Trends in Mathematics

Mats Andersson
Jan Boman
Christer Kiselman
Pavel Kurasov
Ragnar Sigurdsson
Editors

Analysis Meets Geometry

The Mikael Passare Memorial Volume

Birkhäuser

Birkhäuser

Trends in Mathematics

Trends in Mathematics is a series devoted to the publication of volumes arising from conferences and lecture series focusing on a particular topic from any area of mathematics. Its aim is to make current developments available to the community as rapidly as possible without compromise to quality and to archive these for reference.

Proposals for volumes can be submitted using the Online Book Project Submission Form at our website www.birkhauser-science.com.

Material submitted for publication must be screened and prepared as follows:

All contributions should undergo a reviewing process similar to that carried out by journals and be checked for correct use of language which, as a rule, is English. Articles without proofs, or which do not contain any significantly new results, should be rejected. High quality survey papers, however, are welcome.

We expect the organizers to deliver manuscripts in a form that is essentially ready for direct reproduction. Any version of TEX is acceptable, but the entire collection of files must be in one particular dialect of TEX and unified according to simple instructions available from Birkhäuser.

Furthermore, in order to guarantee the timely appearance of the proceedings it is essential that the final version of the entire material be submitted no later than one year after the conference.

More information about this series at http://www.springer.com/series/4961

Mats Andersson • Jan Boman • Christer Kiselman
Pavel Kurasov • Ragnar Sigurdsson
Editors

Analysis Meets Geometry

The Mikael Passare Memorial Volume

Birkhäuser

Editors
Mats Andersson
Department of Mathematics
Chalmers University of Technology
Göteborg, Sweden

Jan Boman
Department of Mathematics
Stockholm University
Stockholm, Sweden

Christer Kiselman
Department of Information Technology
Uppsala University
Uppsala, Sweden

Pavel Kurasov
Department of Mathematics
Stockholm University
Stockholm, Sweden

Ragnar Sigurdsson
Science Institute
University of Iceland
Reykjavik, Iceland

ISSN 2297-0215 ISSN 2297-024X (electronic)
Trends in Mathematics
ISBN 978-3-319-52469-6 ISBN 978-3-319-52471-9 (eBook)
DOI 10.1007/978-3-319-52471-9

Library of Congress Control Number: 2017951453

Mathematics Subject Classification (2010): 01-00, 32-XX, 51-XX, 35-XX

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Preface

Dear Reader,

This volume is dedicated to the life and work of *Mikael Passare*, who unexpectedly passed away in September 2011 at the age of 52. It will give you a chance to study several texts that are inspired by his mathematical interests – either directly so, or in a more indirect way. One chapter is even coauthored by Mikael.

We are indeed happy that so many outstanding scientists have answered our call to contribute a chapter to the book.

Before arriving to these mathematical chapters, you can get an idea about Mikael's life and his far-reaching research, which spans from analysis in several complex variables and complex geometry to amoebas and tropical geometry.

We hope that you will enjoy getting acquainted with the ideas of a great mathematician and a great human being.

Mats Andersson *Jan Boman* *Christer Kiselman*
Pavel Kurasov *Ragnar Sigurdsson*

Mikael Passare 1959–2011

Part I

Memorial Contributions

This part collects several articles describing Mikael Passare, as a mathematician and as a great personality. It contains also Mikael's *Curriculum Vitae*, a list of his publications and the list of all 152 countries that Mikael managed to visit.

2010, Västervik, Sweden

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 3–4

Mikael Passare

Curriculum Vitae

1959-01-01. Kjell Alrik Mikael Pettersson is born in Västerås, Sweden. Mother: Britt Gunvor Emilia Pettersson, later with the family name Elfström. Father: Werner Siems. Stepfathers: Kjell Pettersson and Hans Elfström.

1976-09. Mikael starts his studies at Uppsala University.

1978-06-09. Mikael finishes high school, *Rudbeckianska skolan* in Västerås. In addition to English and French, he has studied Russian during three years in that school. He receives the highest possible marks in all subjects except gymnastics.

1979-06 through 1980-08. Military service at the National Defence Radio Establishment (FRA), where he learns even more Russian.

1979-12-03. Mikael receives a diploma for a Bachelor Degree.

1980-02-14. Mikael is accepted as a graduate student at Uppsala University.

Academic year 1980–81. Mikael studies at Stanford University, Palo Alto.

Academic year 1981–82. Mikael studies at Lomonosov University, Moscow. He is supported by a scholarship from the Swedish Institute.

1982-04-06. Mikael Pettersson and Галина Лепёшкина (Galina Lepjosjkina) marry in Moscow.

1984, Fall Term, through 1986, Spring Term. Mikael holds the Lundström–Åman scholarship for two academic years.

1984-12-15. Mikael Pettersson defends his doctoral thesis at Uppsala University. Opponent: Nils Øvrelid. The members of the grading committee are Jan-Erik Björk, Lennart Carleson, and Björn Engquist.

1984-12-18. The family name *Passare* is approved for Kjell Alrik Mikael Pettersson and Galina Pettersson, née Lepjosjkina.

1985–1986. Lecturer at Stockholm University.

1985–1986, 1987–1990. Research Assistant at Stockholm University.

1985-12-15. Max Petter Passare, Galina's and Mikael's son, is born.

Academic year 1986–87. Mikael is at Université Pierre-et-Marie-Curie (Paris VI) and Université de Paris-Sud (Paris IX), having received a post-doctoral fellowship from the Swedish Natural Science Research Council (NFR).

1988. Mikael receives the title of Docent.

1988. Mikael is awarded the Marcus and Marianne Wallenberg scholarship.

1990–1994. Research Lecturer at the Royal Institute of Technology, Stockholm.

1991. The Royal Society of Sciences, Uppsala, awards the Lilly and Sven Thuréus Prize to Mikael.

1991-08-19. Märta Sofia Passare, Galina's and Mikael's daughter, is born.

Academic year 1992–93. Mikael receives an Alexander von Humboldt fellowship and spends the year at Humboldt-Universität zu Berlin.

1994-10-01. Mikael starts as Professor at Stockholm University. His chair is the one that was created for Sonja Kovalevsky and held by his mathematical grandfather Lars Hörmander.

1997-03. The first Nordan Conference takes place in Trosa at the initiative of Mikael Passare and Peter Ebenfelt.

2000-01-15. Mikael organizes a symposium at Stockholm University to celebrate the 150th anniversary of Sonja Kovalevsky.

2001. Mikael receives the Göran Gustafsson Prize.

2011-09-15. Mikael dies in Oman.

2011-10-28. Mikael's funeral in *Norra begravningsplatsen* north of Stockholm. Celebrant: Noomi Arvas Liljefors. He is buried in a grave with the address Block 21A, number 187, not far from Sonja Kovalevsky's.

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 5–9

Mikael Passare's Publications

1984. Pettersson,[1] Mikael. *Residues, Currents, and Their Relation to Ideals of Holomorphic Currents.* Uppsala: Uppsala University, Department of Mathematics. Report No. 10, November 1984, 94 pp. (PhD Thesis defended on 1984 December 15. Opponent: Nils Øvrelid.)

1985. Passare, Mikael. Produits des courants résiduels et règle de Leibniz. *C. R. Acad. Sci. Paris Sér. I Math.* **301**, no. 15, 727–730.

1986. Passare, Mikael. Ideals of holomorphic functions defined by residue currents. *Complex analysis and applications* '85 (*Varna,* 1985), pp. 511–514. Publ. House Bulgar. Acad. Sci., Sofia.

1987. Passare, Mikael. Courants méromorphes et égalité de la valeur principale et de la partie finie. *Séminaire d'Analyse P. Lelong – P. Dolbeault – H. Skoda, Années* 1985/1986, pp. 157–166. Lecture Notes in Math. 1295. Berlin et al.: Springer-Verlag. (Reviewed[2] by Salomon Ofman.)

1988a. Andersson, Mats; Passare, Mikael. A shortcut to weighted representation formulas for holomorphic functions. *Ark. mat.* **26**, no. 1, 1–12. (Reviewed by Bo Berndtsson.)

1988b. Passare, Mikael. Residue solutions to holomorphic Cauchy problems. *Seminar in Complex Analysis and Geometry* 1987 (*Rende,* 1987), pp. 99–105, Sem. Conf., 1. Rende: EditEl. (Reviewed by E. J. Akutowicz.)

1988c. Passare, Mikael. Residues, currents, and their relation to ideals of holomorphic functions. *Math. Scand.* **62**, no. 1, 75–152. (Reviewed by Alicia Dickenstein.)

1988d. Passare, Mikael. A calculus for meromorphic currents. *J. reine angew. Math.* **392**, 37–56. (Reviewed by Alicia Dickenstein.)

1989. Berndtsson, Bo; Passare, Mikael. Integral formulas and an explicit version of the fundamental principle. *J. Funct. Anal.* **84**, no. 2, 358–372. (Reviewed by R. Michael Range.)

1991a. Andersson, Mats; Passare, Mikael. Complex Kergin interpolation. *J. Approx. Theory* **64**, no. 2, 214–225. (Reviewed by A. G. Law.)

---

[1] This was Mikael's family name from birth and until 1984 December 18, three days after his thesis defense.
[2] These remarks refer to *MathSciNet.*

1991b. Andersson, Mats; Passare, Mikael. Complex Kergin interpolation and the Fantappiè transform. *Math. Z.* **208**, no. 2, 257–271. (Reviewed by Harold P. Boas.)

1991c. Passare, Mikael. A new division formula for complete intersections. *Proceedings of the Tenth Conference on Analytic Functions* (*Szczyrk,* 1990)*. Ann. Polon. Math.* **55**, 283–286. (Reviewed by Gerd Müller.)

1992, 1993a. Passare, M.; Tsikh, A. On the relations between the local structure of holomorphic mappings, multidimensional residues and generalized Mellin transforms (Russian). *Dokl. Akad. Nauk* **325**, no. 4, 664–667; translation in *Russian Acad. Sci. Dokl. Math.* **46** (1993), no. 1, 88–91. (Reviewed by Alexandr M. Kytmanov.)

1993b. Passare, Mikael. On the support of residue currents. *Several complex variables* (*Stockholm,* 1987/1988), pp. 542–549. Math. Notes 38, Princeton Univ. Press, Princeton, NJ. (Reviewed by Salomon Ofman.)

1993c. Passare, Mikael. Halva sanningen om en viktig produkt. Residyer i flera variabler. Föredrag vid Kungl. Vetenskaps-Societetens högtidsdag den 8 november 1991 [Half of the truth about an important product. Residues in several variables.] Lecture at the Solemnity of the Royal Society of Sciences, 1991 November 08. **In:** *Kungl. Vetenskaps-Societens i Uppsala årsbok* 1992, pp. 17–20. Uppsala: The Royal Society of Sciences.

1994. Passare, Mikael; Tsikh, August; Zhdanov, Oleg. A multidimensional Jordan residue lemma with an application to Mellin–Barnes integrals. *Contributions to complex analysis and analytic geometry*, pp. 233–241. Aspects Math., E26. Braunschweig: Vieweg. (Reviewed by Aleksandr G. Aleksandrov.)

1995a. Demailly, Jean-Pierre; Passare, Mikael. Courants résiduels et classe fondamentale. *Bull. Sci. Math.* **119**, no. 1, 85–94. (Reviewed by Mongi Blel.)

1995b. Passare, Mikael; Tsikh, August. Residue integrals and their Mellin transforms. *Canad. J. Math.* **47**, no. 5, 1037–1050. (Reviewed by Mongi Blel.)

1996a. Passare, Mikael; Tsikh, August. Defining the residue of a complete intersection. *Complex Analysis, Harmonic Analysis and Applications* (*Bordeaux,* 1995), pp. 250–267. Pitman Res. Notes Math. Ser., 347, Harlow: Longman. (Reviewed by Carlos A. Berenstein.)

1996b, 1997. Passare, M.; Tsikh, A. K.; Cheshel′, A. A. Iterated Mellin–Barnes integrals as periods on Calabi–Yau manifolds with two modules (Russian). *Teoret. Mat. Fiz.* **109** (1996), no. 3, 381–394; translation in *Theoret. and Math. Phys.* **109**, no. 3, 1544–1555 (1997). (Reviewed by V. V. Chueshev.)

1998. Bykov, Valery; Kytmanov, Alexander; Lazman, Mark. *Elimination Methods in Polynomial Computer Algebra.* Translated from the 1991 Russian original by Kytmanov and revised by the authors. Translation edited and with a preface by Mikael Passare. Mathematics and its Applications, 448. Dordrecht: Kluwer Academic Publishers. xii + 237 pp. ISBN: 0-7923-5240-8. (Reviewed by Harold P. Boas.)

1999. Henkin, Gennadi; Passare, Mikael. Abelian differentials on singular varieties and variations on a theorem of Lie–Griffiths. *Invent. Math.* **135**, no. 2, 297–328. (Reviewed by Reinhold Hübl.)

2000a. Forsberg, Mikael; Passare, Mikael; Tsikh, August. Laurent determinants and arrangements of hyperplane amoebas. *Adv. Math.* **151**, no. 1, 45–70. (Reviewed by Guangfeng Jiang.)

2000b. Passare, Mikael; Tsikh, August; Yger, Alain. Residue currents of the Bandner–Martinelli type. *Publ. Mat.* **44**, no. 1, 85–117. (Reviewed by Harold P. Boas.)

2000c. Aizenberg, Lev; Passare, Mikael. **C**-convexity, convexity in complex analysis. **In:** *Encyclopædia of Mathematics*, Supplement, vol. II, pp. 102–104. Dordrecht: Kluwer Academic Publishers.

2001a. Passare, Mikael. Sesam öppna dig [Open Sesame]. *Nämnaren* **28**(4), 37–39.

2001b. Passare, Mikael. *Complex geometry. Minicourse, June* 30*–July* 3*,* 2001. Lectures at the *Première École d'Été Franco-Nordique de Mathématiques*, EEFN, Lake Erken, Sweden, 2001 June 26–July 03. Institut Mittag-Leffler, Lecture Notes No. 3, 2000/2001, 8 pp.

2002. Passare, Mikael; Rullgård, Hans. Multiple Laurent series and polynomial amoebas. *Actes des Rencontres d'Analyse Complexe* (*Poitiers-Futuroscope,* 1999), pp. 123–129. Poitiers: Atlantique. (Reviewed by Guangfeng Jiang.)

2003a. [Passare, Mikael (Ed.)]. *Nordan Ett.* [Abstracts from *Nordan* 1, held in] Trosa 1997 March 14–16, 15 pp. [Stockholm: Stockholm University 2003.]

2003b. [Passare, Mikael (Ed.)]. *Nordan Två.* [Abstracts from *Nordan* 2, held in] Marstrand 1998 April 24–26, 15 pp. [Stockholm: Stockholm University 2003.]

2004a. Passare, Mikael; Rullgård, Hans. Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope. *Duke Math. J.* **121**, no. 3, 481–507. (Reviewed by A.Yu. Rashkovskiĭ.)

2004b. Andersson, Mats; Passare, Mikael; Sigurdsson, Ragnar. *Complex Convexity and Analytic Functionals.* Progress in Mathematics, 225. Basel: Birkhäuser Verlag. xii + 160 pp. ISBN: 3-7643-2420-1. (Reviewed by Sergey Ivashkovich.)

2004c. Passare, Mikael; Tsikh, August. Algebraic equations and hypergeometric series. **In:** *The Legacy of Niels Henrik Abel*, pp. 653–672. Berlin: Springer. (Reviewed by Allen R. Miller.)

2004d. Passare, Mikael. Amoebas, convexity and the volume of integer polytopes. *Complex analysis in several variables – Memorial Conference of Kiyoshi Oka's Centennial Birthday*, pp. 263–268. Adv. Stud. Pure Math., 42. Tokyo: Math. Soc. Japan. (Reviewed by A.Yu. Rashkovskiĭ.)

2004e. [Passare, Mikael (Ed.)]. *Nordan Tre.* [Abstracts from *Nordan* 3, held in] Saltsjöbaden 1999 April 22–25, 19 pp. [Stockholm: Stockholm University 2004.]

2004f. Passare, Mikael. Amöbor och Laurentserier [Amoebas and Laurent series]. **In:** 2004e:19. (A report on 2000a.)

2005a. Passare, Mikael; Sadykov, Timur; Tsikh, August. Singularities of hypergeometric functions in several variables. *Compos. Math.* **141**, no. 3, 787–810. (Reviewed by A.Yu. Rashkovskiĭ.)

2005b. Passare, Mikael; Tsikh, August. Amoebas: their spines and their contours. **In:** *Idempotent Mathematics and Mathematical Physics*, pp. 275–288. Contemp. Math., 377. Providence, RI: Amer. Math. Soc. (Reviewed by Eugenii Shustin.)

2005c. Leĭnartas, E. K.; Passare, M.; Tsikh, A. K. Asymptotics of multidimensional difference equations (Russian). *Uspekhi Mat. Nauk* **60**, no. 5(365), 171–172; translation in *Russian Math. Surveys* **60**, no. 5, 977–978.

2005d. [Passare, Mikael (Ed.)]. *Nordan Fyra.* [Abstracts from *Nordan* 4, held in] Örnsköldsvik 2000 May 05–07, 18 pp. [Stockholm: Stockholm University 2005.]

2006a. [Passare, Mikael (Ed.)]. *Nordan Fem.* [Abstracts from *Nordan* 5, held in] Voksenåsen, Oslo, 2001 May 04–06, 17 pp. [Stockholm: Stockholm University 2006.]

2006b. Passare, Mikael. Amöbor, Monge–Ampère-mått och trianguleringar av Newton-polytopen [Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope]. **In:** 2006a:6. (A report on 2004a.)

2007a. [Passare, Mikael (Ed.)]. *Nordan Sex.* [Abstracts from *Nordan* 6, held in] Reykjavík 2002 March 08–10, 17 pp. [Stockholm: Stockholm University 2007.]

2007b. [Passare, Mikael (Ed.)]. *Nordan Sju.* [Abstracts from *Nordan* 7, held in] Visby 2003 May 23–25, 17 pp. [Stockholm: Stockholm University 2007.]

2008a. Passare, Mikael. How to compute $\sum 1/n^2$ by solving triangles. *Amer. Math. Monthly* **115**, no. 8, 745–752.

2008b. Leĭnartas, E. K.; Passare, M.; Tsikh, A. K. Multidimensional versions of Poincaré's theorem for difference equations (Russian). *Mat. Sb.* **199**, no. 10, 87–104; translation in *Sb. Math.* **199**, no. 9-10, 1505–1521. (Reviewed by Victor I. Tkachenko.)

2008c. [Passare, Mikael (Ed.)]. *Nordan Åtta.* [Abstracts from *Nordan* 8, held in] Nösund, Orust 2004 May 14–16, 15 pp. [Stockholm: Stockholm University 2008.]

2008d. [Passare, Mikael (Ed.)]. *Nordan Nio.* [Abstracts from *Nordan* 9, held in] Sigtuna 2005 April 22–24, 17 pp. [Stockholm: Stockholm University 2008.]

2008e. Passare, Mikael. Mormors glasögon och räkning modulo nio [Grandma's glasses, and counting modulo nine]. *Nämnaren* **35**(1), 31.

2009a. Passare, Mikael (Ed.). *Complex Analysis and Digital Geometry. Proceedings from the Kiselmanfest,* 2006. Acta Universitatis Upsaliensis. Skrifter rörande Uppsala universitet. C. Organisation och Historia [Publications concerning Uppsala University. C. Organization and History] 86. Uppsala: Uppsala University. 364 pp. ISBN: 978-91-554-7672-4.

2009b. Passare, Mikael. Preface. **In:** Passare (Ed.) 2009a:7–8.

2009c. Passare, Mikael. Christer Kiselman's mathematics. **In:** Passare (Ed.) 2009a:9–26. (Reviewed by Norman Levenberg.)

2009d. [Passare, Mikael (Ed.)]. *Nordan Tio.* [Abstracts from *Nordan* 10, held in] Sundsvall 2006 May 19–21, 14 pp. [Stockholm: Stockholm University 2009.]

2009e. Passare, Mikael. Hypergeometriska serier och integraler (Hypergeometric series and integrals). **In:** 2009d:6.

2010a. Nilsson, Lisa; Passare, Mikael. Discriminant coamoebas in dimension two. *J. Commut. Algebra* **2**, no. 4, 447–471. (Reviewed by Eugenii Shustin.)

2010b. [Passare, Mikael (Ed.)]. *Nordan Elva.* [Abstracts from *Nordan* 11, held in] Oscarsborg, Drøbak, 2007 May 18–20. [Stockholm: Stockholm University 2010.]

2011a. [Passare, Mikael (Ed.)]. *Nordan Tolv.* [Abstracts from *Nordan* 12, held in] Mariehamn 2008 April 18–20. [Stockholm: Stockholm University 2011.]

2011b. Passare, Mikael; Risler, Jean-Jacques. On the curvature of the real amoeba. *Proceedings of the Gökova Geometry-Topology Conference* 2010, pp. 129–134. Somerville, MA: International Press.

