoted by p^* (see Fig.1_b). $\bar{p}$ is the charge density of the fixed and positive ionized atoms of doping pentavalent impurities, having velocity $\mathbf{v}$ (i.e. they are comoving with the body). Thus, we have the following charge conservation laws

$$\rho \dot{P} + j_{i,i}^P = g^P, \qquad \rho \dot{p}^* + j_{i,i}^{p^*} = g^{p^*}, \qquad \rho \dot{\bar{p}} = \bar{g}^p \qquad \rho \dot{p} + j_{i,i}^p = g^p, \quad (9)$$

where $p = P + p^*$ $j_{i,i}^p = j_{i,i}^P + j_{i,i}^{p^*}$, $j_{i,i}^{\bar{p}} = 0$ and $g^p = g^P + g^{p^*}$. Furthermore, we assume that the concentrations $\bar{n}$ and $\bar{p}$ are practically constant. Hence,

$$\dot{\bar{n}} = \dot{\bar{p}} = 0 \qquad \text{and} \qquad \bar{g}^n = \bar{g}^p = 0. \tag{10}$$

g^n and g^p describe the recombination of electrons and holes and satisfy the equation

$$g^n + g^p = 0. \tag{11}$$

Also, we have

the evolution equations for the *electron*, *hole* and *heat fluxes* having the form:

$$\dot{j}_i^n = J_i^n(C), \qquad \dot{j}_i^p = J_i^p(C), \qquad \dot{q}_i = Q_i(C), \tag{12}$$

where $\mathbf{J}^n$, $\mathbf{J}^p$ and $\mathbf{Q}$ are the electron, hole and heat flux sources;

the continuity equation:

$$\dot{\rho} + \rho v_{i,i} = 0, \tag{13}$$

where the mass charge carriers have been neglected compared to ρ (see the final remark about it in Section 3);

the momentum balance:

$$\rho \dot{v}_i - T_{ji,j} - \rho Z \mathcal{E}_i - \varepsilon_{ijk} \left(j_j^n + j_j^p + \overset{\triangle}{P}_j \right) B_k - P_j \mathcal{E}_{i,j} - f_i = 0, \tag{14}$$

where

$$\overset{\triangle}{P}_i = \dot{P}_i + P_i v_{k,k} - P_k v_{i,k}, \quad \mathcal{E}_i = E_i + \varepsilon_{ijk} v_j B_k, \tag{15}$$

T_{ij} denotes the total stress tensor and f_i is the body force;

the momentum of momentum balance:

$$\varepsilon_{ijk} T_{jk} + c_i = 0. \tag{16}$$

In [27] it was demonstrated that the couple c_i for unit volume is vanishing, so that the stress tensor T_{ij} is symmetric;

the internal energy balance:

$$\rho \dot{e} - T_{ji} v_{i,j} - \left(j_j^n + j_j^p \right) \mathcal{E}_j - \rho \mathcal{E}_i \dot{\mathcal{P}}_i + q_{i,i} - \rho r = 0, \tag{17}$$

where v_i are the components of the barycentric velocity of the body, e is the internal energy density, r is the heat source distribution per unit volume, $P_i = \rho \mathcal{P}_i$ and $\mathrm{v}_{i,j}$ is the velocity gradient given by

$$v_{i,j} = L_{ij} = \dot{F}_{ik} \left(F_{kj} \right)^{-1},$$

where F_{ij} denotes the deformation gradient;

the *evolution equations for the dislocation density and the dislocation flux*:

$$\dot{a}_{ij} + \mathcal{V}_{ijk,k} - A_{ij}(C) = 0, \qquad \dot{\mathcal{V}}_{ijk} - \mathcal{V}_{ijk}(C) = 0, \tag{18}$$

where A_{ij} and $\mathcal{V}_{ijk}$ are the dislocation density and the dislocation flux sources.

All the admissible solutions of the proposed evolution equations should be restricted by the following *entropy inequality*:

$$\rho\dot{S} + J_{Sk,k} - \frac{\rho r}{\theta} \geq 0, \tag{19}$$

where S denotes the entropy per unit mass and $\mathbf{J}_S$ is the entropy flux associated with the fields of the set C. $\mathbf{J}_S$ is defined by

$$\mathbf{J}_S = \frac{1}{\theta}\mathbf{q} + \mathbf{k}, \tag{20}$$

with $\mathbf{k}$ an additional term called *extra entropy flux density.*

In [27] in order to close the balance equation system the entropy inequality was analyzed by Liu's theorem [14]. For the entropy extra flux $\mathbf{k}$ the following form was obtained

$$k_k = -q_k + \mu^n j_k^n + \mu^p j_k^p + \pi_{ij}\mathcal{V}_{ijk} + \rho v_k \psi, \tag{21}$$

where $\mu^n \equiv \frac{\partial\psi}{\partial n}$, $\mu^p \equiv \frac{\partial\psi}{\partial p}$ and $\pi_{ij} \equiv \rho\frac{\partial\psi}{\partial a_{ij}}$ are thermodynamical potentials, with $\psi = e - \theta S - \mathcal{E}_i\mathcal{P}_i$ the free energy density. Using Smith's theorem [28], in the case of defective semiconductors only of n-type, isotropic polynomial representations, satisfying the objectivity and material frame indifference principles (see [19] and [21]), were derived for the constitutive functions, where the following forms were assumed for the quantities responsible for the dislocation field

$$a_{ij} = a\delta_{ij}, \quad A_{ij} = A\delta_{ij}, \quad \mathcal{V}_{ijk} = \mathcal{V}_k\delta_{ij},\ V_{ijk} = V_k\delta_{ij}. \tag{22}$$

In this paper, using the results obtained in [27] by Liu's theorem and, applying Smith's theorem, we derive the constitutive relations for n and p type semiconductors in the same above assumptions (22) for the dislocation field. In particular, we have

$$\begin{aligned} T_{ij} \;=\; & \beta_\tau^1\delta_{ij} + \beta_\tau^2\varepsilon_{ij} + \beta_\tau^3\varepsilon_{ik}\varepsilon_{kj} + \beta_\tau^4\mathcal{E}_i\mathcal{E}_j + \beta_\tau^5(\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k + \varepsilon_{ik}\mathcal{E}_j\mathcal{E}_k) + \\ & + \beta_\tau^6(\varepsilon_{jk}\varepsilon_{ks}\mathcal{E}_i\mathcal{E}_s + \varepsilon_{ik}\varepsilon_{ks}\mathcal{E}_s\mathcal{E}_j), \end{aligned} \tag{23}$$

$$P_i = (\beta^1_{\mathfrak{P}}\delta_{ik} + \beta^2_{\mathfrak{P}}\varepsilon_{ik} + \beta^3_{\mathfrak{P}}\varepsilon_{ij}\varepsilon_{jk})\mathcal{E}_k, \tag{24}$$

$$\mu^n = \beta^1_n n + \beta^2_n p + \beta^3_n a + \beta^4_n \theta + \beta^5_n \mathcal{E}_k\mathcal{E}_k + \\ +(\beta^6_n\delta_{ij} + \beta^7_n\varepsilon_{ij} + \beta^8_n\varepsilon_{jk}\varepsilon_{ki} + \beta^9_n\mathcal{E}_i\mathcal{E}_j + \beta^{10}_n\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k)\varepsilon_{ij}, \tag{25}$$

$$\mu^p = \beta^1_p n + \beta^2_p p + \beta^3_p a + \beta^4_p \theta + \beta^5_p \mathcal{E}_k\mathcal{E}_k + \\ +(\beta^6_p\delta_{ij} + \beta^7_p\varepsilon_{ij} + \beta^8_p\varepsilon_{jk}\varepsilon_{ki} + \beta^9_p\mathcal{E}_i\mathcal{E}_j + \beta^{10}_p\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k)\varepsilon_{ij}, \tag{26}$$

$$\pi = \beta^1_\pi n + \beta^2_\pi p + \beta^3_\pi a + \beta^4_\pi \theta + \beta^5_p \mathcal{E}_k\mathcal{E}_k + \\ +(\beta^6_\pi\delta_{ij} + \beta^7_\pi\varepsilon_{ij} + \beta^8_\pi\varepsilon_{jk}\varepsilon_{ki} + \beta^9_\pi\mathcal{E}_i\mathcal{E}_j + \beta^{10}_\pi\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k)\varepsilon_{ij}, \tag{27}$$

$$g^n = \beta^1_{g^n} n + \beta^2_{g^n} p + \beta^3_{g^n} a + \beta^4_{g^n} \theta + \beta^5_{g^n} \mathcal{E}_k\mathcal{E}_k + \\ +(\beta^6_{g^n}\delta_{ij} + \beta^7_{g^n}\varepsilon_{ij} + \beta^8_{g^n}\varepsilon_{jk}\varepsilon_{ki} + \beta^9_{g^n}\mathcal{E}_i\mathcal{E}_j + \beta^{10}_{g^n}\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k)\varepsilon_{ij}, \tag{28}$$

and $g^p = -g^n$, where $\beta^\alpha_\tau,\ \beta^\gamma_{\mathfrak{P}},\ \beta^\varepsilon_n,\ \beta^\varepsilon_p,\ \beta^\varepsilon_\pi,\ \beta^\varepsilon_{g^n}$ $(\alpha = 1, 2, ..., 6,$ $\gamma = 1, 2, 3, \quad \varepsilon = 1, 2, ..., 10)$ can be functions of the following invariants

$$n,\ p,\ \theta, a,\ \mathcal{E}_i\mathcal{E}_i,\ \varepsilon_{kk},\ \varepsilon_{ij}\varepsilon_{ij},\ \varepsilon_{ij}\varepsilon_{jk}\varepsilon_{ki},\ \varepsilon_{ij}\mathcal{E}_i\mathcal{E}_j,\ \varepsilon_{ij}\varepsilon_{jk}\mathcal{E}_i\mathcal{E}_k. \tag{29}$$

Furthermore, we have obtained the following approximated expressions for the evolution equations for the dislocation density, dislocation, electron, hole and heat fluxes

$$\dot{a} + \mathcal{V}_{k,k} = \delta^1_a\varepsilon_{kk} + \delta^2_a n + \delta^3_a p + \delta^4_a\theta + \delta^5_a a + \delta^6_a\mathcal{E}_i + \delta^7_a a_{,i} + \delta^8_a n_{,i} + \\ +\delta^9_a p_{,i} + \delta^{10}_a\theta_{,i} + \delta^{11}_a\mathcal{V}_i + \delta^{12}_a j^n_i + \delta^{13}_a j^p_i + \delta^{14}_a q_i, \tag{30}$$

$$\dot{\mathcal{V}}_k = \delta^1_\upsilon\mathcal{E}_k + \delta^2_\upsilon a_{,k} + \delta^3_\upsilon n_{,k} + \delta^4_\upsilon p_{,k} + \delta^5_\upsilon\theta_{,k} + \delta^6_\upsilon\mathcal{V}_k + \delta^7_\upsilon j^n_k + \delta^8_\upsilon j^p_k + \delta^9_\upsilon q_k, \tag{31}$$

$$\dot{j}^n_k = \delta^1_n\mathcal{E}_k + \delta^2_n a_{,k} + \delta^3_n n_{,k} + \delta^4_n p_{,k} + \delta^5_n\theta_{,k} + \delta^6_n\mathcal{V}_k + \delta^7_n j^n_k + \delta^8_n j^p_k + \delta^9_n q_k, \tag{32}$$

$$\dot{j}^p_k = \delta^1_p\mathcal{E}_k + \delta^2_p a_{,k} + \delta^3_p n_{,k} + \delta^4_p p_{,k} + \delta^5_p\theta_{,k} + \delta^6_p\mathcal{V}_k + \delta^7_p j^n_k + \delta^8_p j^p_k + \delta^9_p q_k, \tag{33}$$

$$\dot{q}_k = \delta^1_q\mathcal{E}_k + \delta^2_q a_{,k} + \delta^3_q n_{,k} + \delta^4_q p_{,k} + \delta^5_q\theta_{,k} + \delta^6_q\mathcal{V}_k + \delta^7_q j^n_k + \delta^8_q j^p_k + \delta^9_q q_k, \tag{34}$$

where $\delta^\zeta_a,\ \delta^\eta_\upsilon,\ \delta^\eta_n,\ \delta^\eta_p,\ \delta^\eta_q$ $(\zeta = 1, 2, ..., 14,\ \eta = 1, 2, ..., 9)$ can depend on invariants built on the set C (see eq. (4)). The laws (23) - (34) are very general, but it is possible to treat special problems describing the physical reality in several situations by some simplifications.

3 A geometric model for extrinsic semiconductors with defects of dislocation

In this Section, following [2], [3], [4], [5], [23], [24] and [25], we deepen the study of the geometric model for the thermodynamics of extrinsic defective semiconductors outlined in [17], where the dynamical system for *simple material elements* of these media, the expressions of the entropy function and the entropy 1-form were obtained. In particular, we derive the transformation induced by the process and, applying the closure conditions for the entropy 1-form, the necessary conditions for the existence of the entropy function.

Consider a material element and define the state space at time t as the set B_t of all the state variables which "fit" the configuration of the element at time t. B_t is assumed to have the structure of a finite dimensional manifold. The "total state space" is the disjoint union $\mathcal{B} = \bigcup_t \{t\} \times B_t$ with a given natural structure of fibre bundle over $\mathbb{R}$ where time flows (see [4] and [5]). We call it the thermodynamic fiber bundle. We consider the case in which the instantaneous state space B_t does not vary in time (i.e. there is an abstract space B such that $B_t \simeq B$ for all instants of time t) and the state space $\mathcal{B}$ has the topology of the Cartesian product $\mathcal{B} \simeq \mathbb{R} \times B$. Furthermore, we consider an abstract space of *processes* (see [2], [3], [23], [24] and [25]) i.e. a set Π of functions

$$P_t^i : [0, t] \to \mathcal{G},$$

where $[0, t]$ is any time interval, the space $\mathcal{G}$ being a suitable target space defined by the problem under consideration, i a label ranging in an unspecified index set for all allowed processes and $t \in \mathbb{R}$ the so called *duration* of the process. For the given state space B we suppose that the set Π is such that the following statements hold:

- $\exists D : P_t^i \in \Pi \to D(P_t^i) \equiv D_t^i \in \mathcal{P}(B)$.
  D is the *domain function*, D_t^i is the *domain* of the i-th process (of duration t) and $\mathcal{P}(B)$ is the set of all the subsets of B;

- $\exists R : P_t^i \in \Pi \to R(P_t^i) \equiv R_t^i \in \mathcal{P}(B)$. R is the *range function* and R_t^i is called the *range* of the i-th process (of duration t);

- considering the restrictions

$$P_t^i = P_t^i \mid_{[0,\tau]} \qquad (\tau \leq t) \tag{35}$$

new processes, called *restricted processes*, are obtained and they satisfy the following condition:

$$\forall \tau < t \qquad D(P_t^i) \subseteq D(P_\tau^i). \tag{36}$$

Incidentally, this implies that $\bigcap_{\tau=0}^{t} D(P_\tau^i) = D(P_t^i)$, where t is the maximal duration.

Then, a continuous function is defined

$$\chi : (t, P_t^i) \in \mathbb{R} \times \Pi \to \rho_t^i \in C^0(B, B) \tag{37}$$

with

$$\rho_t^i : b \in D_t^i \subseteq B \to \rho_t^i(b) = b_t \in R_t^i \subseteq B, \tag{38}$$

so that for any instant of time t and for any process $P_t^i \in \Pi$ a continuous mapping, ρ_t^i, called *transformation induced by the process* is generated, which gives point by point a correspondence between the initial state b and the final state $\rho_t^i(b) = b_t$.
Moreover, if P_t^i and P_s^j are processes such that $D_s^j \cap R_t^i \neq \emptyset$, then the function

$$(P_s^j \circ P_t^i) : [0, t+s] \to \mathcal{G}$$

defined by

$$(P_s^j \circ P_t^i)(\tau) = \begin{cases} P_t^i(\tau), & \tau \in [0, t] \\ P_s^j(\tau - t), & \tau \in]t, t+s] \end{cases} \tag{39}$$

is a process having the following domain

$$D(P_s^j \circ P_t^i) = (\rho_t^i)^{-1}(D_s^j \cap R_t^i) \tag{40}$$

and, $\forall b \in D(P_s^j \circ P_t^i)$, the transformation induced by the process $P_s^j \circ P_t^i$ is defined by

$$\rho_{t+s}^{ij}(b) = [\rho_s^j(\rho_t^i)(b)]. \tag{41}$$

Now, we introduce a function of time

$$\lambda_b^i(\tau) = \begin{cases} b & \text{if } \tau = 0 \\ \rho_t^i(b) & \text{if } \tau \in]0, t] \end{cases} \qquad \text{with} \quad b \in D_t^i \tag{42}$$

such that the transformation for the medium is a function

$$\delta : \tau \in \mathbb{R} \longrightarrow \delta(\tau) = (\tau, \lambda_b^i(\tau)) \in \mathbb{R} \times B. \tag{43}$$

With these positions the transformation is interpreted as a curve δ in the union of all the state spaces such that it intersects the instantaneous state space just once.
Now, we assume that the behavior of extrinsic thermoelastic semiconductors with defects of dislocation is described by the following state variables

$$\mathbf{C} = \{F_{ij}, D_i, B_i, n, p, e, a_{ij}, \mathcal{V}_{ijk}, j_i^n, j_i^p, q_i, n_{,i}, p_{,i}, \theta_{,i}, a_{ij,k}\},$$

where we have taken into consideration the gradient of deformation F_{ij} instead of the strain tensor, following standard methods. The full state space is then

$$B = Lin(\mathfrak{V})\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus\mathbb{R}\oplus\mathbb{R}\oplus\mathbb{R}\oplus\mathcal{W}_1\oplus\mathcal{W}_2\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus\mathfrak{V}\oplus Lin(\mathcal{W}_1),$$

where $\mathfrak{V} \simeq \mathbb{R}^3$, $\mathcal{W}_1$ and $\mathcal{W}_2$ are vector spaces accounting for the internal variables $\mathbf{a}$ and $\boldsymbol{\mathcal{V}}$, respectively.
Moreover, applying the usual method, we assume that for each pair (P_t^i, b) the following dynamical system holds (see [2], [3], [23], [24] and [25])

$$\begin{cases} \dot{\mathbf{F}} &= \mathbf{L}\mathbf{F} \\ \dot{\mathbf{D}} &= \boldsymbol{\mathcal{H}} \\ \dot{\mathbf{B}} &= \boldsymbol{\Xi} \\ \rho\dot{n} &= G^n \\ \rho\dot{p} &= G^p \\ \rho\dot{e} &= \mathbf{T}\cdot\mathbf{L} + h \\ \dot{\mathbf{a}} &= \boldsymbol{\gamma} \\ \dot{\boldsymbol{\mathcal{V}}} &= \mathbf{V} \\ \dot{\mathbf{j}}^n &= \mathbf{J}^n \\ \dot{\mathbf{j}}^p &= \mathbf{J}^p \\ \dot{\mathbf{q}} &= \mathbf{Q} \\ \dot{\nabla n} &= \boldsymbol{\mathcal{N}} \\ \dot{\nabla p} &= \boldsymbol{\mathcal{P}} \\ \dot{\nabla\theta} &= \boldsymbol{\Theta} \\ \dot{\nabla\mathbf{a}} &= \boldsymbol{\Gamma}, \end{cases} \tag{44}$$

where

$$\begin{aligned} &\mathcal{H}_i = \varepsilon_{ijk}H_{k,j} - (j_i^n + j_i^p) - \rho Z v_i, \quad \Xi_i = -\varepsilon_{ijk}E_{k,j}, \quad G^n = g^n - j^n_{i,i}, \\ &G^p = g^p - j^p_{i,i}, \quad h = (j_i^n + j_i^p)\mathcal{E}_i - \frac{\dot{\rho}}{\rho}\mathcal{E}_i P_i + \mathcal{E}_i\dot{P}_i - q_{i,i} + \rho r, \\ &\gamma_{ij} = -\mathcal{V}_{ijk,k} + A_{ij}, \end{aligned} \tag{45}$$

(see eqn.s $(5)_1$, $(6)_1$, $(8)_4$, $(9)_4$, (17), (18) and (23) - (34)) and δ is defined by eq. (43).
The constitutive functions θ, $\mathbf{T}$, $\mathbf{P}$, $\mathbf{J}^n$, $\mathbf{J}^p$, $\mathbf{Q}$, $\mathbf{A}$, $\mathbf{V}$, g^n and g^p are defined in the following way

$$\theta : \mathbb{R} \times B \to \mathbb{R}^{++},\ \mathbf{T} : \mathbb{R} \times B \to Sym(\mathfrak{V}),\ \mathbf{P} : \mathbb{R} \times B \to \mathfrak{V},$$

$$\mathbf{J^n} : \mathbb{R} \times B \to \mathfrak{V},\ \mathbf{J^p} : \mathbb{R} \times B \to \mathfrak{V},\ \mathbf{Q} : \mathbb{R} \times B \to \mathfrak{V},\ \mathbf{A} : \mathbb{R} \times B \to \mathcal{W}_1,$$

$$\mathbf{V} : \mathbb{R} \times B \to \mathcal{W}_2, \quad g^n : \mathbb{R} \times B \to \mathbb{R}, \qquad g^p : \mathbb{R} \times B \to \mathbb{R}.$$

The set $(B, \Pi, \theta, \mathbf{T}, \mathbf{P}, \mathbf{J^n}, \mathbf{J^p}, \mathbf{A}, \mathbf{V}, \mathbf{Q}, g^n, g^p)$ defines the simple material element for defective extrinsic semiconductors (see [24]).