2011c. Passare, Mikael; Rojas, J. Maurice; Shapiro, Boris. New multiplier sequences via discriminant amoebae. *Mosc. Math. J.* **11**, no. 3, 547–560, 631.

2011d. Brändén, Petter; Passare, Mikael; Putinar, Mihai (Eds.). *Notions of Positivity and the Geometry of Polynomials.* Trends in mathematics. Basel: Birkhäuser.

2012. Itenberg, Ilia; Jöricke, Burglind; Passare, Mikael (Eds.). *Perspectives in analysis, geometry, and topology. On the occasion of the* 60*th birthday of Oleg Viro.* Progress in Mathematics, 296. New York: Birkhäuser/Springer. xxxii + 464 pp. ISBN: 978-0-8176-8276-7.

2013a. Passare, Mikael; Pochekutov, Dmitry; Tsikh, August. Amoebas of complex hypersurfaces in statistical thermodynamics. *Math. Phys. Anal. Geom.* **16**, 89–108.

2013b. Passare, Mikael; Sottile, Frank. Discriminant coamoebas through homology. *J. Commut. Algebra* **5**, no. 3, 413–440.

2013c. Nilsson, Lisa; Passare, Mikael. Mellin transforms of multivariate rational functions. *J. Geom. Anal.* **23**, no. 1, 24–46. (Reviewed by Osman Yürekli.)

2014. Berkesch, Christine; Forsgård, Jens; Passare, Mikael. Euler–Mellin integrals and A-hypergeometric functions. *Michigan Math. J.* **63**, no. 1, 101–123.

2017. Nisse, Mounir; Passare, Mikael. Amoebas and Coamoebas of Linear Spaces. This volume, pp. 63–80.

Sources

Data on Mikael's publications are taken from *MathSciNet*, from a CV he wrote in 2000, and from certain other documents.

List of Visited Countries

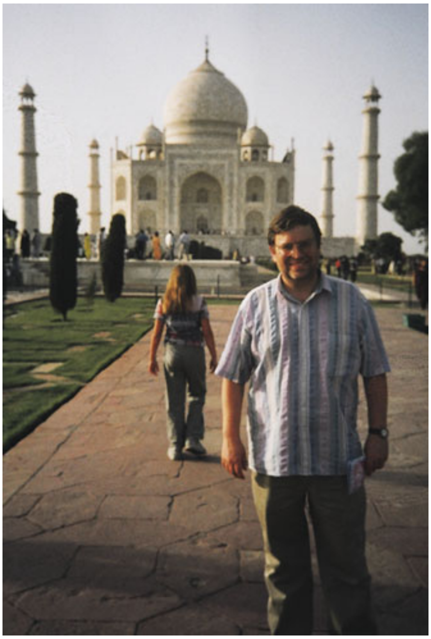
1. Albania
2. Algeria
3. Andorra
4. Angola

2011, Dubai, Arab Emirates

5. Arab Emirates
6. Argentina
7. Armenia
8. Australia
9. Austria
10. Azerbaijan
11. Barbados
12. Belarus
13. Belgium
14. Belize
15. Bolivia
16. Bosnia
17. Botswana
18. Brazil
19. Brunei
20. Bulgaria
21. Burkina Faso
22. Burundi
23. Cambodia
24. Cameroon
25. Canada
26. Cape Verde
27. Central African Republic
28. Chad
29. Czech Republic
30. Chile
31. China
32. Colombia
33. Comoros
34. Congo-Brazzaville
35. Congo-Kinshasa
36. Costa Rica
37. Croatia
38. Cuba
39. Cyprus
40. Denmark
41. East Timor

December 30, 1999,
Pyramids, Egypt

42. Egypt
43. El Salvador
44. Equatorial Guinea
45. Estonia
46. Ethiopia
47. Fiji
48. Finland
49. France
50. Gabon
51. Gambia
52. Georgia
53. Germany
54. Ghana
55. Greece
56. Grenada
57. Guinea
58. Guinea Bissau
59. Honduras
60. Hungary
61. India



1997, Taj Mahal, India

62. Indonesia
63. Ireland
64. Island
65. Israel
66. Italy
67. Japan
68. Jordan
69. Kazakhstan
70. Kenya
71. Kyrgyzstan
72. Kuwait
73. Laos
74. Lebanon
75. Lesotho
76. Latvia
77. Liberia
78. Libya
79. Liechtenstein
80. Lithuania
81. Luxembourg
82. Macedonia
83. Malawi
84. Malaysia
85. Mali
86. Malta
87. Morocco
88. Mauritania
89. Mexico
90. Moldova
91. Monaco
92. Mongolia
93. Montenegro
94. Mozambique
95. Namibia
96. Netherlands
97. Nepal
98. New Zealand
99. Nicaragua

2006, Niger

100. Niger
101. Nigeria
102. North Korea

2009, Svalbard, Norway

103. Norway
104. Oman

2011, Oman

105. Pakistan
106. Panama
107. Paraguay
108. Peru
109. Philippines
110. Poland

1999, Portugal

111. Portugal
112. Qatar
113. Romania
114. Russia
115. Rwanda
116. Samoa Newline (Western Samoa)
117. Saint Lucia
118. Saint Vincent and the Grenadines
119. San Marino
120. Saudi Arabia
121. Senegal
122. Serbia
123. Sierra Leone
124. Singapore
125. Slovakia
126. Slovenia
127. South Africa
128. South Korea
129. Spain
130. Sri Lanka
131. Sudan
132. Swaziland
133. Sweden
134. Switzerland
135. Syria

2003, Warm springs, border to South African Republic

136. Taiwan
137. Tanzania
138. Thailand
139. Togo
140. Tongo
141. Tunisia
142. Turkey
143. Uganda
144. Ukraine
145. United Kingdom
146. United States of America
147. Uruguay
148. Uzbekistan
149. Vatican City
150. Vietnam
151. Zambia
152. Zimbabwe

2001, Vietnam

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 15–33

My Life with Mikael

Galina Passare

My husband Mikael Passare passed away in 2011 during a trip to Oman. He was hiking in the Wadi Shab area and was trying to find a particularly noted cave when he suddenly suffered a heart attack in the strong heat.

His death was quite abrupt and very dramatic for those of us that he left behind on Earth. It was completely unexpected, and according to doctors went so quick that he probably did not feel anything. His death was very beautiful and somehow fitting for his way of life: he was completely healthy, he was on a journey, and when he died he was looking for a cave. Moreover, he actually found it: he was just at the entrance to the cave when he was found.

I must say that it is not at all easy to write about Mikael. He was very meticulous with details, particularly concerning himself. But when you remember someone you have your own view, which can sometimes be different from the person's experience, so I hope that Mikael would not have been disappointed with what I am writing about him. Anyhow, this is how I remember him.

At home (2008, Nockeby, Stockholm, Sweden)

With Galina Passare (2009, Costa Rica)

World

Mikael's relationship to the world was always extremely optimistic. He did not see problems, he saw only solutions, and nothing appeared impossible for him. I was amazed many times when things that I thought belonged to the fantasy world suddenly became a reality with Mikael. When he one day said "I shall visit every country in the world", I thought he was joking. But with 152 countries visited, there were not so many left before he reached the goal. This seemingly impossible task had, as many times before with Mikael, become possible and realistic.

Travel was a very special part of Mikael's life. He always tried to combine travel with his job and he visited universities and mathematical institutions all over the world. In that way, he established new contacts and always helped people with everything imaginable, as in Chad and Niger, where he and his friend Anders Wändahl installed computer programs to provide access to mathematical libraries for local mathematicians. He lectured on every continent, and in several languages.

To document his travels Mikael always sent postcards to himself from the countries he visited. The album with all these postcards from the countries he had visited was very important to him, and he was always pleased to hold it in his hands.

When Mikael was asked which trip was the most enjoyable, it was always the next trip he would make. He could talk forever about future trips – plan, discuss details, imagine how it would be. He always planned his trips thoroughly, and it was very seldom that things did not work out. But if he had to pick some favourites among the trips he had made, then he always mentioned the most arduous and

With Anders Wändahl (2006, Niger)

August 9, 2010, Colombo, Sri Lanka

difficult journeys, as when he travelled alone through Tajikistan on an overcrowded bus without a seat, standing squeezed by other travellers trying to hang on a small handle, almost as if he was hanging in the air. The journey took three days – it was great fun, he thought.

Or when we were looking at the solar eclipse in Nigeria together with Anders Wändahl and wanted to continue to Chad. On the map it looked as if there was a road between Nigeria and Chad, but it turned out that there was not. Of course we wanted to continue anyway. It took us a day to agree with some locals who had cars that could cross the desert. When we finally got under way, it turned out that we had paid 10,000 Swedish kronor to cover just 15 km. It was a very dangerous journey, driving through the desert. On the border to Chad there was only one place to stay for the night – a brothel – and it turned out that there was

Solar eclipse (2006, Niger)

no public communication from there to the capital, so the only option was just sit and wait for a car to pass by.

I even started to think that perhaps we would have to stay in this small village in Chad for ages: Mikael would start a mathematical school there, Anders would fix Internet connection to the world outside and I would be working as a doctor. But the next day, a car caravan arrived and we ended up in these cars completely unexpectedly with our bags, but without any water. It was fine, Mikael said, because the drivers told us it would only take 4 hours to arrive. But it took 24 hours instead. I who whined the whole time and was a woman got at least a little water, but Mikael and Anders began to urinate blood, so it was very lucky that we arrived alive and safe to Chad's capital the following day. Along the way we had to stay overnight in the desert, along with real scorpions. This was an

With Timur Sadykov (1996, Krasnoyarsk, Russia)

example of the type of travel that Mikael loved so much: adventure, experience on the borderline between life and death.

You should not think that Mikael deliberately took great risks or planned difficulties that would put him on the borderline. He always tried to plan in a good and safe way, but the places he wanted to visit were sometimes such that difficulties turned up. And were remembered for the rest of one's life.

Once, Mikael's Achilles tendon broke while playing tennis, shortly before a trip half-way around the world – taking the Trans-Siberian railway through Siberia, Mongolia and China and continuing to North Korea, South Korea, Singapore and Hong Kong. We all were concerned how the trip would go, but Mikael had surgery, got a big bandage around the foot and leg, and hopped on crutches in Krasnoyarsk's rocky nature reserve and on the Great Wall of China. He talked happily to people around him about his accident. Mikael saw this not as a limitation but as a new kind of adventure, a new way to explore the world.

Mikael was phenomenal at orienting himself, even in unfamiliar environments. Perhaps he learned this during his time as a Scout? He loved to find his way using the Sun or stars, regardless of whether we were in the woods, on the sea, or even in a big city. Once we went to Delhi in India, and took a taxi to get to a small railway station. The taxi driver turned out to be a real rascal, and tried to drive us to a different station from which other cars of his could drive us on. I would never have noticed that he was driving in the wrong direction, but Mikael understood that immediately:

– *Where are we going?*
– *Where you requested, sir, we drive to the train station.*
– *Oh no,* said Mikael, you should turn right here, and then left …

It all ended by Mikael taking over and directing the drive, to the great amazement of everyone involved, including me. Mikael had never been to Delhi before.

Mikael had a hard time with beach holidays. If he happened to be staying at such a place, he would lie in a dark room and think about math or solving crossword puzzles. Crosswords, incidentally, were something he particularly loved, both solving them and constructing his own. I think it was because his brain always needed to be occupied. Calm holidays were not for his temperament. Once while we were on a beach holiday in Tunisia, he took the opportunity to fly a one-day round trip to Rome, to send postcards to himself from the Vatican City.

He never demanded any special amenities around him and sometimes appeared to me excessively spartan. He could stay the night anywhere – we spent nights outside a residential area in Ireland on the ground (we did not have enough money to stay in a hotel, this was our first trip outside Sweden), we spent nights in cars, at weird cheap hotels, in the homes of people we did not know before. So it was not always easy to be Mikael's wife. But for me, these difficulties did not bother me much; indeed, I enjoyed such a life.

Languages

It has been said that the centres for mathematics and language in the brain lie close together, and this could truly be seen in Mikael. He had high standards for what it meant to know a language: one should know the whole grammar, speak and write without problems, and of course be able to read thick novels without a dictionary. By this criterion he knew Swedish, English, German, French, Russian

2003, Etosha National Park, Namibia

and Finnish. But he was familiar with many more languages. For example, he knew Italian, Spanish, and Polish well enough to read newspapers and talk to people, and he also learned a little Arabic, Chinese, and Japanese. Wherever he travelled, he would take with him a grammar of the language, and make sure to learn common words and phrases so that he could at least ask directions and order in restaurants.

I was often altogether surprised and awestruck by his linguistic talent. For example, he always denied that he could speak Italian or Spanish, but when we went to some Spanish-speaking countries, he sat on the plane and read thick books such as I mentioned earlier. After that he could speak fluently with people on the streets, and everywhere else. Once we were on a train in France and there were two couples sitting next to us, one from Spain and one from Italy. The couple from Spain said that they could speak Italian, and the couple from Italy that they could speak Spanish, and they really wanted to communicate with each other, but the communication did not work: they simply could not understand each other. In the end, Mikael acted as an interpreter, translating from Spanish to Italian and vice versa. I was completely surprised, and wondered

– *But Mikael, you told me you do not know Spanish and Italian?*

He looked at me and replied:

– *But I was forced to, don't you see?*

January 2008, Bali

Challenges

Mikael always tried to expand his own world, challenge himself and test his limits, whether it concerned his physical limits, his knowledge, his strength or consciousness, everything.

For example he liked to swim between the islands near Ekerö outside Stockholm. Sometimes I did so with him, but I was afraid of big boats suddenly passing by, so I did not do it so often. He did it at least once a summer. Once while swimming from Hässelby to Solviksbadet – a distance of several kilometres – he was attacked by a large gull (of course he was carrying his clothes on his head as he swam), and he laughed about this for a long time. It could have been quite dangerous, because the gull targeted his face.

He had an incredibly competitive spirit, and whenever he played a game he would always naturally try to win. It was quite annoying for his family, his sister

2011, Stockholm

and his mother. But that is how he was – if he played he wanted to win. But if he lost, it was by no means a disaster for him either. *What's the problem? It was just a game, nothing more.* So I never saw him sad, even when he lost.

Mikael was always very hardy, and would go around in only a thin jacket in autumn and even winter. I think it was a matter of inner conviction – he would sometimes say *it is March now*, and go to work without a jacket. And it always went well; he did not feel any cold. And it was the same with heat. He had great endurance. It might also have been about pushing one's boundaries and challenging oneself?

Even in sports, he continued to challenge the world and himself. Almost every day he chose what he wanted to achieve. Biking to work, even to the Institut Mittag-Leffler near Stockholm, swimming 4 km three times a week, playing tennis, skiing alone for days in the mountains, skating many kilometres, often completely alone, even swimming around and between islands, challenging and surpassing himself.

Once we made a skating tour and Mikael ended up in the water. It was quite an unpleasant experience for me, since it was my duty to rescue Mikael. But when he came up completely soaked, at $-10\,^{\circ}\mathrm{C}$, he did not rush to the car and go home, but decided nonetheless to continue and reach the goal that he had planned. Amazingly, he did not feel cold! And he reached the goal, as he always had before.

Home

It was of course great fun to travel with Mikael to many parts of the world, but it was almost as much fun to walk near our own house at home, as he always found new ways to walk and new little things that no-one else had thought of. Walking around the nearby island Kärsön, sitting at a small completely unknown restaurant, eating, almost sitting on the street, at a small kiosk with Thai food in Blackeberg, or going to an unknown exhibition nearby. He always came up with interesting and wholly unexpected ideas. One's whole life became an adventure.

With Galina Passare (1986, Stockholm)

When his nephew about ten years old returned home after visiting us in Stockholm, he was asked what was his most fun and exciting experience in Stockholm. He replied that it was riding in the car with Mikael from home to his work at the university. And I understand exactly what he meant – Mikael's way of driving, choosing routes, talking – everything!

Mikael was quite attached to old things. I think his possessions were like his books – they were alive for him and could tell many stories from the past. It was impossible for him to get rid of such things. This was probably something that he had picked up in his childhood: old things should be used as long as possible. He had his old bike from the 70s, which he always repaired. He was very proud of his old Volvo, a gift from his grandfather, and drove it happily telling everyone that it needed no seat belts because it was too old – it was the apple of his eye.

Mikael asked me many times "*Don't you remember this, you've read it*?" It could be about botany, zoology, religion – anything. If you have read it, surely you remember it; this was his attitude. Unfortunately I didn't. But it seemed he remembered everything he had read, and there was no subject he could not talk about.

2010, Västervik

Mikael was theoretically talented and incredibly intelligent, and eventually became a professor, but that did not mean he was impractical. He could do everything he needed. For example, he made the floor in the basement, removed a large concrete base that had supported a washing machine, and so on. He built his country house almost single-handedly. He was gifted in doing practical things. Doing things with his own hands reflected very well his principle of life: if you do something, you should do it properly or not at all.

Once we had Polish workers with us, rebuilding our garage. They left behind a huge pile of stones and gravel. They asked us if we wanted them to remove it. It would take them at least a week, and we would have to pay for it. No, said Mikael. And the next day the pile was gone – Mikael had removed it himself, working all day and half the night. Imagine how surprised the workers looked when they arrived the following morning!

If we were buying furniture, it had to be the best and the most famous designer. Why? Well, the point was simply that if you do something, then you should do it properly!

Mikael did not care much about his appearance, but if he did, it had to be done properly. When he became a professor, it was almost mandatory to get a good jacket. And what jacket? Obviously a tweed jacket. And where would you buy it? In England, of course!

Books were very important to him. He used to say that he would like to have a wife who was so fond of books that she used special gloves for them. He would be very angry if his books were damaged – his books were almost living things for him.

He drank pure tea, not tea with milk, because tea with milk looks like dishwater, and drinking something should be an aesthetic experience.

August 2011, at home Stockholm, Nockeby

The children were of course very important to Mikael. He tried very hard to give them the best of everything – playing the piano, singing, learning to read music, play tennis, learning languages, mathematics, physics, geography, and much more.

Traditions like Christmas and Midsummer were always very important for Mikael. Christmas we always celebrated with his mother.

With Max Passare (December 2010, San Francisco)

Music

The music centre in the brain is said to be located close to the centres for mathematics and languages. Indeed, music was another of Mikael's talents. He composed music for a play *City blues*, that has been shown much at the theatre in Västerås. He loved listening to a radio programme where one had to guess the names of melodies, and he almost always knew the answer. His knowledge of music was absolutely amazing. Besides this, he played the clarinet and piano.

He liked to sing, and sang in various choirs (including the choir at Moscow University), took private singing lessons, and always sang when we traveled anywhere by car. Singing was his way of keeping awake when he drove at night, which he often did in order to save time when driving long distances.

Projects

Mikael's life always consisted of a number of projects – small and big tasks that he wanted to accomplish in his life. They could be very different things, and often they could barely be described. Swimming around Lindö outside Västervik, visiting all the countries in the world, learning how to play a particular piece on the clarinet

With Märta and Galina Passare (December 2010, Berkeley, California, USA)

or piano, learning Japanese characters, writing diary every day, driving in one day to Bulgaria, and so on.

A project that was particularly important to him in his last years was learning Finnish. He started a small group, consisting of three people; besides him, there were a professor and a graduate student in mathematics. He spoke of the group as the Finnish club. Once, he learned that his favourite mystery writer Reijo Mäki's book *Fallen Angel* was to become a film. Then he immediately planned an excursion for the Finnish club – they would go to Pori, where the action takes place. And they did. They went to Finland, drank the special liqueur with caramel, just like in the book, visited all the places that the writer described, and finally saw the movie itself. It was typical of Mikael: organising, and fully carrying out, something that becomes an experience you will never forget.

One of the dreams that he did not have time to carry out was making an underground passage. In the beginning I thought that it was just a boyhood dream, nothing more, but after 152 countries visited, I have suddenly understood that with his attitude, determination, and working capacity everything in this world is possible – simply everything. Unfortunately, he could not build his underground passage.

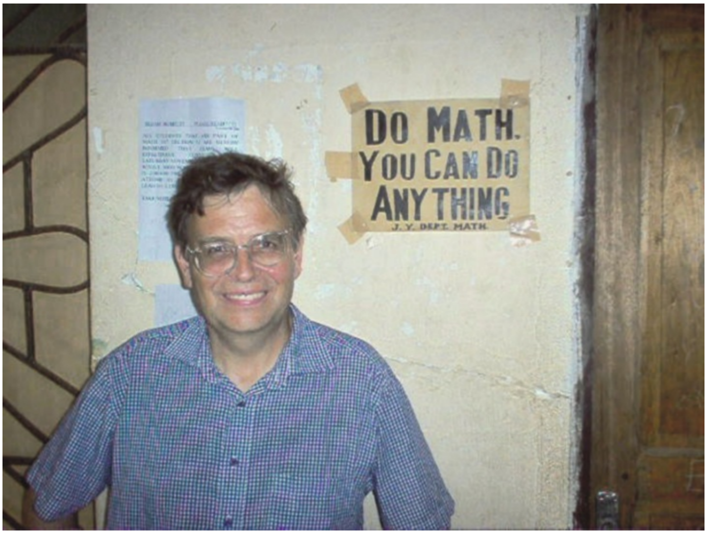
Personality

I have never heard or seen Mikael being envious of anyone. Never. He did not understand how one could be. A basic rule for him was always: you should never pay attention to what you don't have or what you have lost, but just think of what you have.

There are many things in the world that are just frustrating – paying bills, for example, standing in line at a border, or paying a fine. Unpleasant things, as I saw

With Märta Passare (1996, Monte Carlo)

it, but Mikael liked to take such matters into his own hands and deal with them. And when you asked him, but why must one come to such awfully tedious things in the world, he would say, "*Galya, that's how the world works, it's just a game. And this is a matter of just accepting such rules. To play the game, you must do it, stand in the queue, for example, and be happy because it's just a game.*"



2010, Camerun

At Nordan meeting in Mariehamn (2008, Mariehamn, Åland)

Mikael often said: if there is something that you do not like but cannot change, then you should accept it, and instead work on changing your relationship to it.

Mikael always had his own perspective on things. But he was always open and curious, ready to expand his own boundaries. He was never religious, for instance, but knew the Bible very well, and questions that could not be decided he always left open. Even great religious questions. He was really not a conservative character. He was open to all sorts of questions and discussions.

Mikael was always able to talk with people on their own level. He talked with alcoholics in Russia, with professors in France and the USA, with cleaning ladies in Stockholm, with Polish workers, film directors, and representatives of TV and radio. It was always at the appropriate level, and the other person would never

September 2011, Oman (the last photo)

feel that he was talking with a professor, although that is what Mikael was. He could explain complex things in a very simple way, a talent that never ceased to amaze me. Our daughter Märta, who studied mathematics for a while, always said that her father was the best teacher.

Mikael wrote a diary, and indeed had done so since high school. Day after day. It helped him to remember people he met, addresses, events, weather. He never failed to describe a single day in his life, until the end.

When we celebrated 25 years together, we went to the marvellous Easter Island. And for another anniversary, we traveled to Bali, walked in the woods there, and experienced wonderful things. But the most amazing thing was that the very day we arrived in Bali, there was a celebration called Gala Young, almost like my name. And again I marvelled at his ability to come up with amazing ideas at the right moment.

Mikael was an incredible man, but perhaps a little difficult in that it was not easy to find his weaknesses, where other people could beat him. I think in this way he irritated some people with his great intellect, his incredible capacity for concentration and strength, an inner strength that could also be expressed as physical strength.

Mikael had a temper and could sometimes react strongly to things he did not like, but he could not be angry for long, and only after a few minutes he would be back to normal.

He was a very good listener, but never understood why people would talk about their problems – if you have a problem, it should be solved, not talked about! He did not like superficial conversations, where people talk for long periods about nothing. He proposed that we marry one week after we met. Why should one talk more if it was not needed?

Galina Passare *Lost in the depths of cosmos* (2015)

Thanks

Mikael found time to accomplish incredibly many things in his relatively short life. But one remarkable thing was that he always had time. It seemed like he was everywhere at once. I think that his students felt they could talk to him at any time; his children may even have felt that he interfered too much. He had time to be with his friends and of course with me. How did he manage all that? I do not understand at all.

Mikael was the most amazing and incredible and kind-hearted person I have met in my life. I am deeply thankful to my Destiny for sending him. And I am very grateful to Mikael for his talent at finding inspired solutions in all possible situations.

Mikael really took care of me, which I am very grateful for, but also made sure that I did not always have it too easy. And he forced me to take care of myself. For instance, he might leave me in the middle of New York alone and say we would meet at a certain place at a later date, and let me take care of myself. Or just drop me suddenly at the Metro station in a completely foreign city and say he was on his way to a meeting and I had to find the way back to the hotel on my own. It was not particularly dangerous, really, but quite unexpected. He disappeared so quickly that I was always quite taken aback.

And he did so now, too, left me alone on this earth, suddenly and unexpectedly, and I must try to find the way back – to myself. I have to do it. And I will find the way, Mikael. I promise.

Galina Passare
e-mail: passaregalina@hotmail.com

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 35–57

Mikael Passare (1959–2011)

Christer O. Kiselman

Mikael Passare was a brilliant mathematician who died much too early. In this chapter we present a sketch of his work and life.

Mikael was born in Västerås, Sweden, on 1959 January 01, and pursued a fast and brilliant career as a mathematician. He started his studies at Uppsala University in the fall of 1976 while still a high-school student, merely seventeen and a half. He finished high school in June 1978 at the Rudbeckianska skolan in Västerås, gave his first seminar talk in November 1978 at Uppsala University, where he got his Bachelor Degree in 1979, and where he also was an assistant.

Mikael Pettersson (age 24), Jean Francois Colombeau,
Leif Abrahamsson, and Urban Cegrell
(November 1983, Uppsala; Sweden)
(photo Christer Kiselman)

He was accepted as a graduate student at Uppsala University on 1980 February 14 with me as advisor, and presented his thesis on 1984 December 15. The opponent was Nils Øvrelid.

The thesis was written by a certain Mikael Pettersson, the family name Pettersson being his from birth. However, on 1984 December 18, three days after the thesis presentation, the Swedish Patent and Registration Office approved the family name *Passare*[1] for Kjell Alrik Mikael Pettersson and Galina Pettersson, née Lepjosjkina. As a consequnece, the diploma for the Degree of Doctor of Philosophy was issued on 1984 December 19 to Kjell Alrik Mikael Passare.

Mikael was a research assistant (half-time) and lecturer (half-time) at Stockholm University from January through June 1985, research assistant financed by the Swedish Natural Science Research Council (NFR) July 1985–August 1986 and later research assistant, July 1987 through 1990.

With Christer Kiselman (November 8, 1991, Uppsala, Sweden)
(photo Galina Passare)

He received the title of *Docent* (corresponding to the *Habilitation* in some countries) on 1988 January 28. He was a senior university lecturer (full time) from July 1988, from time to time on leave of absence. Later he was a research lecturer at the Royal Institute of Technology, July 1990–1994. He was appointed full professor at Stockholm University from 1994 October 01.

In addition to English and French, he studied Russian during three years in high school (Rudbeckianska skolan 1978), and later did his military service at the National Defence Radio Establishment (the Swedish national authority for signals

[1] The Swedish word *passare* means 'a pair of compasses'. So the classical task of compass-and-straightedge construction in Euclidean geometry receives a new meaning in Swedish: 'construction using Passare and ruler'.

intelligence) from June 1979 through August 1980. There he learned even more Russian.

He spent four academic years in four different countries: during the academic year 1980-81 he was at Stanford University; during 1981–82 at Lomonosov University in Moscow; during 1986–87 at Université Pierre et Marie Curie (Paris VI) and Université de Paris-Sud (Paris IX), having received a post-doctoral fellowship from the Swedish Natural Science Research Council (NFR). During 1992–93 he was at Humboldt-Universität zu Berlin on an Alexander von Humboldt fellowship. He was a guest professor in France on several occasions: at Toulouse (June 1988), Grenoble (April 1992), Bordeaux (May 1992), Paris VII (March 1993), Lille (April 1999), and Bordeaux again (June 2000).

Mikael was much appreciated as a researcher and teacher, and was very active outside the university. He was Head of the Department of Mathematics at Stockholm University from January 2005 through August 2010, and then Director of the newly created Stockholm Mathematics Center, common to Stockholm University and the Royal Institute of Technology. When Burglind Juhl-Jöricke and Oleg Viro had resigned from Uppsala University on 2007 February 08, he arranged for a guest professorship for Burglind at Stockholm University, and was one of the organizers of a big conference in honor of Oleg, *Perspectives in Analysis, Geometry, and Topology*, at Stockholm University during seven days, 2008 May 19–25.

As president of the Swedish National Committee for Mathematics, he led the Swedish delegation to the General Assembly of the International Mathematical Union in Bangalore, Karnataka, India, in August 2010. In 2011 he invited, as president of the National Committee, Bernd Sturmfels to lecture in Linköping, Lund, and Göteborg.

Mikael Passare was Deputy Director for Institut Mittag-Leffler, Djursholm, Sweden, from 2010. He was very much appreciated for his activity there, which included organizing the Felix Klein Days for teachers and a research school for high-school students.

Starting in July 2001, he served during ten years as one of the editors of the *Arkiv för matematik* (Ari Laptev, personal communication 2011-10-19). During the period 2004 April 01–2009 June 25 he was one of the Associate Editors for the *Journal of Mathematical Analysis and Applications* (Don Prince, personal communication 2011-10-13).

Mikael was a member of the Swedish Committee for Mathematics Education (SKM) from January 1997, when SKM started its activity, until December 2004. Mikael's efforts in SKM can only be explained by his firm dedication to mathematics education in the schools. He participated actively by organizing meetings, authoring reports to policy makers, and influencing politicians and officials at the Ministry of Education and the Swedish National Agency for Education. (Gerd Brandell, personal communication 2011-10-17.)

The activity of SKM to which he devoted most of his energy was the international competition *International Mathematical Kangaroo*, originally *Kangourou sans frontières*, in Swedish *Kängurun – Matematikens Hopp*. He took the initiative to start, with SKM as organizer, a Swedish version of this competition in 1999. He translated the problems, which arrived in English or French, up to 2009. He checked also that the mathematical content was correct after the necessary adaption to Swedish traditions in problem formulation and other circumstances. He participated, at least during the first five to six years, in the choice of problems to the Swedish edition, and he continued his commitment to the competition even after his time in SKM. The competition is run by SKM in cooperation with the National Centre for Mathematics Education, NCM. In 2010, more than 80,000 students at all levels participated. (Gerd Brandell, personal communication 2011-10-17; Karin Wallby, personal communication 2011-10-19.)

Mikael was a member of the Steering Group for the National Graduate School in Mathematics Education from March 2000, until it ceased in December 2006. The school, which had about twenty PhD students, was financed by the Bank of Sweden Tercentenary Foundation (*Riksbankens Jubileumsfond*, RJ) and the Swedish Research Council (*Vetenskapsrådet*, VR). He actively and constructively contributed to shaping the education of this Graduate School, both in his role as member of the Steering Group and by participating in many meetings between PhD students and advisors. He was project leader for the school's activity at the Department of Mathematics at Stockholm University. Of the school's PhD students, two were at Stockholm: Kirsti Löfwall Hemmi and Andreas Ryve, who both got their PhDs in 2006. (Gerd Brandell, personal communication 2011-10-17.)

The Sonja Kovalevsky School in Stockholm, a private elementary school, started its activity in the Fall of 1999. Its profile includes chess, mathematics, and Russian. The aim was, among other things, to benefit from educational experience from Russia. Mikael was a member of the school's Board from the beginning.

At the time of his death, Mikael was President of the Swedish Mathematical Society and also a member of the Committee for Developing Countries (CDC) of the European Mathematical Society. His activity for mathematics in Africa is described in a later section.

Mikael died from a sudden cardiac arrest in Oman in the evening of 2011 September 15.[2] His next of kin are his wife Galina Passare, his son Max, and his daughter Märta.

Mikael's nine PhD students

Mikael served as advisor of nine PhD students who successfully completed their degrees. They are registered in the *Mathematics Genealogy Project* and are:

[2] The cause of death has been established to be a complete occlusion of the right coronary artery leading to an acute myocardial infarction and an immediate death; there are no injuries whatsoever that would indicate a fall into a canyon (The Swedish National Board of Forensic Medicine (2011)).

Yang Xing, 1992, Stockholm University: *Zeros and Growth of Entire Functions of Several Variables, the Complex Monge–Ampère Operator and Some Related Topics.* Now Senior Lecturer at Lund University.

Mikael Forsberg, 1998, The Royal Institute of Technology: *Amoebas and Laurent Series.* Now Senior Lecturer at Gävle University College.

Lars Filipsson, 1999, The Royal Institute of Technology: *On Polynomial Interpolation and Complex Convexity.* Now Senior Lecturer at the Royal Institute of Technology, Stockholm.

Timur Sadykov, 2002, Stockholm University: *Hypergeometric Functions in Several Complex Variables.* Although not mentioned in the Genealogy Project, August Tsikh served as a coadvisor (Timur Sadykov, personal communication 2011-11-26; August Tsikh, personal communication 2011-12-06). Now Timur is Full Professor at the Department of Mathematics at the Russian Plekhanov University, Moscow.

Hans Rullgård, 2003, Stockholm University: *Topics in Geometry, Analysis and Inverse Problems.* Now at Comsol Group, Stockholm, a company providing software solutions for multiphysics modelling.

Johan Andersson, 2006, Stockholm University: *Summation Formulae and Zeta Functions.* Now Senior Lecturer at Mälardalen University, Campus Västerås.

Alexey Shchuplev, 2007, Stockholm University: *Toric Varieties and Residues.* August Tsikh was second advisor. Now Assistant Professor and Head of Laboratory at the Siberian Federal University in Krasnoyarsk.

David Jacquet, 2008, Stockholm University: *On Complex Convexity.* Now Specialist Consultant in quantitative analysis and CEO of his company Mathsolutions Sweden AB.

Lisa Nilsson, 2009, Stockholm University: *Amoebas, Discriminants, and Hypergeometric Functions.* August Tsikh was second advisor. Now she is employed at the insurance company If Skadeförsäkring AB in Stockholm as risk analyst within capital modelling.

Mikael's mathematics

Residue theory

Mikael soon became known as an eminent researcher in complex analysis in several variables, where his thesis was an important breakthrough with new results in residue theory. Its title was *Residues, Currents, and Their Relation to Ideals of Holomorphic Functions* [1984], and it was later published in [1988c].[3]

Residue theory in several variables is a notoriously difficult part of complex analysis. Mikael's work was inspired by that of Miguel E. M. Herrera (1938–1984). Miguel and I were together at the Institute for Advanced Study in Princeton

[3] Years in brackets refer to the list of Mikael Passare's publications. Years in parentheses refer to publications listed at the end of this chapter.

during the academic year 1965–66, and it was there that I learned about residues from him.[4] His results, which culminated in the paper by Herrera and Lieberman (1971) and the much quoted book by Coleff and Herrera (1978), were well known long before these publications. I could somehow serve as mediator to Mikael for this interest without doing much research on residues myself.

Also Alicia Dickenstein, who was a student of Miguel and got her PhD at Buenos Aires in 1982, knew this theory very well and soon came into contact with Mikael. As for integral formulas, Mikael took advice from Bo Berndtsson, already then a renowned expert in that field.

Another important person for Mikael's mathematical development was Gennadi Henkin (1942–2016). They met in Moscow during the academic year 1981–82, afterwards several times in the period 1983–1990, and then in France and Sweden during the period 1991–2010, for example in Trosa in 1997, Saltsjöbaden in 1999, and in Uppsala in 2006. During these meetings they discussed, in particular, integral formulas of Cauchy–Leray type and applications from the papers of Gennadi and Bo (starting with Henkin (1969)).

While residues in one complex variable have been well understood for a long time, the situation is quite different in several variables. There were pioneers like Henri Poincaré (1854–1912) and Jean Leray (1906–1998). Alexandre Grothendieck (1928–2014) developed a residue theory in higher dimensions, but it was quite abstract. Through work by Miguel Herrera, François Norguet (1929–2010), and Pierre Dolbeault (1924–2015), the theory could be linked to distribution theory, developed by Laurent Schwartz (1915–2002), and that was the road that Mikael continued to follow. He worked intensively with August Tsikh, on residue theory as well as on amoebas.

Residues in one complex variable. In one complex variable we can observe that there is a lot of symmetry:

$$\int_{\varepsilon<|z|<r} z^j \bar{z}^k f(|z|) dx \wedge dy = 0, \qquad j,k \in \mathbf{Z}, \quad j \neq k.$$

This means that heavy masses are balanced, and implies that, when calculating residues, it is enough to work with the principal value, PV (*valeur principale*, VP); we need not use the more difficult and unstable construction of the finite part, FP (*partie finie*, PF). (In real analysis, the finite part inevitably appears: the distribution on the real axis given by the function $\log|x|$, $x \in \mathbf{R}$, has the derivative $\mathrm{PV}(1/x)$ and the second derivative $-\mathrm{FP}(1/x^2)$.)

If we write a smooth function φ as $\varphi(z) = P(z) + R(z)$, where P is a polynomial in z and $\bar{z}$ of degree at most $m-1$, $m \in \mathbf{N}$, $m \geq 1$, and $R(z)/z^m$ is bounded near the origin, it follows from the symmetry mentioned above that P does not

[4] He also introduced me to a great writer: Ursula LeGuin.

influence the following integral at all.

$$\int\limits_{\varepsilon<|z|<r} \frac{\varphi(z)}{z^m} dx \wedge dy = \int\limits_{\varepsilon<|z|<r} \frac{R(z)}{z^m} dx \wedge dy.$$

As ε tends to 0, the last integral tends to

$$\int\limits_{|z|<r} \frac{R(z)}{z^m} dx \wedge dy.$$

We define the *principal value* $\mathrm{PV}(1/z^m)$ of $1/z^m$ by

$$\left\langle \mathrm{PV}\left(\frac{1}{z^m}\right), \varphi \right\rangle = \mathrm{PV}\int\limits_{\mathbf{C}} \frac{\varphi(z)}{z^m} dx \wedge dy = \lim_{\varepsilon\to 0} \int\limits_{\varepsilon<|z|} \frac{\varphi(z)}{z^m} dx \wedge dy,$$

which exists for all test functions $\varphi \in \mathscr{D}(\mathbf{C})$. If now f/g is meromorphic with a pole at the origin, we obtain

$$\left\langle \mathrm{PV}\left(\frac{f}{g}\right), \varphi \right\rangle = \mathrm{PV}\int\limits_{\mathbf{C}} \frac{f(z)}{g(z)} \varphi(z) dx \wedge dy, \qquad \varphi \in \mathscr{D}(\mathbf{C}),$$

and we define the *residue* $\mathbf{res}(f/g)$ of f/g as

$$\mathbf{res}\left(\frac{f}{g}\right) = \frac{\partial}{\partial \bar z} \mathrm{PV}\left(\frac{f}{g}\right) \in \mathscr{D}'(\mathbf{C}).$$

Residues in several variables. Let f and g be holomorphic functions of n complex variables. The *principal value* $\mathrm{PV}(f/g)$ of f/g is a distribution defined by the formula

$$\left\langle \mathrm{PV}\left(\frac{f}{g}\right), \varphi \right\rangle = \lim_{\varepsilon\to 0} \int\limits_{|g|>\varepsilon} \frac{f\varphi}{g} = \lim_{\varepsilon\to 0} \int \frac{\chi f \varphi}{g}, \qquad \varphi \in \mathscr{D}(\mathbf{C}^n),$$

where $\chi = \chi(|g|/\varepsilon)$ and χ is a smooth function on the real axis satisfying $0 \le \chi \le 1$ and $\chi(t) = 0$ for $t \le 1$, $\chi(t) = 1$ for $t \ge 2$ (in [1985:727] when $f = 1$; in [1988:39] in general).

The *residue current* is $\bar\partial\,\mathrm{PV}(f/g)$. It is natural to ask if there exist interesting algebras of these currents, for instance whether the products

$$(\mathrm{PV}(f_1/g_1))(\mathrm{PV}(f_2/g_2)), \quad \left(\overline{\partial}(\mathrm{PV}(f_1/g_1))\right)(\mathrm{PV}(f_2/g_2))$$

and other similar products can be defined.

Schwartz proved (1954) that it is in general impossible to multiply two distributions while respecting the associative law. He indicated three distributions $u, v, w \in \mathscr{D}'(\mathbf{R})$ where uv, vw, $(uv)w$ and $u(vw)$ all have a good meaning, but where $(uv)w \ne u(vw)$. He took $u = \mathrm{PV}(1/x)$, the principal value of $1/x$; v as the identity, i.e., the smooth function $v(x) = x$, which can be multiplied to any distribution; and $w = \delta$, the Dirac measure placed at the origin. Then we have $uv = 1$, $(uv)w = \delta$, while $vw = 0$, $u(vw) = 0$. Hence there is no associative multiplication.

Mikael's construction of residue currents goes as follows. Take

$$f = (f_1, \dots, f_{p+q}), \quad g = (g_1, \dots, g_{p+q}),$$

two $(p+q)$-tuples of holomorphic functions, and consider the limit

$$\lim_{\varepsilon_j \to 0} \frac{f_1}{g_1} \cdots \frac{f_{p+q}}{g_{p+q}} \bar\partial\chi_1 \wedge \cdots \wedge \bar\partial\chi_p \cdot \chi_{p+1} \cdots \chi_{p+q},$$

where $\chi_j = \chi(|g_j|/\varepsilon_j)$, and the ε_j tend to zero in some way.

Coleff and Herrera (1978:35–36) took $q = 0$ or 1, and assumed that ε_j tends to zero much faster than ε_{j+1}, which in this context means that $\varepsilon_j/\varepsilon_{j+1}^m \to 0$ for all $m \in \mathbf{N}$ and $j = 1, \dots, p+q-1$; thus it is almost an iterated limit. This gives rise to the strange situation that, in general, the limit depends on the order of the functions (and is not just an alternating product).