Following standard procedures (see [3], [4] and [5]), in this geometrical structure we are able to introduce an action s, called "*entropy function*", which is related to a reversible transformation between the initial and the actual states b and b_t, respectively, in the following way:

$$s(\rho_t^i, b, t) = -\int_0^t \frac{1}{\rho} \nabla \cdot \mathbf{J}_S d\tau, \tag{46}$$

where $\mathbf{J}_S$ is defined according to equation (20). Then, we get

$$s = \int_0^t -\frac{1}{\rho\theta} \nabla \cdot \mathbf{q} d\tau + \int_0^t \frac{1}{\rho\theta^2} \mathbf{q} \cdot \nabla\theta d\tau - \int_0^t \frac{1}{\rho} \nabla \cdot \mathbf{k} d\tau. \tag{47}$$

Using the internal energy balance equation and the relation $\mathbf{L} = \nabla \mathbf{v} = \dot{\mathbf{F}}\mathbf{F}^{-1}$, we obtain the following expression for $\nabla \cdot \mathbf{q}$

$$\nabla \cdot \mathbf{q} = -\rho\dot{e} + \mathbf{T} \cdot (\dot{\mathbf{F}}\mathbf{F}^{-1}) + (\mathbf{j}^n + \mathbf{j}^p) \cdot \boldsymbol{\mathcal{E}} - \frac{\dot{\rho}}{\rho} \boldsymbol{\mathcal{E}} \cdot \mathbf{P} + \boldsymbol{\mathcal{E}} \cdot \dot{\mathbf{P}}, \tag{48}$$

so that the final expression for the entropy function is calculated as an integral along a path into the space $\mathbb{R} \times B$ of all thermodynamic variables together with the independent time variable

$$s(\rho_t^i, b, t) = \int_\delta \Omega, \quad \text{with}$$

$$\Omega = -\frac{1}{\rho\theta}(\mathbf{T}\mathbf{F}^{-T}) \cdot d\mathbf{F} - \frac{1}{\rho\theta} \boldsymbol{\mathcal{E}} \cdot d\mathbf{D} + \frac{1}{\theta} de + \left[\frac{1}{\rho\theta^2} \mathbf{q} \cdot \nabla\theta - \frac{1}{\rho\theta} (\mathbf{j}^n + \mathbf{j}^p) \cdot \boldsymbol{\mathcal{E}} \right.$$

$$+\frac{1}{\rho^2\theta}\dot{\rho}\boldsymbol{\mathcal{E}}\cdot\mathbf{P}+\frac{\varepsilon_0}{\rho\theta}\boldsymbol{\mathcal{E}}\cdot\dot{\mathbf{E}}-\frac{1}{\rho}\nabla\cdot\mathbf{k}\Big]\,d\tau, \tag{49}$$

where we have used the relation $\mathbf{T}\cdot(\dot{\mathbf{F}}\mathbf{F}^{-1})=(\mathbf{T}\mathbf{F}^{-T})\cdot\dot{\mathbf{F}}$ (being $\mathbf{F}^{-T}=(\mathbf{F}^{-1})^T$ and T denoting matrix transposition). In eq.(49) the entropy function defines a 1-form Ω in $\mathbb{R}\times B$ called the *entropy 1-form*. In components the entropy 1-form Ω becomes:

$$\Omega=\omega_\mu d\mathfrak{q}^\mu+\omega_0 dt=\omega_A d\mathfrak{q}^A\quad(A=1,2,...,16),$$

where

$$\mathfrak{q}^A=(\mathbf{F},\mathbf{D},\mathbf{B},n,p,e,\mathbf{a},\boldsymbol{\mathcal{V}},\mathbf{j}^n,\mathbf{j}^p,\mathbf{q},\nabla n,\nabla p,\nabla\theta,\nabla\mathbf{a},t)$$

and

$$\omega_A=\Big[\left(-\frac{1}{\rho\theta}\mathbf{T}\mathbf{F}^{-T}\right),\left(\frac{1}{\rho\theta}\boldsymbol{\mathcal{E}}\right),0,0,0,\left(\frac{1}{\theta}\right),0,0,0,0,0,0,0,$$

$$0,0,0,\left(\frac{1}{\rho\theta^2}\mathbf{q}\cdot\nabla\theta-\frac{1}{\rho\theta}(\mathbf{j}^n+\mathbf{j}^p)\cdot\boldsymbol{\mathcal{E}}+\frac{1}{\rho^2\theta}\dot{\rho}\boldsymbol{\mathcal{E}}\cdot\mathbf{P}+\frac{\varepsilon_0}{\rho\theta}\boldsymbol{\mathcal{E}}\cdot\dot{\mathbf{E}}-\frac{1}{\rho}\nabla\cdot\mathbf{k}\right)\Big].$$

Thus, by external differentiation, a 2-form is obtained:

$$d\Omega=\frac{\partial w_A}{\partial\mathfrak{q}^B}d\mathfrak{q}^B\wedge d\mathfrak{q}^A\qquad(A,B=1,2,...,16).$$

Since $d\Omega$ can be written in the following form

$$d\Omega=\sum_{B<A}\frac{\partial w_A}{\partial\mathfrak{q}^B}d\mathfrak{q}^B\wedge d\mathfrak{q}^A+\sum_{B>A}\frac{\partial w_A}{\partial\mathfrak{q}^B}d\mathfrak{q}^B\wedge d\mathfrak{q}^A$$

$$=\sum_{B<A}\left(\frac{\partial w_A}{\partial\mathfrak{q}^B}-\frac{\partial w_B}{\partial\mathfrak{q}^A}\right)d\mathfrak{q}^B\wedge d\mathfrak{q}^A,$$

applying the closure conditions for the entropy 1-form, we obtain the necessary conditions for the existence of the entropy function s during the processes under consideration setting

$$\frac{\partial w_A}{\partial\mathfrak{q}^B}=\frac{\partial w_B}{\partial\mathfrak{q}^A}.$$

In our case we have

$$\partial_e\left(-\tfrac{1}{\rho\theta}\mathbf{T}\mathbf{F}^{-T}\right)=\partial_{\mathbf{F}}\left(\tfrac{1}{\theta}\right),\quad\partial_{\mathbf{D}}\left(-\tfrac{1}{\rho\theta}\mathbf{T}\mathbf{F}^{-T}\right)=\partial_{\mathbf{F}}\left[-\tfrac{1}{\rho\theta}\boldsymbol{\mathcal{E}}\right],$$

$$\partial_{\mathbf{D}}\left(\tfrac{1}{\theta}\right)=\partial_e\left[-\tfrac{1}{\rho\theta}\boldsymbol{\mathcal{E}}\right],\quad \tfrac{\partial\omega_A}{\partial\mathfrak{q}^B}=0\ \ (A=1,2,6,16;\ B=3,4,5,7,...,15),$$

$$\partial_t\left(-\tfrac{1}{\rho\theta}\mathbf{T}\mathbf{F}^{-T}\right)=\partial_{\mathbf{F}}\left[\tfrac{1}{\rho\theta^2}\mathbf{q}\cdot\nabla\theta-\tfrac{1}{\rho\theta}(\mathbf{j}^n+\mathbf{j}^p)\cdot\boldsymbol{\mathcal{E}}+\tfrac{1}{\rho^2\theta}\dot{\rho}\boldsymbol{\mathcal{E}}\cdot\mathbf{P}+\tfrac{1}{\rho\theta}\epsilon_0\boldsymbol{\mathcal{E}}\cdot\dot{\mathbf{E}}-\tfrac{1}{\rho}\nabla\cdot\mathbf{k}\right],$$

$$\partial_t\left(\tfrac{1}{\theta}\right)=\partial_e\left[\tfrac{1}{\rho\theta^2}\mathbf{q}\cdot\nabla\theta-\tfrac{1}{\rho\theta}(\mathbf{j}^n+\mathbf{j}^p)\cdot\boldsymbol{\mathcal{E}}+\tfrac{1}{\rho^2\theta}\dot{\rho}\boldsymbol{\mathcal{E}}\cdot\mathbf{P}+\tfrac{1}{\rho\theta}\epsilon_0\boldsymbol{\mathcal{E}}\cdot\dot{\mathbf{E}}-\tfrac{1}{\rho}\nabla\cdot\mathbf{k}\right],$$

$$\partial_t\left[-\tfrac{1}{\rho\theta}\boldsymbol{\mathcal{E}}\right]=\partial_{\mathbf{D}}\left[\tfrac{1}{\rho\theta^2}\mathbf{q}\cdot\nabla\theta-\tfrac{1}{\rho\theta}(\mathbf{j}^n+\mathbf{j}^p)\cdot\boldsymbol{\mathcal{E}}+\tfrac{1}{\rho^2\theta}\dot{\rho}\boldsymbol{\mathcal{E}}\cdot\mathbf{P}+\tfrac{1}{\rho\theta}\epsilon_0\boldsymbol{\mathcal{E}}\cdot\dot{\mathbf{E}}-\tfrac{1}{\rho}\nabla\cdot\mathbf{k}\right].$$

We remark that in semiconductor crystals ρ is practically constant, so that all results derived in the paper containing the time derivative of ρ can be disregarded. The above relations give the necessary conditions characterizing a sort of "irrotationality" of the entropy 1-form during the analyzed transformation. If the entropy 1-form in eq. (49) is closed and its coefficients are regular, this form is exact and the existence of an upper-potential satisfying relation $S(b_t)-S(b)\geq s$ is ensured, for all $P_t^i\in\Pi$, with $b_t=\rho_t^i(b)$ [3]. Starting from the entropy 1-form it is possible to introduce and investigate an extended thermodynamical phase space in a suitable way [26].

References

[1] F. Barrile and L. Restuccia. Thermodynamics of type-II high Tc Superconductors. In *Series on Advances in Mathematics for Applied Sciences, Applied and Industrial Mathematics in Italy* III, World Scientific, 2009.

[2] B. D. Coleman and M. E. Gurtin, Thermodynamics with internal state variables. *J. Chem. Phys.* 47: 597, 1967.

[3] B. D. Coleman and D. R. Owen. A mathematical foundation for thermodynamics. *Arch. Rat. Mech. Anal.* 54: 1, 1974.

[4] M. Dolfin, M. Francaviglia and P. Rogolino. A geometric perspective on irreversible thermodynamics with internal variables. *Journal of Non-Equilibrium Thermodynamics* 23: 250-263, 1998.

[5] M. Dolfin, M. Francaviglia and P. Rogolino. A geometric model for the thermodynamics of simple materials. *Arch. Rat. Mech. Anal.* 43: 29-36, 1999.

[6] M. Dolfin, M. Francaviglia and L. Restuccia. Thermodynamics of deformable dielectrics with a non-Euclidean structure as internal variable. *Technische Mechanik.* 24: 137-145, 2004.

[7] M. Dolfin, M. Francaviglia, S. Preston and L. Restuccia. Material element model and the geometry of the entropy form. *International Journal of Geometric Methods in Modern Physics* 7: 1021-1042, 2010.

[8] M. Francaviglia, L. Restuccia and P. Rogolino. Entropy production in polarizable bodies with internal variables. *Journal of Non-Equilibrium Thermodynamics* 29: 221-235, 2004.

[9] D. Germanò and L. Restuccia. Thermodynamics of piezoelectric media with dislocations, In *Series on Advances in Mathematics for Applied Sciences, Applied and Industrial Mathematics in Italy II*, World Scientific, 2007.

[10] D. Hull. *Introduction to Dislocations.* Pergamon Press, London, Oxford, 1975.

[11] C. Kittel. *Introduction to Solid State Physics.* John Wiley and Sons, 3rd Ed., New York, 1966.

[12] J. Kubik. A macroscopic description of geometrical pore structure of porous solids. *Int. J. Engng. Sci.* 24: 971-980, 1986.

[13] Landolt-Börnstein. *Numerical Data and Functional Relationships in Science and Technology.* NS III/17a, Springer, 1982.

[14] I. S. Liu. The Method of Lagrange multipliers for exploitation of the entropy principle. *Arch. Rat. Mech. Anal.* 46: 131, 1972.

[15] M. E. Malaspina and L. Restuccia. A geometrical model for a fluid flow in porous structures. *Communications to SIMAI congress.* ISSN: 1827-9015, DOI: 10.1685/CSC09328, 3: 1-12, 2009.

[16] B. Maruszewski. On a dislocation core tensor. *Phys. stat. sol. (b)* 168: 59, 1991.

[17] M. P. Mazzeo and L. Restuccia. Thermodynamics of extrinsic semiconductors with dislocations. *Communication to SIMAI Congress.* ISSN: 1827-9015, DOI: 10.1685/CSC06113: 1-5, 2006.

[18] M. P. Mazzeo and L. Restuccia. Thermodynamics of semiconductors with impurities. *Atti Accademia Peloritana dei Pericolanti.* ISSN: 1825-1242, LXXXVI, DOI: 10.1478/C1A0802006, 2008.

[19] W. Muschik and L. Restuccia. Changing the observer and moving materials in continuum physics: objectivity and frame-indifference. *Technische Mechanik* 22: 2, 152, 2002.

[20] W. Muschik and L. Restuccia. Terminology and classifications of special versions of continuum thermodynamics. *Communications to SIMAI Congress.* ISSN: 1827-9015, DOI: 10.1685/CSC06120:1-5, 2006.

[21] W. Muschik and L. Restuccia. Systematic remarks on Objectivity and frame-indifference, liquid crystal theory as an example. *Archive of Applied Mechanics* ISSN: 0939-1533, doi: 10.1007/s00419-007-0193-2: 1-18, 2008.

[22] F. R. N. Nabarro. *Theory of Crystal Dislocations.* Clarendon Press, Oxford, 1967.

[23] W. Noll. A mathematical theory of the mechanical behaviour of continuous media. *Arch. Rat. Mech. Anal.* 2: 197, 1958.

[24] W. Noll. A new mathematical theory of simple materials. *Arch. Rat. Mech. Anal.* 48: 1-50, 1972.

[25] D. R. Owen. *A first course in the mathematical foundations of thermodynamics.* Springer-Verlag, New York, 1984.

[26] S. Preston and J. Vargo, Indefinite metric of R. Mrugala and the geometry of thermodynamical phase state. *Atti Accademia Peloritana dei Pericolanti.* ISSN: 1825-1242, LXXXVI-Suppl.; DOI:10.1478/C1S0801019, 2008.

[27] L. Restuccia and B. Maruszewski. Interactions between electronic field and dislocations in a deformable semiconductor. *Int. Journal of Applied Electromagnetics and Mechanics* 6: 139-154, 1995.

[28] G. F. Smith. On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors. *Int. J. Engng. Sci.* 9: 899, 1971.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

A NEW LOOK AT THE LYAPUNOV INEQUALITY*

Constantin P. Niculescu†

Abstract

Given a Banach space E, it is proved that any function u in $C^2([a,b],E)$ verifies the inequality

$$\max\left\{\|u(a)\|,\|u(b)\|\right\}+\frac{b-a}{4}\int_a^b\|u''(t)\|\,dt\geq\sup_{t\in[a,b]}\|u(t)\|.$$

The constant $(b-a)/4$ is sharp. Several applications are included.

MSC: Primary 26D10, 34B24; Secondary 26A24, 26A45, 46B20.

keywords: Sturm-Liouville problem, function of bounded variation, differentiable function.

1 Introduction

The well-known Lyapunov inequality states that if $q:[a,b]\to\mathbb{R}$ is a continuous function, then a necessary condition for the boundary value problem

$$\begin{cases} u''+qu=0 \\ u(a)=u(b)=0, \end{cases} \tag{1}$$

*Accepted for publication on December 30, 2010.

†cniculescu47@yahoo.com, University of Craiova, Department of Mathematics, A.I. Cuza Street 13, Craiova 200585, ROMANIA.

to have nontrivial solutions is that

$$\int_a^b |q(t)|\, dt > \frac{4}{b-a}. \tag{2}$$

See the monograph [11] and the survey [3] (which also includes an excellent account on the history of this result).

The following equivalent version of the Lyapunov inequality was proved by Borg [2] (who attributes it to Beurling): for every twice continuously differentiable function $u : [a, b] \to \mathbb{R}$ such that $u(a) = u(b) = 0$ and $u(t) > 0$ for $t \in (a, b)$, we have

$$\int_a^b \frac{|u''(t)|}{u(t)} dt > \frac{4}{b-a}. \tag{3}$$

The aim of this paper is to embed (3) into a stronger inequality that relates the values of a differentiable function on an interval, the values at the endpoints and the total variation of its derivative:

Theorem 1. *Let $u : [a, b] \to \mathbb{R}^N$ be a function which admits an integrable second derivative. Then*

$$\max\{\|u(a)\|, \|u(b)\|\} + \frac{b-a}{4} \int_a^b \left\|u''(t)\right\| dt \geq \sup_{t\in[a,b]} \|u(t)\|.$$

As usually, $\mathbb{R}^N$ denotes here the Euclidean N-dimensional space.

The restriction to the case of functions taking values in $\mathbb{R}^N$ is not essential. A similar result works for all functions taking values in an arbitrary Banach space. This will be discussed in Section 4.

Theorem 1 has a very natural kinematic interpretation: Suppose a point moves in the Euclidean space according to the law of motion $u = u(t)$. Then the difference between the maximum deviation from the origin during an interval of time $[a, b]$ and the maximum deviation at the endpoints of this interval does not exceed

$$\frac{1}{4}\,(\text{elapsed time}) \times \text{total variation of speed}.$$

Recall that every differentiable function $v : [a, b] \to \mathbb{R}^N$ with integrable derivative has bounded total variation and this is given by the formula

$$\bigvee_a^b v = \int_a^b \left\|v'(t)\right\| dt.$$

See [1], p. 104.

The proof of Theorem 1 will make clear that we can deal with other boundary conditions and more general second order differential operators. Some important remarks concerning the case of Neumann boundary conditions can be found in [5].

Also, instead of the L^1 norm in the left hand side and the sup norm in the right hand side we may consider other pairs of L^p norms (with $p \in [1, \infty]$). All these questions will be discussed elsewhere.

2 Consequences of the main result

Theorem 1 has a number of interesting consequences even in the 1-dimensional case. We start with the following stronger version of the inequality of Lyapunov:

Corollary 1. (*A. Wintner* [14]). *Let $q = q(t)$ be a real-valued continuous function defined on an interval $[a, b]$. A necessary condition for the equation $u'' + q(t)u = 0$ to have a nontrivial solution possessing (at least) two zeros is that*

$$\int_a^b q^+(t)dt > \frac{4}{b-a}.$$

Here $q^+ = \sup\{q, 0\}$ denotes the positive part of q.

Proof: By Sturm's Separation Theorem, since $q^+ \geq q$, the equation $v'' + q^+(t)v = 0$ is a Sturm majorant for the equation $u'' + q(t)u = 0$, and hence has a nontrivial solution v with two zeros $\alpha < \beta$ in $[a, b]$. See [8], Corollary 3.1, p. 335. Lyapunov's result follows now from Theorem 1, applied to the restriction of v to $[\alpha, \beta]$. In fact,

$$\begin{aligned}\sup_{t\in[\alpha,\beta]} |v(t)| &< \frac{\beta-\alpha}{4} \int_\alpha^\beta q^+(t)\,|v(t)|\,dt \\ &\leq \frac{b-a}{4} \left(\sup_{t\in[\alpha,\beta]} |v(t)|\right) \int_a^b q^+(t)dt,\end{aligned}$$

and it remains to simplify both sides by $\sup_{t\in[\alpha,\beta]} |v(t)|$. ■

Using a change of variable due to Hille [9], one can extend easily Corollary 1 to all second-order differential equations of the form

$$u'' + g(t)u' + f(t)u = 0,$$

where f is continuous and g is continuously differentiable. In fact, the corresponding equation for $v = u \exp\left(\frac{1}{2}\int_a^t g(s)ds\right)$ is in normal form,

$$v'' + q(t)v = 0,$$

where $q(t) = f(t) - \frac{1}{4}g^2(t) - \frac{1}{2}g'(t)$.

Theorem 1 imposes an obstruction on the nonzero eigenvalues of the operator $Du = -u'' + qu$ with Dirichlet boundary conditions:

Corollary 2. *Suppose that $q : [a,b] \to \mathbb{R}$ is a continuous function, and $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ is a continuous function which admits an estimate of the form $|f(t,u)| \leq \varphi(t)\,|u|$ for a suitable $\varphi \in C([a,b],\mathbb{R})$ with $\varphi > 0$ on (a,b). Then every eigenvalue of the regular Sturm-Liouville problem,*

$$\begin{cases} -u'' + qu = \lambda f(t,u) \\ u(a) = u(b) = 0, \end{cases} \tag{4}$$

admits an estimate of the form

$$|\lambda| \geq \left(\frac{4}{b-a} - \int_a^b |q|\,dt\right)\left(\int_a^b \varphi dt\right)^{-1}.$$

The linear case of the Sturm-Liouville problem (4) (that is, when $f(t,u) = \varphi(t)u$) is presented in many books, for example in [8] and [13]. In this case the spectrum $-u'' + qu$ consists of an increasing sequence of positive eigenvalues λ_n with $\lambda_n \to \infty$.

Notice that Corollary 2 also works in the vector case (when u and f take values in $\mathbb{R}^N$).

Theorem 1 provides useful to establish Weierstrass type criteria of convergence:

Corollary 3. *Let $(u_n)_n$ be a sequence of real-valued twice differentiable functions defined on an interval $[a,b]$. If:*

i) this sequence is convergent at the endpoints; and

ii) the derivatives of second order u_n'' are integrable and

$$\lim_{m,n\to\infty} \int_a^b \left|u_m''(t) - u_n''(t)\right| dt = 0,$$

then the sequence $(u_n)_n$ is uniformly convergent.

Moreover, if u is the limit of the sequence $(u_n)_n$, and all derivatives u_n'' are bounded, then u is differentiable and

$$u' = \lim_{n\to\infty} u_n' \quad \textit{uniformly.}$$

Proof: The first part is a direct consequence of Theorem 1. The second part follows from an old result due to Hadamard [7] (see also [12]): Let I be an interval and let $f : I \to \mathbb{R}$ be a twice differentiable bounded function, with bounded second derivative. Then f' is also bounded and

$$\left\|f'\right\|_\infty \le \begin{cases} \dfrac{2\left\|f\right\|_\infty}{m(I)} + \dfrac{m(I)}{2}\left\|f''\right\|_\infty, & \text{if } m(I) \le 2\sqrt{\left\|f\right\|_{L^\infty} / \left\|f''\right\|_\infty} \\ 2\sqrt{\left\|f\right\|_\infty \cdot \left\|f''\right\|_\infty}, & \text{if } m(I) \ge 2\sqrt{\left\|f\right\|_{L^\infty} / \left\|f''\right\|_\infty} \text{ and } I \ne \mathbb{R} \\ \sqrt{2\left\|f\right\|_\infty \cdot \left\|f''\right\|_\infty}, & \text{if } I = \mathbb{R}. \end{cases}$$

Here $m(I)$ denotes the length of I. ■

3 The scalar case of Theorem 1

The scalar case of Theorem 1 is a consequence of the following more general result:

Theorem 2. *Let $u : [a,b] \to \mathbb{R}$ be a real-valued differentiable function whose derivative has bounded variation. Then*

$$\max\left\{\left|u(a)\right|, \left|u(b)\right|\right\} + \frac{b-a}{4}\bigvee_a^b u' > \sup_{t\in[a,b]} \left|u(t)\right|,$$

except for the affine functions, where equality holds true.

Proof: Step 1. We first consider the case where

$$u(a) = u(b) = 0. \tag{5}$$

In this case (by replacing u by $-u$, if necessary) we may assume that $|u|$ attains its maximum at a point $c \in (a,b)$ and

$$\sup_{t\in[a,b]} \left|u(t)\right| = u(c).$$

Then by the Lagrange mean value theorem there are points $t_1 \in (a,c)$ and $t_2 \in (c,b)$ such that

$$u(c) = u(c) - u(a) = u'(t_1)(c-a)$$

and

$$u(c) = u(c) - u(b) = -u'(t_2)(b-c).$$

Therefore

$$\begin{aligned} \bigvee_a^b u' &\geq \sup_{a<s_1<c<s_2<b} \left|u'(s_1) - u'(s_2)\right| \\ &\geq u'(t_1) - u'(t_2) \\ &= \left(\frac{1}{c-a} + \frac{1}{b-c}\right) u(c) \\ &\geq \frac{4}{b-a} \sup_{t\in[a,b]} |u(t)|, \end{aligned} \tag{6}$$

the last step being a consequence of the arithmetic mean - harmonic mean inequality.

Step 2. We prove next (under the condition (5)) that the equality

$$\frac{b-a}{4} \bigvee_a^b u' = \sup_{t\in[a,b]} |u(t)| \tag{7}$$

occurs only for the function u identically zero. In fact, it suffices to show that $u|_{[a,c]}$ equals the affine function g joining $(a,0)$ and $(c,u(c))$ and $u|_{[c,b]}$ equals the affine function h joining $(c,u(c))$ and $(b,0)$. These equalities yield

$$g'(c) = u'_-(c) = u'_+(c) = h'(c)$$

whence $\frac{u(c)}{c-a} = -\frac{u(c)}{b-c}$. Therefore $u(c) = 0$ and this forces $u \equiv 0$.

The equality $u|_{[a,c]} = g$ (as well as the equality $u|_{[c,b]} = h$) can be proved by reductio ad absurdum. For example, if $u(d) < g(d)$ for some point $d \in (a,c)$, then by the Lagrange mean value theorem there is a $t' \in (d,c)$ such that

$$\begin{aligned} u'(t') &= \frac{u(c) - u(d)}{c-d} > \frac{u(c) - g(d)}{c-d} \\ &= \frac{g(c) - g(d)}{c-d} = \frac{u(c)}{c-a} = g'(t_1) = u'(t_1). \end{aligned}$$

This yields to a contradiction since

$$\begin{aligned}\bigvee_a^b u' &= u'(t_1) - u'(t_2) < u'(t') - u'(t_2)\\ &= \left|u'(t') - u'(t_2)\right| \le \bigvee_a^b u';\end{aligned}$$

the first equality is a consequence of (6) and (7).

The case where $u(d) > g(d)$ for some point $d \in (a, c)$ can be treated similarly.

Step 3. In the general case we have to represent u as

$$u = (u - \varphi) + \varphi,$$

where φ is the affine function joining the points $(a, u(a))$ and $(b, u(b))$. Then $u - \varphi$ vanishes at the endpoints and the result established at Step 1 applies. Therefore

$$\begin{aligned}\sup_{t\in[a,b]} |u(t)| &\le \sup_{t\in[a,b]} |(u-\varphi)(t)| + \sup_{t\in[a,b]} |\varphi(t)|\\ &\le \frac{b-a}{4} \bigvee_a^b (u-\varphi)' + \max\{|u(a)|, |u(b)|\}\\ &= \frac{b-a}{4} \bigvee_a^b u' + \max\{|u(a)|, |u(b)|\},\end{aligned}$$

the equality being possible only when $u - \varphi \equiv 0$. ■

4 The case of vector-valued functions

The proof of Theorem 1 can be reduced to the scalar case by *linearization*, taking into account that

$$\left(\sum_{k=1}^N u_k^2\right)^{1/2} = \sup\left\{\sum_{k=1}^N \alpha_k u_k : \sum_{k=1}^N \alpha_k^2 \le 1\right\}.$$

Indeed, by assuming that Theorem 1 works in the case of scalar functions, for every $x \in [a, b]$ and every family $(\alpha_k)_{k=1}^N$ of real numbers such that

$\sum_{k=1}^{N} \alpha_k^2 \leq 1$ we have

$$\left|\sum_{k=1}^{N} \alpha_k u_k(x)\right| \leq \frac{b-a}{4} \int_a^b \left(\sum_{k=1}^{N} |\alpha_k| \left|u_k''(t)\right|\right) dt + \max\left\{\sum_{k=1}^{N} |\alpha_k| |u_k(a)|, \sum_{k=1}^{N} |\alpha_k| |u_k(b)|\right\} \leq \frac{b-a}{4} \int_a^b \left\|u''(t)\right\| dt + \max\left\{\|u(a)\|, \|u(b)\|\right\},$$

that yields the conclusion of Theorem 1 in the Euclidean case.