Mikael took instead $\varepsilon_j = \varepsilon^{s_j}$ for fixed $s_1, \dots, s_{p+q}$. The limit, which will be written as $R^pP^q[f/g](s)$, where we now write $[\dots]$ for the principal value, does not exist for arbitrary s_j. But he proved [1985:728] that, if we remove finitely many hyperplanes, then $R^pP^q[f/g](s)$ is locally constant in a finite subdivision of the simplex

$$\Sigma = \left\{ s \in \mathbf{R}^{p+q};\ s_j > 0,\ \ \sum s_j = 1 \right\},$$

so that the mean value

$$R^pP^q\left[\frac{f}{g}\right] = \fint_\Sigma R^pP^q\left[\frac{f}{g}\right](s) = \bar\partial\left[\frac{f_1}{g_1}\right] \wedge \cdots \wedge \bar\partial\left[\frac{f_p}{g_p}\right] \cdot \left[\frac{f_{p+1}}{g_{p+1}}\right] \cdots \left[\frac{f_{p+q}}{g_{p+q}}\right]$$

exists (Definition A in [1987]). This is the product of p residue currents and q principal-value distributions.

In the little paper [1993c], based on his talk when accepting the Thuréus Prize in 1991, he discusses the possibility of defining the product $\mathrm{PV}(1/x)\delta$ on the real axis, and finds that it should be $-\frac{1}{2}\delta'$, which is the mean value of $-\delta'$ and zero. This is an analogue in real analysis to the mean value over Σ which he considered in the complex case.

Leibniz' rule for the derivative of a product and some other rules of calculus hold; for example we have [1988d:43]:

$$\left[\frac{1}{z_1}\right]\left[\frac{z_1}{z_2}\right] = \left[\frac{1}{z_2}\right],$$

which yields

$$\left(\bar\partial\left[\frac{1}{z_1}\right]\right)\left\{\left[\frac{1}{z_1}\right]\left[\frac{z_1}{z_2}\right]\right\} = \left(\bar\partial\left[\frac{1}{z_1}\right]\right)\left[\frac{1}{z_2}\right],$$

while

$$\left\{\left(\bar\partial\left[\frac{1}{z_1}\right]\right)\left[\frac{1}{z_1}\right]\right\}\left[\frac{z_1}{z_2}\right] = \frac{1}{2}z_1\left(\bar\partial\left[\frac{1}{z_1^2}\right]\right)\left[\frac{1}{z_2}\right] = \frac{1}{2}\left(\bar\partial\left[\frac{1}{z_1}\right]\right)\left[\frac{1}{z_2}\right].$$

Thus the associative law does not hold.

We saw in Schwartz' example that an associative multiplication is impossible in general; the example shown here makes us wonder whether it is possible to define an associative multiplication in some algebra of principal-value distributions and residue currents. We may also ask if there is an interesting non-associative algebra of principal-value distributions and residue currents.

For complete intersections, i.e., when the set of common zeros of f_1, f_2, ..., f_p has maximal codimension, Mikael established a division formula with remainder term:

$$h = \sum_1^p g_j f_j + h \cdot \mathbf{res},$$

where **res** is the residue current, which is a factor in the remainder term $h \cdot \mathbf{res}$ and has the property that $f_j \cdot \mathbf{res} = 0$ for all j. This implies that h belongs to the ideal generated by $f_1, \ldots, f_p$ if and only if $h \cdot \mathbf{res}$ vanishes. This is a beautiful characterization of the ideals of holomorphic functions and explains the choice of title in the papers [1984, 1986, 1988c]. The characterization of the ideals with the help of residues was proved independently and at about the same time by Alicia Dickenstein and Carmen Sessa (1985:424).

This characterization of ideals enabled Mikael and Bo to formulate an elegant and explicit variant of Leon Ehrenpreis' Fundamental Principle; it was published in a joint paper with Bo [1989]. Later, Mats Andersson and Elizabeth Wulcan (2007) could define a residue without the assumption of a complete intersection. In their work, an important role was played by a paper by Mikael, August, and Alain Yger, viz. [2000b].

For complete intersections we have, according to [1988d:42, Theorem 4 iii)],

$$g_j R^p P^1[1/g] = 0, \qquad g = (g_1, \ldots, g_{p+1}), \quad j = 1, \ldots, p,$$

where $R^p P^1[1/g]$ is the mean value over the simplex Σ defined above. However, Mikael showed in [1988d:43, Example 3] that this is not true when we do not have complete intersections: with $n = 2$, $p = 2$, $q = 0$, $f_1 = f_2 = 1$, $g_1(z) = z_1 z_2$, $g_2(z) = z_2$ we get

$$R^2[1/g] = \overline{\partial}\left[\frac{1}{z_1 z_2}\right] \wedge \overline{\partial}\left[\frac{1}{z_2}\right] = \frac{1}{2}\,\overline{\partial}\left[\frac{1}{z_1}\right] \wedge \overline{\partial}\left[\frac{1}{z_2^2}\right];$$

when we multiply this current by g_2, we get

$$g_2 R^2[1/g] = \frac{1}{2} z_2 \overline{\partial}\left[\frac{1}{z_1}\right] \wedge \overline{\partial}\left[\frac{1}{z_2^2}\right] = \frac{1}{2}\,\overline{\partial}\left[\frac{1}{z_1}\right] \wedge \overline{\partial}\left[\frac{1}{z_2}\right] \neq 0.$$

Other examples in this calculus are [1988d:43]:

$$\left[\frac{1}{z_1}\right]\bar{\partial}\left[\frac{1}{z_2^2}\right] = 2\left[\frac{1}{z_1 z_2}\right]\bar{\partial}\left[\frac{1}{z_2}\right] \text{ and } \bar{\partial}\left[\frac{1}{z_1}\right] \wedge \bar{\partial}\left[\frac{1}{z_2^2}\right] = 2\bar{\partial}\left[\frac{1}{z_1 z_2}\right] \wedge \bar{\partial}\left[\frac{1}{z_2}\right].$$

The original definition and the definition which uses meromorphic extension agree [1987:159]:

$$R^p P^q \left[\frac{1}{g}\right] = \lim_{\varepsilon \to 0} \frac{\bar\partial |g_1|^\varepsilon}{g_1} \wedge \cdots \wedge \frac{\bar\partial |g_p|^\varepsilon}{g_p} \cdot \frac{|g_{p+1}|^\varepsilon}{g_{p+1}} \cdots \frac{|g_{p+q}|^\varepsilon}{g_{p+q}}.$$

Here the left-hand side is defined according to Definition A already mentioned, while the right-hand side, called Definition B, is the one which comes from meromorphic extension. In fact, when $\Re\varepsilon$ is sufficiently large, the expression following the lim operator defines a current. It has a meromorphic extension which is holomorphic near $\varepsilon = 0$, and the right-hand side is its value at $\varepsilon = 0$.

In a CV which Mikael wrote in 2000 he mentions a book project with August Tsikh as coauthor and which had the title *Multidimensional Residues and Toric Varieties.* He gives a detailed table of contents of the five chapters in the book. Later they abandoned this project, since amoebas and tropical geometry became more interesting for them, and they aimed at writing a book on amoebas (August Tsikh, personal communication 2011-10-06).

Lineal convexity

André Martineau (1930–1972) gave a couple of seminars on lineal convexity (*convexité linéelle*) in Nice during the academic year October 1967 through September 1968, when I was there. This is a kind of complex convexity which is stronger than pseudoconvexity and weaker than convexity. Since I was of the opinion that the results for this convexity property were too scattered in the literature and did not always have optimal proofs, I suggested that Mikael write a survey article on the topic.

On the one hand, this piece of advice was certainly very good, for he found a lot of results in cooperation with his friends Mats Andersson and Ragnar Sigurðsson (Mikael's mathematical uncle). On the other hand, it was perhaps not such a good suggestion, for the survey just kept growing, and two preprints were circulating starting in 1991[5] – and by then they had been busy writing for a long time already. The article became a book, and it did not appear until 2004 [2004b]. Anyway, it is thanks to André Martineau that lineal convexity came to be studied in the Nordic countries – and the book has become a standard reference.

In this book, the authors study in detail a property which Martineau called strong lineal convexity (*convexité linéelle forte*), and which he did not characterize geometrically. This notion, in the book called **C**-*convexity*, is not linked to any cleistomorphism (closure operator), since the intersection of two strongly lineally convex sets need not have the property. Therefore it has a different character than

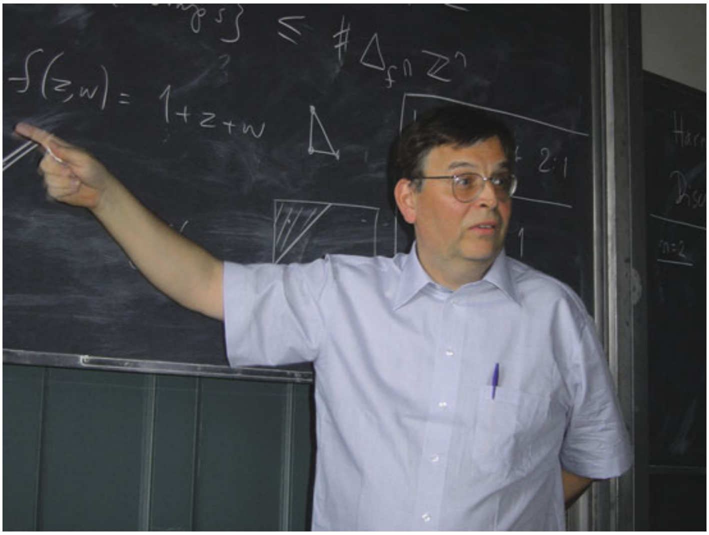
[5]I no longer possess any documentation about preprints from 1991, but in the CV that Mikael wrote in 2000, two are mentioned: Andersson, Mats; Passare, Mikael; Sigurdsson, Ragnar (1995), Complex convexity and analytic functionals I, Reykjavík, 71 pp.; and (2000), Complex convexity and analytic functionals II, Reykjavík and Sundsvall, 103 pp. The book [2004b] came to comprise xii + 160 pages.

lineal convexity and usual convexity, which, as is well known, are the fixed points of cleistomorphisms.

An important characterization of strong lineal convexity has been obtained recently by Gennadi Henkin and Peter Polyakov: a lineally convex compact set is strongly lineally convex if and only if it can be approximated in the Hausdorff metric by lineally convex compact sets with smooth boundaries (Henkin & Polyakov (2012: Proposition 2.4)). For related question see also Kiselman (2016) and the references mentioned there.

Amoebas and tropical geometry

Mikael's later work is concerned with amoebas and coamoebas – the first publications in this field were Mikael Forsberg's thesis (1998) and their joint paper [2000a]. The spine of an amoeba – in mathematical zoology, amoebas are vertebrates – is a tropical hypersurface. Tropical mathematics is a rather new branch of mathematics, where addition and multiplication is replaced by the maximum operation and addition, somewhat similar to taking the logarithm of a sum and a product.[6] Mikael's interest in tropical mathematics was a break with his earlier work on complex analysis, which he once compared with my switching to digital geometry.



At Institut Mittag-Leffler (2008, Djursholm, Sweden)
(photo Ragnar Sigurdsson)

An amoeba is a set in $\mathbf{R}^n$ defined as follows. We define a mapping

$$\mathrm{Log}\colon (\mathbf{C} \smallsetminus \{0\})^n \to \mathbf{R}^n \quad \text{by} \quad \mathrm{Log}(z) = (\log|z_1|, \log|z_2|, \ldots, \log|z_n|).$$

[6] It seems that the first use of the adjective *tropical* in this sense in the title of a publication was in Simon (1988).

If f is a function defined in $(\mathbf{C} \smallsetminus \{0\})^n$, then its *amoeba* is the image under Log of its set of zeros. The term was introduced by Gelfand et al. (1994). For more recent developments see, e.g., Viro (2011).

One can of course study the image in $\mathbf{R}^n$ of any set, but zero sets of certain functions have interesting properties. An amoeba is typically a closed semianalytic subset of $\mathbf{R}^n$ with tentacles which go out to infinity and separate the components of the complement of the amoeba. The number of such components is at most equal to the number of integer points in the Newton polytope for f if f is a Laurent polynomial; in certain cases equal to the latter number [2000a].

An easy example, which Mikael himself used in his lectures, is the zero set of the polynomial $P(z, w) = 1 + z + w$ of degree one. A zero $(z, w) \in \mathbf{C}^2$ must satisfy $1 \le |z| + |w|$; $|z| \le |w| + 1$; and $|w| \le 1 + |z|$. It is easy to see that any point $(p, q) \in \mathbf{R}^2$ which satisfies the three inequalities $1 \le p + q$; $p \le q + 1$; and $q \le 1 + p$ is equal to $(|z|, |w|)$ for some zero (z, w) of P. (A useful observation here is the fact that the corresponding strict inequalities are the exact conditions under which there exists a triangle with side lengths 1, p and q.) The amoeba of P is then given by the three inequalities $1 \le e^x + e^y$; $e^x \le e^y + 1$; and $e^y \le 1 + e^x$.

Of course one can study the zero set directly without taking the logarithm. That it nevertheless has interesting consequences to take the logarithm was shown by Mikael in [2008a]: it is about area preserving.

A *coamoeba* is defined analogously but with the mapping Log replaced by the mapping $\operatorname{Arg}(z) = (\arg z_1, \arg z_2, \ldots, \arg z_n)$. Mikael wanted to establish formally the duality between amoebas and coamoebas, and he started to write a paper with Mounir Nisse, which Mounir has now finished (this volume, pp. 63–80). For other relevant papers, see Nisse (2009) and Nisse & Sottile (2013a; 2013b).

Jens Forsgård and Petter Johansson have continued the work on coamoebas and published two papers (2014; 2015) on the subject.

A straight line in the plane can be described by an equation

$$ax + by + c = 0,$$

and hence as the fold line of the convex function

$$f(x, y) = (ax + by + c) \vee 0, \qquad (x, y) \in \mathbf{R}^2,$$

where the maximum operation is denoted by $\vee$: $s \vee t = \max(s, t)$, $s, t \in \mathbf{R}$. To *tropicalize* means to replace addition by the maximum operation and multiplication by addition. By this procedure we obtain

$$g(x, y) = (a + x) \vee (b + y) \vee c, \qquad (x, y) \in \mathbf{R}^2.$$

A tropical straight line can therefore be defined as the union of the fold lines for the function g, which consists of three rays. They emanate from the point $(c - a, c - b)$ in the directions $(1, 1)$, $(-1, 0)$ and $(0, -1)$. For example, the amoeba of the polynomial $P(z, w) = 1 + z + w$ mentioned earlier contains the tropical line emanating from $(0, 0)$, which is its spine.

If $p = (p_1, p_2)$ and $q = (q_1, q_2)$ are two points in the plane with

$$q_1 \neq p_1,\ q_2 \neq p_2 \text{ and } q_2 - q_1 \neq p_2 - p_1,$$

then it is easy to see that there is one and only one tropical straight line through p and q. If one of the conditions is not satisfied, there exist infinitely many tropical straight lines through p and q, but Mikael explained that one should accept only those lines that are stable under small perturbations; then you get a single line. In the same way, two distinct tropical straight lines meet in a single point if we only accept intersections that are stable under small perturbations.

Just like in spherical geometry there do not exist any distinct parallel lines. We can go on and ask about all the axioms of Euclidean geometry.

The similarity between the procedure of tropicalization and taking the logarithm is based on the formulas

$$\log(x \times y) = \log x + \log y, \qquad x, y > 0, \text{ and}$$

$$\log x \vee \log y \leq \log(x + y) \leq \log 2 + (\log x \vee \log y), \qquad x, y > 0,$$

where the error, no larger than $\log 2$, is relatively small if x or y is large.

In the little paper [2008a], which is indeed a gem, Mikael shows how the concept of an amoeba can be used to show the well-known formula $\zeta(2) = \sum_1^\infty 1/n^2 = \pi^2/6 \approx 1.644934$ (the so-called Basel problem).

The Pluricomplex Seminar

I started a seminar series in Uppsala in the 1970s. In the beginning it was more like a study group, and had no name, since I thought that a name could be hampering. But later I discovered that almost everything was about several complex variables, and during a visit to Strasbourg I saw that Jean-Pierre Ramis had used the name *Séminaire pluricomplexe*. That sounded mysterious enough, and I borrowed it to Uppsala. During the Fall Semester of 1980 the title was *Pluricomplex Analysis and Geometry*; in the Spring Semester of 1981 it was *Pluricomplex Analysis*, and from the Spring Semester of 1982 on, the name was *The Pluricomplex Seminar*.

Mikael's gave his first lecture in the seminar during the Fall Semester of 1978. He reported on chosen sections of the little book by Lev Isaakovič Ronkin (1931–1998) entitled *The Elements of the Theory of Analytic Functions of Several Variables* (1977), which had been published in Russian in 2,700 copies in Kiev the year before and cost 93 kopecks. The task was a part of the examination for the course *Mathematics D*.

Lectures held by Mikael Passare. Except in four cases, the lectures listed here were given by Mikael at the Pluricomplex Seminar.

1978-11-13. *Analytisk fortsättning* [*Analytic continuation*]. (Report on a special project for the advanced course *Mathematics D*.)

1982-11-01. *Henkin–Ramirez formulas for weight factors* (*according to Bo Berndtsson and Mats Andersson*).

1983-01-24. *Godtyckliga områden som projektioner av pseudokonvexa områden* [*Arbitrary domains as projections of pseudoconvex domains*].

1983-04-18. *Integraloperatorer för att lösa Cauchy–Riemanns ekvationer* (*efter R. Michael Range*) [*Integral operators for solving the Cauchy–Riemann equations* (*after R. Michael Range*)].

1983-06-15. *Samband mellan mängder av Newtonkapacitet noll och pluripolära mängder* (*efter Azim Sadullev*) [*Links between sets of Newton capacity zero and pluripolar sets* (*after Azim Sadullaev*)].

1984-05-14. *Ideal i ringen av holomorfa funktioner definierade medelst strömmar, I* [*Ideals in the ring of holomorphic functions defined by means of currents, I*].

1984-05-21. *Ideal i ringen av holomorfa funktioner definierade medelst strömmar, II* [*Ideals in the ring of holomorphic functions defined by means of currents, II*].

1984-12-10. *Residuer, strömmar och deras relation till ideal av holomorfa funktioner* [*Residues, currents, and their relation to ideals of holomorphic functions*] (cf. [1984], the thesis which was to be defended five days later).

1985-03-20. *Produkter av residuströmmar* [*Products of residue currents*] (cf. [1985]).

1986-01-17. *A new proof for integral representation formulas without boundary term.*

1986-04-14. *Principal values of meromorphic functions.*

1986-09-18. 1. *Shortcut to weighted representation formulas for holomorphic functions* (cf. [1988a]). 2. *Impressions from the IntFernational Congress of Mathematicians, Berkeley.*

1988-11-07. *Kergin interpolation of entire functions* (cf. [1991a; 1991b]).

1989-02-22. *Continuity of residue integrals in codimension two.*

1989-05-24. *Integralformler och residuer på komplexa mångfalder* [*Integral formulas and residues on complex manifolds*].

1989-10-04. *Kergin interpolation on* **C**-*convex sets.*

1991-02-18. *Mathematical impressions from Krasnoyarsk:* 1. *Holomorphic extension from a part of the boundary.* 2. *Toric varieties.*

1993-05-05. *Projektiv konvexitet* [*Projective convexity*].

1993-09-24. A lecture at a meeting of the Swedish Math. Society at the Royal Institute of Technology.

1994-05-19. *Holomorphic differential forms on analytic sets.*

1998-09-07. *Amoebas and Laurent determinants* (cf. [2000a; 2004a]).

2000-03-14. *Constant terms in powers of a Laurent polynomial.*

2001-05-04. *Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope.* Lecture at the Nordan Meeting in Oslo.

2001-10-16. *Complex convexity – recent results of Kiselman and Hörmander.*

2002-05-07. *Discriminant amoebas.*

2002-11-19. *Algebraic equations and hypergeometric functions.*

2003-03-04. *The Lee–Yang circle theorem and geometry of amoebas.*

2003-10-21. *Amöbor, polytoper och tropisk geometri* [*Amoebas, polytopes, and tropical geometry*].

2004-01-10. *Koamöbor och Mellin-transformer av rationella funktioner* [*Coamoebas and Mellin transforms of rational functions*].

2006-05-19. A lecture at the Nordan Meeting in Sundsvall.

2006-07. A minicourse on amoebas given at Institut de Mathématiques de Jussieu, Paris.

2010-03-09. (*Co*)*amoebas of linear spaces.*

2010-10-19. *Mellin transforms and hypergeometric functions.*

Originally, the seminars took place at Uppsala with a lecture in general every week. From the Spring Semester of 1999 on, when Mikael had become well established as a professor at Stockholm, they became a joint activity for Uppsala University, Stockholm University, and the Royal Institute of Technology (KTH), with an alternating venue. To minimize travel we had two lectures every second week. From 2007, when I had switched to digital geometry, mathematical morphology, and discrete optimization, and Burglind Juhl-Jöricke had left Uppsala University, it became an activity exclusively in Stockholm.

The *Nordan* Meetings

Together with Mats Andersson and Peter Ebenfelt, Mikael Passare initiated a series of encounters on complex analysis in the five Nordic countries. Mikael and Peter organized the first conference, which took place in Trosa, Sweden, March 14–16, 1997; Mats the second, in Marstrand, Sweden, April 24–26, 1998. Following a voting procedure at the end of the first meeting, these yearly meetings were named *Nordan*[7] – a clear reference to *Les Journées complexes du Sud*, which during a long time have taken place in the south of France.

Mikael edited abstracts in Swedish of the lectures – which had all been given in English. These brochures were published with a delay of a few years. Twelve of them have come out; he was preparing the thirteenth, which was to report on *Nordan* 13 held in Borgarfjordur in 2009, and asked Ragnar Sigurðsson on 2011 September 10, to write a preface in Icelandic (Ragnar Sigurðsson, personal communication 2011-10-04).

Lars Filipsson emphasizes (personal communication 2011-10-06) that Mikael wrote these brochures in Swedish to develop Swedish terms in higher mathematics, especially in complex analysis – otherwise the Swedish mathematical terms reach up to the first, possibly the second, university year only.

[7] This is the name in Swedish of a chilly wind from the north, but also reminds us of the original purpose: to promote Nordic Analysis.

Nordic meetings like these were something that Mikael and Mats had discussed and planned during many years; both of them wanted to create a forum with a more relaxed atmosphere, where Nordic complex analysts, in particular the young ones, could feel more at home than at big international conferences, and which would give those that worked in the Nordic countries occasion to get to know each other better.

And the initiative turned out to be a long-lasting success: the fifteenth encounter took place in Röstånga in southern Sweden, 2011 May 06–08; the sixteenth in Kiruna in northern Sweden, 2012 May 11–13; the seventeenth in Svolvær, Norway, 2013 May 24–26; the eighteenth at CIRM, Luminy, France, 2014 March 24–29, as a joint session of Nordan and the Komplex Analysis Winter school And workshop (KAWA); and the nineteenth in Reykjavík, Iceland, 2015 April 25–26. The twentieth Nordan took place in Stockholm, 2016 March 16–20 as a session of the 27th Nordic Congress of Mathematicians.

Africa

Mikael Passare was a Member of the Board of the International Science Programme (ISP), Uppsala, and a Member of the Board of the Pan-African Centre for Mathematics (PACM) in Dar es-Salaam, Tanzania. He was a driving force in the creation of this Pan-African Centre, which is a collaborative project between Stockholm University and the University of Dar es-Salaam.

Mohamed E. A. El Tom, Chairman of the Board of PACM and a member of the Reference Group for Mathematics of ISP, says that he is confident that had it not been for Mikael, PACM would have remained a mere idea in the head of its initiator, i.e., in Mohamed's head; see El Tom (2011). Mikael started working with great conviction and enthusiasm on the idea when Mohamed first suggested it to him while they were walking on a Meroetic archeological site near Khartoum in April 2004. (Mohamed El Tom, personal communication 2011-10-17.)

Mikael took an early, informal contact with the Vice-Chancellor (*Rektor*) of Stockholm University, who expressed his approval in principle (Mohamed El Tom, personal communication 2011-10-20).

In October 2008, Mikael and Mohamed discussed the idea with Anders Karlhede, Dean of the Division, and asked whether Stockholm University could be a partner in the project. Anders immediately took the question to Stefan Nordlund, Dean of the Faculty. The latter proved to be very positive, which was decisive for the coming commitment of Stockholm University to PACM. (Anders Karlhede, personal communication 2011-10-19.)

Mikael then presented the idea to the Department of Mathematics at Stockholm University. While the department did not object to the idea, it was only natural that some members raised many significant issues that required clarification. Mikael maintained correspondence with Mohamed about these and related isues for more than two years, at the end of which he managed to secure the approval of the department to collaborate in establishing the Centre at some suitable university in Africa. Later he was an influential member of the committee that short-listed African universities for hosting PACM. Subsequently, he was a member of a delegation led by Stefan Nordlund which visited some of the short-listed universities and made appropriate recommendations to the Vice-Chancellor of Stockholm University.

Mikael never ceased devoting of his precious time to the Centre. His last assignment was to chair and constitute a search committee for the Director of the Centre, a process he initiated before he was asked by the Board of the Centre to undertake it. Such was Mikael, ahead of others in thinking, and working to realize important objectives without being asked to do so. When Mohamed conveyed to him the Board's decision regarding the search committee, he responded promptly, accepting the charge, and promised to respond with detailed ideas upon his return from the trip to Dubai, Oman and Iran that he was planning to undertake.

Mikael's commitment and enthusiasm for the Centre was unsurpassed. He was confident that the grand objective of establishing a world-class Centre of Mathematics in Africa is attainable. (Mohamed El Tom, personal communication 2011-10-17; this note applies to the last three paragraphs.)

Sonja Kovalevsky

The chair which Mikael Passare held was the one which was created for Sonja Kovalevsky (1850 January 03/15–1891 February 10). An earlier incumbent during seven years, 1957–1964, was Lars Hörmander (1931–2012), Mikael's mathematical grandfather. Mikael was proud of having been appointed to Sonja's chair. He is buried not far from her grave.

Exactly 150 years after Sonja's birth, on 2000 January 15, Mikael organized a symposium to her memory. It was held in the *Aula Magna* of Stockholm University. Among the invited speakers were Agneta Pleijel, Roger Cooke and Ragni Piene.

Languages

In the section on the *Nordan* meetings, I have already mentioned that Mikael was interested in developing Swedish mathematical terms. He knew many languages. His Russian was "really perfect!" according to Timur Sadykov (personal communication 2011-10-13); "he spoke Russian perfectly, so it was totally impossible to recognize his Swedish origin" (Andrei Khrennikov, personal communication 2011-11-26). He took a course in French corresponding to 30 ECTS credit points at Stockholm University before going to Paris in 1986-87 (diploma dated 1985-09-03). He learned some Fijian when he visited the Republic of Fiji (Timur Sadykov, personal communication 2011-10-16).

His knowledge of German was very good although he had not studied that language in high school. He studied also Finnish and spoke the language so well that he was interviewed in the Finnish-language *Sisuradio* in Sweden.

Spanish and Italian he knew enough to get along. He was in Italy and Spain with Anders Wändahl, and never talked English when visiting a restaurant or when asking for directions in the street. He could also speak some Polish and Bulgarian.

Finally, he studied Arabic and could at least read that language. Maybe Arabic would have been his next project. (Anders Wändahl, personal communication 2011-10-19; this remark applies to this paragraph and the preceding one.)

An extraordinary curiosity

Andrei Khrennikov writes:

> I would like to mention Mikael's extraordinary curiosity, which was extended to a large variety of fields. In particular, he discussed with excitement the possibility of mathematical modeling of cognition, human psychological behavior, and consciousness. I met Mikael and Galina the last time in July 2010, in Stockholm, and during one evening we discussed a large variety of topics: complex and p-adic analysis, mathematical foundations of quantum physics, quantum nonlocality, Bell's inequality and experiments [...] (Andrei Khrennikov, personal communication 2011-11-26)

Music

Mikael loved classical music; in his teens he sold his bicycle in order to buy a piano. He played clarinet and flute. He composed a piece for clarinet, which was played in a theater in Stockholm. His last love was an instrument called theremin.[8] He dreamed about being able to play it.[9]

He also loved to sing and was a member in a choir and learned to sing solo both in Stanford and in Moscow. (Galina Passare, personal communication 2011-10-17; this applies to all of this section.)

A "Swedish Classic"

Mikael swam several times a week, at least 2 km. He loved the mountains and skied long distances (sometimes 90–130 km) spending the night in cottages. He swam between islands in Lake Mälaren close to Stockholm. (Galina Passare, personal communication 2011-10-17.)

[8] Терменвокс, which was invented by Лев Сергеевич Термен, Léon Theremin (1896–1993).

[9] At his funeral on 2011 October 28, *Dance in the Moon* was played on CD; the performer was Lydia Kavina, a leading thereminist.

He ran the Stockholm Marathon. To Yûsaku Hamada's guest lectures in Uppsala on 2002 September 10, he went by bike from Stockholm (Yûsaku Hamada, personal communication 2011-09-19). Another time he skated on Lake Mälaren to the seminar in Uppsala. (After the seminar, however, he went back to Stockholm by train.)

The Viking Run (in Swedish: *Vikingarännet*) is the world's biggest regular skating event on natural ice, and is arranged yearly since 1999. It usually starts at Skarholmen in Uppsala and finishes at some place in or near Stockholm (depending on ice conditions; sometimes the ice is so bad that the competition has to be canceled). Mikael participated in the Viking Run several times.

Mikael also performed what is known as a "Swedish Classic" in 1989. It consists of four parts, which have to be done within a twelve-month period: (1) A ski run, either the Engelbrekt Run, 60 km, or the Vasa Run / Open Track, 90 km; (2) Going around Lake Vättern on bicycle, 300 km; (3) The Vansbro Swim, 3 km; and (4) The Lidingö Run, 30 km. Mats Andersson (personal communication 2011-10-12) remembers that he claimed the cycling to be the most painful of the four, noting the chafing after so many hours on the saddle.

He loved bandy (a sport similar to ice hockey but played with a ball and on an ice field the size of a soccer field) and missed only one single Swedish Championship Final since 1980, viz. the one in 1983 (Anders Wändahl, personal communication 2011-10-12). He was considerate also of other bandy fans. When Magnus Carlehed, his mathematical nephew, was traveling in 1990 around the globe in areas where ice rinks are not so common, he sent to Magnus a big envelope to the address *Poste restante*, Denpasar (Bali), Indonesia. It contained a video recording showing the Swedish Bandy Championship Final. (Magnus Carlehed, personal communication 2011-12-09.) He went to Arkhangelsk to see the Bandy World Championship there (probably in 1999, when Russia won over Sweden in the final). He also traveled to Oulu with Björn Ivarsson, his mathematical younger brother, to see the Championship in 2001 (when Russia won over Sweden again in the final; Björn Ivarsson, personal communication 2011-12-14).

A passionate traveler

Mikael saw at least three total solar eclipses: the one which took place on 1999 August 11 he saw in Turkey (although it would perhaps have been easier to see it in Bulgaria); the eclipse of 2002 December 04 he saw in Mozambique; and on 2006 March 29 he was in Niger with Anders Wändahl (although the southern coast of Turkey would have been easier to reach from Sweden and had a greater chance of a clear sky without sand storms). After that they continued to Chad.

Mikael was a passionate traveler. He visited 152 countries. When he and I, together with several other Swedish mathematicians, were invited in September 2006 to celebrate the twentieth anniversary of the *Groupe Inter-Africain de Recherche en Analyse, Géométrie et Applications* (GIRAGA) and after that to participate

2010, India (photo Galina Passare)

in the *First African-Swedish Conference on Mathematics*, both in Yaoundé, Cameroon, he first visited the Central African Republic and continued afterwards to Equatorial Guinea and Gabon (Anders Wändahl, personal communication 2011-11-14); thus he got four new countries on his list – assuming that he had not been in any of these before – while I got only one.

The United Arab Emirates and Oman turned out to be the last ones. Land number 153 should have been Iran: he planned to arrive at Tehran Imam Khomeini International Airport on September 17 at 21:25 (Mikael Passare, electronic letter 2011-09-15 to mathematicians in Tehran). Siamak Yassemi, Head of the School of Mathematics, University of Tehran, was ready to meet him there.

Finally

Mikael's significance goes much beyond his own research. Many persons have testified to his positive view of life, his humor, and to his genuine interest in people he met. He was an unusually stimulating partner in discussions; listening, inspiring, and supportive, in professional situations as well as private ones.

For Mikael's friends and colleagues around the world his unexpected departure is a severe loss.

For me personally, Mikael's disappearance seems unreal. He was always there for me. I shall remember him with joy and gratitude as long as I live.

Two proposals

At a meeting at Stockholm University to Mikael's memory on 2011 September 28, arranged by Tom Britton, I ended my speech by presenting two proposals.

The first proposal was that Stockholm University organize a conference to his memory, where his many mathematical achievements could be presented and discussed. It has now been realized (at least in part) as a yearly event, *Mikael Passare's Day*, organized at Stockholm University in September or October each of the six years 2011 through 2016.

Since, as far as I know, Mikael has not published all his ideas on tropical geometry, I proposed, secondly, that his former students write a survey article about these ideas (and of course other mathematical ideas). Alicia Dickenstein (2011-09-24), August Tsikh (2011-10-03), Alexey Shchuplev (2011-10-06), and Hans Rullgård (2011-10-11) have all spontaneously approved of this and want to contribute to this project.

The second proposal can of course be realized as a part of the first, viz. if the survey article is published in the conference proceedings.

See also my paper "Questions inspired by Mikael Passare's mathematics" (2012, 2014).

References

Andersson, Mats; Wulcan, Elizabeth. 2007. Residue currents with prescribed annihilator ideals. *Ann. Sci. École Norm. Sup.* (4) **40**, no. 6, 985–1007.

Coleff, Nicolás R.; Herrera, Miguel E. 1978. *Les courants résiduels associés à une forme méromorphe.* Lecture Notes in Mathematics 633. X + 211 pp. Berlin: Springer.

Dickenstein, A[licia]; Sessa, C[armen]. 1985. Canonical representatives in moderate cohomology. *Invent. Math.* **80**, no. 3, 417–434.

El Tom, Mohamed E. A. 2011. Towards a new collaborative model for strengthening capacity in mathematical research in Africa. **In:** Kiselman, Christer (Ed., 2011), *Regional and Interregional Cooperation to Strengthen Basic Sciences in Developing Countries. Addis Ababa,* 1–4 *September* 2009, pp. 173–187. Acta Universitatis Upsaliensis. Skrifter rörande Uppsala universitet. C. Organisation och historia [Publications concerning Uppsala University. C. Organization and History] 88. Uppsala: Uppsala University. 423 pp. ISBN: 978-91-554-7910-7.

Forsberg, Mikael. 1998. *Amoebas and Laurent Series.* PhD Thesis, defended on 1998 May 18, The Royal Institute of Technology.

Forsgård, Jens. 2015. *Tropical aspects of real polynomials and hypergeometric functions.* Doctoral Dissertation, defended on 2015 June 04 at Stockholm University.

Forsgård, Jens; Johansson, Petter. 2014. Coamoebas and line arrangements in dimension two. *Math. Z.* **278**, no. 1–2, 25–38.

Forsgård, Jens; Johansson, Petter. 2015. On the order map for hypersurface coamoebas. *Ark. mat.* **53**, No. 1, 79–104.

Gelfand, I. M.; Kapranov, M. M.; Zelevinsky, A. V. 1994. *Discriminants, Resultants, and Multidimensional Determinants.* Math. Theory Appl., x + 523 pp. Boston: Birkhäuser.

Henkin, G. M. 1969. Integral representation of functions which are holomorphic in strictly pseudoconvex regions, and some applications (Russian). *Mat. Sb.* (*N.S.*) **78** (120), 611–632.

Henkin Gennadi M.; Polyakov Peter L. 2012. Residual $\overline{\partial}$-cohomology and the complex Radon transform on subvarieties of $\mathbf{C}P^n$. *Math. Ann.* **354**, no. 2, 497–527.

Herrera, M.; Lieberman, D. 1971. Residues and principal values on complex spaces. *Math. Ann.* **194**, 259–294.

Kiselman, Christer O. 2011. Mikael Passare 1959–2011. *Medlemsutskicket från Svenska Matematikersamfundet*, October 2011, pp. 17–35. (Obituary in Swedish, published by the Swedish Mathematical Society on 2011-10-25.)

Kiselman, Christer O. 2012, 2014. Questions inspired by Mikael Passare's mathematics. *Afr. Mat.*, published online 2012-10-27. Print version in *Afr. Mat.* **25**, issue 2 (2014), 271–288.

Kiselman, Christer O. 2016. Weak lineal convexity. **In:** Białas-Cież, Leokadia; Kosek, Marta (Eds.), *Constructive Approximation of Functions*. Banach Center Publications, Polish Academy of Sciences, volume **107**, pp. 159–174.

Nisse, Mounir. 2009. Geometric and combinatorial structure of hypersurface coamoebas. arXiv:0906.2729.

Nisse, Mounir; Sottile, Frank. 2013a. Non-Archimedean coamoebae. **In:** Amini, O.; Baker, M.; Faber, X. (Eds.), *Tropical and non-Archimedean geometry*, pp. 73–91. Contemporary Mathematics, **605**. Providence, RI: American Matematical Society.

Nisse, Mounir; Sottile, Frank. 2013b. The phase limit of a variety. *Algebra Number Theory* **7** (2) 339–352.

Persson, Ulf. 2011. Mikael Passare 1959–2011 [in Swedish with a summary in English]. *Normat* **59**, no. 2, 49–50, 96.

Ронкин, Лев Исаакович. 1977. *Элементы теории аналитических функций многих переменных.* Kiev: Издательство ≪Наукова Думка≫. 168 pp.

Rudbeckianska skolan. 1978. *Avgångsbetyg för Mikael Pettersson* 1978-06-09. *Elevkort. Terminsbetyg* [Graduation Certificate when leaving high school on 1978 June 09].

Schwartz, Laurent. 1954. Sur l'impossibilité de la multiplication des distributions. *C. R. Acad. Sci. Paris* **239**, 847–848.

Simon, Imre. 1988. Recognizable sets with multiplicities in the tropical semiring. **In:** *Mathematical foundations of computer science*, (Carlsbad, 1988), pp. 107–120. Lecture Notes in Computer Science, **324**, Berlin: Springer.

The Swedish National Board of Forensic Medicine. 2011. Statement A11-1041 of 2011 December 13 (on the cause of Mikael Passare's death).

Viro, Oleg Ya. 2011. On basic concepts of tropical geometry. (In Russian.) *Tr. Mat. Inst. Steklova* **273**, 271–303. Sovremennye problemy matematiki. Translated in *Proc. Steklov Inst. Math.* **273**, No. 1, 252–282.

Sources

Many persons have contributed important information on Mikael's life. I would like to thank especially Galina Passare, Mats Andersson, Jan Boman, Gerd Brandell, Mohamed El Tom, Lars Filipsson, Gennadi Henkin, Anders Karlhede, Mounir Nisse, Karin Wallby, and Anders Wändahl.

The sources are in certain cases messages that I have received during the writing process, and, if so, reported as a personal communication at the end of a sentence or a paragraph. Otherwise I have relied on documents that I have saved, notes that I have made – and my memory.

Christer O. Kiselman
Uppsala University
Department of Information Technology
P. O. Box 337
SE-751 05 Uppsala, Sweden
e-mail: kiselman@it.uu.se, christer@kiselman.eu

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 59–60

Mikael Passare

Magnus Carlehed

I got to know Mikael in the eighties, when I was a PhD student at Stockholm University. I had completed a number of graduate courses and became interested in complex analysis. Mikael was a young lecturer in Stockholm, and our overlapping interest in that subject brought us into each other's orbits. As far as I know, I was his very first student. During the academic year 1987–88, Several Complex Variables (SCV) was the theme of the Mittag-Leffler Institute, and once a week I took the commuter train to Djursholm for the seminars. Many prominent researchers in SCV visited the institute, and it struck me how Mikael seemed to have a personal relationship with all of them. He was a highly social person and introduced me to all of his maths friends. At the time, integral formulas and residues were among the hot trends in SCV. Mikael had done some important research in this field, and he explained to me how the previous algebraic approach to the subject was making place for a more analytical and computational one.

For personal reasons I quit my PhD studies in 1989, and took them up again only years later at another university. Hence, from 1989 onwards, my relation with Mikael became a purely private one. He was a good friend, who always had time for a chat, and although our contact became sparser in the later years, we kept in touch until his untimely death in 2011. Mikael valued the simple life; he had no admiration for the consumption society. For many years he made a point of using an extremely old Volvo as the family car. But he compensated for this with a great interest in travelling; at the time of his death, he had succeeded to visit 152 countries. To Mikael, mathematics was always an activity in a social context, and he attended as many conferences as he could. Nationalism and borders were alien concepts to Mikael, who treated all people equally and was a true cosmopolite. He worked internationally with mathematics in Africa, amongst other places.

Mikael had many interests besides mathematics: such as politics, music, and sport. We shared an interest in bandy, a team winter sport that is played outdoors (mainly in Nordic countries and Russia) by skaters on a field as large as a soccer field. Bandy is particularly popular in certain areas of Sweden, and we both grew up in those parts of the country. When it came to bandy, Mikael left behind his normally balanced manners, and he could almost be described as fanatical. To him bandy was more than a sport, it was history, it was culture. During the 1980s the dominating team in Sweden was Boltic; a team that Mikael didn't recognize as authentic, as it lacked tradition and many of its players were transferred from other teams. With something between joviality and seriousness, he expressed contempt for that team more than he supported his own. We used to go to games together and it created a special feeling of brotherhood to stand in the freezing cold for 90 minutes, cheering and shouting. When I made a world trip as a backpacker in 1990, he sent me a large envelope to a Poste Restante address in Bali. Upon opening it, I found a VHS cassette with the full game of the Swedish championship final. Mikael also went to almost all the World Cup finals, whether they were in Sweden or in Siberia.