It is worth to mention that Theorem 1 actually works in the general framework of Banach spaces.

Theorem 3. *Given a Banach space E, every twice differentiable function $u : [a,b] \to E$ whose second derivative is (Bochner) integrable verifies the inequality*

$$\max\left\{\|u(a)\|, \|u(b)\|\right\} + \frac{b-a}{4} \int_a^b \left\|u''(t)\right\| dt \geq \sup_{t\in[a,b]} \|u(t)\|.$$

The constant $(b-a)/4$ is sharp.

Proof: In fact, according to a classical result due Weierstrass, there exists a point $t_0 \in [a,b]$ such that

$$\|u(t_0)\| = \sup_{t\in[a,b]} \|u(t)\|.$$

Then, by Theorem 1, for every norm-1 linear functional x' in the dual space E' we have

$$\begin{aligned} \left|x'(u(t_0))\right| &\leq \max\left\{\left|x'\left(u(a)\right)\right|, \left|x'\left(u(b)\right)\right|\right\} + \frac{b-a}{4} \int_a^b \left|x'\left(u''(t)\right)\right| dt \\ &\leq \max\left\{\|u(a)\|, \|u(b)\|\right\} + \frac{b-a}{4} \int_a^b \left\|u''(t)\right\| dt. \end{aligned}$$

The proof ends by taking the least upper bound in both sides over all $x' \in E'$ with $\|x'\| = 1$, and using the following well-known consequence of the Hahn-Banach extension theorem:

$$\sup_{x'\in E',\ \|x'\|=1} \left|x'(u(t_0))\right| = \|u(t_0)\|.$$

See [15], Corollary 2, p. 108. ■

5 Some open questions

The literature concerning the analogues of Lyapunov inequality for partial differential equations already counts some important contributions. See for example [4], [5] and [6]. It is thus natural to ask whether Theorem 1 admits an extension to the case of functions of several variables.

Suppose that Ω is a bounded open subset Ω of $\mathbb{R}^N$. Does there exist a second order differential operator A (which in the case of functions of one real variable coincide with the second derivative) and a positive constant $c(\Omega)$ (that depends only on the geometry of the domain Ω) such that every real-valued continuous function $u \in C(\bar{\Omega}) \cap C^2(\Omega)$ with Au integrable verify the inequality

$$\max_{x\in\partial\Omega} |u(x)| + c\,(\Omega) \int_\Omega \|Au(x)\|\, dx \geq \sup_{x\in\bar{\Omega}} |u(x)|? \tag{8}$$

Adrian Tudorascu (oral communication) provided a simple counterexample showing that the natural candidate for A, the Laplacian of u,

$$\Delta u = \sum_{k=1}^{N} \frac{\partial^2 u}{\partial x_k^2},$$

fails even in the case where Ω is the unit ball in $\mathbb{R}^2$. However, the status of (8) is open for $Au = \operatorname{Hess} u$, where

$$\operatorname{Hess} u = \left(\frac{\partial^2 u}{\partial x_j \partial x_k}\right)_{j,k=1}^{N}$$

represents the Hessian matrix of u. Adrian Tudorascu and I have found some consequences that make plausible a positive answer.

A final open question comes in connection with Corollary 3 above. We do not know if the hypothesis regarding the boundedness of the derivatives of second order is essential or not.

Acknowledgement. Research partially supported by CNCSIS Grant 420/2008. We acknowledge helpful correspondence from Florin Popovici and Adrian Tudorascu.

References

[1] R. G. Bartle, *A Modern Theory of Integration*, Graduate Studies in Mathematics vol. **32**, American Mathematical Society, Providence, Rhode Island, 2001.

[2] G. Borg, Über die Stabilität gewisser Klassen von linearen Differentialgleichungen, *Arkiv för Matematik, Astronomi och Fysik,* **31** (1945), 1-31.

[3] R. C. Brown and D. B. Hinton, Lyapunov inequalities and their applications. In vol. *Survey in Classical Inequalities* (T. Rassias ed.), pp. 1-25, Kluwer Academic Publishers, 2000.

[4] A. Cañada, J. A. Montero, S. Villegas, Lyapunov-type Inequalities and Applications to PDE, in: *Proceedings of the 5th European Conference on Elliptic and Parabolic Problems*: *A Special Tribute to the Work of Haim Brezis*, in: Progress Nonlinear Differential Equations Appl., vol. **63**, Birkhäuser, Boston, MA, 2005, pp. 103–110.

[5] A. Cañada, J. A. Montero, S. Villegas, Lyapunov-type inequalities and Neumann boundary value problems at resonance, *Math. Inequal. Appl.* **8** (2005), 459-476.

[6] A. Cañada, J. A. Montero and S. Villegas, Lyapunov inequalities for partial differential equations, *Journal of Functional Analysis* **237** (2006), 176–193.

[7] J. Hadamard, Sur le module maximum d'une fonction et ses dérivées, *Comptes Rendus des séances de la Societé Mathematique de France*, 1914, pp. 68–72.

[8] P. Hartman, *Ordinary differential equations*, John Wiley & Sons, New York, 1964.

[9] E. Hille, Über die Nulstellen der Hermiteschen Polynome, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **44** (1933), 162–165.

[10] A. Lyapunov, Problème général de la stabilité du mouvement, *Ann. Fac. Sci. Univ. Toulouse* **9** (1907), 203-475. (Reproduced in *Ann. Math. Study* **17**, Princeton, 1947).

[11] D. S. Mitrinović, J. E. Pečarić, A. M. Fink, *Inequalities Involving Functions and Their Integrals and Derivatives*, Kluwer Academic Publishers, Dordrecht, 1991.

[12] C. P. Niculescu and C. Buse, The Hardy-Landau-Littlewood Inequalities with less Smoothness, *JIPAM* **4** (2003), Issue 3, article 51.

[13] M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations*, Springer Verlag, 1993.

[14] Wintner A., On the nonexistence of conjugate points, *Amer. J. Math.* **73** (1951), 368-380.

[15] K. Yosida, *Functional Analysis*, 6th edition, Springer-Verlag, 1980.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

GLOBAL RANDOM WALK SIMULATIONS FOR SENSITIVITY AND UNCERTAINTY ANALYSIS OF PASSIVE TRANSPORT MODELS*

Nicolae Suciu† Călin Vamoş‡ Harry Vereecken§ Peter Knabner¶

Abstract

The Global Random Walk algorithm (GRW) performs a simultaneous tracking on a fixed grid of huge numbers of particles at costs comparable to those of a single-trajectory simulation by the traditional Particle Tracking (PT) approach. Statistical ensembles of GRW simulations of a typical advection-dispersion process in groundwater systems with randomly distributed spatial parameters are used to obtain reliable estimations of the input parameters for the upscaled transport model and of their correlations, input-output correlations, as well as full probability distributions of the input and output parameters.

MSC: 65M75, 82C70, 65C05

*Accepted for publication on November 15, 2010.

†suciu@am.uni-erlangen.de, Chair for Applied Mathematics I, Friedrich-Alexander University Erlangen-Nuremberg, Germany, and nsuciu@ictp.acad.ro, Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy, Cluj Napoca, Romania.

‡cvamos@ictp.acad.ro, Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy, Cluj Napoca, Romania.

§h.vereecken@fz-juelich.de, Agrosphere Institute IBG-3, Research Center Jülich, Germany.

¶knabner@am.uni-erlangen.de, Chair for Applied Mathematics I, Friedrich-Alexander University Erlangen-Nuremberg, Germany.

keywords: Probabilistic particle methods, Transport processes, Monte Carlo methods, Groundwater contamination

1 Introduction

Models of passive scalar transport in highly heterogeneous media, such as groundwater systems, turbulent atmosphere, or plasmas, are often based on a stochastic partial differential equation for the concentration field $c(\mathbf{x}, t)$,

$$\partial_t c + \mathbf{V}\nabla c = D\nabla^2 c, \tag{1}$$

with space variable drift $\mathbf{V}(\mathbf{x})$ which is a sample of a random velocity field, and a local diffusion coefficient D which is assumed constant [9, 10, 14, 15, 7]. The normalized concentration solving (1) for the initial condition $c(\mathbf{x}, 0) = \delta(\mathbf{x}-\mathbf{x}_0)$ is the probability density function of the diffusion process described by the Itô stochastic ordinary differential equation

$$X_i(t) = x_{0i} + \int_0^t V_i[\mathbf{X}(t')]dt' + W_i(t), \tag{2}$$

where $i = 1, 2, 3$, $x_{0i} = X_i(0)$ are deterministic initial positions and W_i are the components of a Wiener process of mean zero and variance $2Dt$ [5].

In this paper we consider contaminant transport in saturated groundwater systems. The time-stationary random velocity field $\mathbf{V}(\mathbf{x})$ is, in this case, the solution of the continuity and Darcy equations

$$\nabla\mathbf{V} = 0, \ \mathbf{V} = -K\nabla h, \tag{3}$$

where $K(\mathbf{x})$ is the hydraulic conductivity of the medium and h is the piezometric head [7]. Dirichlet boundary conditions, consisting of constant heads at the inlet and outlet boundaries of the domain, ensure the stationarity in time of the velocity field $\mathbf{V}$. The hydraulic conductivity K is supplied by various interpretations of field-scale measurements in form of a spatially distributed random parameter (random field) [2].

If the random velocity field, obtained by solving (3) for an ensemble of realizations of the K field, has a finite correlation range then it can be shown that, under certain conditions, the ensemble mean concentration is described asymptotically by an upscaled model of form (1), with drift coefficient given by the mean velocity and enhanced diffusion coefficients proportional with

the velocity correlation lengths [6, 4]. Under less restrictive conditions, with the only assumption that the first two spatial moments of the concentration are finite at finite times, the mean concentration can still be described by an equivalent Gaussian distribution with time variable diffusion coefficients [15], referred to as the "macrodispersion" model in the hydrological literature [2]. Root-mean-square deviations of the solutions to (1), for fixed realizations of the velocity field, from the predictions of the upscaled model are often used to quantify the uncertainty in stochastic modeling of transport in random environments [9, 12, 13, 14]. When the estimated mean-square uncertainty is acceptably small, one considers that "ergodic conditions" are met and the macrodispersion model can be successfully used to describe the transport in a single realization of the groundwater formation [9]. Nevertheless, for contamination risk assessments mean-square uncertainty assessments are not enough and extreme values of the stochastic predictions are also required. Such a task can be carried out by assessing the correlations and the full probability distributions of the input/output parameters [1].

When solving advection-dominated transport problems associated to (1), like the one considered here, with Péclet numbers Pe$= U\lambda/D = 100$, where U is the amplitude of the mean velocity and λ a correlation length, the challenge is to ensure the stability of the solutions and to avoid the numerical diffusion [7]. Therefore, numerical solutions to the Itô equation (2), implemented in so called Particle Tracking (PT) algorithms, are often used to simulate trajectories of computational particles and to estimate concentrations by particles densities. PT methods are stable, free of numerical diffusion, thus suitable for advection-dominated transport problems. However, since the computational costs increase linearly with the number of particles, the estimated concentrations are too inaccurate for large-scale simulations of transport in groundwater. Overcoming the limitations of the sequential PT procedure, the Global Random Walk (GRW) has no limitations as concerning the number of particles [9, 16]. As shown in Sect. 2.2 below, GRW provides accurate simulations of the concentration field at costs comparable to those of a single-trajectory PT simulation.

The paper is organized as follows. After recalling basic notions about Euler schemes and PT methods in Section 2.1, we introduce in Section 2.2 the GRW algorithm as a weak numerical scheme for the Itô equation and in Section 2.3 we present a two-dimensional GRW algorithm. A Monte Carlo approach based on GRW is described in Section 3.1. Finally, in Section

3.2 we demonstrate the ability of the GRW approach to produce a detailed sensitivity and uncertainty numerical analysis of the macrodispersion model.

2 Numerical simulations of diffusion processes

2.1 Itô equation and Particle Tracking

Let us consider the one-dimensional Itô equation (2) and an equidistant time discretization $0 < \delta t < \cdots < k\delta t < \cdots < K\delta t = T$. In most of its implementations, the PT simulation of the particle's trajectory consists of an Euler approximation Y_t of the solution $X(t)$, which is a continuous time process satisfying the iterative scheme

$$Y_{k+1} = Y_k + V_k \delta t + \delta W_k, \tag{4}$$

where $Y_k = Y_{k\delta t}$, $V_k = V(Y_k)$, and $\delta W_k = W_{k+1} - W_k$ is the increment of the Wiener process. While the *strong convergence* of order $\beta > 0$ of the Euler scheme requires

$$\lim_{\delta t \longrightarrow 0} E\left(|X_t - Y_t|\right) \leq C\delta t^{\beta},$$

where E denotes the expectation, for the *weak convergence* of order $\beta > 0$, it suffices that

$$\lim_{\delta t \longrightarrow 0} \left|E\left(g(X_t)\right) - E\left(g(Y_t)\right)\right| \leq C\delta t^{\beta},$$

for some functionals $g(X_t)$ (e.g. moments $E(X_t^m)$, $m \geq 1$).

For strong pathwise convergence, the Euler scheme (4) has to consider the Wiener process specified in the Itô equation (2). For weak convergence, when only the probability distribution is approximated, the increments of the Wiener process can be replaced by random variables ξ with similar moments. For weak Euler scheme of order $\beta = 1$ the first three moments of ξ have to satisfy, for some constant M, the condition [5, Sect. 5.12]

$$|E(\xi)| + \left|E(\xi^3)\right| + \left|E(\xi^2) - \delta t\right| \leq M\delta t^2.$$

Easily generated noise increments satisfying the above condition are the two-states random variables

$$\xi : \Omega \longrightarrow \{-\sqrt{2D\delta t}, +\sqrt{2D\delta t}\},\ P\{\xi = \pm\sqrt{2D\delta t}\} = \frac{1}{2}. \tag{5}$$

2.2 Global Random Walk

As far as one approximates probability distributions and their moments the trajectories of the weak Euler scheme are in fact not necessary. The probability distribution of the surrogate random increments of the Wiener process (5) is the limit over a large number of trials N of the relative frequency n/N of occurrence of n heads or tails of an unbiased coin. This can also be thought of as probability that a random walker takes unbiased left/right jumps of constant length $\delta x = \sqrt{2D\delta t}$ on a lattice,

$$P\{\leftarrow\} = P\{\rightarrow\} = \lim_{N\longrightarrow\infty} \frac{n^{\leftarrow}}{N} = \lim_{N\longrightarrow\infty} \frac{n^{\rightarrow}}{N} = \frac{1}{2}, \tag{6}$$

where $n^{\leftarrow}$ and $n^{\rightarrow}$ are the number of walkers jumping to the first-neighbor left site and to the first-neighbor right site, respectively.

The evaluation of the moments $E(X_t^m)$ within the numerical implementation of the weak Euler scheme consists of an arithmetic average, over an ensemble of trajectories (4), of the position of the particles at a given time, which approximates the stochastic average with respect to the probability distribution, $E(X_t) = \int x^m P(t, dx)$. The latter average can also be estimated by discretizing the integral on a regular grid of length L and space step δx as a sum $\sum_{i=1}^{L} (i\delta x)^m P(i\delta x)$, where the probability distribution at a fixed time $P(i\delta x)$ can be approximated by the relative frequency of occupation of the i-th lattice site, n_i/N. Since, according to (5), the walkers cannot be trapped at lattice sites, the occupancy number n_i is the sum of numbers of wlakers reaching the site i from the left, $n_i^{\rightarrow}$, and from the right, $n_i^{\leftarrow}$, i.e. $n_i = n_i^{\rightarrow} + n_i^{\leftarrow}$. One obtains thus the estimation of the m-th order moment of X_t given by

$$E(X_t^m) = \sum_{i=1}^{L} (i\delta x)^m \left(\frac{n_i^{\rightarrow}}{N} + \frac{n_i^{\leftarrow}}{N} \right). \tag{7}$$

For large N, the random variables $n_i^{\rightarrow}$ and $n_i^{\leftarrow}$ occurring in (6-7) can be well approximated as follows. If the number n_i of walkers at the grid site i is even then half of them jump to the left and half to the right, $n_i^{\leftarrow} = n_i^{\rightarrow} = n/2$. If n_i is odd then one walker is allocated to either $n_i^{\leftarrow}$ or to $n_i^{\rightarrow}$ with the same probability, $P\{\leftarrow\} = P\{\rightarrow\} = 1/2$. One obtains in this way a GRW algorithm for the Wiener process, described by equation (2) without drift term [16]. Figure 1 illustrates the evolution of the number n_i of random walkers over the first three simulation steps, obtained with a straightforward

MATLAB implementation of the above one-dimensional GRW algorithm. The concentration at a given time (solution of (1)) can be simply estimated as $c(i\delta x) = n_i/\delta x$.

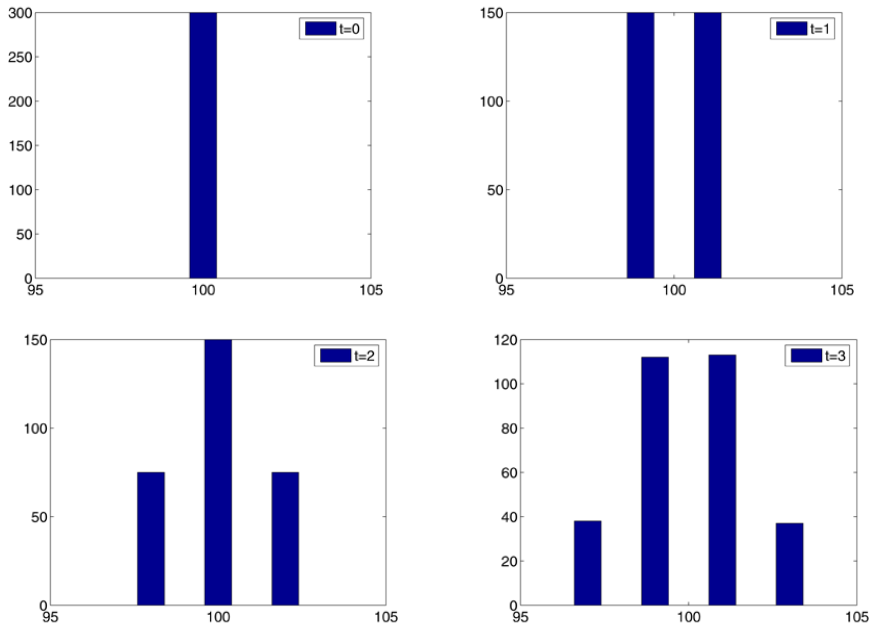


Figure 1: Distribution of $N = 300$ random walkers after the first three time steps of the GRW simulation.

Unlike the discrete-time grid-free weak Euler scheme, the GRW algorithm is a discrete time-space stochastic scheme. As follows from (5) the constant amplitude δx of the random jumps ξ is related to the time step δt and the diffusion coefficient D by

$$D = \frac{\delta x^2}{2\delta t}. \tag{8}$$

Since the numerical scheme is constrained by the relation (8), GRW is not affected by numerical diffusion. GRW is also stable because the number of random walkers N is conserved. Figure 2 shows the estimated mean $M = E(X_t)$ and diffusion coefficient $D = [E(X_t^2) - E(X_t)^2]/(2t)$, computed according to (7), as well as the final distribution of n_i for a diffusion process with $D = 1$ resulted from a GRW simulation with $\delta x = 1$ and $\delta t = 0.5$.

It is also possible to simplify the GRW algorithm by completely removing the randomness from the scheme. This is done by setting $n_i^{\leftarrow}$ and $n_i^{\rightarrow}$ to the exact value of $n/2$. In this case N has no longer the meaning of a number of random walkers and can be taken as an arbitrary positive real number, for instance equal to 1. This deterministic GRW scheme is equivalent to the finite-difference scheme for the heat equation and converges as δx^2 for $\delta x \to 0$ [16]. Since according to relation (8) $\delta x^2 \sim \delta t$, the deterministic GRW has the same order of convergence with the time step as the weak Euler scheme of order $\beta = 1$. The convergence of the stochastic GRW simulation reaches the same order only if the number of random walkers N is large enough to smooth out the random fluctuations of n_i. Figure 3 shows the dependence on N of the absolute error $eD(t) = |D_{grw}(t) - D|$ and the convergence of the norm $\|D_{grw} - D\|$ defined by

$$\|D_{grw} - D\|^2 = \sum_{k=1}^{T/\delta t} \left[D_{grw}(k\delta t) - D\right]^2 .$$

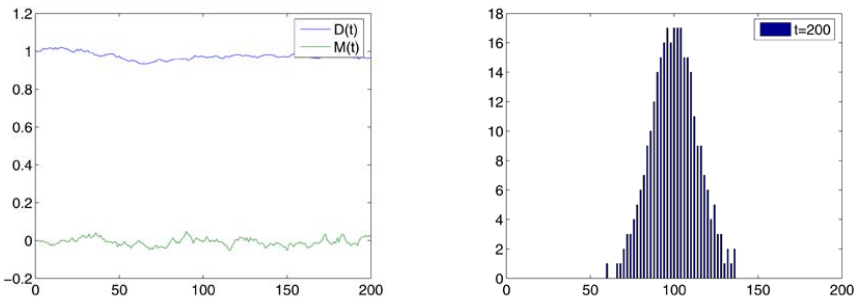


Figure 2: Estimation of the diffusion coefficient $D(t)$ and of the mean $M(t)$ (left) and distribution of $N = 300$ random walkers after 200 time steps in the GRW simulation (right).

Note that the GRW scheme described above is practically insensitive to the number of random walkers N. Assuming that all L grid points contain random walkers at all the computation time steps, one needs LT calls of a uniformly-distributed random-numbers generator for the entire simulation. Hence, the total computation time is of the order of that for the simulation of a single trajectory of the Itô process by the weak Euler scheme. Since for $N =$

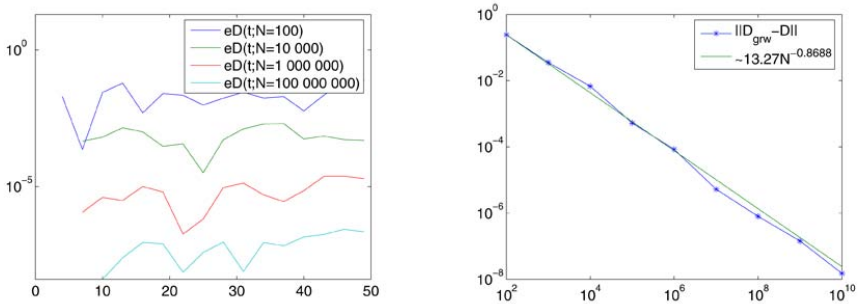


Figure 3: Errors for the estimations of diffusion coefficients for increasing N (left) and the convergence of the error norm (right).

1 the output of the simulation is the trajectory of a single random walker, GRW can be thought of as a superposition of particle tracking procedures for arbitrary large numbers of particles. Since the computational cost of a simulation for N trajectories with the Euler scheme is of the order of NT, the GRW algorithm achieves a speed-up of computations, with respect to PT, of the order N/L. For example, while the convergence investigations with GRW presented in Figure 3 were performed in about one second, similar investigations with the Euler scheme required several minutes on the same computer. In case of realistic simulations of diffusion processes, when very large numbers of particles should be considered, e.g. $N = 10^{24}$ (Avogadro's number), as well as large grids of the order of $L = 10^6$ nodes, a huge speed-up of computations by a factor of 10^{18} can be achieved by using the GRW algorithm.

2.3 Two-dimensional GRW algorithm

For a two-dimensional transport problem, the solution of the parabolic equation (1) is simulated with N particles undergoing advective displacements and diffusive jumps according to the random walk law on a regular grid. The concentration at a given time $t = k\delta t$ and a point $(x_1, x_2) = (i_1\delta x_1, i_2\delta x_2)$ is given by

$$c(x_1, x_2, t) = \frac{1}{N\Delta_1\Delta_2} \sum_{i'_1=-s_1}^{s_1} \sum_{i'_2=-s_2}^{s_2} n(i_1 + i'_1, i_2 + i'_2, k), \tag{9}$$

where $\Delta_l = 2s_l\delta x_l$, $l = 1, 2$, are the lengths of the symmetrical intervals centered at x_l and $n(i_1, i_2, k)$ is the number of particles which at the time step k lie at the grid point (i_1, i_2).

For constant diffusion coefficient D, the two-dimensional simulation consists of repeating the one-dimensional procedure on each of the two spatial directions [16, 11]. The one-dimensional GRW algorithm, which generalizes the algorithm presented in Section 2.2 to account for advective displacements, describes the scattering of the $n(i, k)$ particles from (x_i, t_k) by

$$n(j, k) = \delta n(j, j + v_j, k) + \delta n(j + v_j - d, j, k) + \delta n(j + v_j + d, j, k), \quad (10)$$

where $v_j = V_j\delta t/\delta x$ are discrete displacements produced by the velocity field and d describes the diffusive jumps. The quantities δn introduced in (10) are Bernoulli random variables and describe respectively, the number of particles which remain at the same grid site after an advective displacement and the number of particles jumping to the left and to the right of the advected position $j + v_j$. The distribution of the particles at the next time $(k + 1)\delta t$ is given by

$$n(i, k + 1) = \sum_j \delta n(i, j, k).$$

The average number of particles undergoing diffusive jumps and the average number of particles remaining at the same node after the displacement v_j are given by the relations

$$\overline{\delta n(j + v_j \pm d, j, k)} = \frac{1}{2} r \; \overline{n(j, k)},$$

$$\overline{\delta n(j, j + v_j, k)} = (1 - r) \; \overline{n(j, k)},$$

where $0 \le r \le 1$. The diffusion coefficient D is related to the grid steps by the relation

$$D = r\frac{(d\delta x)^2}{2\delta t},$$

which generalizes (8) and ensures that the scheme does not produce numerical diffusion.

Particularizing the above one-dimensional GRW algorithm for genuine diffusion, i.e. letting $v_j = 0$ in (10), one can easily see that the evolution of the mean number of particles is described by

$$\overline{n(i, k + 1)} = \frac{r}{2}\overline{n(i + d, k)} + (1 - r)\overline{n(i, k)} + \frac{r}{2}\overline{n(i - d, k)}. \quad (11)$$

which has the form of the explicit scheme for the heat equation. Since the scheme (11) is consistent and, by condition $r \leq 1$ (von Neumann's criterion), it is also stable, it converges with the order $O(\delta x^2)$. Moreover, as demonstrated numerically in [16], the un-averaged GRW solution $n(i,k)$ converges as $\mathcal{O}(\delta x^2) + \mathcal{O}(N^{-1/2})$. Thus, for sufficiently large numbers of particles GRW has the same order of convergence as the stable finite differences scheme.

It is worth noting that while for constant drift coefficients V_j the GRW algorithm is still equivalent to a finite difference scheme, the equivalence fails for space variable V_j. Indeed, in the latter case to the site i contribute not only particles jumping from two symmetrical left and right sites, like in the finite difference scheme (11), but also particles coming from distances $v_j \pm d$ which depend on the variable drift coefficient V_j. However, GRW remains equivalent to a superposition of many PT schemes and this makes it suitable for simulating advection-diffusion processes described by the parabolic equation (1). In fact, as shown in Section 2.2 above, GRW is a weak scheme for solving Itô equations, which approximates the true probability distribution (concentration) at all grid points and time steps, without solving for individual trajectories. This is the essential feature which considerably increases the performance of the GRW algorithm with respect to PT, where, after the sequential simulation of particles trajectories, a post-processing is required to count the contribution of the computational particles to the concentration, estimated at given points in space and time steps.

The "reduced fluctuations" GRW algorithm generalizes the simple procedure described in Section 2.2 by

$$\delta n(j + v_j - d, j, k) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ [n/2] + \theta & \text{if } n \text{ is odd,} \end{cases}$$

where $n = n(j,k) - \delta n(j, j + v_j, k)$, $[n/2]$ is the integer part of $n/2$ and θ is a variable taking the values 0 and 1 with probability 1/2. Further, the number of particles jumping in the opposite direction, $\delta n(j, j + v_j + d, k)$ is determined by (10). This algorithm is appropriate for large scale problems, for two reasons. Firstly, the diffusion front does not extend beyond the limit concentration defined by one particle at a grid point, keeping a physical significant shape (unlike in finite differences schemes, where a pure diffusion front has a cubic shape of side $\sim \sqrt{2Dt}$). Secondly, the reduced fluctuations algorithm requires only a minimum number of calls of the random number generator.

A comparison with a PT code (done for the diffusion over ten time steps of N particle starting at the center of a cubic grid) shows that while for the GRW algorithm there were practically no limitations concerning the total number of particles and the computation time was of about one second, PT simulations for $N = 10^9$ particles already required a computing time of about one hour and 256 processors on a CRAY T3E parallel machine [16].

To compute moments, as for instance the variance of particle displacements $s_{ll}^2 = E(X_l^2) - E(X_l)^2$, $l = 1, 2$, a more accurate result is obtained if instead of the concentration (9) one uses the point density of the number of particles $n(i_1, i_2, k)$:

$$\frac{1}{(\delta x)^2} s_{ll}^2(k\delta t) = \frac{1}{N} \sum_{i_1,i_2} i_l^2 n(i_1, i_2, k) - \left[\frac{1}{N} \sum_{i_1,i_2} i_l \; n(i_1, i_2, k) \right]^2 .$$

With this, the effective diffusion coefficients will be computed as

$$D_{ll}^{eff}(k\delta t) = s_{ll}^2/(2k\delta t). \tag{12}$$

Let us consider N_{X_0} points uniformly distributed inside the initial plume, N/N_{X_0} particles at each initial point and let $n(i_1, i_2, k; i_{01}, i_{02})$ be the distribution of particles at the time step k given by the GRW procedure for a diffusion process starting at $(i_{01}\delta x_1, i_{02}\delta x_2)$. Writing the distribution for the extended plume as

$$n(i_1, i_2, k) = \sum_{i_{01},i_{02}} n(i_1, i_2, k; i_{01}, i_{02}),$$

the averages defining the first two moments can be rewritten in the form

$$\frac{1}{N} \sum_{i_1,i_2} \alpha n(i_1, i_2, k) = \frac{1}{N_{X_0}} \sum_{i_{01},i_{02}} \left(\frac{N_{X_0}}{N} \sum_{i_1,i_2} \alpha n(i_1, i_2, k; i_{01}, i_{02}) \right), \tag{13}$$

where α stands for i_l and i_l^2 respectively. As follows from (13), the first two moments $E(X_l)$, and $E(X_l^2)$, as well as the effective diffusion coefficients (12) are averages over the trajectories of the diffusion process starting at given initial positions and over the distribution of the initial positions.

3 Sensitivity and uncertainty analysis

3.1 Monte Carlo simulations

To enable the simulation of large ensembles of transport realizations, a linearization of the flow equation (3) was considered and the velocity samples were generated, for given statistics of the hydraulic conductivity K, by the Kraichnan's randomization method [8], which has been successfully used in numerical investigations on large scale behavior of the passive transport in aquifers [3, 9, 10]. We considered a log-normally distributed conductivity K, i.e. a normal lnK field with variance σ^2 and exponential isotropic correlation $\rho(|\mathbf{x}_1 - \mathbf{x}_2|) = \sigma^2 exp(-|\mathbf{x}_1 - \mathbf{x}_2|/\lambda)$, where λ is the correlation length. For a given pressure gradient between the inlet and outlet boundaries, which fixes the value of the ensemble mean velocity $U = |\langle \mathbf{V} \rangle|$, the incompressible Darcy flow, solution of equations (3), was approximated by a superposition of N_p periodic modes

$$V_i(\mathbf{x}) = U\delta_{i1} + U\sigma \sqrt{\frac{2}{N_p}} \sum_{l=1}^{N_p} p_i(\mathbf{q}_l) \sin(\mathbf{q}_l \cdot \mathbf{x} + \alpha_l). \tag{14}$$

The wave vectors $\mathbf{q}_l$ are mutually independent random variables, with probability distribution proportional with the spectral density of the lnK field, and the phases α_l are random variables uniformly distributed in the interval $[0, 2\pi]$. The functions p_l are projectors which ensure the incompressibility of the flow. It has been shown that V_i tends to a Gaussian random field when $N_p \to \infty$ [8]. It was also found that $N_p = 6400$, which we fix in the following, provides reliable approximations of the velocity field at the problem's spatial scale considered here [9, 3].

The mean velocity occurring in (14), which can be freely chosen, was set to a typical value of $U = 1$ m/day. We also have chosen a typical local-scale diffusion coefficient in (1), $D = 0.01$ m^2/day, and $\lambda = 1$ m for the correlation length of the lnK field, so that the Péclet number was set to Pe$= U\lambda/D = 100$. We conducted Monte Carlo simulations for two cases, corresponding to two extreme degrees of heterogeneity: $\sigma^2 = 0.1$, for which the approximation (14) of the velocity field is accurate and the macrodispersion model is expected to provide a reliable description of the mean behavior of the transport process, and $\sigma^2 = 6$, an extremely large value, for which (14) is no longer close to the true solution of flow equations (3) but can however

serve to illustrate the situation when the macrodispersion model might be inadequate.

The behavior of a passive tracer, initially uniformly distributed in slabs of dimensions $100\lambda \times \lambda$ perpendicular to the mean flow direction, was simulated over 2000 days for the low heterogeneity case $\sigma^2 = 0.1$, in 1024 realizations of the random field (14), and over 300 days, in 100 realizations in the highly heterogeneous case $\sigma^2 = 6$. The plume's shapes in the two extreme cases are compared in Figure 4. (Note that the spatial simulation domain was, in all cases, large enough to avoid the influence of the boundaries.)

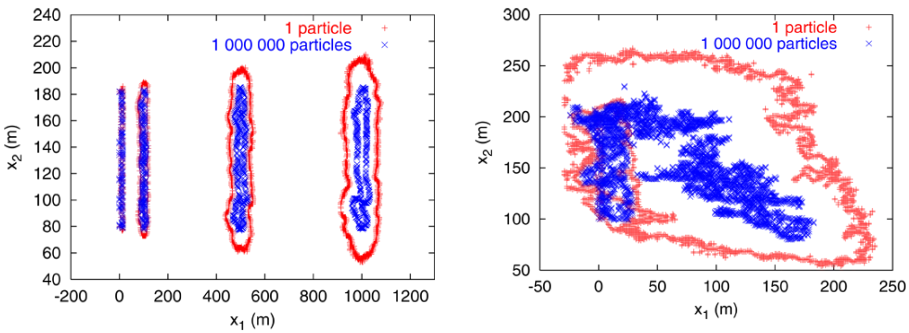


Figure 4: Plume contours for $\sigma^2 = 0.1$ at $t = 0$, 100, 500 and 1000 days (left panel) and for $\sigma^2 = 6$ at $t = 0$, 10, and 100 days (right panel).

Monte Carlo estimates, by equal-weight (arithmetic) averages over the corresponding ensembles of realizations, hereafter denoted by $\langle \cdots \rangle$, were computed for the set of input parameters of the macrodispersion model, consisting of longitudinal $u = E(X_1)/t$ and transverse $v = E(X_2)/t$ components of the center of mass velocity, longitudinal $D_x = D_{11}^{eff}$ and transverse $D_y = D_{22}^{eff}$ effective diffusion coefficients (12), for the only output parameter considered here, consisting of the cross-section space average concentration at the center of mass (hereafter denoted by c), as well as for their cross-correlations, $\langle uv \rangle$, $\langle uD_x \rangle$, $\langle uD_y \rangle$, $\langle vD_x \rangle$, $\langle vD_y \rangle$, $\langle D_xD_y \rangle$, $\langle uc \rangle$, $\langle vc \rangle$, $\langle D_xc \rangle$, and $\langle D_yc \rangle$. Probability densities of the parameters, approximated by histograms, were summed-up to estimate cumulative probability distributions.

3.2 Results

The left panel of Figure 5 shows that for low heterogeneity ($\sigma^2 = 0.1$) the only input-input relevant correlation is that between the longitudinal velocity of the center of mass and the transverse effective diffusion coefficient. The sensitivity of the transverse dispersion to the mean longitudinal flow indicates the increased role of the transverse dispersion for small mean flow velocity. The results for the highly heterogeneous case ($\sigma^2 = 6$) from the right panel of Figure 5 show stronger correlations between the input parameters, which are expected to facilitate the uncertainty propagation and to reduce the reliability of the macrodispersion model.

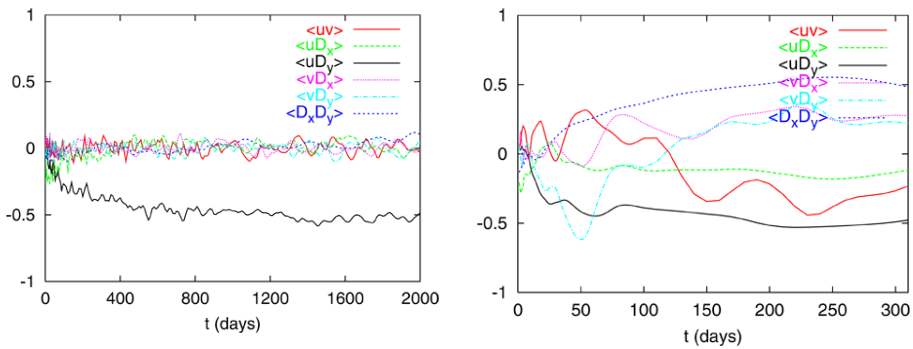


Figure 5: Correlations between input parameters of the macrodispersion model (velocity components of center of mass, u and v, and dispersion coefficients, D_x and D_y) for $\sigma^2 = 0.1$ (left panel) and $\sigma^2 = 6$ (right panel).

As expected, for low heterogeneity (left panel of Figure 6) there is a strong correlation between the longitudinal effective diffusion coefficient and the cross-section averaged concentration. This suggests that, when the only output parameter of interest is the cross-section concentration, the macrodispersion model can be trusted as reliable for single-realizations of the transport process, in agreement with other observations that the cross-section concentration can be modeled as an one-dimensional advection-diffusion process governed by the longitudinal effective diffusion coefficient [9]. The situation is different for high heterogeneity (right panel of Figure 6), where the cross-section concentration is also strongly correlated with the transverse effective diffusion coefficient. Again, this result renders questionable the applicability of the macrodispersion model to highly heterogeneous media.

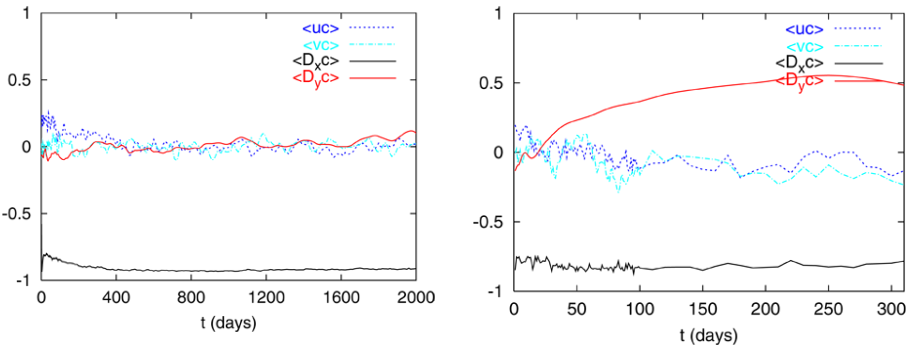


Figure 6: Correlations between input parameters u, v, D_x, and D_y, and the output parameter c (the cross-section space average concentration at the center of mass) for $\sigma^2 = 0.1$ (left panel) and $\sigma^2 = 6$ (right panel).

To illustrate the capability of the Monte Carlo approach based on GRW simulations to produce a full statistical description of the transport process, we present in Figure 7 the estimated cumulative probability distributions of the cross section concentration at the plumes center of mass and of the longitudinal velocity of the center of mass. In a forthcoming work, these probability distributions will be used as reference data in developing a probability density function method similar to those used in modeling turbulent transport [1]. The novelty of the new approach will consists of a three-dimensional GRW solution of the equations governing the evolution of the concentration probability density in the cartesian product between the physical space and the concentration domain.

Acknowledgement. This work was supported by the Deutsche Forschungsgemeinschaft under Grant SU 415/1-2, Jülich Supercomputing Centre Project No. JICG41, and Romanian Ministry of Education and Research under Grant 2-CEx06-11-96

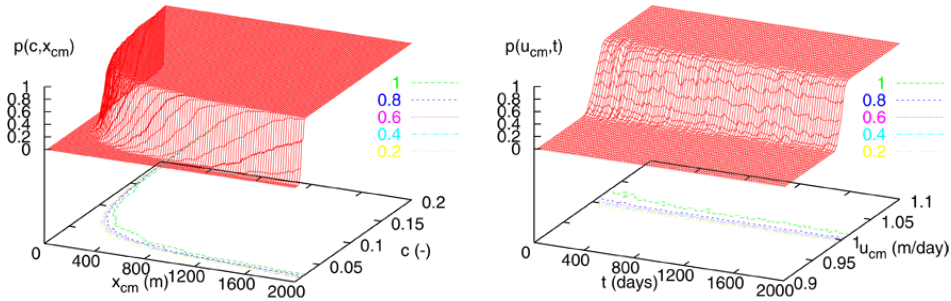


Figure 7: Probability distributions of the concentration estimated along the longitudinal component of the center of mass $c(x_{cm})$ (left panel) and of the longitudinal component of the center of mass velocity as function of time $u_{cm}(t)$ (right panel), for $\sigma^2 = 0.1$.

References

[1] P. J. Colucci, F. A. Jaberi, P. Givi. Filtered density function for large eddy simulation of turbulent reacting flows. *Phys. Fluids* 10:499-515, 1998.

[2] G. Dagan. *Flow and Transport in Porous Formations.* Springer, Berlin, 1989.

[3] J. Eberhard, N. Suciu, C. Vamoş. On the self-averaging of dispersion for transport in quasi-periodic random media. *J. Phys. A: Math. Theor.* 40:597-610, doi: 10.1088/1751-8113/40/4/002, 2007.

[4] A. Fannjiang, T. Komorowski. Turbulent diffusion in Markovian flows, *Ann. Appl. Probab.* 9:591-610, 1999.

[5] P. E. Kloeden, E. Platen. *Numerical Solutions of Stochastic Differential Equations.* Springer, Berlin, 1999.

[6] T. Komorowski, G. Papanicolaou. Motion in a Gaussian incompressible flow, *Ann. Appl. Probab.* 7:229-264, 1997.

[7] F. A. Radu, N. Suciu, J. Hoffmann, A. Vogel, O. Kolditz, C.-H. Park, S. Attinger. Accuracy of numerical simulations of contaminant transport in heterogeneous aquifers: a comparative study. *Adv. Water Resour.*, doi:10.1016/j.advwatres.2010.09.012, 2010 (in press).

[8] K. Sabelfeld. *Monte Carlo methods in boundary value problems.* Springer, Berlin, 1991.

[9] N. Suciu, C. Vamoş, J. Vanderborght, H. Hardelauf, H. Vereecken. Numerical investigations on ergodicity of solute transport in heterogeneous aquifers. *Water Resour. Res.* 42:W04409, doi:10.1029/2005WR004546, 2006.

[10] N. Suciu, C. Vamoş, J. Eberhard. Evaluation of the first-order approximations for transport in heterogeneous media. *Water Resour. Res.* 42:W11504, doi:10.1029/2005WR004714, 2006.

[11] N. Suciu, C. Vamoş, I. Turcu, C.V.L. Pop, and L. I. Ciortea. Global random walk modeling of transport in complex systems. *Comput. Visual. Sci.*, 12:77-85, doi:10.1007/s00791-007-0077-6, 2007.

[12] N. Suciu, C. Vamoş, H. Vereecken, K. Sabelfeld, P. Knabner. Memory effects induced by dependence on initial conditions and ergodicity of transport in heterogeneous media. *Water Resour. Res.* 44:W08501, doi:10.1029/2007WR006740, 2008.

[13] N. Suciu, P. Knabner. Comment on 'Spatial moments analysis of kinetically sorbing solutes in aquifer with bimodal permeability distribution' by M. Massabo, A. Bellin, and A. J. Valocchi. *Water Resour. Res.* 45:W05601, doi:10.1029/2008WR007498, 2009.

[14] N. Suciu, C. Vamos, F. A. Radu, H. Vereecken, P. Knabner. Persistent memory of diffusing particles. *Phys. Rev. E* 80:061134, doi:10.1103/PhysRevE.80.061134, 2009.

[15] N. Suciu. Spatially inhomogeneous transition probabilities as memory effects for diffusion in statistically homogeneous random velocity fields. *Phys. Rev. E* 81:056301, doi:10.1103/PhysRevE.81.056301, 2010.

[16] C. Vamoş, N. Suciu, H. Vereecken. Generalized random walk algorithm for the numerical modeling of complex diffusion processes. *J. Comput. Phys.* 186:527-544, doi:10.1016/S0021-9991(03)00073-1, 2003.

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

DEGENERATED HOPF BIFURCATIONS IN A MATHEMATICAL MODEL OF ECONOMICAL DYNAMICS *

Laura Ungureanu†

Abstract

It is assumed that the dynamics of the capital of a firm is governed by a Cauchy problem for a system of two nonlinear ordinary differential equations containing three real parameters. In this paper we determine a $k \geq 3$ order degenerated Hopf bifurcation point for this economical model. To this aim the normal form technique is used.

MSC: 37L10, 37G05, 91B55

keywords: nonlinear dynamics, Hopf bifurcation, normal form, Liapunov coefficients

1 Introduction

The nonlinear dynamics theory enables us to understand and develop more realistic processes and methods in economic models. The development of the theory of singularities and the theory of bifurcation has completed the multitude of ways at our disposal to analyze and represent more and more complex dynamics, giving us the possibility of analyzing some systems which were hard, if not impossible to approach by traditional methods. The study

*Accepted for publication on January 11, 2011.

†Spiru Haret University, European Centre of Managerial and Business Studies, Craiova ungureanu@lycos.com

of nonlinear dynamics is of outmost interest because the economical systems are by excellence nonlinear systems. Many of these contain multiple discontinuities and incorporate inherent instability being permanently under shock actions, extern and intern perturbations.The classical methods based on continuity, linearity and stability have been proven unstable for representing economic phenomena and processes with a higher degree of complexity. The researchers are bound to follow these processes in a dynamic way, to study qualitatively the changes that interfere with the implicated economic variables as well as the results obtained with their help.There are several models describing microeconomical dynamics. One of them is shown by the subsequent model consisting in the Cauchy problem $x(0) = x_0$, $y(0) = y_0$ for the system o.d.e. in $\mathbf{R}^2$.

1.1 Mathematical model

Let K_t be the capital of a firm at the time t and let L_t be the number of workers. Then the production force reads $y_t = F(K_t, L_t)$. The dynamics of the capital depends on the politics of firm development involving the net profit π_t, the dividends covering by shareholders δ_t (where $\delta_t\pi_t$ represents the dividends and $(1 - \delta_t)\pi_t$ are the remaining investments), the capital depreciation by a coefficient μ_t and the income obtained by liquidation of the depreciated assets at the revenue costs λ_t. Let γ_t be the rate of change of the capital, such that $\pi_t = \gamma_t y_t$. Then, according to Oprescu [6], Ungureanu [7]

$$\begin{cases} \dot{K}(t) = \gamma_t(1 - \delta_t)F(K_t, L_t) - \mu_t(1 - \lambda_t)K_t \\ \dot{L}(t) = \alpha_1 K_t + \alpha_2 L_t - \alpha_0 \end{cases}$$

where the dot over quantities represents the differentiation with respect to time. Within this system K and $L : R \to R$ are unknown functions depending on independent variable t (time), $K-$ the capital of a firm and L- the number of workers.

This study is made according to the simplifying assumption that the parameters are considered constant $\mu_t = \mu$, $\delta_t = \delta$, $\gamma_t = \gamma$, $\lambda_t = \lambda$. If $y_t = VK_t^{\alpha}L_t^{\beta}$ and the production has an increasing physical efficiency, i.e. $\alpha + \beta > 1$, the above equations become

$$\begin{cases} \dot{x} = cx^2y + bx \\ \dot{y} = x + \alpha_2 y - 1 \end{cases} \tag{1.1.1}$$

where we choose $\alpha = 2$, $\beta = 1$, $x = \beta_1 K_t$, $y = \beta_2 L_t$, $\beta_1 = \alpha_1/\alpha_0$, $\beta_2 = 1/\alpha_0$ for $\alpha_0 \neq 0$, $\alpha_1 \neq 0$, $a = V\gamma(1-\delta)$, $b = -\mu(1-\lambda)$, $c = a\alpha_0^2/\alpha_1$. In this way the new state functions x and y are proportional to the capital and working force respectively. In addition, the number of parameters was reduced from eight to three.