Mikael's contribution to Swedish mathematics community is immense; first and foremost because of his research, but also because of his deep engagement and enthusiasm for everything he did. He had many successful students over the years. He truly believed in the value of mathematics for society, and it was very natural that he chaired the Swedish Mathematical Society and the National Committee for Mathematics for some time. In the latter role, he called me in 2011 to ask me if I was interested to become a member of the committee. It turned out to become our last contact. A very fine personality has left us.

Magnus Carlehed
e-mail: carlehed@outlook.com

Part II

Research Articles

This part contains original research articles written by Mikael's mathematical friends and/or inspired by Mikael's contribution to mathematics.

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 63–80

Amoebas and Coamoebas of Linear Spaces

Mounir Nisse and Mikael Passare

Abstract. We give a complete description of amoebas and coamoebas of k-dimensional very affine linear spaces in $(\mathbb{C}^*)^n$. This include an upper bound of their dimension, and we show that if a k-dimensional very affine linear space in $(\mathbb{C}^*)^n$ is generic, then the dimension of its (co)amoeba is equal to $\min\{2k, n\}$. Moreover, we prove that the volume of its coamoeba is equal to π^{2k}. In addition, if the space is generic and real, then the volume of its amoeba is equal to $\pi^{2k} / 2^k$.

Mathematics Subject Classification (2010). 14T05, 32A60.

Keywords. Very affine linear spaces, amoebas, coamoebas, logarithmic Gauss map.

1. Introduction

Amoebas and coamoebas are very fascinating notions in mathematics, the first has been introduced by I. M. Gelfand, M. M. Kapranov and A. V. Zelevinsky in 1994 [3], and the second by the second author in a talk in 2004. They are natural projections of complex varieties, and which turn out to have relations to several other fields: tropical geometry, real algebraic geometry, generalized hypergeometric functions, mirror symmetry, and others (e.g., [6], [7], [13], [12], [17], [19]). More precisely, the amoebas (respectively coamoebas) of complex algebraic and generally analytic varieties in the complex algebraic torus $(\mathbb{C}^*)^n$ are defined as their image under the logarithmic mapping $\mathrm{Log} : (z_1, \ldots, z_n) \mapsto (\log|z_1|, \ldots, \log|z_n|)$ (respectively the argument mapping $\mathrm{Arg} : (z_1, \ldots, z_n) \mapsto (\frac{z_1}{|z_1|}, \ldots, \frac{z_1}{|z_1|})$). Amoebas (respectively coamoebas) are the link between classical complex algebraic geometry and tropical (respectively complex tropical) geometry. More precisely, amoebas degenerate to piecewise-linear objects called tropical varieties (see [13], and [19]), and comoebas degenerate to a non-Archimedean coamoebas which are the coamoebas of some lifting in the complex algebraic torus of tropical varieties. See [18] for

Research of the first author is partially supported by NSF MCS grant DMS-0915245.

more details about non-Archimedean coamoebas, and [16] about this degeneration in case of hypersurfaces. Whereas the theory of (co)amoebas of complex hypersurfaces is by now reasonably well understood (see, e.g., [2], [11], [16], and [19]), much less is known about the structure of (co)amoebas coming from varieties of higher codimension. A natural first step in this direction is to explore the case of linear spaces.

Being of a logarithmic nature, it is natural that coamoebas are closely related to the exponents of the defining functions of V, and to the associated Newton polytopes. This connection is extensively explored in the thesis of the first author [3], [15], and [19]. Another important connection is to the currently very active field of tropical geometry, a piecewise linear incarnation of classical algebraic geometry where the varieties can be seen as non-Archimedean versions of amoebas (see [7], [12], [13] and others).

A fundamental theorem was shown by K. Purbhoo [20] for the general study of amoebas that do not come from hypersurfaces. The theorem states that the amoeba of an algebraic variety V is equal to the intersection of all hypersurface amoebas corresponding to functions in the defining ideal $\mathcal{I}(V)$ of the variety V. We give a simple proof of this theorem with an extension to coamoebas.

Theorem 1.1. *Let $V \subset (\mathbb{C}^*)^n$ be an algebraic variety with defining ideal $\mathcal{I}(V)$. Then the amoeba (respectively coamoeba) of V is given as follows:*

$$\mathscr{A}(V) = \bigcap_{f \in \mathcal{I}(V)} \mathscr{A}(V_f) \qquad \text{and} \qquad co\mathscr{A}(V) = \bigcap_{f \in \mathcal{I}(V)} co\mathscr{A}(V_f).$$

In [19], Rullgård and the second author showed that the area of complex plane curve amoebas is finite and the bound is given in terms of the Newton polygon of the defining polynomial. They, also compute the area of the amoeba of a plane line. It was shown by Mikhalkin and Rullgård that this bound is always sharp [14]. In [8], Madani and the first author generalized this result and showed that the volume of the amoeba of a k-dimensional algebraic variety of $(\mathbb{C}^*)^n$ with $n \geq 2k$ is finite. Moreover, they proved in [9] that the finiteness of the volume of the amoeba of a generic analytic variety is equivalent to the variety being algebraic. Theorem 1.1 and Proposition 3.1 was shown separately and in the same time by Petter Johansson in [4].

Let V be a variety in the projective space $\mathbb{CP}^n$. We choose homogeneous coordinates $[Z_0 : \cdots : Z_n]$ so that V is transverse to coordinate hyperplanes $Z_j = 0$ and all their intersections. The complement of the arrangement of coordinate hyperplanes in $\mathbb{CP}^n$ is $(\mathbb{C}^*)^n$. Then the variety $\mathscr{V} = V \cap (\mathbb{C}^*)^n$ is called a *very affine variety*, and in the case where $P(k)$ is a k-dimensional linear subspace of $\mathbb{CP}^n$ we say that $\mathscr{P}(k) = P(k) \cap (\mathbb{C}^*)^n$ is a *very affine linear space*, and by abuse of language we will call it just affine linear space. Moreover, $\mathscr{P}(k)$ can be presented as a complete intersection of hyperplanes given by first degree equations $f_1(z) = \cdots = f_{n-k}(z) = 0$, where $z = (z_1, \ldots, z_n) = (Z_1/Z_0, \ldots, Z_n/Z_0)$ stands for the affine coordinates in $(\mathbb{C}^*)^n$.

Theorem 1.2. *Let $\mathscr{P}(k)$ be a generic affine linear subspace of $(\mathbb{C}^*)^{2k}$. Then we have the following:*

(i) *The volume of the coamoeba $co\mathscr{A}(\mathscr{P}(k))$ is equal to π^{2k};*
(ii) *Moreover, if $\mathscr{P}(k)$ is real, then the volume of its amoeba $\mathscr{A}(\mathscr{P}(k))$ is equal to $\frac{\pi^{2k}}{2^k}$.*

The present paper is organized as follows. We give definitions, background, and some known results in connection with this paper in Section 2. We prove Theorem 1.1 in Section 3, and detailed description of amoebas and coamoebas of lines in n-dimensional complex algebraic torus in Section 4.1 for any $n \geq 2$. We prove Theorem 1.2 in Section 5.

Remark. My first meeting and mathematical discussion with Michael was during the summer school in Paris in 2006 where he gave a series of lectures on amoebas. We talked a lot on the geometric and topological properties of these objects in particular the solidness of some of them. Moreover, at Stockholm University, when I visited him in the same year, we discussed their similarity to other objects called coamoebas. At that time we did not know exactly what kind of similarities because the ambient spaces of these two objects are different: one is compact and the other is not compact. Amoebas are closed subsets in the Euclidean space but coamoebas are not closed and not open subsets of the real torus. However, both of them have a similar (dual in some sense) combinatorial properties, and strongly related to the combinatorial type of the Newton polytopes of the defining polynomial in the hypersurface case. At that time we did not know a lot of things in higher codimension. This work was started on June 2011, but after the tragic death of Mikael Passare on 15 September 2011, the completion and writing of this paper was done by the first author.

2. Preliminaries

In this section, we review some known results related to this paper, and give some notations and definitions. Let V be an algebraic variety in $(\mathbb{C}^*)^n$. The *amoeba* $\mathscr{A}$ of V is by definition the image of V under the logarithmic map defined as follows (see M. Gelfand, M. M. Kapranov and A. V. Zelevinsky [3]):

$$\begin{array}{rccl} \mathrm{Log} : & (\mathbb{C}^*)^n & \longrightarrow & \mathbb{R}^n \\ & (z_1, \ldots, z_n) & \longmapsto & (\log|z_1|, \ldots, \log|z_n|). \end{array}$$

The argument map is the map defined as follows:

$$\begin{array}{rccl} \mathrm{Arg} : & (\mathbb{C}^*)^n & \longrightarrow & (S^1)^n \\ & (z_1, \ldots, z_n) & \longmapsto & (\frac{z_1}{|z_1|}, \ldots, \frac{z_1}{|z_1|}). \end{array}$$

The *coamoeba* of V, denoted by $co\mathscr{A}$, is its image under the argument map (defined by the second author in 2004).

Purbhoo shows that the amoeba of an algebraic variety V is equal to the intersection of all hypersurface amoebas corresponding to functions in the defining ideal $\mathcal{I}(V)$ of the variety V (see [20], Corollary 5.2). Passare and Rullgård prove the following (see [19]):

Theorem 2.1 (Passare–Rullgård, (2000)). *Let f be a Laurent polynomial in two variables. Then the area of the amoeba of an algebraic plane curve with defining polynomial f is not greater than π^2 times the area of the Newton polytope of f.*

In [14], Mikhalkin and Rullgård showed that up to multiplication by a constant in $(\mathbb{C}^*)^2$, the algebraic plane curves with Newton polygon Δ with maximal amoeba area are defined over $\mathbb{R}$. Furthermore, their real loci are isotopic to the so-called Harnack curves (possibly singular with ordinary real isolated double points). Moreover, Rullgård and the second author compute the area of the amoeba of a line in the plane.

Madani and the first author showed that if the dimension n of the ambient space is at least the double of the dimension of V (i.e., $n \geq 2\dim_{\mathbb{C}}(V) = 2k$), then the map $\mathrm{Log} \circ \mathrm{Arg}^{-1}$ conserves the $2k$-volume, i.e., the absolute value of the determinant of its Jacobian, when it exists, is equal to one (see [9], Proposition 3.1). Moreover, the same proposition shows that the set of critical points of the logarithmic and the argument maps restricted to V coincide. Hence, if the argument map restricted to the set of regular points is injective, and the cardinality d of the inverse image under the logarithmic map of a regular value in the amoeba is constant, then the volume of the amoeba will be the volume of the coamoeba divided by d. So, first we show that if V is a generic k-dimensional linear space in $(\mathbb{C}^*)^{2k}$, then the argument map restricted to the set of regular points is injective, and we compute the volume of its coamoeba. Moreover, if the linear space is real, we show that the cardinality of the inverse image under the logarithmic map of a regular value in the amoeba is constant and equal to 2^k. Finally, we compute the amoeba volume using the conservation of the volume by the map $\mathrm{Log} \circ \mathrm{Arg}^{-1}$.

In the following paragraph, we will recall the definition of the logarithmic Gauss map for hypersurface, and its generalization. We will present some known relations between this map and (co)amoebas. Let $V \subset (\mathbb{C}^*)^n$ be an algebraic hypersurface with defining polynomial f, and denote by V_{reg} the subset of its smooth points. The *logarithmic Gauss map* of the hypersurface V is the holomorphic map defined by (see Kapranov [5]):

$$\begin{array}{rccl} \gamma: & V_{\mathrm{reg}} & \longrightarrow & \mathbb{CP}^{n-1} \\ & z & \longmapsto & \gamma(z) = [v(z)], \end{array}$$

where $[v(z)] = [z_1 \frac{\partial f}{\partial z_1}(z) : \cdots : z_n \frac{\partial f}{\partial z_n}(z)]$ denotes the class of the vector $v(z)$ in $\mathbb{CP}^{n-1}$.

Madani and the first author generalize this map to any codimension, and extract some relations between the set of its critical points and (co)amoebas, and they generalized an earlier result of Mikhalkin [11] on critical points of the logarithmic map (see [10]). More precisely, let $V \subset (\mathbb{C}^*)^n$ be an algebraic variety of dimension k with defining ideal $\mathcal{I}(V)$ generated by $\{f_1, \ldots, f_l\}$. A holomorphic map γ_G from the set of smooth points of V to the complex Grassmannian $\mathbb{G}_{n-k,\,n}$ was defined as follows: If we denote by V_{reg} the subset of smooth points of V as before, and $M(l \times n)$ denotes the set of $l \times n$ matrices. Let g_G be the following map:

$$\begin{array}{cccc} g_G & : & V_{\text{reg}} & \longrightarrow & M(l \times n) \\ & & z = (z_1, \ldots, z_n) & \longmapsto & \begin{pmatrix} z_1 \frac{\partial f_1}{\partial z_1}(z) & \ldots & z_n \frac{\partial f_1}{\partial z_n}(z) \\ \vdots & \vdots & \vdots \\ z_1 \frac{\partial f_l}{\partial z_1}(z) & \ldots & z_n \frac{\partial f_l}{\partial z_n}(z) \end{pmatrix}. \end{array}$$

Since z is a smooth point of V, then the complex vector space L_z generated by the rows of the matrix $g_G(z)$ is of dimension $n-k$, and orthogonal to the tangent space to V at z. Indeed, the problem is local and V_{reg} is locally a complete intersection. Moreover, the tangent space to V at a regular point is contained in the tangent space of all the hypersurfaces defined by the polynomials f_i, and each row vector of index i is orthogonal to the hypersurface defined by the polynomial f_i which contains V. This means that the image of V_{reg} by g_G is contained in the subvariety of $M(l \times n)$ consisting of $l \times n$ matrices of rank $n - k$, which we map to the complex Grassmannian $\mathbb{G}_{n-k,\,n}$. Composing this identification with g_G we obtain the desired map:

$$\gamma_G : V_{\text{reg}} \to \mathbb{G}_{n-k,\,n}$$

called the *generalized logarithmic Gauss map.*

If $V \subset (\mathbb{C}^*)^n$ is a hypersurface, Mikhalkin showed that the set of critical points of $\text{Log}_{|V}$ coincides with $\gamma_G^{-1}(\mathbb{RP}^{n-1})$ (see Lemma 3 in [11], and Lemma 4.3 in [12]). This result was generalized by Madani and the first author for higher codimension in [10].

Throughout all this paper, the genericity of an algebraic variety $V \subset (\mathbb{C}^*)^n$ is defined as follows:

Definition 2.1. An irreducible algebraic variety $V \subset (\mathbb{C}^*)^n$ of dimension k is generic if it satisfies the following:

(1) The variety V contains an open dense subset U such that the Jacobian of the restriction to U of the logarithmic map $\text{Jac}(\text{Log}_{|U})$ has maximal rank, i.e., $\min\{2k, n\}$;
(2) The variety V lies in no affine subgroup, otherwise we may replace $(\mathbb{C}^*)^n$ by the smallest affine subgroup containing V.

We denote by $\mathscr{L}og_{|V}$ the complex logarithmic map, and Re the real part of a complex vector. In this case, we have $\text{Log}_{|V} = \text{Re} \circ \mathscr{L}og_{|V}$. This means that the amoeba of V is the real part of $\mathscr{L}og_{|V}(V)$ (by taking the imaginary part we obtain the same conclusion for the coamoeba),

$$\begin{tikzcd}
V \subset (\mathbb{C}^*)^n \arrow[rr, "\mathscr{L}og_{|V}"] \arrow[dr, "\mathrm{Log}_{|V}"'] & & \mathbb{C}^n \supset \mathscr{L}og(V) \arrow[dl, "\mathrm{Re}"] \\
& \mathscr{A}(V) \subset \mathbb{R}^n. &
\end{tikzcd}$$

We can check that for any $r \in \mathbb{R}^n$, the set $T_r := \mathrm{Log}^{-1}(r)$ is an n-dimensional real torus, and $r \in \mathscr{A}(V)$ if and only if $T_r \cap V \neq \emptyset$.

3. (Co)amoebas of complex algebraic varieties

In this section, we describe the amoeba (respectively coamoeba) of a complex variety V with defining ideal $\mathcal{I}(V)$ as the intersection of the amoebas (respectively coamoebas) of the complex hypersurfaces with defining polynomials in $\mathcal{I}(V)$.

The first part of Theorem 1.1 concerning amoebas was shown by Purbhoo in 2008 (see Corollary 5.2 in [20]). We present a very simple proof of this fact, and extend it to coamoebas.

Our first observation, is the following proposition about the dimension of (co)amoebas:

Proposition 3.1. *Let $V \subset (\mathbb{C}^*)^n$ be an irreducible algebraic variety of dimension k. Then, the dimension of the (co)amoeba $\mathscr{A}(V)$ of V satisfies the following:*

$$\dim((co)\mathscr{A}(V)) \leq \min\{2k, n\}.$$

In particular, if V is generic, then the dimension of its amoeba is $\min\{2k, n\}$.

Proof. The rank of the Jacobian of the logarithmic (respectively argument) map restricted to V at a regular point is equal to $\min\{2k, n\}$. So, the dimension of the (co)amoeba cannot exceed $\min\{2k, n\}$. Moreover, if the dimension of the amoeba (respectively coamoeba) of a k-dimensional irreducible variety V in $(\mathbb{C}^*)^n$ is strictly less than $\min\{2k, n\}$, then the map Re is not an immersion (respectively submersion) if $n \geq 2k$ (respectively $n < 2k$). Hence, the set of critical points of the logarithmic (respectively argument) map is equal to all the variety (see [10] for more details about critical values of the logarithmic Gauss map in higher codimension case). □

Let $V_f \subset (\mathbb{C}^*)^n$ be a hypersurface with defining polynomial f. Then, by definition, the amoeba of V_f is the image by the logarithmic map of the subset $\mathscr{S}_f$ of $(\mathbb{R}_+^*)^n$ defined as follows:

$$\mathscr{S}_f := \{(x_1, \ldots, x_n) \in (\mathbb{R}_+^*)^n \,|\, \exists\, z \in (\mathbb{C}^*)^n \text{ such that } x_i = |z_i|, \text{ and } f(z) = 0\}.$$

Since $\mathscr{L}og : (\mathbb{R}_+^*)^n \to \mathbb{R}^n$ is a diffeomorphism, we have the following:

$$\bigcap_{f\in\mathcal{I}(V)} \mathrm{Log}\,(\mathscr{S}_f) = \mathrm{Log}\left(\bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f\right),$$

where $\mathrm{Log}\,(\mathscr{S}_f)$ is used with abuse of notation.

Lemma 3.1. *We have the following equality:*

$$\bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f = \{(x_1,\dots,x_n) \in (\mathbb{R}_+^*)^n |\ x_i = |z_i|,\ \text{and}\ (z_1,\dots,z_n) \in V\}.$$

Proof. Let r be in

$$(\mathbb{R}_+^*)^n \setminus \{(x_1,\dots,x_n) \in (\mathbb{R}_+^*)^n |\, x_i = |z_i| \text{ and } (z_1,\dots,z_n) \in V\},$$

and T_r be the real torus $\mathrm{Log}^{-1}(r)$. So, $T_r \cap V$ is empty. Let $f \in \mathcal{I}(V)$ with $f(z) = \sum c_\alpha z^\alpha$ and g be the Laurent polynomial defined by $g(z) = \sum \overline{c}_\alpha w^\alpha$ with $w = (\frac{r_1^2}{z_1},\dots,\frac{r_n^2}{z_n})$ where the r_j's are the coordinates of r, and $\overline{c}_\alpha$ denotes the conjugate of the coefficient c_α. The value of the Laurent polynomial $h(z) = f(z)g(z)$ is equal to the value of $|f(z)|^2$ for every $z \in T_r$. By construction, the hypersurface V_h with defining polynomial h contains V (because $h \in \mathcal{I}(V)$). Let $\langle f_1,\dots,f_s\rangle$ be a set of generators of the ideal $\mathcal{I}(V)$, and for any j let g_j be the Laurent polynomial defined as before. We can check the hypersurface defined by the polynomial $G = \sum f_j g_j$ contains V and does not intersect the torus T_r. This proves that $r \in (\mathbb{R}_+^*)^n \setminus \bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f$. Hence, we have the inclusion:

$$\bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f \subset \{(x_1,\dots,x_n) \in (\mathbb{R}_+^*)^n \,|\, x_i = |z_i|,\ \text{and}\ (z_1,\dots,z_n) \in V\}.$$

Now let $(x_1,\dots,x_n) \in (\mathbb{R}_+^*)^n$ such that $x_i = |z_i|$ and $(z_1,\dots,z_n) \in V$, then for all $f \in \mathcal{I}(V)$ we have $f(z_1,\dots,z_n) = 0$. This means that $(x_1,\dots,x_n) \in \bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f$. $\square$

Proof of Theorem 1.1. The first equality of Theorem 1.1 is a consequence of Lemma 3.1. In fact, by applying the logarithmic map to both sides of the equality of Lemma 3.1 we obtain: $\mathrm{Log}\left(\bigcap_{f\in\mathcal{I}(V)} \mathscr{S}_f\right) = \mathscr{A}(V)$, and then

$$\mathscr{A}(V) = \bigcap_{f\in\mathcal{I}(V)} \mathscr{A}(V_f).$$

Let us prove the second equality of Theorem 1.1. Let $w \in \bigcap_{f\in\mathcal{I}(V)} co\mathscr{A}(V_f)$, then there exists a fundamental domain $\mathscr{D} = ([a; a+2\pi[)^n$ in the universal covering of the real torus $(S^1)^n$ and a unique $\widetilde{w} \in \mathscr{D}$ such that $w = \exp(i\widetilde{w})$. In this domain, the exponential map is a diffeomorphism between $\mathscr{D}$ and $(S^1)^n \setminus (S^1)^{n-1} \wedge \cdots \wedge (S^1)^{n-1}$ where $(S^1)^{n-1} \wedge \cdots \wedge (S^1)^{n-1}$ denotes the bouquet of n tori of dimension $n-1$. Let us define the subset $co\mathscr{S}_f$ of $\mathscr{D}$ as follow:

$$co\mathscr{S}_f := \{\theta \in \mathscr{D} |\ \text{there exists } z \in V_f \text{ and } \exp(i\theta) = \mathrm{Arg}\,(z)\}.$$

So, we have:

$$\bigcap_{f\in\mathcal{I}(V)} \exp\left(ico\mathscr{S}_f\right) = \exp\left(i \bigcap_{f\in\mathcal{I}(V)} co\mathscr{S}_f\right)$$

because the exponential map is a diffeomorphism from $\mathscr{D}$ into its image. Moreover, $\widetilde{w}$ is contained in the intersection $\bigcap_{f\in\mathcal{I}(V)} co\mathscr{S}_f$. But the last intersection, using the same argument as in Lemma 3.1, can be described as follows:

$$\begin{aligned}\bigcap_{f\in\mathcal{I}(V)} co\mathscr{S}_f &= \bigcap_{f\in\mathcal{I}(V)} \{\theta\in\mathscr{D} \mid \text{ there exists } z\in V_f \text{ and } \exp(i\theta)=\operatorname{Arg}(z)\}\\ &= \{\theta\in\mathscr{D} \mid \text{ there exists } z\in V \text{ and } \exp(i\theta)=\operatorname{Arg}(z)\}.\end{aligned}$$

Indeed, to prove the last equality, let $e^{i\theta}\notin co\mathscr{A}(V)$, and for each generator $f_j(z)=\sum c_\alpha z^\alpha$ of $\mathcal{I}(V)$ we define the polynomial g_j as follows:

$$g_j(z)=\sum \overline{c}_\alpha (e^{-2i\theta})^\alpha z^\alpha.$$

If $z\in\operatorname{Arg}^{-1}(e^{i\theta})$, then we have $f_jg_j(z)=|f_j(z)|^2$. The polynomial $G=\sum_j f_jg_j$ is in $\mathcal{I}(V)$, but $e^{i\theta}\notin co\mathscr{A}(V_G)$ because $|f_j(z)|^2>0$ and hence $G(z)=\sum_j f_jg_j(z)>0$ for every j and every $z\in\operatorname{Arg}^{-1}(e^{i\theta})$. Namely, we have the following inclusion:

$$\bigcap_{f\in\mathcal{I}(V)} co\mathscr{S}_f \subset \{\theta\in\mathscr{D} \mid \text{ there exists } z\in V \text{ and } \exp(i\theta)=\operatorname{Arg}(z)\}.$$

In other words, $\bigcap_{f\in\mathcal{I}(V)} co\mathscr{A}_f \subset co\mathscr{A}(V)$. $\square$

4. (Co)Amoebas of linear spaces

Throughout this section, $\mathscr{P}:=P(k)\cap(\mathbb{C}^*)^{k+m}$ where $P(k)$ is the k-dimensional affine linear subspace of $\mathbb{C}^{k+m}$ given by the parametrization ρ as follows:

$$\begin{array}{cccc}\rho \;: & \mathbb{C}^k & \longrightarrow & \mathbb{C}^{k+m}\\ & (t_1,\dots,t_k) & \longmapsto & (t_1,\dots,t_k,f_1(t_1,\dots,t_k),\dots,f_m(t_1,\dots,t_k)),\end{array} \tag{1}$$

where $f_j(t_1,\dots,t_k)=b_j+\sum_{i=1}^k a_{ji}t_i$, and a_{ji}, b_j are complex numbers for $i=1,\dots,k$, and $j=1,\dots,m$. By abuse of language, we call $\mathscr{P}$ an affine linear space instead of very affine linear space. First of all, if $\mathscr{P}$ is generic then all the coefficients b_j are different than zero. Otherwise $\mathscr{P}$ will be contained in an affine subgroup of $(\mathbb{C}^*)^{k+m}$. Indeed, if there exits j such that $b_j=0$, then there is an action of $\mathbb{C}^*$ on $\mathscr{P}$, and then $\mathscr{P}$ can be viewed as a product of $\mathbb{C}^*$ with an affine linear space of dimension $k-1$. Namely, $\mathscr{P}$ lies in no affine subgroup, i.e., $\rho(\mathbb{C}^k)$ meets each of the n coordinate hyperplanes of $\mathbb{C}^n$ in distinct hyperplanes, otherwise we may replace $(\mathbb{C}^*)^n$ by the smallest affine subgroup containing $\mathscr{P}$.

Lemma 4.1. *If $\mathscr{P}$ is generic, then we can assume that $f_1(t_1,\dots,t_k)=1+\sum_{i=1}^k t_i$.*

Proof. In fact, if we make a translation by $\frac{1}{b_1}$ in the algebraic multiplicative torus $(\mathbb{C}^*)^{k+m}$, we get

$$\left(\frac{t_1}{b_1},\dots,\frac{t_1}{b_1},\frac{f_1(t_1,\dots,t_k)}{b_1},\dots,\frac{f_m(t_1,\dots,t_k)}{b_1}\right).$$

We translate again by $a=(a_{11},a_{21},\dots,a_{1k},1,\dots,1)$ to obtain:

$$\left(\frac{a_{11}t_1}{b_1},\dots\frac{a_{1k}t_k}{b_1},1+\sum_{i=1}^{k}\frac{a_{1i}t_i}{b_1},\frac{f_2(t_1,\dots,t_k)}{b_1},\dots,\frac{f_m(t_1,\dots,t_k)}{b_1}\right).$$

For any point z in $(\mathbb{C}^*)^{k+m}$, we denote by τ_z the translation by z in the multiplicative group $(\mathbb{C}^*)^{k+m}$, and denote by ρ' the required parametrization, i.e.,

$$\rho'(t_1,\dots,t_k)=\left(t_1,\dots,t_k,1+\sum_{i=1}^{k}t_i,f_2(t_1,\dots,t_k),\dots,f_m(t_1,\dots,t_k)\right).$$

Hence, we obtain $\tau_a\circ\tau_{\frac{1}{b_1}}\circ\rho=\rho'\circ\tau_c$, where $c=(\frac{a_{11}}{b_1},\dots,\frac{a_{1k}}{b_1})$, and then, for any $(t_1,\dots,t_k)$ in $(\mathbb{C}^*)^k$ we have:

$$\operatorname{Arg}\Big(\rho(t_1,\dots,t_k)\Big)-\operatorname{Arg}(b_1)+\operatorname{Arg}(a)=\operatorname{Arg}\Big(\rho'(\tau_c(t_1,\dots,t_k))\Big).$$

We obtain the same relation if we replace the argument map by the logarithmic map. This means that the amoeba (respectively coamoeba) of a generic complex affine linear space $\mathscr{P}$ given by the parametrization (1) is the translation in the real space $\mathbb{R}^{k+m}$ (respectively the real torus $(S^1)^{k+m}$) by a vector v in $\mathbb{R}^{k+m}$ (respectively a point in the real torus) of an affine linear space given by a parametrization such that $f_1(t_1,\dots,t_k)=1+\sum_{i=1}^{k}t_i$. Hence, $co\mathscr{A}(\mathscr{P})=\tau_v\circ co\mathscr{A}(\mathscr{P}_{\rho'})$ where $\mathscr{P}_{\rho'}$ is the affine linear space given by the required parametrization, and we have a similar equality for their amoebas. In the last formula, v is the argument of the vector $b_1^{-1}a$. $\square$

To be more precise, $\mathscr{P}$ can be seen as the image by ρ of the complement in $\mathbb{C}^k$ of an arrangement of n hyperplanes $\mathscr{H}:=\cup_{i=1}^{k}\{t_i=0\}\cup_{j=1}^{m}\{f_j=0\}$.

4.1. (Co)Amoebas of lines in $(\mathbb{C}^*)^{1+m}$

In this subsection we give a complete description of (co)amoebas of generic lines in $(\mathbb{C}^*)^{1+m}$ (we mean a complex subvariety of complex dimension one defined by an ideal generated by polynomials of degree one). Moreover, we describe the (co)amoebas of real lines, i.e., lines those are invariant under the involution given by the conjugation of complex numbers. In other word, lines given by a parametrization with real coefficients. But first, let L be a generic line in $(\mathbb{C}^*)^{1+m}$ parametrized as follows:

$$\begin{array}{rcccl}\rho & : & \mathbb{C}^* & \longrightarrow & (\mathbb{C}^*)^{1+m}\\ & & t & \longmapsto & (t,t+1,a_2t+b_2,\dots,a_mt+b_m),\end{array}\qquad(2)$$

where a_j and b_j are non-vanishing complex numbers.

Lemma 4.2. *There are two types of amoebas of lines in* $(\mathbb{C}^*)^{1+m}$ *for* $m \geq 3$. *There are amoebas with boundary and other without boundary (we mean topological boundary). The amoebas of generic lines given by the parametrization* (2) *have boundary if and only if* $\frac{a_i}{b_i} \in \mathbb{R}^*$ *for all* $j = 2, \ldots, m$.

Proof. Since the boundary of an amoeba is a subset of the set of critical values of the logarithmic map, then an amoeba has a boundary means that the set of critical points of the logarithmic map restricted to the variety is nonempty (see [10], and [11] for more details about the critical points). The Jacobian of the logarithmic map restricted to the line L is given by:

$$\mathrm{Jac}(\mathrm{Log}_{|L})(t) = \frac{\partial \mathrm{Log}}{\partial(t,\bar{t})} = \frac{1}{2}\begin{pmatrix} 1/t & 1/\bar{t} \\ 1/(t+1) & 1/(\bar{t}+1) \\ a_2/(a_2t+b_2) & \bar{a}_2/(\overline{a_2t+b_2}) \\ \vdots & \vdots \\ a_m/(a_mt+b_m) & \bar{a}_m/(\overline{a_mt+b_m}) \end{pmatrix}.$$

Hence, a point $\rho(t)$ is critical for $\mathrm{Log}_{|L}$ if and only if all the 2×2-minors of the Jacobian matrix have determinant equal to zero. Let us write down these relations. The determinant of the 2×2-minor given by the two first rows:

$$\frac{1}{2}\begin{pmatrix} 1/t & 1/\bar{t} \\ 1/(t+1) & 1/(\bar{t}+1) \end{pmatrix}$$

is equal to zero, means the following equality holds:

$$\frac{1}{t}\frac{1}{\bar{t}+1} = \frac{1}{\bar{t}}\frac{1}{t+1}.$$

This implies that t should be real. For all $i = 2, \ldots, m$, the 2×2-minor:

$$\frac{1}{2}\begin{pmatrix} 1/t & 1/\bar{t} \\ a_i/(a_it+b_i) & \bar{a}_i/(\overline{a_it+b_i}) \end{pmatrix}$$

gives the following relation:

$$\frac{1}{t}\frac{\bar{a}_i}{(\overline{a_it+b_i})} = \frac{1}{\bar{t}}\frac{a_i}{a_it+b_i}.$$

But t is real, so $\bar{a}_i(a_it+b_i) = a_i(\overline{a_it+b_i})$, and hence $\frac{a_i}{b_i} = \overline{(\frac{a_i}{b_i})}$, i.e., $\frac{a_i}{b_i} \in \mathbb{R}^*$ (because L is generic, all the coefficients are different than zero). So, if $\frac{a_i}{b_i} \in \mathbb{R}^*$ for $i = 2, \ldots, m$, then the set of critical points of $\mathrm{Log}_{|L}$ is the image under ρ of the real part of $\mathbb{C}^*$, where this image intersects $(m+2)$ quadrants of $\mathbb{R}^{1+m}$ because L is generic. Moreover, this shows that the set of critical values of $\mathrm{Log}_{|L}$ is the image under $\mathrm{Log}\circ\rho$ of the real part of $\mathbb{C}^*$, and the number of its connected components is $(m+2)$. So, a generic complex line given by the parametrization (2) with $\frac{a_i}{b_i} \in \mathbb{R}^*$ for $i = 2, \ldots, m$ is real up to a translation by a complex number, and its amoeba is a surface with boundary, and the boundary has $(m+2)$ connected components. Also, we can check in this case that the cardinality of the inverse image of a regular (respectively critical) value is two (respectively one). $\square$

This motivates the following definition (see [14] for real plane curves):

Definition 4.1. A generic affine line given by the following parametrization:

$$\begin{array}{rcccl} \rho & : & \mathbb{C}^* & \longrightarrow & (\mathbb{C}^*)^{1+m} \\ & & t & \longmapsto & (t, a_1t+b_1, a_2t+b_2, \ldots, a_mt+b_m), \end{array} \tag{3}$$

where a_j and b_j are in $\mathbb{C}^*$ is called real up to a translation by a vector in $(\mathbb{C}^*)^{1+m}$ if and only if $[\frac{a_1}{b_1} : \cdots : \frac{a_m}{b_m}] \in \mathbb{RP}^{m-1}$.

If a line L in $(\mathbb{C}^*)^{1+m}$ with $m \geq 2$ is not real, then its amoeba is a surface without boundary homeomorphic to the Riemann sphere without $(m+2)$ points (see proof of Lemma 4.2), and the map $\operatorname{Log}_{|L}$ is a one-to-one map.

The following lemma gives a description of the coamoeba of a generic line in $(\mathbb{C}^*)^{1+m}$ with $m \geq 1$

Lemma 4.3. *Let $L \subset (\mathbb{C}^*)^{1+m}$ be a generic line given by the parametrization* (3)*. The restriction of the argument map to the set of its regular points in L is injective, and the inverse image under the argument map of a critical value has real dimension one.*

Proof. To see injectivity, let $(e^{i\theta}, e^{i\psi_1}, \ldots, e^{i\psi_m})$ be a fixed regular value in $co\mathscr{A}(L)$. In other word, we have $t = |t|e^{i\theta}$, and $f_j(t) = (a_jt + b_j) = |a_jt + b_j|e^{i\psi_1}$ for $j = 1, \ldots, m$, and consider a_jt, b_j, and $f_j(t)$ as a vectors in the complex plane. Hence, for each $j = 1, \ldots, m$ we obtain a parallelogram with vertices the origin, and the extremities of the three vectors a_jt, b_j, and $f_j(t)$. If one of these vectors is fixed, and the arguments of the two others are fixed (which is our case, because b_j is given and the arguments of a_jt and $f_j(t)$ are fixed by assumption), then there exists at most one parallelogram with those vertices. This implies the injectivity.

The second part of the lemma comes from the fact that the set of critical points of the logarithmic map and the argument map coincide (see Proposition 3.1 in [9]). Indeed, the set of critical points is equal to $(m+2)$ connected components of dimension one (each one corresponds to the intersection of the real part of L with some quadrant of $(\mathbb{R}^*)^{m+1}$). $\square$

The set of critical points of the argument map restricted to L given by the parametrization (3) is the image by ρ of the real part of $\mathbb{C}^*$ translated by $(1, b_1, \ldots, b_m)$ in $(\mathbb{C}^*)^{1+m}$ as a multiplicative group. So, the set of critical values consists of the translation by $(1, \frac{b_1}{|b_1|}, \ldots, \frac{b_m}{|b_m|})$ of $(m+2)$ points in the real torus $(S^1)^{1+m}$ from the 2^{m+1} real points corresponding to the arguments of the 2^{m+1} quadrants of $\operatorname{Re}((\mathbb{C}^*)^{1+m}) = (\mathbb{R}^*)^{1+m}$. The closure of the coamoeba of L contains an arrangement of $(m+1)$ geodesic circles. Each circle corresponds to an end of the line (i.e., where L meets a coordinate axis). The union of these circles is the set of accumulation points of arguments of sequences in L with unbounded logarithm, and is called the phase limit set of L (see [17] for more details). It is the counterpart of the logarithmic limit set introduced by Bergman in 1971 (see [1] and [7] for more details), which consists of $(m+2)$ points is our case. In Figure 1

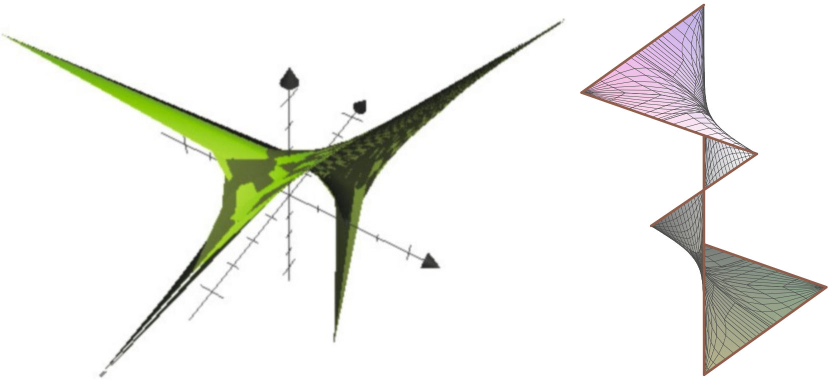


FIGURE 1. The amoeba and the coamoeba of the real line in $(\mathbb{C}^*)^3$ given by the parametrization $\rho(z) = (z, z + \frac{1}{2}, z - \frac{3}{2})$. The amoeba is topologically the closed disk without four points of its boundary.

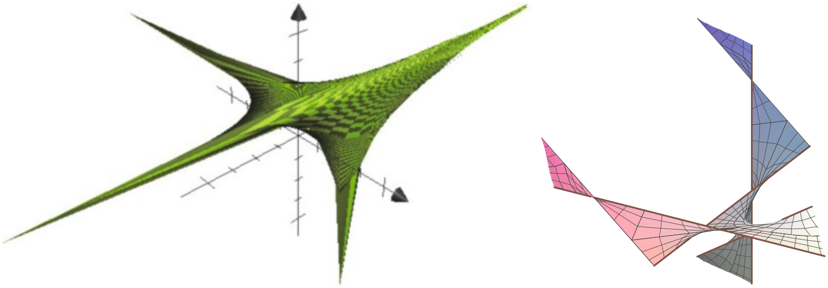


FIGURE 2. The amoeba and the coamoeba of the complex line (i.e., not real) in $(\mathbb{C}^*)^3$ given by the parametrization $\rho(z) = (z, z+1, z-2i)$. The amoeba is topologically the Riemann sphere without four points.

(respectively Figure 2), we draw the amoeba and the coamoeba of a real (respectively non real) line in $(\mathbb{C}^*)^3$. The coamoebas in Figure 1, and Figure 2 are made with collaboration with F. Sottile.

5. Volume of (co)amoebas of k-dimensional very affine linear spaces in $(\mathbb{C}^*)^{2k}$

It was shown by Rullgård and the second author in [19] that the area of the amoeba of a complex algebraic plane curve is always finite, and the bound is given in terms of the area of the Newton polygon of the defining polynomial. Mikhalkin and Rullgård proved that this bound is always sharp for (possibly singular) Harnack curves (see [14]). It was shown by Madani and the first author in [8] that the volume of the amoeba of a k-dimensional algebraic variety in $(\mathbb{C}^*)^n$ with $n \geq 2k$ is finite. This generalizes the result of Rullgård and the second author about the finiteness of the volume of the amoeba of plane curves. In this section, we compute the volume of the amoeba of a generic real k-dimensional very affine linear space in $(\mathbb{C}^*)^{2k}$. We will proceed as follows: (i) We show that the argument map restricted to the subset of regular points in the very affine linear space is injective; (ii) We compute the volume of the coamoeba of any k-dimensional very affine linear space in $(\mathbb{C}^*)^{2k}$; (iii) We compute the cardinality of the inverse image under the logarithmic map of any regular value in the amoeba of a real affine space, and prove that this cardinality is a constant and equal to 2^k; (iv) We use that the map $\mathrm{Log} \circ \mathrm{Arg}^{-1}$ conserves the volume, i.e., the determinant of its Jacobian has absolute value equal one (see Proposition 3.1 in [9]), and finally we compute the volume of the amoeba, which is equal to the coamoeba volume divided by 2^k if the plane is real. We will use the following lemma proved in [10], which is a generalization of Mikhalkin's Lemma 4.3 in [12] for hypersurface:

Lemma 5.1 (Madani–Nisse). *Let $V \subset (\mathbb{C}^*)^n$ be a k-dimensional algebraic variety, and z be a smooth point of V. Then z is a critical point for the map $\mathrm{Log}_{|V}$ if and only if the image of the tangent space T_zV to V at z by the derivative of the complex logarithm $d\mathscr{L}og$ contains at least s purely imaginary linearly independent vectors with $s = \max\{1, 2k - n + 1\}$.*

Also, we will use the following proposition proved in [10]:

Proposition 5.1 (Madani–Nisse). *Let $\mathscr{P} \subset (\mathbb{C}^*)^n$ be a generic k-dimensional very affine linear space with $n \geq 2k$. Suppose that the complex dimension of $\mathscr{P} \cap \overline{\mathscr{P}}$ is equal to l, with $0 \leq l \leq k$. Then, for any regular value x in the amoeba $\mathscr{A}(\mathscr{P})$ of $\mathscr{P}$, the cardinality of $\mathrm{Log}^{-1}(x)$ is at least 2^l.*

Let $\mathscr{P} \subset (\mathbb{C}^*)^{2k}$ be a generic k-dimensional very affine linear space. Suppose $\mathscr{P}$ is given by the parametrization ρ:

$$\begin{array}{rccl} \rho \ : & (\mathbb{C}^*)^k & \longrightarrow & (\mathbb{C}^*)^{2k} \\ & (t_1, \dots, t_k) & \longmapsto & (t_1, \dots, t_k, f_1(t_1, \dots, t_k), \dots, f_k(t_1, \dots, t_k)), \end{array} \tag{4}$$

with $f_j(t_1, \dots, t_k) = b_j + \sum_{i=1}^k a_{ji} t_i$, where a_{ji}, and b_j are complex numbers for $i = 1, \dots, k$ and $j = 1, \dots, k$. Since the space $\mathscr{P}$ is generic, then there is no $b_j = 0$.