1.2 Equilibrium points

Here α_2,b,$c \in \mathbf{R}$ are constant economical parameters and x and y are two economical state functions which are proportional to the capital and working force respectively.

The dynamics generated by (1.1.1) strongly depends on the three parameters. However, it is qualitatively unchanged for parameters lying in some areas of the parameter space. Correspondingly, for various points in these areas, the phase portraits are topologically equivalent.

In phase portraits formation a particular influence is exercised by the equilibria. They are the starting points in the study of the dynamical bifurcation (understood as a negation of the structural stability).

In the (x, y) phase plane they correspond to the equilibrium points denoted by $\overline{u}$.

The following cases hold:

a) $b = c = \alpha_2 = 0 \Rightarrow$ (1.1.1) has an infinity of equilibria $\overline{u} = (1, y_0)\,, \forall y_0 \in \mathbf{R}$ possessing the eigenvalues $s_1 = s_2 = 0$;

b) $b = c = 0$, $\alpha_2 \neq 0 \Rightarrow$ (1.1.1) has an infinity of equilibria $\overline{u} = (1 - \alpha_2 y_0, y_0)$, $\forall\, y_0 \in \mathbf{R}$ possessing the eigenvalues $s_1 = 0$, $s_2 = \alpha_2$;

c) $b = \alpha_2 = 0$, $c \neq 0 \Rightarrow$ (1.1.1) has a unique equilibrium $\overline{u} = (1, 0)$ possessing the eigenvalues $s_{1,2} = \pm\sqrt{c}$ for $c > 0$ and $s_{1,2} = \pm i\sqrt{-c}$ for $c < 0$;

d) $c = \alpha_2 = 0$, $b \neq 0 \Rightarrow$ (1.1.1) has no equilibrium;

e) $c = 0$, $b\alpha_2 \neq 0 \Rightarrow$ (1.1.1) has an equilibrium $\overline{u} = \left(0, \alpha_2^{-1}\right)$ possessing the eigenvalues $s_1 = b$, $s_2 = \alpha_2$;

f) $b = 0$, $c\alpha_2 \neq 0 \Rightarrow$ (1.1.1) has two equilibria $\overline{u}_1 = \overline{u}_2 = \left(0, \alpha_2^{-1}\right)$and $\overline{u}_3 = (1, 0)$ possessing the eigenvalues $s_1 = 0$,$s_2 = \alpha_2$ and $s_{1,2} = \left(\alpha_2 \pm \sqrt{\alpha_2^2 + 4c}\right)/2$, respectively;

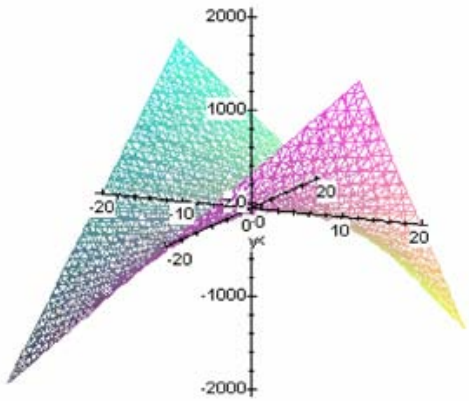


Figure 1: The surface S

g) $\alpha_2 = 0$, $bc \neq 0 \Rightarrow$ (1) has an equilibrium $\overline{u} = (1, -b/c)$ possessing the eigenvalues $s_{1,2} = \left(-b \pm \sqrt{b^2 + 4c}\right)/2$;

h) $\alpha_2 bc \neq 0 \Rightarrow$ (1.1.1) has three equilibria $\overline{u}_1 = \left(0, \alpha_2^{-1}\right)$,
$\overline{u}_2 = \left(\frac{c+\sqrt{c^2+4bc\alpha_2}}{2c}, \frac{c-\sqrt{c^2+4bc\alpha_2}}{2c\alpha_2}\right)$, $\overline{u}_3 = \left(\frac{c-\sqrt{c^2+4bc\alpha_2}}{2c}, \frac{c+\sqrt{c^2+4bc\alpha_2}}{2c\alpha_2}\right)$.

In the general case h), the three equilibria can never coincide, neither in the limit b, c, $\alpha_2 \to \pm\infty$. However, two of them can coincide at the points of the parameter space situated on a surface S (Figure 1). More exactly if $\overline{u}_1 = \overline{u}_2 = (1/2, 1/2\alpha_2)$. Therefore S is a hyperboloid with two sheets. It has the equation $c = -4b\alpha_2$, where $b\alpha_2 \neq 0$ and , and its sheets are situated in the octants characterized by $c > 0$, $b\alpha_2 < 0$, and $c < 0$, $b\alpha_2 > 0$, respectively.

In the domain determined by sheets of S and the plane on which it is supported,(b, α_2) , the system (1.1.1) has an equilibrium point. Outside this domain, at the points which do not belong to S or the three planes $b = 0$, $c = 0$, $\alpha_2 = 0$, the system (1.1.1) possesses three equilibria.

We recall that on the sheets of S (1.1.1) possesses two equilibria and S has no point in the plans of coordinates on the parameter space. We can have two equilibria only on S and in the $b = 0$ plane without axes, one equilibrium is double, namely $\overline{u}_2 = \overline{u}_3 = \left(0, \alpha_2^{-1}\right)$, and another one $\overline{u}_2 = (1, 0)$ simple. Let us notice that in this case $c \neq 0$.

Let us define the domains D_1 and D_2 determined by the sheets of S and the $c = 0$ plane ($b > 0$, $\alpha_2 < 0$ and $b < 0$, $\alpha_2 > 0$, respectively). The domains D_1 and D_2 do not contain Oc axis. There are three equilibria only for points of the parameter space situated outside the domains D_1 and D_2 .

System (1.1.1) can have one equilibrium only in the following three cases:

1) Points situated on the Oc axis without origin. In this case the equilibrium is $\overline{u}_2 = (1, 0)$;

2) The $c = 0$ plane without axis. In this case the equilibrium point is $\overline{u}_1 = \left(0, \alpha_2^{-1}\right)$;

3) The $\alpha_2 = 0$ plane. In this case the equilibrium is $\overline{u}_2 = (1, -b/c)$.

To points of the Ob axis without origin no equilibrium corresponds. For points of the $O\alpha_2$ axis including the origin there are an infinity of equilibria situated on the straight-line $x + \alpha_2 y - 1 = 0$. On the $O\alpha_2$ axis without origin the corresponding equilibria have the form $(x_0, (1 - x_0)/\alpha_2)$. Among them there is $\overline{u}_2 = (1, 0)$ (corresponding to $x_0 = 1$), $\overline{u}_1 = \overline{u}_3 = \left(0, \alpha_2^{-1}\right)$ (corresponding to $x_0 = 0$) and $\overline{u}_2 = \overline{u}_3 = (1/2, 1/2\alpha_2)$ (corresponding to $x_0 = 1/2$). It follows that to the points of the $O\alpha_2$ axis without origin the same equilibria $\overline{u}_1$and $\overline{u}_2 = \overline{u}_3$ correspond as for the points of S.

The half axes $\alpha_2 > 0$ and $\alpha_2 < 0$ consist of accumulation points for S. This is true both when S is considered as a topologic subspace of R^3 and when S possesses the above property concerning the equilibria (i.e. $\overline{u}_2 = \overline{u}_3$).

However, for points of the $O\alpha_2$ axis without origin, apart from the equilibria $\overline{u}_1$ and $\overline{u}_2 = \overline{u}_3$, there exists an infinity of other equilibria depending on the initial datum.

Finally, to the origin of the parameters space an infinity of equilibria of the form $(1, y_0)$corresponds. Among them, there is also the point $\overline{u}_2 = (1, 0)$ (corresponding to $y_0 = 0$).

2 Nonhyperbolic singularities of Hopf type

2.1 Normal forms

Using the eigenvalues and the eigenvectors of the nonhyperbolic point of equilibrium $\overline{\mathbf{u}}_3$ corresponding to the values of parameters $\alpha_2 = b$, $c < -4b^2$, we put the system (1.1.1) in the normal form and emphasises that it corresponds to a degenerated Hopf singularity.

Proposition 2.2.1. *Up to terms of degree greater than three system*

$$\begin{cases} \dot{x} = cx^2y + bx, \\ \dot{y} = x + by - 1 \end{cases} \tag{2.1.1}$$

has around the equilibrium $\overline{\mathbf{u}}_3 = (\frac{c-\sqrt{\Delta}}{2c}, \frac{c+\sqrt{\Delta}}{2bc}) \overset{not}{=} (u, v)$, *where* $\Delta = c^2 + 4b^2c$, *the normal form*

$$\begin{cases} \dot{x}_5 = irx_5 + Cx_5^2y_5, \\ \dot{y}_5 = -iry_5 + \overline{C}x_5y_5^2. \end{cases}$$

Proof. We carry the point $\overline{\mathbf{u}}_3$ at the origin by means of the change of coordinates $x_1 = x - u$, $y_1 = y - v$. Then (2.1.1) becomes

$$\begin{cases} \dot{x}_1 = -bx_1 + cu^2y_1 + cvx_1^2 + 2cux_1y_1 + cx_1^2y_1, \\ \dot{y}_1 = x_1 + by_1. \end{cases} \tag{2.1.1}'$$

The matrix associated to the system linearized around the point $(x_1, y_1) = (0, 0)$ is $Q = \begin{pmatrix} -b & cu^2 \\ 1 & b \end{pmatrix}$ and it admits the purely imaginary eigenvalues $s_{1,2} = \pm i\sqrt{\frac{-c-4b^2+\sqrt{c^2+4bc}}{2}} \overset{not}{=} \pm ir$. Hence $\overline{\mathbf{u}}_3$ is a Hopf singularity. Let $\overline{\mathbf{p}} = (s_1 - b, 1)$ be an eigenvector of Q corresponding to the eigenvalue $s_1 = ir$. Then, $\overline{\mathbf{p}}$ may be written in the form $\overline{\mathbf{p}} = \overline{\mathbf{q}} + i\overline{\mathbf{t}}$ where $\overline{\mathbf{q}} = (-b, 1)$,and $\overline{\mathbf{t}} = (r, 0)$. The matrix $P = \begin{pmatrix} r & -b \\ 0 & 1 \end{pmatrix}$ is nonsingular and so, we may perform the transformation

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \frac{1}{r} \begin{pmatrix} 1 & b \\ 0 & r \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

and obtain a system in (x_2, y_2). As the linearized system corresponding at (x_2, y_2) has not a matrix in a diagonal form we perform the change $\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = M_c \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ where $M_c = \begin{pmatrix} 0,5 & 0,5 \\ -0,5i & 0,5i \end{pmatrix}$. Therefore, $\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = M_cP^{-1} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$. As M_c is a matrix in the complex field it follows that $x_3, y_3 \in \mathbf{C}$ namely $x_3 = \overline{y}_3$, ($\overline{y}_3$ is the complex conjugate of y_3). We have $x_3 = \frac{1}{r}[x_1 + (b + ir)y_1]$, $y_3 = \frac{1}{r}[x_1 + (b - ir)y_1]$ or, $x_1 = \frac{r+ib}{2}x_3 + \frac{r-ib}{2}y_3$, $y_1 = \frac{i}{2}(y_3 - x_3)$. Thus, (2.1.1) becomes

$$\begin{cases} \dot{x}_3 = irx_3 + \frac{T}{2}, \\ \dot{y}_3 = -iry_3 + \frac{T}{2}, \end{cases} \tag{2.1.2}$$

where $T = a_{20}x_3^2 + \overline{a}_{20}y_3^2 + a_{11}x_3y_3 + a_{30}x_3^3 + \overline{a}_{30}y_3^3 + a_{21}x_3^2y_3 + \overline{a}_{21}x_3y_3^2$, $a_{20} = \frac{bc-5b\sqrt{\Delta}}{4r} + i\sqrt{\Delta}$, $a_{11} = \frac{b(\sqrt{\Delta}-c)}{2r}$, $a_{21} = \frac{ci(c-2b^2-\sqrt{\Delta}+4irb)}{8r}$, $a_{30} = \frac{-ci(r+ib)^2}{4r}$.

In order to eliminate the nonresonant terms of second degree it is necessary to complete the following table, where $\Lambda_{m,i} = (\mathbf{m} \cdot s) - s_i$ and $h_{m,1} = \frac{X_{m,i}}{(\mathbf{m}\cdot s)-s_i}$, $s = (s_1, s_2)$ [1].

Table 2.1

m_1	m_2	$X_{m,1}$	$X_{m,2}$	$\Lambda_{m,1}$	$\Lambda_{m,2}$	$h_{m,1}$	$h_{m,2}$
2	0	$\frac{a_{20}}{2}$	$\frac{a_{20}}{2}$	ir	$3ir$	$\frac{a_{20}}{2ir}$	$\frac{a_{20}}{6ir}$
1	1	$\frac{a_{11}}{2}$	$\frac{a_{11}}{2}$	$-ir$	ir	$-\frac{a_{11}}{2ir}$	$\frac{a_{11}}{2ir}$
0	2	$\frac{\overline{a}_{20}}{2}$	$\frac{\overline{a}_{20}}{2}$	$-3ir$	$-ir$	$-\frac{\overline{a}_{20}}{6ir}$	$-\frac{\overline{a}_{20}}{2ir}$

It follows the transformation

$$\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} + \begin{pmatrix} \frac{a_{20}}{2ir}x_4^2 - \frac{a_{11}}{2ir}x_4y_4 - \frac{\overline{a}_{20}}{6ir}y_4^2 \\ \frac{a_{20}}{6ir}x_4^2 + \frac{a_{11}}{2ir}x_4y_4 - \frac{a_{20}}{2ir}y_4^2 \end{pmatrix},$$

which introduced in (2.1.2), leads to the system

$$\begin{cases} \dot{x}_4 = irx_4 + Ax_4^3 + \overline{A}y_4^3 + Cx_4^2y_4 + \overline{C}x_4y_4^2, \\ \dot{y}_4 = -iry_4 + Ax_4^3 + \overline{A}y_4^3 + Cx_4^2y_4 + \overline{C}x_4y_4^2, \end{cases} \tag{2.1.3}$$

where $A = \frac{6a_{20}^2+a_{11}a_{20}+6ira_{30}}{12ir}$ and $C = \frac{2a_{20}\overline{a}_{20}-3a_{11}a_{20}+3a_{11}^2+6ira_{21}}{12ir}$.

In order to reduce the nonresonant terms of order three (2.1.3) we use the table.

Table 2.2

m_1	m_2	$X_{m,1}$	$X_{m,2}$	$\Lambda_{m,1}$	$\Lambda_{m,2}$	$h_{m,1}$	$h_{m,2}$
3	0	A	A	$2ir$	$4ir$	$\frac{A}{2ir}$	$\frac{A}{4ir}$
2	1	C	C	0	$2ir$	–	$\frac{C}{2ir}$
1	2	$\overline{C}$	$\overline{C}$	$-2ir$	0	$-\frac{\overline{C}}{2ir}$	–
0	3	$\overline{A}$	$\overline{A}$	$-4ir$	$-2ir$	$-\frac{\overline{A}}{4ir}$	$-\frac{\overline{A}}{2ir}$

Thus we obtain the transformation

$$\begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} x_5 \\ y_5 \end{pmatrix} + \begin{pmatrix} \frac{A}{2ir}x_5^3 - \frac{\overline{C}}{2ir}x_5y_5^2 - \frac{\overline{A}}{4ir}y_5^3 \\ \frac{A}{4ir}x_5^3 + \frac{C}{2ir}x_5^2y_5 - \frac{\overline{A}}{2ir}y_5^3 \end{pmatrix}$$

leading to the system

$$\begin{cases} \dot{x}_5 = irx_5 + Cx_5^2y_5, \\ \dot{y}_5 = -iry_5 + \overline{C}x_5y_5^2. \end{cases} \tag{2.1.4}$$

In this system we retained terms up to the third degree. Thus we obtained the normal form in **C**. Obviously the second equation is the conjugate of the first, therefore, up to terms of the third degree the normal form is $(2.1.4)_1$.

Theorem 2.1.1 *The Hopf singularity* $\overline{\mathbf{u}}_3$ *is degenerated of order* $k \geq 2$.

Proof. Taking into account the expressions of a_{20}, a_{11}, a_{21}, r, Δ a direct computation leads us to the expression of C :

$C = -\frac{ic}{48r^3}\left(16b^4 + 5b^2c - c^2 + c\sqrt{\Delta} - 7b^2\sqrt{\Delta}\right)$. Since $c < -4b^2$ it follows that $16b^4 + 5b^2c - c^2 + c\sqrt{\Delta} - 7b^2\sqrt{\Delta} < -4b^2 - c^2 - 3b^2\sqrt{\Delta} < 0$, hence $C \neq 0$ and C is purely imaginary. Then (2.1.4) has the follow normal form, according to Arrowsmith [1]:

$$\begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sum_{k=1}^{[(N-1)/2]} (y_1^2 + y_2^2)^k \left\{ a_k \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + b_k \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} \right\} +$$

$O\left(|y|^{N+1}\right)$, $\beta = \sqrt{\det A}, N \geq 3$, [.] represents integer part and a_k, $b_k \in \mathbf{R}$

where $a_1 = 0$ and $b_1 = \operatorname{Im} C \neq 0$, whence the conclusion of the theorem.(Figure 1)

Corollary 2.1.1 *The first Liapunov coefficient associated to system (2.1.1)′ is null (*$\operatorname{Re} C = 0$*).*

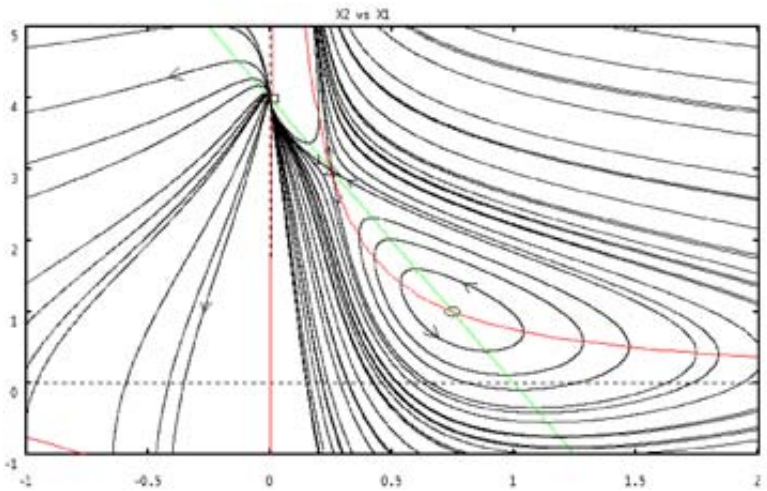


Figure 2: Local phase portrait in the degenerated Hopf bifurcation, for $\alpha_2 = b = -0.1$, $c = -0.04$.

2.2 Computation of the Liapunov coefficients

Proposition 2.2.1. *The system (2.1.1)′ is topologically equivalent to system*

$$\begin{cases} \dot{x}_2 = -ry_2 + cvrx_2^2 + (2cu - 2vb)x_2y_2 + \frac{cvb^2 - 2cub}{r}y_2^2 + crx_2^2y_2 \\ -2cbx_2y_2^2 + \frac{cb^2}{r}y_2^3, \\ \dot{y}_2 = rx_2. \end{cases} \tag{2.2.1}$$

Proof. The transformation of coordinates

$$\begin{cases} x_1 = rx_2 - by_2, \\ y_1 = y_2 \end{cases}$$

carries system (2.1.1)′ in (2.2.1). In this conditions, according to Chow and Wang [2] there exists a smooth function $F(x) = \frac{r}{2}(x_2^2 + y_2^2) + O(|x, y|^3)$ such that

$$\langle \text{gradF}, \text{X}_0 \rangle = \sum_{i=1}^{m} V_i(x_2^2 + y_2^2)^{i+1} + O(|x, y|^{m+1}) \tag{2.2.2}$$

where X_0 is the vector field corresponding to (2.2.1), and V_i are the Liapunov coefficients.

Proposition 2.2.2. *For the system (2.2.1) we have* $V_1 = 0$ *and*
$V_2 = -\frac{b^2c^2}{24r^2}\Big[13c^2 + 78b^2c + 5bc + 104b^4 - 90b^3 - (3b + 52b^2 + 13c)\sqrt{c^2 + 4b^2c}\Big]$.

Proof. We look for F to the form $F(x) = \frac{r}{2}(x_2^2 + y_2^2) + \sum_{i+j=k}\sum_{k>3} c_{ij}x_2^i y_2^j$.

Therefore, we have

$$\begin{aligned}
\langle \mathrm{gradF}, \mathrm{X_0}\rangle \;=\; & r^2x_2y_2 + r\sum_{i+j=3} jc_{ij}x_2^{i+1}y_2^{j-1} - r^2x_2y_2 - r\sum_{i+j=4} ic_{ij}x_2^{i-1}y_2^{j+1} + \\
& cvr^2x_2^3 + cvr\sum_{i+j=3} ic_{ij}x_2^{i+1}y_2^j + 2c(u - vb)rx_2^2y_2 + \\
& 2c(u - vb)\sum_{i+j=4} ic_{ij}x_2^i y_2^{j+1} + \frac{cvb^2 - 2cub}{r}rx_2y_2^2 + \\
& \frac{cvb^2 - 2cub}{r}y_2^2\sum_{i+j=3} ic_{ij}x_2^{i-1}y_2^{j+2} + cr^2x_2^3y_2 + \\
& cr\sum_{i+j=4} ic_{ij}x_2^{i+1}y_2^{j+1} - 2cbrx_2^2y_2^2 - 2cb\sum_{i+j=3} ic_{ij}x_2^i y_2^{j+2} + \\
& cb^2x_2y_2^3 + \frac{cb^2}{r}\sum_{i+j=4} ic_{ij}x_2^{i-1}y_2^{j+3}
\end{aligned}$$

Identifying the monomials of degree three coefficients in (2.2.2) we obtain the following system in unknowns c_{ij}, $i + j = 3$:

$$\begin{cases}
rc_{21} + cvr^2 = 0, \\
-rc_{12} = 0, \\
3rc_{03} - 2rc_{21} + cvb^2 - 2cub = 0, \\
2rc_{12} - 3rc_{30} + 2cru - 2cvbr = 0.
\end{cases}$$

The solution of this system reads $c_{12} = 0$, $c_{21} = -cvr$, $c_{03} = \frac{2cub - cvb^2 - 2r^2cv}{3r}$, $c_{30} = \frac{2cu - 2cvb}{3}$.

Identifying the monomials of degree four coefficients in (2.2.2) we obtain the following system in unknowns c_{ij}, $i + j = 4$ and V_1

$$\begin{cases} rc_{31}+3cvrc_{30}=V_1, \\ -rc_{13}+\frac{cvb^2-2cub}{r}c_{12}=V_1, \\ 4rc_{04}-2rc_{22}+2c(u-vb)c_{12}+2\frac{cvb^2-2cub}{r}c_{21}+cb^2=0, \\ 3rc_{13}-3rc_{31}+cvrc_{12}+4c(u-vb)c_{21}+3\frac{cvb^2-2cub}{r}c_{30}-2cbr=2V_1, \\ 2rc_{22}-4rc_{40}+2cvrc_{21}+6c(u-vb)c_{30}+cr^2=0, \end{cases}$$

the solution of which is $V_1=0$, $c_{13}=0$, $c_{31}=-2c^2uv+2c^2v^2b$, $c_{04}=\frac{c_{22}}{2}+\frac{2c^2v^2b^2-4c^2uvb-cb^2}{4r}$, $c_{40}=\frac{c_{22}}{2}+\frac{4c^2u^2-8c^2uvb+4c^2v^2b^2-2c^2v^2r^2+cr^2}{4r}$.

Remark 2.2.1. The result $V_1=0$ represents a new proof for Theorem 2.1.1.

Identifying the monomials of degree five coefficients in (2.2.2) we obtain the following system in unknowns c_{ij}, $i+j=5$

$$\begin{cases} rc_{41}+4cvrc_{40}=0, \\ rc_{14}+\frac{cvb^2-2cub}{r}c_{13}+\frac{cb^2}{r}c_{12}=0, \\ 5rc_{05}-2rc_{23}+2c(u-vb)c_{13}+2\frac{cvb^2-2cub}{r}c_{22}-2bcc_{12}+\frac{2cb^2}{r}c_{21}=0 \\ 4rc_{14}-3rc_{32}+cvrc_{13}+4c(u-vb)c_{22}+3\frac{cvb^2-2cub}{r}c_{31}+crc_{12} \\ -4cbc_{21}+\frac{3cb^2}{r}c_{30}=0, \\ 3rc_{23}-4rc_{41}+2cvrc_{22}+6c(u-vb)c_{31}+4\frac{cvb^2-2cub}{r}c_{40} \\ +2crc_{21}-6bcc_{30}=0, \\ 2rc_{32}-5rc_{50}+3cvrc_{31}+8c(u-vb)c_{40}+3crc_{30}=0 \end{cases}$$

whence

$c_{14}=0$,

$c_{41}=-4cvc_{40}$,

$c_{23}=\left[\frac{1}{3r}(-16cvr-4\frac{cvb^2-2cub}{r})c_{40}-2cvrc_{22}-6c(u-vb)c_{31}-2crc_{21}\right.$
$\left.+6cbc_{30}\right]$,

$c_{32}=\left[\frac{1}{3r^2}4cr(u-vb)c_{22}+6c^3v^3b^3-18c^3v^2ub^2+12c^3vu^2b+4c^2vbr^2\right.$
$\left.+2c^2ub^2-2c^2vb^3\right]$,

$c_{50}=\frac{1}{5r}\left[2rc_{32}+3cvrc_{31}+8c(u-vb)c_{40}+3crc_{30}\right]$,

$c_{05}=\frac{2}{5}c_{23}-\frac{2cb^2}{5r^2}c_{21}-\frac{2}{5r^2}(cvb^2-2cub)c_{22}$.

By identifying the coefficients of the monomials of degree six in (2.2.2) we obtain the following system in unknowns V_2 and c_{ij}, $i+j=6$

$$\begin{cases} rc_{51}+5cvrc_{50}=V_2, \\ -rc_{15}+\frac{cvb^2-2cub}{r}c_{14}+\frac{cb^2}{r}c_{13}=V_2 \\ 2rc_{42}-6rc_{60}+4cvrc_{41}+10c(u-vb)c_{50}+4crc_{40}=0 \\ 3rc_{33}-5rc_{51}+3cvrc_{32}+8c(u-vb)c_{41}+5\frac{cvb^2-2cub}{r}c_{50} \\ +3crc_{31}-8cbc_{40}=3V_2, \\ 4rc_{24}+2cvrc_{23}+6c(u-vb)c_{32}-4rc_{42}+4\frac{cvb^2-2cub}{r}c_{41}+ \\ 2crc_{22}-6bcc_{31}+\frac{4cb^2}{r}c_{40}=0, \\ 5rc_{15}-3rc_{33}+cvrc_{14}+4c(u-vb)c_{23}+3\frac{cvb^2-2cub}{r}c_{32}+crc_{13} \\ -4cbc_{22}+\frac{3cb^2}{r}c_{31}=3V_2, \\ 6rc_{06}-2rc_{24}+2c(u-vb)c_{14}+2\frac{cvb^2-2cub}{r}c_{23}-2bcc_{13}+\frac{2cb^2}{r}c_{22}=0. \end{cases}$$

Since $c_{15}=-\frac{V_2}{r}$, $c_{51}=\frac{V_2}{r}-5cvrc_{50}$, by replacing the found value for $3rc_{33}$ from the sixth equation in the fourth equation and taking into account the found values for c_{ij}, $i+j=5$, we have

$$\begin{aligned} V_2 &= -\frac{b^2c^2}{24r^2}13c^2+78b^2c+5bc+104b^4-90b^3 \\ &\quad -(3b+52b^2+13c)\sqrt{c^2+4b^2c}. \end{aligned}$$

The set $V_2=0$ intersects the domain $\left\{\alpha_2=b, c<-4b^2\right\}$ along a curve γ_3 the existence of which is proved by studying the sign of V_2 in the domain considered, and also by numerical methods (figure 3).

It is important to remark that the equilibrium $\bar{\mathbf{u}}_3$ exists also on the half-axis $c<0$ and in this case $b=0$ implies $V_2=0$. The curve γ_3 and the negative half-axis $c<0$ divide the interior of the parabola $\alpha_2=b$, $c=-4b^2$ in three regions:

$U_1=\left\{(b,b,c)|\ c<0, -\frac{\sqrt{-c}}{2}<b<0\right\}$,

$U_2=\left\{(b,b,c)|\ c<0, 0<b<b(c)\right\}$,

$U_3=\left\{(b,b,c)|\ c<0, b(c)<b<\frac{\sqrt{-c}}{2}\right\}$,

where $\alpha_2=b$, $b=b(c)$ are the equations of the curve γ_3. It is easy to see that $V_2<0$ on $U_1\cup U_2$, $V_2>0$ on U_3. As $V_2=0$ on γ_3 and on the half-axis $c<0$ it follows

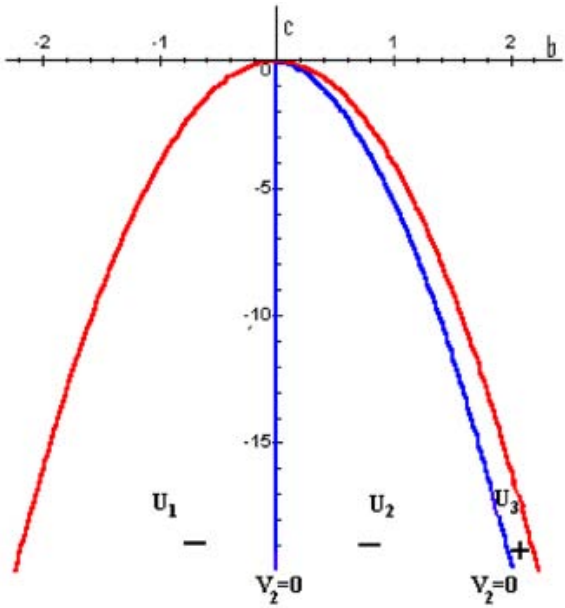


Figure 3: The sign of V_2 in $\{\alpha_2 = b, c < -4b^2\}$.

Theorem 2.2.1 *The point of equilibrium* $\bar{\mathbf{u}}_3$ *is locally a Bautin bifurcation with the Liapunov coefficient* $V_2 < 0$ *for* $(\alpha_2, b, c) \in U_1 \cup U_2$*, a Bautin bifurcation with the Liapunov coefficient* $V_2 > 0$ *for* $(\alpha_2, b, c) \in U_3$*, and a degenerated Hopf bifurcation of order* $k \geq 3$ *for a point of* γ_3 *or of the negative half-axis* $c < 0$.

2.3 Conclusions

From economic point of view, the variation of capital K and labor L over time it can be observed, starting from the initial significant data corresponding to some points in the parameter space. Therefore, there are situations when the system considered enable a periodic solution appropriate to a cyclical economic evolution. Negative phenomena such as production shortage and increase of unemployment rate and also the positive ones, featured by refurbishment of production capacities that could induce the growth of demand for consumption goods and determination of employment level, can be relieved.

Acknowledgement. The presented work has been conducted in the context of the GRANT-PCE-II-2008-2-IDEAS-450, The Management of Knowledge for the Virtual Organization. Possibilities to Support Management using Systems Based on Knowledge funded by The National University Research Council from Romania, under the contract 954/2009.

References

[1] Arrowsmith A.K., Place C.M., *An introduction to dynamical systems*, Cambridge University Press, 1990;

[2] Chow S.N., Li C., Wang D., Normal forms and bifurcation of planar vector fields, Cambridge University Press, 1994;

[3] Ermentrout B, XPPAUT, http://www.math.pitt.edu/bard/xpp/html

[4] Ermentrout B, Simulating, analyzing, and animating dynamical systems: a guide to xppaut for researchers and students, SIAM 2002

[5] Linch S, Dynamical Systems with Applications using Maple 2nd Edition, companion software, Birkhäuser 2009, http://www.maplesoft.com/applications

[6] Oprescu Gh., *Mathematics for economists*, Ed. Fundatiei "Romania de Maine", Bucharest, 1996 (Romanian);

[7] Ungureanu L. - *Structural Stability and Bifurcation in two Models of Economic Dynamics*, Pitesti University Press, Pitesti, 2004 (Romanian).

Annals of the Academy of Romanian Scientists
Series on Mathematics and its Applications
ISSN 2066 - 6594 Volume 3, Number 1 / 2011

In Memoriam Adelina Georgescu

UNIVERSAL REGULAR AUTONOMOUS ASYNCHRONOUS SYSTEMS: $\omega-$LIMIT SETS, INVARIANCE AND BASINS OF ATTRACTION*

Şerban Vlad†

Abstract

The asynchronous systems are the non-deterministic real time-binary models of the asynchronous circuits from electrical engineering. Autonomy means that the circuits and their models have no input. Regularity means analogies with the dynamical systems, thus such systems may be considered to be the real time dynamical systems with a 'vector field' $\Phi : \{0,1\}^n \to \{0,1\}^n$. Universality refers to the case when the state space of the system is the greatest possible in the sense of the inclusion. The purpose of this paper is that of defining, by analogy with the dynamical systems theory, the $\omega-$limit sets, the invariance and the basins of attraction of the universal regular autonomous asynchronous systems.

MSC: 94C10

keywords: asynchronous system, $\omega-$limit set, invariance, basin of attraction

*Accepted for publication on December 16, 2010.

†serban_e_vlad@yahoo.com, Str. Zimbrului, Nr. 3, Bl. PB68, Ap. 11, 410430, Oradea, Romania, web page: www.serbanvlad.ro

1 Introduction

We denote by $\mathbf{B} = \{0, 1\}$ the binary Boole algebra, endowed with the discrete topology and with the usual algebraical laws:

	$-$		$\cdot$	0	1		$\cup$	0	1		$\oplus$	0	1
0	1	,	0	0	0	,	0	0	1	,	0	0	1
1	0		1	0	1		1	1	1		1	1	0

Table 1

The real numbers set $\mathbf{R}$ is the time set and $t \in \mathbf{R}$ are the time instants.

The $\mathbf{R} \to \mathbf{B}$ functions give the deterministic[1] real time-binary models of the digital electrical signals and they are not studied in literature. An asynchronous circuit without input, considered as a collection of n signals, should be deterministically modelled by a function $x : \mathbf{R} \to \mathbf{B}^n$ called state. We have however several parameters related with the asynchronous circuit that are either unknown, or perhaps variable or simply ignored in modeling such as the temperature, the tension of the mains and the delays that occur in the computation of the Boolean functions. For this reason, instead of a function x we have in general a set X of functions x, called state space, or non-deterministic[2] autonomous asynchronous system, where each function x represents a possibility of modeling the circuit. When X is constructed by making use of a 'vector field' $\Phi : \mathbf{B}^n \to \mathbf{B}^n$, the system X is called regular. The universal regular autonomous asynchronous systems are the Boolean dynamical systems and they can be identified with Φ.

We give in Figure 1 at a) the example of the NAND gate defined by $\phi : \mathbf{B}^2 \to \mathbf{B}$, $\forall(\mu_1, \mu_2) \in \mathbf{B}^2, \phi(\mu_1, \mu_2) = \overline{\mu_1\mu_2}$ and at b) the example of an autonomous circuit made with two such devices and characterized by $\Phi : \mathbf{B}^2 \to \mathbf{B}^2$, $\forall(\mu_1, \mu_2) \in \mathbf{B}^2, (\Phi_1(\mu_1, \mu_2), \Phi_2(\mu_1, \mu_2)) = (\overline{\mu_2}, \overline{\mu_1\mu_2})$.

The dynamics of these asynchronous systems[3] is described by the so called state portraits, see Figure 1 c) where the arrows show the increase of time. For any $i \in \{1, 2\}$, the coordinate μ_i is underlined if $\Phi_i(\mu_1, \mu_2) \neq \mu_i$ and it is called unstable, or enabled, or excited in this case. The coordinates μ_i

[1] 'Deterministic' means that each signal is modeled by exactly one $\mathbf{R} \to \mathbf{B}$ function.

[2] 'Non-deterministic' means that each signal is modeled by several $x_i : \mathbf{R} \to \mathbf{B}$ functions or, equivalently, that each circuit is modeled by several functions $x \in X$.

[3] The systems are (vaguely) the models of the circuits.

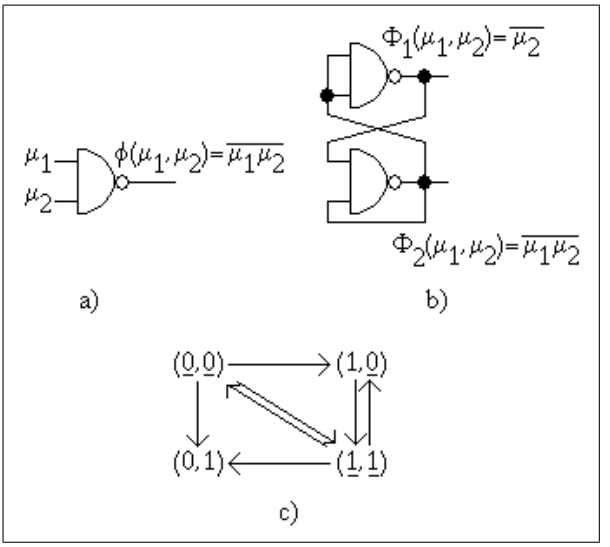


Figure 1: a) The NAND gate, b) Example of system using the NAND gate, c) The state portrait of the system from b)

that are not underlined satisfy by definition $\Phi_i(\mu_1, \mu_2) = \mu_i$ and are called stable, or disabled, or not excited. Three arrows start from the point $(0, 0)$ where both coordinates are unstable, showing the fact that $\Phi_1(0, 0)$ may be computed first, $\Phi_2(0, 0)$ may be computed first or $\Phi_1(0, 0), \Phi_2(0, 0)$ may be computed simultaneously and similarly for the point $(1, 1)$. Note that the two possibilities of defining the system, state portrait and formula, are equivalent. Note also that the system was identified with the function Φ.

The existence of several possibilities of changing the state of the system (three possibilities in $(0, 0)$ and $(1, 1)$, one possibility in $(1, 0)$, no possibility in $(0, 1)$) is the key characteristic of asynchronicity, as opposed to synchronicity where the coordinates $\Phi_i(\mu)$ are always computed simultaneously, $i \in \{1, ..., n\}$ for all $\mu \in \mathbf{B}^n$ and the system's run is: $\mu, \Phi(\mu), (\Phi \circ \Phi)(\mu), ..., (\Phi \circ ... \circ \Phi)(\mu), ...$

Our present aim is to show how the well-known concepts of $\omega-$limit set, invariance and basin of attraction from the dynamical systems theory, by real to binary translation, may be integrated in the asynchronous systems theory.

2 Preliminaries

Notation 1. *$\chi_A : \mathbf{R} \to \mathbf{B}$ is the notation of the characteristic function of the set $A \subset \mathbf{R}$: $\forall t \in \mathbf{R}$, $\chi_A(t) = \begin{cases} 1, t \in A \\ 0, t \notin A \end{cases}$.*

Notation 2. *We denote by Seq the set of the sequences $t_0 < t_1 < ... < t_k < ...$ of real numbers that are unbounded from above.*

Definition 3. *The sequence $\alpha : \mathbf{N} \to \mathbf{B}^n, \forall k \in \mathbf{N}, \alpha^k = \alpha(k)$ is called **progressive** if the sets $\{k | k \in \mathbf{N}, \alpha_i^k = 1\}$ are infinite for all $i \in \{1, ..., n\}$. We denote the set of the progressive sequences by Π_n.*

Definition 4. *The functions $\rho : \mathbf{R} \to \mathbf{B}^n$ of the form $\forall t \in \mathbf{R}$,*

$$\rho(t) = \alpha^0 \chi_{\{t_0\}}(t) \oplus \alpha^1 \chi_{\{t_1\}}(t) \oplus ... \oplus \alpha^k \chi_{\{t_k\}}(t) \oplus ... \tag{1}$$

*where $\alpha \in \Pi_n$ and $(t_k) \in Seq$ are called **progressive** and their set is denoted by P_n.*

Definition 5. *Let be the function $\Phi : \mathbf{B}^n \to \mathbf{B}^n$. i) For $\nu \in \mathbf{B}^n$ we define $\Phi^\nu : \mathbf{B}^n \to \mathbf{B}^n$ by $\forall \mu \in \mathbf{B}^n$, $\Phi^\nu(\mu) = (\overline{\nu_1}\mu_1 \oplus \nu_1 \Phi_1(\mu), ..., \overline{\nu_n}\mu_n \oplus \nu_n \Phi_n(\mu))$.*

ii) The functions $\Phi^{\alpha^0 ... \alpha^k} : \mathbf{B}^n \to \mathbf{B}^n$ are defined for $k \in \mathbf{N}$ and $\alpha^0, ..., \alpha^k \in \mathbf{B}^n$ iteratively: $\forall \mu \in \mathbf{B}^n$, $\Phi^{\alpha^0 ... \alpha^k \alpha^{k+1}}(\mu) = \Phi^{\alpha^{k+1}}(\Phi^{\alpha^0 ... \alpha^k}(\mu))$.

*iii) The function $\Phi^\rho : \mathbf{B}^n \times \mathbf{R} \to \mathbf{B}^n$ that is defined in the following way $\Phi^\rho(\mu, t) = \mu \chi_{(-\infty, t_0)}(t) \oplus \Phi^{\alpha^0}(\mu) \chi_{[t_0, t_1)}(t) \oplus \Phi^{\alpha^0 \alpha^1}(\mu) \chi_{[t_1, t_2)}(t) \oplus ... \oplus \Phi^{\alpha^0 ... \alpha^k}(\mu) \chi_{[t_k, t_{k+1})}(t) \oplus ...$ is called **flow**, **motion** or **orbit** (of $\mu \in \mathbf{B}^n$). We have assumed that $\rho \in P_n$ is like at (1).*

*iv) The set $Or_\rho(\mu) = \{\Phi^\rho(\mu, t) | t \in \mathbf{R}\}$ is also called **orbit** (of μ).*

Remark 6. *The function Φ^ν shows how an asynchronous iteration of Φ is made: for any $i \in \{1, ..., n\}$, if $\nu_i = 0$ then Φ_i is not computed, since $\Phi_i^\nu(\mu) = \mu_i$ and if $\nu_i = 1$ then Φ_i is computed, since $\Phi_i^\nu(\mu) = \Phi_i(\mu)$.*

The definition of $\Phi^{\alpha^0 ... \alpha^k}$ generalizes this idea to an arbitrary number $k+1$ of asynchronous iterations, with the supplementary request that each coordinate Φ_i is computed infinitely many times in the sequence $\mu, \Phi^{\alpha^0}(\mu), \Phi^{\alpha^0 \alpha^1}(\mu), ..., \Phi^{\alpha^0 ... \alpha^k}(\mu), ...$ whenever $\alpha \in \Pi_n$.

The sequences $(t_k) \in Seq$ make the pass from the discrete time $\mathbf{N}$ to the continuous time $\mathbf{R}$ and each $\rho \in P_n$ shows, in addition to $\alpha \in \Pi_n$, the time instants t_k when Φ is computed (asynchronously). Thus $\Phi^\rho(\mu, t), t \in \mathbf{R}$ is

the continuous time computation of the sequence μ, $\Phi^{\alpha^0}(\mu)$, $\Phi^{\alpha^0\alpha^1}(\mu)$, ..., $\Phi^{\alpha^0...\alpha^k}(\mu)$, ... *made in the following way: if* $t < t_0$ *nothing is computed, if* $t \in [t_0, t_1)$, $\Phi^{\alpha^0}(\mu)$ *is computed, if* $t \in [t_1, t_2), \Phi^{\alpha^0\alpha^1}(\mu)$ *is computed, ..., if* $t \in [t_k, t_{k+1}), \Phi^{\alpha^0...\alpha^k}(\mu)$ *is computed, ...*

When α *runs in* Π_n *and* (t_k) *runs in Seq we get the 'unbounded delay model' of computation of the Boolean function* Φ*, represented in discrete time by the sequences* $\mu, \Phi^{\alpha^0}(\mu), \Phi^{\alpha^0\alpha^1}(\mu), ..., \Phi^{\alpha^0...\alpha^k}(\mu), ...$ *and in continuous time by the orbits* $\Phi^\rho(\mu, t)$ *respectively. We shall not insist on the non-formalized way that the engineers describe this model; we just mention that the 'unbounded delay model' is a reasonable way of starting the analysis of a circuit in which the delays occurring in the computation of the Boolean functions* Φ *are arbitrary positive numbers. If we restrict suitably the ranges of* α *and* (t_k) *we get the 'bounded delay model' of computation of* Φ *and if both* α, (t_k) *are fixed, then we obtain the 'fixed delay model' of computation of* Φ*, determinism.*

Theorem 7. *Let* $\alpha \in \Pi_n, (t_k) \in Seq$ *be arbitrary and the function* $\rho(t) = \alpha^0\chi_{\{t_0\}}(t) \oplus \alpha^1\chi_{\{t_1\}}(t) \oplus ... \oplus \alpha^k\chi_{\{t_k\}}(t) \oplus ..., \rho \in P_n$*. The following statements are true:*

a) $\{\alpha^k | k \geq k_1\} \in \Pi_n$ *for any* $k_1 \in \mathbf{N}$;

b) $(t_k) \cap (t', \infty) \in Seq$ *for any* $t' \in \mathbf{R}$;

c) $\rho\chi_{(t',\infty)} \in P_n$ *for any* $t' \in \mathbf{R}$;

d) $\forall \mu \in \mathbf{B}^n, \forall \mu' \in \mathbf{B}^n, \forall t' \in \mathbf{R}, \Phi^\rho(\mu, t') = \mu' \Longrightarrow \forall t \geq t', \Phi^\rho(\mu, t) = \Phi^{\rho\chi_{(t',\infty)}}(\mu', t)$.

Proof. a) If $\{k | k \in \mathbf{N}, \alpha_i^k = 1\}$ is infinite, then $\{k | k \geq k_1, \alpha_i^k = 1\}$ is also infinite, $\forall i \in \{1, ..., n\}$.

b) If $t_0 < t_1 < t_2 < ...$ is unbounded from above, then any sequence of the form $t_{k_1} < t_{k_1+1} < t_{k_1+2} < ...$ is unbounded from above, $k_1 \in \mathbf{N}$.

c) This is a consequence of a) and b).

d) We presume that $t' < t_0$. In this situation $\mu = \mu', \rho = \rho\chi_{(t',\infty)}$ and the statement is obvious, so that we may assume now that $t' \geq t_0$. In this case, some $k_1 \in \mathbf{N}$ exists with $t' \in [t_{k_1}, t_{k_1+1})$ and $\mu' = \Phi^{\alpha^0...\alpha^{k_1}}(\mu)$. Because

$$\rho\chi_{(t',\infty)}(t) = \alpha^{k_1+1}\chi_{\{t_{k_1+1}\}}(t) \oplus \alpha^{k_1+2}\chi_{\{t_{k_1+2}\}}(t) \oplus ...,$$

$$\Phi^{\rho\chi_{(t',\infty)}}(\mu', t) = \mu'\chi_{(-\infty,t_{k_1+1})}(t) \oplus \Phi^{\alpha^{k_1+1}}(\mu')\chi_{[t_{k_1+1},t_{k_1+2})}(t)$$
$$\oplus \Phi^{\alpha^{k_1+1}\alpha^{k_1+2}}(\mu')\chi_{[t_{k_1+2},t_{k_1+3})}(t) \oplus ...$$

we get

$\forall t \in [t', t_{k_1+1})$,

$$\Phi^{\rho}(\mu, t) = \Phi^{\alpha^0 \dots \alpha^{k_1}}(\mu),$$

$$\Phi^{\rho\chi_{(t',\infty)}}(\mu', t) = \mu' = \Phi^{\alpha^0 \dots \alpha^{k_1}}(\mu);$$

$\forall t \in [t_{k_1+1}, t_{k_1+2})$,

$$\Phi^{\rho}(\mu, t) = \Phi^{\alpha^0 \dots \alpha^{k_1} \alpha^{k_1+1}}(\mu),$$

$$\Phi^{\rho\chi_{(t',\infty)}}(\mu', t) = \Phi^{\alpha^{k_1+1}}(\mu') = \Phi^{\alpha^{k_1+1}}(\Phi^{\alpha^0 \dots \alpha^{k_1}}(\mu)) = \Phi^{\alpha^0 \dots \alpha^{k_1} \alpha^{k_1+1}}(\mu);$$

...

The statement of the Theorem holds. □

Theorem 8. *Let be $\mu \in \mathbf{B}^n, \rho \in P_n$ and $\tau \in \mathbf{R}$. The function $\rho'(t) = \rho(t-\tau)$ is progressive and we have $\Phi^{\rho'}(\mu, t) = \Phi^{\rho}(\mu, t-\tau)$.*

Proof. We put ρ under the form

$$\rho(t) = \alpha^0 \chi_{\{t_0\}}(t) \oplus ... \oplus \alpha^k \chi_{\{t_k\}}(t) \oplus ...,$$

$\alpha \in \Pi_n, (t_k) \in Seq$ and we note that

$$\rho'(t) = \rho(t-\tau) = \alpha^0 \chi_{\{t_0+\tau\}}(t) \oplus ... \oplus \alpha^k \chi_{\{t_k+\tau\}}(t) \oplus ...$$

where $(t_k + \tau) \in Seq$. We infer

$$\Phi^{\rho'}(\mu, t) = \mu\chi_{(-\infty, t_0+\tau)}(t) \oplus \Phi^{\alpha^0}(\mu)\chi_{[t_0+\tau, t_1+\tau)}(t) \oplus ...$$

$$... \oplus \Phi^{\alpha^0 \dots \alpha^k}(\mu)\chi_{[t_k+\tau, t_{k+1}+\tau)}(t) \oplus ... = \Phi^{\rho}(\mu, t-\tau).$$

□

Definition 9. *The **universal regular autonomous asynchronous system** that is generated by $\Phi : \mathbf{B}^n \to \mathbf{B}^n$ is by definition $\Xi_\Phi = \{\Phi^\rho(\mu, \cdot) | \mu \in \mathbf{B}^n, \rho \in P_n\}$; any $x(t) = \Phi^\rho(\mu, t)$ is called **state** (of Ξ_Φ), μ is called **initial value** (of x), or **initial state** (of Ξ_Φ) and Φ is called **generator function** (of Ξ_Φ).*

Remark 10. *The asynchronous systems are non-deterministic in general, due to the uncertainties that occur in the modeling of the asynchronous circuits. Non-determinism is produced, in the case of Ξ_Φ, by the fact that the initial state μ and the way ρ of iterating Φ are not known.*

Definition 11. *Let $v : \mathbf{N} \to \mathbf{B}^n, x : \mathbf{R} \to \mathbf{B}^n$ be some functions. If $\exists k' \in \mathbf{N}, \forall k \geq k', v(k) = v(k')$, we say that **the limit** $\lim\limits_{k\to\infty} v(k)$ **exists** and we use the notation $\lim\limits_{k\to\infty} v(k) = v(k')$. Similarly, if $\exists t' \in \mathbf{R}, \forall t \geq t', x(t) = x(t')$,we say that **the limit** $\lim\limits_{t\to\infty} x(t)$ **exists** and we denote $\lim\limits_{t\to\infty} x(t) = x(t')$. Sometimes $\lim\limits_{k\to\infty} v(k)$, $\lim\limits_{t\to\infty} x(t)$ are called the **final values** of v, x.*

Theorem 12. *[7] $\forall \mu \in \mathbf{B}^n, \forall \mu' \in \mathbf{B}^n, \forall \rho \in P_n, \lim\limits_{t\to\infty} \Phi^\rho(\mu, t) = \mu' \Longrightarrow \Phi(\mu') = \mu'$, if the final value of $\Phi^\rho(\mu, \cdot)$ exists, it is a fixed point of Φ.*

Proof. Let $\mu \in \mathbf{B}^n, \mu' \in \mathbf{B}^n, \rho \in P_n$ be arbitrary and fixed. The hypothesis states the existence of $t' \in \mathbf{R}$ with

$$\forall t \geq t', \Phi^\rho(\mu, t) = \mu'$$

thus, from Theorem 7 d),

$$\forall t \geq t', \Phi^{\rho\chi_{(t',\infty)}}(\mu', t) = \mu'.$$

We infer that $\forall i \in \{1, ..., n\}, \exists t'' > t'$ such that

$$\rho_i(t'') = \rho_i\chi_{(t',\infty)}(t'') = 1,$$

$$\Phi_i^{\rho\chi_{(t',\infty)}}(\mu', t'') = \Phi_i(\mu') = \mu'_i.$$

□

Theorem 13. *[7] $\forall \mu \in \mathbf{B}^n, \forall \mu' \in \mathbf{B}^n, \forall \rho \in P_n, (\Phi(\mu') = \mu'$ and $\exists t' \in \mathbf{R}, \Phi^\rho(\mu, t') = \mu') \Longrightarrow \forall t \geq t', \Phi^\rho(\mu, t) = \mu'$,meaning that if the fixed point μ' of Φ is accessible, then it is the final value of $\Phi^\rho(\mu, \cdot)$.*

Proof. Let $\mu \in \mathbf{B}^n, \mu' \in \mathbf{B}^n, \rho \in P_n$ be arbitrary and fixed. From the hypothesis and Theorem 7 d) we infer

$$\forall t \geq t', \Phi^\rho(\mu, t) = \Phi^{\rho\chi_{(t',\infty)}}(\mu', t)$$

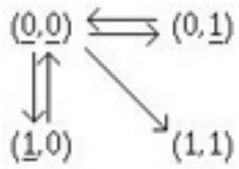


Figure 2: $\exists \rho \in P_2, \omega_\rho((1,0)) = \{(0,0),(0,1)\}$ and $\exists \rho' \in P_2, \omega_{\rho'}((1,0)) = \{(1,1)\}$

thus $\forall i \in \{1,...,n\}, \exists \varepsilon > 0, \forall t \in [t', t'+\varepsilon), \Phi_i^{\rho\chi_{(t',\infty)}}(\mu', t)$ can take one of the values μ'_i and $\Phi_i(\mu')$. But $\mu'_i = \Phi_i(\mu')$, wherefrom the previous property takes place for arbitrary ε and

$$\forall t \geq t', \Phi^\rho(\mu, t) = \mu'.$$

□

Corollary 14. $\forall \mu \in \mathbf{B}^n, \forall \rho \in P_n, \Phi(\mu) = \mu \Longrightarrow \forall t \in \mathbf{R}, \Phi^\rho(\mu, t) = \mu$.

Proof. From Theorem 13, with $\mu = \mu'$, where t' may be chosen such that $\forall t < t', \rho(t) = 0$. □

3 $\omega-$limit sets

Definition 15. *For $\mu \in \mathbf{B}^n$ and $\rho \in P_n$, the set $\omega_\rho(\mu) = \{\mu' | \mu' \in \mathbf{B}^n, \exists (t_k) \in Seq, \lim_{k\to\infty} \Phi^\rho(\mu, t_k) = \mu'\}$ is called the $\omega-$**limit set** of the orbit $\Phi^\rho(\mu, \cdot)$.*

Remark 16. *The previous definition agrees with the usual definitions of the $\omega-$limit sets of the real time or discrete time dynamical systems see [2] page 5, [5] page 26, [1] page 20.*

Example 17. *In Figure 2, we consider*

$$\rho(t) = (1,1)\chi_{\{0\}}(t) \oplus (0,1)\chi_{\{1\}}(t) \oplus (1,1)\chi_{\{2\}}(t) \oplus (0,1)\chi_{\{3\}}(t) \oplus ...,$$

$$\rho'(t) = (1,1)\chi_{\{0\}}(t) \oplus (1,1)\chi_{\{1\}}(t) \oplus (1,1)\chi_{\{2\}}(t) \oplus ...$$

and we have

$$\Phi^\rho((1,0),t) = (1,0)\chi_{(-\infty,0)}(t) \oplus (0,0)\chi_{[0,1)}(t) \oplus (0,1)\chi_{[1,2)}(t)$$

$$\oplus(0,0)\chi_{[2,3)}(t)\oplus(0,1)\chi_{[3,4)}(t)\oplus ...,$$

$$\Phi^{\rho'}((1,0),t)=(1,0)\chi_{(-\infty,0)}(t)\oplus(0,0)\chi_{[0,1)}(t)\oplus(1,1)\chi_{[1,\infty)}(t),$$

thus $\omega_\rho((1,0))=\{(0,0),(0,1)\},\omega_{\rho'}((1,0))=\{(1,1)\}$.

Theorem 18. *For any* $\mu\in\mathbf{B}^n$ *and any* $\rho\in P_n$, *we have:*

a) $\omega_\rho(\mu)\neq\emptyset$;

b) $\forall t'\in\mathbf{R}$, $\omega_\rho(\mu)\subset\{\Phi^\rho(\mu,t)|t\geq t'\}\subset Or_\rho(\mu)$;

c) $\exists t'\in\mathbf{R},\omega_\rho(\mu)=\{\Phi^\rho(\mu,t)|t\geq t'\}$ *and any* $t''\geq t'$ *fulfills* $\omega_\rho(\mu)=\{\Phi^\rho(\mu,t)|t\geq t''\}$;

d) $\forall t'\in\mathbf{R},\forall t''\geq t',\{\Phi^\rho(\mu,t)|t\geq t'\}=\{\Phi^\rho(\mu,t)|t\geq t''\}$ *implies* $\omega_\rho(\mu)=\{\Phi^\rho(\mu,t)|t\geq t'\}$;

e) we presume that $\omega_\rho(\mu)=\{\Phi^\rho(\mu,t)|t\geq t'\},t'\in\mathbf{R}$. *Then* $\forall\mu'\in\omega_\rho(\mu),\forall t''\geq t'$, *if* $\Phi^\rho(\mu,t'')=\mu'$ *we get* $\omega_\rho(\mu)=\{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t\geq t''\}=Or_{\rho\chi_{(t'',\infty)}}(\mu')=\omega_{\rho\chi_{(t'',\infty)}}(\mu')$.

Proof. We put $\rho\in P_n$ under the form

$$\rho(t)=\alpha^0\chi_{\{t_0\}}(t)\oplus ...\oplus\alpha^k\chi_{\{t_k\}}(t)\oplus ...$$

where $\alpha\in\Pi_n$ and $(t_k)\in Seq$. We ask, without loosing the generality, that $\alpha^0=(0,...,0)\in\mathbf{B}^n$, hence $\Phi^\rho(\mu,t_0)=\mu$ and $Or_\rho(\mu)=\{\Phi^\rho(\mu,t_k)|k\in\mathbf{N}\}$.

a) If $Or_\rho(\mu)=\{\mu^1,...,\mu^p\},p\in\{1,...,2^n\}$, we denote with $I_1,...,I_p\subset\mathbf{N}$ the sets

$$I_j=\{k|k\in\mathbf{N},\Phi^\rho(\mu,t_k)=\mu^j\},j=\overline{1,p}.$$

Because $I_1\cup...\cup I_p=\mathbf{N}$, some of these sets are infinite, let them be without loosing the generality $I_1,...,I_{p'},p'\leq p$. We infer $\omega_\rho(\mu)=\{\mu^1,...,\mu^{p'}\}$.

b) For $t'\in\mathbf{R}$, we define

$$k_1=\begin{cases}0,t'<t_0\\ k,t'\in[t_k,t_{k+1})\end{cases}$$

and we obtain

$$\omega_\rho(\mu)=\{\mu^1,...,\mu^{p'}\}=\{\Phi^\rho(\mu,t_k)|k\in I_1\cup...\cup I_{p'}\}$$

$$=\{\Phi^\rho(\mu,t_k)|k\in(I_1\cup...\cup I_{p'})\cap[k_1,\infty)\}$$

$$\subset\{\Phi^\rho(\mu,t_k)|k\in(I_1\cup...\cup I_p)\cap[k_1,\infty)\}=\{\Phi^\rho(\mu,t)|t\geq t'\}$$

$$\subset \{\Phi^\rho(\mu, t_k)|k \in I_1 \cup ... \cup I_p\} = \{\mu^1, ..., \mu^p\} = Or_\rho(\mu).$$

c) If $p' = p$, then $\forall t' \in \mathbf{R}$, $\omega_\rho(\mu) = \{\Phi^\rho(\mu, t)|t \geq t'\} = Or_\rho(\mu)$ from b) and the property holds, thus we can assume that $p' < p$. In this case we define

$$\begin{aligned} k'' &= \min\{k|k \in \mathbf{N}, \forall k' \geq k, k' \in I_1 \cup ... \cup I_{p'}\} \\ &= 1 + \max(I_{p'+1} \cup ... \cup I_p) \end{aligned}$$

for which we have

$$(I_{p'+1} \cup ... \cup I_p) \cap [k'', \infty) = \emptyset$$

and $t' = t_{k''}$ fulfills

$$\omega_\rho(\mu) = \{\mu^1, ..., \mu^{p'}\} = \{\Phi^\rho(\mu, t_k)|k \in I_1 \cup ... \cup I_{p'}\}$$

$$= \{\Phi^\rho(\mu, t_k)|k \in (I_1 \cup ... \cup I_{p'}) \cap [k'', \infty)\}$$

$$= \{\Phi^\rho(\mu, t_k)|k \in (I_1 \cup ... \cup I_p) \cap [k'', \infty)\} = \{\Phi^\rho(\mu, t)|t \geq t'\};$$

any $t'' \geq t'$ gives

$$\omega_\rho(\mu) \overset{b)}{\subset} \{\Phi^\rho(\mu, t)|t \geq t''\} \subset \{\Phi^\rho(\mu, t)|t \geq t'\} = \omega_\rho(\mu).$$

d) Let be $t' \in \mathbf{R}$ such that $\forall t'' \geq t'$,

$$\{\Phi^\rho(\mu, t)|t \geq t'\} = \{\Phi^\rho(\mu, t)|t \geq t''\} \tag{2}$$

and we claim that in this case we have

$$\forall \mu' \in \{\Phi^\rho(\mu, t)|t \geq t'\}, \exists (t'_k) \in Seq, \forall k \in \mathbf{N}, \Phi^\rho(\mu, t'_k) = \mu'. \tag{3}$$

We assume against all reason that (3) is false, meaning that

$$\exists \mu' \in \{\Phi^\rho(\mu, t)|t \geq t'\}, \text{ the set } \{t_k|k \in \mathbf{N}, \Phi^\rho(\mu, t_k) = \mu'\} \text{ is finite.}$$

Then $\exists t'' > \max\{\max\{t_k|k \in \mathbf{N}, \Phi^\rho(\mu, t_k) = \mu'\}, t'\}$ that fulfills $\mu' \in \{\Phi^\rho(\mu, t)|t \geq t'\} \setminus \{\Phi^\rho(\mu, t)|t \geq t''\}$, contradiction with (2). The truth of (3) shows that $\mu' \in \omega_\rho(\mu)$, i.e. $\{\Phi^\rho(\mu, t)|t \geq t'\} \subset \omega_\rho(\mu)$. For all $t'' \geq t'$ we have then

$$\omega_\rho(\mu) \overset{b)}{\subset} \{\Phi^\rho(\mu, t)|t \geq t''\} = \{\Phi^\rho(\mu, t)|t \geq t'\} \subset \omega_\rho(\mu).$$

e) We note that for $t'' \geq t'$ and $\Phi^\rho(\mu, t'') = \mu'$ we can write

$$\omega_\rho(\mu) = \{\Phi^\rho(\mu,t)|t \geq t'\} \overset{c)}{=} \{\Phi^\rho(\mu,t)|t \geq t''\}$$
$$\overset{Theorem\ 7\ d)}{=} \{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t \geq t''\} = \{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t \in \mathbf{R}\}$$
$$= Or_{\rho\chi_{(t'',\infty)}}(\mu').$$

The fact that $\forall t''' \geq t''$,

$$\{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t \geq t''\} \overset{Theorem\ 7\ d)}{=} \{\Phi^\rho(\mu,t)|t \geq t''\} \overset{c)}{=} \{\Phi^\rho(\mu,t)|t \geq t'\}$$
$$\overset{c)}{=} \{\Phi^\rho(\mu,t)|t \geq t'''\} \overset{Theorem\ 7\ d)}{=} \{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t \geq t'''\}$$

shows, by taking into account d), that

$$\{\Phi^{\rho\chi_{(t'',\infty)}}(\mu',t)|t \geq t''\} = \omega_{\rho\chi_{(t'',\infty)}}(\mu').$$

□

Remark 19. *If in Theorem 18 e) we take $t'' \in \mathbf{R}$ arbitrarily, the equation*

$$\omega_\rho(\mu) = \omega_{\rho\chi_{(t'',\infty)}}(\Phi^\rho(\mu,t'')) \tag{4}$$

is still true. Indeed, for sufficiently great t''', the terms in (4) are equal with

$$\{\Phi^\rho(\mu,t)|t \geq t'''\} = \{\Phi^{\rho\chi_{(t'',\infty)}}(\Phi^\rho(\mu,t''),t)|t \geq t'''\}.$$

Theorem 20. *For arbitrary $\mu \in \mathbf{B}^n$,$\rho \in P_n$ the following statements are true:*

a) $\lim_{t\to\infty} \Phi^\rho(\mu,t)$ exists $\Longleftrightarrow$ $card(\omega_\rho(\mu)) = 1$;

b) if $\exists \mu' \in \mathbf{B}^n, \omega_\rho(\mu) = \{\mu'\}$, then $\lim_{t\to\infty} \Phi^\rho(\mu,t) = \mu'$ and $\Phi(\mu') = \mu'$;

c) if $\exists \mu' \in \mathbf{B}^n, \Phi(\mu') = \mu'$ and $\mu' \in Or_\rho(\mu)$, then $\omega_\rho(\mu) = \{\mu'\}$.

Proof. a) Let $\mu \in \mathbf{B}^n$,$\rho \in P_n$ be arbitrary. We get

$$\lim_{t\to\infty} \Phi^\rho(\mu,t) \text{ exists} \Longleftrightarrow \exists \mu' \in \mathbf{B}^n, \exists t' \in \mathbf{R}, \forall t \geq t', \Phi^\rho(\mu,t) = \mu'$$
$$\Longleftrightarrow \exists \mu' \in \mathbf{B}^n, \exists t' \in \mathbf{R}, \{\Phi^\rho(\mu,t)|t \geq t'\} = \{\mu'\}$$
$$\Longleftrightarrow \exists \mu' \in \mathbf{B}^n, \omega_\rho(\mu) = \{\mu'\} \Longleftrightarrow card(\omega_\rho(\mu)) = 1.$$

b) We assume that $\exists \mu' \in \mathbf{B}^n, \omega_\rho(\mu) = \{\mu'\}$, i.e. $\exists \mu' \in \mathbf{B}^n, \exists t' \in \mathbf{R}, \{\Phi^\rho(\mu,t)|t \geq t'\} = \{\mu'\}$ in other words $\lim_{t\to\infty} \Phi^\rho(\mu,t) = \mu'$. The fact that $\Phi(\mu') = \mu'$ results from Theorem 12.

c) This is a consequence of Theorem 13. □

Theorem 21. *Let be* $\mu \in \mathbf{B}^n, \rho \in P_n, \tau \in \mathbf{R}$. *The function* $\rho' \in P_n, \rho'(t) = \rho(t-\tau)$ *fulfills* $\omega_\rho(\mu) = \omega_{\rho'}(\mu)$.

Proof. We use Theorem 8 and we infer the existence of $t' \in \mathbf{R}$ such that

$$\omega_\rho(\mu) = \{\Phi^\rho(\mu, t) | t \geq t'\} = \{\Phi^\rho(\mu, t-\tau) | t - \tau \geq t'\}$$

$$= \{\Phi^{\rho'}(\mu, t) | t \geq t' + \tau\} = \omega_{\rho'}(\mu).$$

□

4 P-invariant and n-invariant sets

Theorem 22. *We consider the function* $\Phi : \mathbf{B}^n \to \mathbf{B}^n$ *and let be the set* $A \in P^*(\mathbf{B}^n)$. *For any* $\mu \in A$, *the following properties are equivalent*

$$\exists \rho \in P_n, Or_\rho(\mu) \subset A, \tag{5}$$

$$\exists \rho \in P_n, \forall t \in \mathbf{R}, \Phi^\rho(\mu, t) \in A, \tag{6}$$

$$\exists \alpha \in \Pi_n, \forall k \in \mathbf{N}, \Phi^{\alpha^0 \ldots \alpha^k}(\mu) \in A \tag{7}$$

and the following properties are also equivalent

$$\forall \rho \in P_n, Or_\rho(\mu) \subset A, \tag{8}$$

$$\forall \rho \in P_n, \forall t \in \mathbf{R}, \Phi^\rho(\mu, t) \in A, \tag{9}$$

$$\forall \alpha \in \Pi_n, \forall k \in \mathbf{N}, \Phi^{\alpha^0 \ldots \alpha^k}(\mu) \in A, \tag{10}$$

$$\forall \lambda \in \mathbf{B}^n, \Phi^\lambda(\mu) \in A. \tag{11}$$

Proof. (9)$\Longrightarrow$(11) Let $\mu \in A, \lambda \in \mathbf{B}^n$ and the function $\rho \in P_n$ be arbitrary,

$$\rho(t) = \alpha^0 \cdot \chi_{\{t_0\}}(t) \oplus \ldots \oplus \alpha^k \cdot \chi_{\{t_k\}}(t) \oplus \ldots \tag{12}$$

with $\alpha \in \Pi_n$ and $(t_k) \in Seq$. We define

$$\rho'(t) = \lambda \cdot \chi_{\{t'\}}(t) \oplus \alpha^0 \cdot \chi_{\{t'+t_0\}}(t) \oplus \ldots \oplus \alpha^k \cdot \chi_{\{t'+t_k\}}(t) \oplus \ldots$$

where $t' \in \mathbf{R}$ is arbitrary and we can see that $\rho' \in P_n$. (9) implies $\Phi^\lambda(\mu) = \Phi^{\rho'}(\mu, t') \in A$.

(11)$\Longrightarrow$(9) Let $\mu \in A$ and $\rho \in P_n$ be arbitrary, given by (12), with $\alpha \in \Pi_n, (t_k) \in Seq$. We get by induction on k :

$t < t_0$: $\Phi^\rho(\mu, t) = \mu \in A$,

$t \in [t_0, t_1)$: $\Phi^\rho(\mu, t) = \Phi^{\alpha^0}(\mu) \in A$ from (11),

...

$t \in [t_{k-1}, t_k)$: $\Phi^{\alpha^0 \dots \alpha^{k-1}}(\mu) \in A$ due to the hypothesis of the induction,

$t \in [t_k, t_{k+1})$: $\Phi^\rho(\mu, t) = \Phi^{\alpha^k}(\Phi^{\alpha^0 \dots \alpha^{k-1}}(\mu)) \in A$ from (11),

...

The rest of the implications are obvious. □

Definition 23. *The set $A \in P^*(\mathbf{B}^n)$ is called a **p-invariant** (or **p-stable**) **set** of the system Ξ_Φ if it fulfills for any $\mu \in A$ one of (5),..., (7) and it is called an **n-invariant** (or n-**stable**) **set** of Ξ_Φ if it fulfills $\forall \mu \in A$ one of (8),..., (11).*

Remark 24. *In the previous terminology, the letter 'p' comes from 'possibly' and the letter 'n' comes from 'necessarily'. Both 'p' and 'n' refer to the quantification of ρ. Such kind of p-definitions and n-definitions recalling logic are caused by the fact that we translate 'real' concepts into 'binary' concepts and the former have no ρ parameters, thus after translation ρ may appear quantified in two ways. The obvious implication is n-invariance $\Longrightarrow$ p-invariance.*

Example 25. *Let $\Phi : \mathbf{B}^2 \to \mathbf{B}^2$ be defined by $\forall \mu \in \mathbf{B}^2, \Phi(\mu_1, \mu_2) = (\overline{\mu_1}, \overline{\mu_2})$ and $\rho(t) = (1,1) \cdot \chi_{\{0,1,2,\dots\}}(t)$. The set $A = \{(0,1), (1,0)\}$ fulfills $\forall \mu \in A, \forall t \in \mathbf{R}, \Phi^\rho(\mu, t) \in A$ i.e. it satisfies (6):*

$$\Phi^\rho((0,1), t) = (0,1) \cdot \chi_{(-\infty,0)}(t) \oplus (1,0) \cdot \chi_{[0,1)}(t) \oplus$$

$$\oplus (0,1) \cdot \chi_{[1,2)}(t) \oplus (1,0) \cdot \chi_{[2,3)}(t) \oplus \dots$$

$$\Phi^\rho((1,0), t) = (1,0) \cdot \chi_{(-\infty,0)}(t) \oplus (0,1) \cdot \chi_{[0,1)}(t) \oplus$$

$$\oplus (1,0) \cdot \chi_{[1,2)}(t) \oplus (0,1) \cdot \chi_{[2,3)}(t) \oplus \dots$$

see Figure 3; $A = \{(0,0), (1,1)\}$ satisfies the same invariance property.

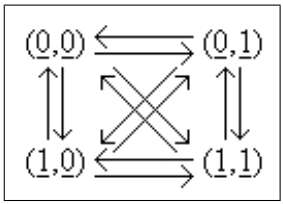


Figure 3: The sets $\{(0,1),(1,0)\}$ and $\{(0,0),(1,1)\}$ are p-invariant

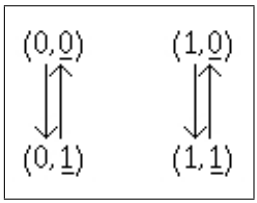


Figure 4: The sets $\{(0,0),(0,1)\}$ and $\{(1,0),(1,1)\}$ are n-invariant

Example 26. *We define the function* $\Phi : \mathbf{B}^2 \to \mathbf{B}^2$ *by* $\forall \mu \in \mathbf{B}^2$, $\Phi(\mu_1, \mu_2) = (\mu_1, \overline{\mu_2})$, *see Figure 4. We notice that the sets* $A = \{(0,0),(0,1)\}$ *and* $A = \{(1,0),(1,1)\}$ *are n-invariant, as they fulfill* $\forall \mu \in A, \forall \rho \in P_2, Or_\rho(\mu) = A$.

Theorem 27. *Let be* $\mu \in \mathbf{B}^n$ *and* $\rho' \in P_n$.

a) If $\Phi(\mu) = \mu$, *then* $\{\mu\}$ *is an n-invariant set and the set Eq of the fixed points of* Φ *is also n-invariant;*

b) the set $Or_{\rho'}(\mu)$ *is p-invariant and* $\bigcup_{\rho \in P_n} Or_\rho(\mu)$[4] *is n-invariant;*

c) the set $\omega_{\rho'}(\mu)$ *is p-invariant.*

Proof. a) From Corollary 14 we have that

$$\forall \rho \in P_n, \forall t \in \mathbf{R}, \Phi^\rho(\mu, t) = \mu \in \{\mu\}.$$

Furthermore, we infer $\forall \mu' \in Eq$, $\forall \rho \in P_n$, $\forall t \in \mathbf{R}$,

$$\Phi^\rho(\mu', t) = \mu' \in Eq.$$

[4] $\bigcup_{\rho \in P_n} Or_\rho(\mu) = \{\mu' | \exists \rho \in P_n, \mu' \in Or_\rho(\mu)\}$.

b) Let be $\mu' \in Or_{\rho'}(\mu)$, thus $t' \in \mathbf{R}$ exists such that $\mu' = \Phi^{\rho'}(\mu, t')$. Then $\forall t \in \mathbf{R}$,

$$\Phi^{\rho' \cdot \chi_{(t',\infty)}}(\mu', t) = \begin{cases} \Phi^{\rho'}(\mu, t), t > t' \\ \mu', t \leq t' \end{cases} \in Or_{\rho'}(\mu).$$

We have proved that $Or_{\rho'}(\mu)$ is p-invariant.

We remark the equality

$$\bigcup_{\rho \in P_n} Or_\rho(\mu) = \bigcup_{\alpha \in \Pi_n} \{\Phi^{\alpha^0 \dots \alpha^k}(\mu) | k \in \mathbf{N}\}$$

and let us take an arbitrary $\mu' \in \bigcup_{\rho \in P_n} Or_\rho(\mu)$. If $\mu' = \mu$ then the statement of the theorem is proved, thus we can assume that $\mu' \neq \mu, \mu' = \Phi^{\alpha^0 \dots \alpha^k}(\mu)$, $\alpha^0, ..., \alpha^k \in \mathbf{B}^n$. For any $\rho'' \in P_n$,

$$\rho'' = \beta^0 \cdot \chi_{\{t'_0\}} \oplus ... \oplus \beta^k \cdot \chi_{\{t'_k\}} \oplus ...$$

$\beta \in \Pi_n, (t'_k) \in Seq$ and any $t \in \mathbf{R}$, we have that $\Phi^{\rho''}(\mu', t)$ is an element of the sequence $\Phi^{\alpha^0 \dots \alpha^k}(\mu)$, $\Phi^{\alpha^0 \dots \alpha^k \beta^0}(\mu)$, ..., $\Phi^{\alpha^0 \dots \alpha^k \beta^0 \dots \beta^{k'}}(\mu)$, ... where $\alpha^0, ..., \alpha^k, \beta^0, ..., \beta^{k'}, ... \in \Pi_n$. The conclusion is that $\Phi^{\rho''}(\mu', t) \in \bigcup_{\rho \in P_n} Or_\rho(\mu)$.

c) This is a consequence of Theorem 18 e). □

5 The basin of p-attraction and the basin of n-attraction

Theorem 28. *We consider the set $A \in P^*(\mathbf{B}^n)$. For any $\mu \in \mathbf{B}^n$, the following statements are equivalent*

$$\exists \rho \in P_n, \omega_\rho(\mu) \subset A, \tag{13}$$

$$\exists \rho \in P_n, \exists t' \in R, \forall t \geq t', \Phi^\rho(\mu, t) \in A, \tag{14}$$

$$\exists \alpha \in \Pi_n, \exists k' \in \mathbf{N}, \forall k \geq k', \Phi^{\alpha^0 \dots \alpha^k}(\mu) \in A \tag{15}$$

and the following statements are equivalent, too

$$\forall \rho \in P_n, \omega_\rho(\mu) \subset A, \tag{16}$$

$$\forall \rho \in P_n, \exists t' \in R, \forall t \geq t', \Phi^\rho(\mu, t) \in A, \tag{17}$$

$$\forall \alpha \in \Pi_n, \exists k' \in \mathbf{N}, \forall k \geq k', \Phi^{\alpha^0 \dots \alpha^k}(\mu) \in A. \tag{18}$$

Proof. (13)$\Longrightarrow$(14) We presume that (13) is true. Some t' exists with

$$\omega_\rho(\mu) = \{\Phi^\rho(\mu, t) | t \geq t'\}$$

and we conclude that $\forall t \geq t'$,

$$\Phi^\rho(\mu, t) \in \omega_\rho(\mu) \subset A.$$

(14)$\Longrightarrow$(13) As $t'' \in \mathbf{R}$ exists with

$$\omega_\rho(\mu) = \{\Phi^\rho(\mu, t) | t \geq t''\},$$

from the truth of (14) we have that

$$\omega_\rho(\mu) \subset \{\Phi^\rho(\mu, t) | t \geq \max\{t', t''\}\} \subset A.$$

□

Definition 29. *The **basin** (or **kingdom**, or **domain**) **of p-attraction** or the **p-stable set** of the set $A \in P^*(\mathbf{B}^n)$ is given by $\overline{W}(A) = \{\mu | \mu \in \mathbf{B}^n, \exists \rho \in P_n, \omega_\rho(\mu) \subset A\}$; the **basin** (or **kingdom**, or **domain**) **of n-attraction** or the **n-stable set** of the set A is given by $\underline{W}(A) = \{\mu | \mu \in \mathbf{B}^n, \forall \rho \in P_n, \omega_\rho(\mu) \subset A\}$.*

Remark 30. *Definition 29 makes use of the properties (13) and (16). We can make use also in this Definition of the other equivalent properties from Theorem 28.*

In Definition 29, one or both basins of attraction $\overline{W}(A), \underline{W}(A)$ may be empty.

Theorem 31. *We have:*
i) $\overline{W}(\mathbf{B}^n) = \underline{W}(\mathbf{B}^n) = \mathbf{B}^n$;
ii) if $A \subset A'$, then $\overline{W}(A) \subset \overline{W}(A')$ and $\underline{W}(A) \subset \underline{W}(A')$ hold.

Definition 32. *When $\overline{W}(A) \neq \emptyset$, A is said to be **p-attractive** and for any non-empty set $B \subset \overline{W}(A)$, we say that A is **p-attractive** for B and that B is **p-attracted** by A; A is by definition **partially p-attractive** if $\overline{W}(A) \notin \{\emptyset, \mathbf{B}^n\}$ and **totally p-attractive** whenever $\overline{W}(A) = \mathbf{B}^n$.*

*The fact that $\underline{W}(A) \neq \emptyset$ makes us say that A is **n-attractive** and in this situation for any non-empty $B \subset \underline{W}(A)$, A is called **n-attractive** for B and B is called to be **n-attracted** by A; we use to say that A is **partially n-attractive** if $\underline{W}(A) \notin \{\emptyset, \mathbf{B}^n\}$ and **totally n-attractive** if $\underline{W}(A) = \mathbf{B}^n$.*

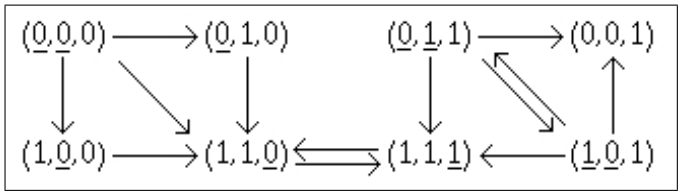


Figure 5: Invariant sets and basins of attraction

Example 33. *We consider the system from Figure 5. The set* $A = \{(0,0,0)\}$ *is neither p-invariant, nor n-invariant:* $\overline{W}(A) = \underline{W}(A) = \emptyset$.

The set $A = \{(0,0,0),(1,1,0),(1,1,1)\}$ *is p-invariant but not n-invariant:* $\overline{W}(A) = \mathbf{B}^3 \setminus \{(0,0,1)\}$, $\underline{W}(A) = \{(0,0,0),(0,1,0),(1,0,0),(1,1,0),(1,1,1)\}$.

We take $A = \{(1,1,0),(1,1,1),(0,0,1)\}$ *which is both p-invariant and n-invariant. A is totally p-attractive,* $\overline{W}(A) = \mathbf{B}^3$ *and it is not totally n-attractive, since* $\underline{W}(A) = \mathbf{B}^3 \setminus \{(0,1,1),(1,0,1)\}$.

The set $A = \{(1,1,0),(1,1,1),(0,1,1),(0,0,1),(1,0,1)\}$ *is p-invariant, n-invariant, totally p-attractive and totally n-attractive because* $\overline{W}(A) = \underline{W}(A) = \mathbf{B}^3$.

Example 34. *The set* $\mathbf{B}^n$ *is totally p-attractive and totally n-attractive (Theorem 31 i)).*

Theorem 35. *Let* $A \in P^*(\mathbf{B}^n)$ *be some set. If A is p-invariant, then* $A \subset \overline{W}(A)$ *and A is also p-attractive; if A is n-invariant, then* $A \subset \underline{W}(A)$ *and A is also n-attractive.*

Proof. Let $\mu \in A$ be arbitrary. The existence of $\rho \in P_n$ such that $Or_\rho(\mu) \subset A$ (from the p-invariance of A) and the inclusion $\omega_\rho(\mu) \subset Or_\rho(\mu)$ show that $\omega_\rho(\mu) \subset A$, thus $\mu \in \overline{W}(A)$. As μ was arbitrary, we get that $A \subset \overline{W}(A)$ and finally that $\overline{W}(A) \neq \emptyset$. A is p-attractive. □

Remark 36. *The previous Theorem shows the connection that exists between invariance and attractiveness. If A is p-attractive, then* $\overline{W}(A)$ *is the greatest set that is p-attracted by A and the point is that this really happens when A is p-invariant. The other situation is dual.*

Theorem 37. *Let be* $A \in P^*(\mathbf{B}^n)$. *If A is p-attractive, then* $\overline{W}(A)$ *is p-invariant and if A is n-attractive, then* $\underline{W}(A)$ *is n-invariant.*

Proof. If A is p-attractive then $\overline{W}(A) \neq \emptyset$ and we prove that $\overline{W}(A)$ is p-invariant. Let $\mu \in \overline{W}(A)$ be arbitrary and fixed. From the definition of $\overline{W}(A)$ some $\rho \in P_n$ exists with the property that $\omega_\rho(\mu) \subset A$. We show that

$$\forall t' \in \mathbf{R}, \Phi^\rho(\mu, t') \in \overline{W}(A),$$

i.e.

$$\forall t' \in \mathbf{R}, \exists \rho' \in P_n, \omega_{\rho'}(\Phi^\rho(\mu, t')) \subset A.$$

Indeed, we fix arbitrarily some $t' \in \mathbf{R}$. With

$$\rho' = \rho\chi_{(t',\infty)}$$

we can write, from Remark 19, equation (4) that

$$\omega_{\rho'}(\Phi^\rho(\mu, t')) = \omega_{\rho\chi_{(t',\infty)}}(\Phi^\rho(\mu, t')) = \omega_\rho(\mu) \subset A.$$

We prove now that $\underline{W}(A)$, which is non-empty from the n-attractiveness of A, is also n-invariant. The property

$$\forall \mu' \in \underline{W}(A), \forall \rho' \in P_n, Or_{\rho'}(\mu') \subset \underline{W}(A),$$

that is equivalent with

$$\forall \mu' \in \underline{W}(A), \forall \rho' \in P_n, \forall \mu'' \in Or_{\rho'}(\mu'), \mu'' \in \underline{W}(A)$$

and with

$$\forall \mu' \in \mathbf{B}^n, \forall \rho \in P_n, \omega_\rho(\mu') \subset A \Longrightarrow$$
$$\Longrightarrow \forall \rho' \in P_n, \forall \mu'' \in Or_{\rho'}(\mu'), \forall \rho'' \in P_n, \omega_{\rho''}(\mu'') \subset A,$$

means the following. Let $\mu' \in \mathbf{B}^n$ and $\rho'' \in P_n$ be arbitrary and fixed. The hypothesis states that for any

$$\rho = \alpha^0 \cdot \chi_{\{t_0\}} \oplus ... \oplus \alpha^k \cdot \chi_{\{t_k\}} \oplus ...$$

$\alpha \in \Pi_n, (t_k) \in Seq$ we have

$$\exists k_1 \in \mathbf{N}, \{\Phi^{\alpha^0 ... \alpha^k}(\mu') | k \geq k_1\} (= \omega_\rho(\mu')) \subset A. \tag{19}$$

We consider arbitrarily the function $\rho' \in P_n$,

$$\rho' = \alpha'^0 \cdot \chi_{\{t'_0\}} \oplus ... \oplus \alpha'^k \cdot \chi_{\{t'_k\}} \oplus ...$$

$\alpha' \in \Pi_n, (t'_k) \in Seq$ and the point $\mu'' \in Or_{\rho'}(\mu')$, thus $k' \in \mathbf{N}$ exists with the property

$$\mu'' = \Phi^{\alpha'^0 \dots \alpha'^{k'}}(\mu').$$

We put ρ'' under the form

$$\rho'' = \alpha''^0 \cdot \chi_{\{t''_0\}} \oplus \dots \oplus \alpha''^k \cdot \chi_{\{t''_k\}} \oplus \dots$$

$\alpha'' \in \Pi_n, (t''_k) \in Seq$. The sequence

$$\Phi^{\alpha''^0 \dots \alpha''^k}(\mu'') = \Phi^{\alpha''^0 \dots \alpha''^k}(\Phi^{\alpha'^0 \dots \alpha'^{k'}}(\mu')) = \Phi^{\alpha'^0 \dots \alpha'^{k'} \alpha''^0 \dots \alpha''^k}(\mu'),$$

$k \in \mathbf{N}$ fulfills the property (19), thus

$$\exists k_2 \in \mathbf{N}, \{\Phi^{\alpha''^0 \dots \alpha''^k}(\mu'') | k \geq k_2\} (= \omega_{\rho''}(\mu'')) \subset A.$$

□

Corollary 38. *If the set $A \in P^*(\mathbf{B}^n)$ is p-invariant, then $\overline{W}(A)$ is p-invariant and if A is n-invariant, then the basin of n-attraction $\underline{W}(A)$ is n-invariant.*

Proof. These result from Theorem 35 and Theorem 37. □

6 Discussion

Some notes on the terminology:

- universality means the greatest in the sense of inclusion. Any $X \subset \Xi_\Phi$ is a system, but we did not study such systems in the present paper;

- regularity means the existence of a generator function Φ, i.e. analogies with the dynamical systems theory;

- autonomy means here that no input exists. We mention the fact that autonomy has another non-equivalent definition also, a system is called autonomous if its input set has exactly one element;

- asynchronicity refers (vaguely) to the fact that we work with real time and binary values. Its antonym synchronicity means that 'discrete time' (and binary values) in which the iterates of Φ are: $\Phi, \Phi \circ \Phi, ..., \Phi \circ ... \circ \Phi, ...$ i.e. in the sequence $\Phi^{\alpha^0}, \Phi^{\alpha^0 \alpha^1}, ..., \Phi^{\alpha^0 \dots \alpha^k}, ...$ all α^k are $(1, ..., 1), k \in \mathbf{N}$. That is the discrete time of the dynamical systems.

Our concept of invariance from Definition 23 reproduces the point of view expressed in [4], page 11, where the dynamical system $S = (T, X, \Phi)$ is given, with $T = \mathbf{R}$ the time set, X the state space and $\Phi : T \times X \to X$ the flow: the set $A \subset X$ is said to be invariant for the system S if $\forall x \in A, \forall t \in T, \Phi_t(x) \in A$. This idea coincides with the one from [5], page 27 where the state space X is a differentiable manifold M.

In [3], page 92 the set $A \subset X$ is called globally invariant via Φ if $\forall t \in T, \Phi_t(A) = A$, recalling the situation of Example 26 and Figure 4. In [6], page 3, the global invariance and the invariance of $A \subset X$ are defined like at [3] and [4].

We mention also the definition of invariance from [1], page 19. Let $P = (T, X, \Phi)$ be a process, where $T = \mathbf{R}$, X is the state space and $\Phi : \overline{T} \times X \to X$ is the flow of P; we have denoted $\overline{T} = \{(t', t) | t', t \in T, t \leq t'\}$. Then $A \subset X$ is invariant relative to Φ if $\Phi_{t',t}(A) \subset A$ for any $(t', t) \in \overline{T}$. This last definition agrees itself with ours in the special case when $t' = 0$ but it is more general since it addresses systems which are not time invariant.

Stability is defined in [5], page 27 where M is a differentiable manifold and the evolution operator $\Phi_t : M \to M, t \in T$ is given. The subset $A \subset M$ is stable for Φ if for any sufficiently small neighborhood U of A a neighborhood V of A exists such that $\forall x \in V, \forall t \geq 0, \Phi_t(x) \in U$. In our case when $M = \mathbf{B}^n$ has the discrete topology, $A \subset \mathbf{B}^n$ and $U = V = A$, this comes to the invariance of A.

In [4], page 16 the closed invariant set $A \subset X$ is called stable for (T, X, Φ) if i) for any sufficiently small neighborhood $U \supset A$ there exists a neighborhood $V \supset A$ such that $\forall t > 0, \forall x \in V, \Phi_t(x) \in U$ and ii) there exists a neighborhood $W \supset A$ such that $\forall x \in W, \Phi_t(x) \to A$ as $t \to \infty$. We see that i) is the same request like at [5] and ii) brings nothing new (item i) means $Or_\rho(\mu) \subset A$, thus a stronger request than item ii) which is $\omega_\rho(\mu) \subset A$ in our case).

In a series of works ([5], page 27), either the set $A \subset M$ is called asymptotically stable if it is stable and attractive, where M is a differentiable manifold, or ([3], page 112, [6], page 5) the fixed point $x_0 \in X$ is called asymptotically stable if it is stable and attractive. We interpret stability as invariance and stating that A or x_0 is stable and attractive means that it is invariant and a weaker property than invariance takes place (see Theorem 35) and finally asymptotic stability means invariance too.

In [2], page 132 the statement is made that many times, in applications, by stability is understood attractiveness. This would mean, in the conditions

of Theorem 35, weakening of the invariance request and we cannot accept this point of view.

In literature, [2] defines at page 6 the basin of attraction of a chaotic attractor $A \subset X$ as the set of the points whose $\omega-$limit set is contained in A. This was reproduced at (13) and (16), where $A \in P^*(\mathbf{B}^n)$ was considered however arbitrary.

The work [3] defines at page 124 the kingdom of attraction of an attractive set $A \subset X$ as the greatest set of points of X whose dynamic ends (for $t \to \infty$) in A; when the kingdom of attraction is an open set, it is called basin of attraction. For us, all the subsets $A \subset \mathbf{B}^n$ are open in the discrete topology of $\mathbf{B}^n$.

In [3], page 123 the invariant set $A \subset X$ is called attractive set for $B \subset X$ if the distance between A and $\Phi_t(B)$ tends to 0 for $t \to \infty$; a set A is attractive if $B \neq \emptyset$ exists that is attracted by A. A slightly different idea is expressed in [6], page 4 where the invariant set A is called attractive for B if $\lim\limits_{t\to\infty} \Phi_t(B) = A$. Unlike these definitions, in Definition 32 the set $A \subset \mathbf{B}^n$ is not required to be invariant and the statement $B \subset \overline{W}(A)$ showing that B is p-attracted by A, i.e. $\forall \mu \in B, \exists \rho \in P_n, \omega_\rho(\mu) \subset A$, reproduces the fact that the distance between A and $\Phi_t(B)$ tends to 0 for $t \to \infty$.

In [5], page 27 M is a differentiable manifold and the subset $A \subset M$ is called attractive for Φ if a neighborhood U of A exists such that $\forall x \in U, \lim\limits_{t\to\infty} \Phi_t(x) \in A$; in this case we say that U is attracted by A. We have reached (13), (16) and the requests of attractiveness $\overline{W}(A) \neq \emptyset, \underline{W}(A) \neq \emptyset$ from Definition 32.

In [2], page 5 a closed invariant set $A \subset X$ is called attractive if a neighborhood U of A exists such that $\forall x \in U, \forall t \geq 0, \Phi_t(x) \in U$ and $\Phi_t(x) \to A$ when $t \to \infty$. Then the set $\bigcup\limits_{t\leq 0} \Phi_t(U)$ is called the basin (the domain) of attraction of the set A.

In [6], page 4 the open set $W(A) \subset X$ representing the greatest set of points of X which is attracted by the attractive set A is called basin of attraction. This definition represents exactly $\overline{W}(A), \underline{W}(A)$ from Definition 29 in the circumstances that (Definition 32) the attractiveness of A means that the previous sets are non-empty.

We have the definition of the basin of attraction from [5], page 27: the maximal set attracted by an attractor $A \subset X$ (invariant set, attractive for one of its neighborhoods) is called the kingdom of attraction of A; when the

kingdom of attraction is an open set, it is called basin of attraction. We conclude, related with the real to binary translation of this definition, that if $A \in P^*(\mathbf{B}^n)$ is p-invariant, then it is p-attractive for itself and thus an 'attractor'; its basin of attraction $\overline{W}(A)$ is non-empty in this case and it is the maximal set attracted by A.

We note that the stable manifold of the equilibrium point $x_0 \in X$ is defined in [6], page 4 and [3], page 93 for the dynamical system (T, X, Φ) by $W(x_0) = \{x \in X | \lim_{t\to\infty} \Phi_t(x) = x_0\}$. In [4], page 46 the terminology of stable set is used for this concept and [6] mentions this terminology too. Thus, by replacing $x_0 \in X$ with $A \subset \mathbf{B}^n$ and $\lim_{t\to\infty} \Phi_t(x) = x_0$ with $\omega_\rho(\mu) \subset A$ we get for $\overline{W}(A), \underline{W}(A)$ the alternative terminology of stable sets (i.e. invariant sets) of A.

References

[1] C. D. Constantinescu. *Haos, fractali şi aplicaţii.* Editura the Flower Power, Piteşti, 2003.

[2] M. F. Danca. *Funcţia logistică, dinamică, bifurcaţie şi haos.* Editura Universităţii din Piteşti, 2001.

[3] A. Georgescu, M. Moroianu, I. Oprea. *Teoria Bifurcaţiei, Principii şi Aplicaţii.* Editura Universităţii din Piteşti, 1999.

[4] Yu. A. Kuznetsov. *Elements of Applied Bifurcation Theory, Second Edition.* Springer, 1997.

[5] M. Sterpu. *Dinamică şi bifurcaţie pentru două modele van der Pol generalizate.* Editura Universităţii din Piteşti, 2001.

[6] M. P. Trifan. *Dinamică şi bifurcaţie în studiul matematic al cancerului.* Editura Pământul, Piteşti, 2006.

[7] Ş. E. Vlad. Boolean dynamical systems. *Romai Journal.* Vol. 3, Nr. 2: 277-324, 2007.