Definition 5.1. A generic k-dimensional very affine linear space $\mathscr{P}(k) \subset (\mathbb{C}^*)^{2k}$ given by the parametrization (4) is said to be real up to a translation by a complex vector in the multiplicative group $(\mathbb{C}^*)^{k+m}$ if and only if the $(m \times k)$-matrix given by

$$\begin{pmatrix} \frac{a_{11}}{b_1} & \cdots & \frac{a_{1k}}{b_1} \\ \vdots & \vdots & \vdots \\ \frac{a_{k1}}{b_k} & \cdots & \frac{a_{kk}}{b_k} \end{pmatrix}$$

has rank k and all of its entries are real.

Let $\mathbb{Z}_2 := \{\pm 1\}$ be the real subgroup of the multiplicative group $\mathbb{C}^*$, and $\mathbb{Z}_2^{2k}$ be the finite real subgroup of $(\mathbb{C}^*)^{2k}$. For each $s \in \mathbb{Z}_2^{2k}$, let ρ_s be the parametrization given by $\rho_s(t_1, \ldots, t_k) = s.\rho(t_1, \ldots, t_k)$ where

$$s.(z_1, \ldots, z_{2k}) = (s_1 z_1, \ldots, s_{2k} z_{2k})$$

for any $(z_1, \ldots, z_{2k}) \in (\mathbb{C}^*)^{2k}$, and $s = (s_1, \ldots, s_{2k}) \in \mathbb{Z}_2^{2k}$. Let $\mathscr{P}_s$ be the k-dimensional very affine linear space in $(\mathbb{C}^*)^{2k}$ parametrized by ρ_s. Let us denote by $\mathrm{Reg}(co\mathscr{A}(\mathscr{P}_s))$ the set of regular values of $co\mathscr{A}(\mathscr{P}_s)$. Remark that if 1 denotes the identity element of the group $\mathbb{Z}_2^{2k}$, then $\mathscr{P} = \mathscr{P}_1$.

Let $u \in \mathbb{Z}_2^{2k}$ and denote by $\mathrm{Reg}(co\mathscr{A}(\mathscr{P}_u))$ the set of regular values of the coamoeba $co\mathscr{A}(\mathscr{P}_u)$.

Proposition 5.2. *With the above notations, the following statements hold:*

(i) *For all s, the argument map from the subset of regular points of $\mathscr{P}_s$ to the set of regular values of its coamoeba $co\mathscr{A}(\mathscr{P}_s)$ is injective;*
(ii) *Let s and r in $\mathbb{Z}_2^{2k}$ with $s \neq r$, then the set*

$$\mathrm{Reg}(co\mathscr{A}(\mathscr{P}_s)) \cap \mathrm{Reg}(co\mathscr{A}(\mathscr{P}_r))$$

is empty;
(iii) *The union $\bigcup_{s\in\mathbb{Z}_2^{2k}} \mathrm{Reg}(co\mathscr{A}(\mathscr{P}_s))$ is an open dense subset of the real torus $(S^1)^{2k}$.*

First of all, we denote by $z := (z_1, \ldots, z_{2k})$ the coordinates of $\mathbb{C}^{2k}$. So, if z is a point in $\mathscr{P}$, then $z_i = t_i$ and $z_{k+i} = f_i(z_1, \ldots, z_k)$ for $1 \leq i \leq k$. Let $\Theta = (e^{i\theta_1}, \ldots, e^{i\theta_k}, e^{i\psi_1}, \ldots, e^{i\psi_k})$ be a point in the set of regular values of $co\mathscr{A}(\mathscr{P})$. This means that the linear system (E) of $2k$ equations and $2k$ variables $(x_1, \ldots, x_k, y_1, \ldots, y_k)$ in $(\mathbb{R}_+^*)^{2k}$:

$$\left\{ \begin{array}{lcl} \mathrm{Re}(b_j + \sum_{l=1}^{k} a_{jl} x_l e^{i\theta_l}) & = & \mathrm{Re}\,(y_j e^{i\psi_j}) \\ \mathrm{Im}(b_j + \sum_{l=1}^{k} a_{jl} x_l e^{i\theta_l}) & = & \mathrm{Im}\,(y_j e^{i\psi_j}) \end{array} \right. \qquad (E)$$

with $j = 1, \ldots, k$, has a solution in $(\mathbb{R}_+^*)^{2k}$. Moreover, if $\mathbb{Z}_2^{2k}$ is viewed as a subgroup of the real torus $(S^1)^{2k}$, then $s.\Theta \in \bigcup_{u\in\mathbb{Z}_2^{2k}} \mathrm{Reg}(co\mathscr{A}(\mathscr{P}_u(k)))$ means that the system (E) has a solution in $(\mathbb{R}^*)^{2k}$.

Since the matrix $A(z)$ defined by:

$$A(z) = \begin{pmatrix} a_{11}z_1 & a_{12}z_2 & \ldots & a_{1k}z_k & -z_{k+1} & 0 & 0 & \ldots & 0 \\ a_{21}z_1 & a_{22}z_2 & \ldots & a_{2k}z_k & 0 & -z_{k+2} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1}z_1 & a_{k2}z_2 & \ldots & a_{kk}z_k & 0 & 0 & 0 & \ldots & -z_{2k} \end{pmatrix}$$

is the image under the logarithmic Gauss map of the point z in $\mathscr{P}$, and the matrix $A(z)$ has rank k when z is a regular point.

Claim I. If $\overline{A}$ denotes the matrix conjugate to A, then for any regular point z of $\mathscr{P}$ the matrix $\widehat{A}(z) = \left(\begin{array}{c} A(z) \\ \overline{A}(z) \end{array} \right)$ is of rank $2k$.

Proof. In fact, the rows of the matrix $A(z)$ form a basis of the orthogonal space to $\mathscr{L}og(\mathscr{P})$ at the point $\mathscr{L}og(z)$. So, if the rank of $\widehat{A}(z)$ is less than $2k$, then the orthogonal space to $\mathscr{L}og(\mathscr{P})$ at $\mathscr{L}og(z)$ contains at least one real vector v different than zero. This is equivalent to saying that the tangent space to $\mathscr{L}og(\mathscr{P})$ at $\mathscr{L}og(z)$ contains at least one purely imaginary vector. Indeed, since v is a vector different than zero orthogonal to both $T_{\mathscr{L}og(z)}(\mathscr{L}og(\mathscr{P}))$ and $\mathrm{Im}(\mathbb{C}^{2k})$, then $T_{\mathscr{L}og(z)}(\mathscr{L}og(\mathscr{P})) \cap \mathrm{Im}(\mathbb{C}^{2k})$ must be of dimension at least one. By Lemma 5.1, this implies that z is a critical point for the logarithmic map, which is in contradiction with our assumption on z. $\square$

The matrix defining the system (E) is $\widetilde{B}(\Theta) = \left(\begin{array}{c} \mathrm{Re}B(\Theta) \\ \mathrm{Im}B(\Theta) \end{array} \right)$ where $B(\Theta)$ is

$$\left(\begin{array}{ccccccccc} a_{11}e^{i\theta_1} & a_{12}e^{i\theta_2} & \dots & a_{1k}e^{i\theta_k} & -e^{i\psi_1} & 0 & 0 & \dots & 0 \\ a_{21}e^{i\theta_1} & a_{22}e^{i\theta_2} & \dots & a_{2k}e^{i\theta_k} & 0 & -e^{i\psi_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1}e^{i\theta_1} & a_{k2}e^{i\theta_2} & \dots & a_{kk}e^{i\theta_k} & 0 & 0 & 0 & \dots & -e^{i\psi_k} \end{array} \right).$$

We can check that the rank of $\widetilde{B}(\Theta)$ is the same as the rank of the matrix $\widetilde{A}(z) = \left(\begin{array}{c} \mathrm{Re}A(z) \\ \mathrm{Im}A(z) \end{array} \right)$ with $z = (x_1e^{i\theta_1}, \dots, x_ke^{i\theta_k}, y_1e^{\psi_1}, \dots, y_ke^{i\psi_k})$, because the variables x_i and y_j are non zero for all $i, j = 1, \dots, k$.

Claim II. The rank of the matrix $\widetilde{A}(z)$ is equal to $2k$.

Proof. Suppose we have a non trivial linear combination of the rows of the matrix $\widetilde{A}(z)$ that is equal to zero. Hence, there exist a real numbers λ_l, and μ_l not all equal to zero, with $l = 1, \dots, k$ such that:

$$\sum_{l,j=1}^{k} \frac{\lambda_l}{2}\Big((z_ja_{lj} + \bar{z}_j\bar{a}_{lj}) - (z_{k+l} + \bar{z}_{k+l})\Big) + \frac{\mu_l}{2i}\Big((z_ja_{lj} - \bar{z}_j\bar{a}_{lj}) - (z_{k+l} - \bar{z}_{k+l})\Big) = 0.$$

We get:

$$\sum_{l=1}^{k}\left(\frac{\lambda_l - i\mu_l}{2}\right)\left(\sum_{j=1}^{k} z_ja_{lj} - z_{k+l}\right) + \sum_{l=1}^{k}\left(\frac{\lambda_l + i\mu_l}{2}\right)\left(\sum_{j=1}^{k} \bar{z}_j\bar{a}_{lj} - \bar{z}_{k+l}\right) = 0.$$

Since the matrix $\widehat{A}(z)$ is of rank $2k$ by Claim I, this implies that $\lambda_l - i\mu_l = 0$, and $\lambda_l + i\mu_l = 0$ for all $l = 1, \dots, k$. This means that all the λ_l's and the μ_l's vanish. This contradicts the fact that some of the real numbers λ_l's and μ_l's are different than zero by hypothesis. Hence, the real rank of the matrix $\widetilde{A}(z)$ is equal to $2k$. $\square$

Proof of Proposition 5.2. Since the k-dimensional linear space $\mathscr{P}$ is generic, the coefficients b_j are different than zero, and the system (E) is consistent. Claim II shows that the system (E) has a unique solution for any Θ in the set of regular values of $co\mathscr{A}(\mathscr{P})$, which proves the first and the second statements of the proposition. The third statement comes from the fact that the set of $\Theta = (\theta_1, \dots, \theta_k, \psi_1, \dots, \psi_k)$ for which the determinant of $\widetilde{B}(\Theta)$ vanishes is a hypersurface in the real torus and then its $2k$-volume is zero. In other words, the union $\bigcup_{s\in\mathbb{Z}_2^{2k}} \mathrm{Reg}(co\mathscr{A}(\mathscr{P}_s))$ is an open dense subset of the real torus $(S^1)^{2k}$. □

Corollary 5.1. *The volume of the coamoeba of any generic k-dimensional linear space in $(\mathbb{C}^*)^{2k}$ is equal to π^{2k}.*

Proof. By Proposition 5.2 (iii), the volume of the disjoint union

$$\bigcup_{s\in\mathbb{Z}_2^{2k}} \mathrm{Reg}(co\mathscr{A}(\mathscr{P}_s))$$

is equal to the volume of all the real torus $(S^1)^{2k}$. Moreover, they have the same volume, because they are obtained from each other by translation (i.e., isometry of the real torus equipped with the flat metric). So, the volume of one of them must be equal to $(2\pi)^{2k}/2^{2k} = \pi^{2k}$. □

We compute the cardinality of the inverse image under the logarithmic map of any regular value in the amoeba of a generic k-dimensional real very affine linear space $\mathscr{P}(k) \subset (\mathbb{C}^*)^{2k}$.

Proposition 5.3. *Let $\mathscr{P}$ be a generic real affine k-dimensional linear subspace of $(\mathbb{C}^*)^{2k}$, and x be a regular value of its amoeba. Then, the cardinality of* $\mathrm{Log}^{-1}(x)$ *is equal* 2^k.

Proof. We assume that $\mathscr{P}$ is given by a parametrization ρ as in (4), where all the coefficients are real numbers. The matrix A defined by:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

is invertible, otherwise the image of ρ is a linear space of dimension strictly less than k. The following diagram is commutative:

$$\begin{tikzcd} (\mathbb{C}^*)^k \arrow[r, "\rho"] \arrow[d, "A"'] & (\mathbb{C}^*)^{2k} \arrow[d, "A\times Id"] \\ (\mathbb{C}^*)^k \arrow[r, "\rho'"] & (\mathbb{C}^*)^{2k}, \end{tikzcd}$$

where ρ' is the parametrization given by:

$$\begin{array}{rcccl} \rho' & : & (\mathbb{C}^*)^k & \longrightarrow & (\mathbb{C}^*)^{2k} \\ & & (T_1, \dots, T_k) & \longmapsto & (T_1, \dots, T_k, b_1 + T_1, \dots, b_k + T_k). \end{array}$$

Each regular value of the amoeba of the k-dimensional linear space $\mathscr{L} := \rho'((\mathbb{C}^*)^k)$ is covered 2^k times under the logarithmic mapping. Indeed, $\mathscr{L}$ is a product of lines $L_1, \ldots, L_k$ in $\mathbb{C}^2$. The matrix A is real, so the image of the set of critical points of the logarithmic mapping restricted to $\mathscr{P}$ is the set of critical points of the logarithmic mapping restricted to $\mathscr{L}$. By Lemma 5.1, if z is a critical in $\mathscr{P}$, then the tangent space to $\mathscr{L}og(\mathscr{P})$ at $\mathscr{L}og(z)$ contains at least one purely imaginary vector v. Since A is real, then the image of v in the tangent space to $\mathscr{L}og(\mathscr{L})$ at $\mathscr{L}og((A \times Id)(z)$ is also purely imaginary tangent vector, and then, the point $(A \times Id)(z)$ is critical. Let $\mathrm{Critp}(\mathrm{Log}_{|\mathscr{P}})$ and $\mathrm{Critp}(\mathrm{Log}_{|\mathscr{L}})$ be the set of critical points of the restriction of the logarithmic map to $\mathscr{P}$ and $\mathscr{L}$ respectively. Since the volume of their amoebas is finite (see [8]), this means that the set of critical values in their amoebas contains a subset of dimension $2k-1$ (at least the topological boundary of the amoeba). Hence, the number of connected components of $\mathscr{P} \setminus \mathrm{Critp}(\mathrm{Log}_{|\mathscr{P}})$ is equal to the number of connected components of $\mathscr{L} \setminus \mathrm{Critp}(\mathrm{Log}_{|\mathscr{L}})$. The fact that the set of critical points of the argument and the logarithmic maps coincide (see, e.g., [9]), and by Proposition 5.2, the restriction of the argument map to the set of regular points is injective, then, the cardinality of $\mathrm{Log}^{-1}(x)$ is at most 2^k. Since $\mathscr{P}$ is real, then by Proposition 5.1, for any regular value $x \in \mathscr{A}(\mathscr{P}(k))$, the cardinality of $\mathrm{Log}^{-1}(x)$ is at least 2^k. Hence, the cardinality of the inverse image of a regular value is equal to 2^k. □

Proof of Theorem 1.2. The first statement of Theorem 1.2 is Corollary 5.1. The second statement of Theorem 1.2 is because the cardinality of the inverse image of a regular value in the amoeba is constant and equal to 2^k, and the map $\mathrm{Log} \circ \mathrm{Arg}^{-1}$ conserve the volume (see [9], Proposition 3,1). Hence, the volume of the amoeba in this case is equal to the volume of the coamoeba divided by 2^k.

References

[1] G. M. Bergman, The logarithmic limit-set of an algebraic variety, Trans. Amer. Math. Soc. **157** (1971), 459–469.

[2] M. Forsberg, M. Passare, and A. Tsikh, Laurent determinants and arrangements of hyperplane amoebas, Advances in Math. **151** (2000), 45–70.

[3] I. M. Gelfand, M. M. Kapranov and A. V. Zelevinski, Discriminants, resultants and multidimensional determinants, Birkhäuser Boston 1994.

[4] P. Johansson, Coamoebas, Licentiate Thesis, Stockholm University, 2010.

[5] M. M. Kapranov, A characterization of A-discriminantal hypersurfaces in terms of the logarithmic Gauss map, Math. Ann. **290**, (1991), no. 2, 277–285.

[6] I. Itenberg, G. Mikhalkin and E. Shustin, *Tropical Algebraic Geometry*, Oberwolfach Seminars, Volume 35, Birkhäuser Basel-Boston-Berlin 2007.

[7] D. Maclagan and B. Sturmfels, Introduction to tropical geometry. Graduate Studies in Math., Vol **161**, American Math. Soc. (2015).

[8] F. Madani and M. Nisse, On the volume of complex amoebas, Proceedings of the AMS, **141** (2013), 1113–1123.

[9] F. Madani and M. Nisse, Analytic varieties with finite volume amoebas are algebraic, J. reine angew. Math. Vol. **706**, (2015), 67–81.

[10] F. Madani and M. Nisse, Generalized logarithmic Gauss map and its relation to (co)amoebas, Math. Nachr. **286**, No. 14–15, (2013), 1510–1513.

[11] G. Mikhalkin, Real algebraic curves, moment map and amoebas, Ann. of Math. **151** (2000), 309–326.

[12] G. Mikhalkin, Decomposition into pairs-of-pants for complex algebraic hypersurfaces, Topology **43** (2004), 1035–1065.

[13] G. Mikhalkin, Enumerative Tropical Algebraic Geometry In $\mathbb{R}^2$, J. Amer. Math. Soc. **18**, (2005), 313–377.

[14] G. Mikhalkin and H. Rullgård, Amoebas of maximal area, Int. Math. Res. Notices **9**, (2001), 441–451.

[15] M. Nisse, Sur la Géométrie et la Topologie des Amibes et Coamibes des Variétés Algébriques Complexes, Thèse de Doctorat de l'Université Pierre et Marie Curie – Paris 6, 2010.

[16] M. Nisse, Geometric and Combinatorial Structure of Hypersurface Coamoebas, preprint, arXiv:0906.2729.

[17] M. Nisse and F. Sottile, The phase limit set of a variety, Algebra & Number Theory, **7**-2, (2013), 339–352.

[18] M. Nisse and F. Sottile, Non-Archimedean coAmoebae, Contemporary Mathematics of the AMS Vol. **605**, (2013), 73–91.

[19] M. Passare and H. Rullgård, Amoebas, Monge–Ampère measures and triangulations of the Newton polytope, Duke Math. J. **121**, (2004), 481–507.

[20] K. Purbhoo, A Nullstellensatz for amoebas, Duke Math. J. **141**, (2008) no. 3, 407–445.

Mounir Nisse and Mikael Passare
School of Mathematics KIAS
85 Hoegiro Dongdaemun-gu
Seoul 02455, Republic of Korea
e-mail: mounir.nisse@gmail.com

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 81–90

One Parameter Regularizations of Products of Residue Currents

Mats Andersson, Håkan Samuelsson Kalm, Elizabeth Wulcan, and Alain Yger

Dedicated to the memory of Mikael Passare

Abstract. We show that Coleff–Herrera type products of residue currents can be defined by analytic continuation of natural functions depending on one complex variable.

Mathematics Subject Classification (2010). 32A26, 32A27, 32B15, 32C30.

1. Introduction

Let f be a holomorphic function defined on a domain in $\mathbb{C}^n$. It is proved in [15] using Hironaka's desingularization theorem that if φ is a test form then

$$\lim_{\epsilon \to 0^+} \int_{|f|^2 > \epsilon} \varphi / f$$

exists and defines the action of a current, denoted $1/f$. The $\bar{\partial}$-image, $\bar{\partial}(1/f)$, is the residue current of f and it has the useful property that it is annihilated by a holomorphic function g if and only if g is in the ideal generated by f. If $f_1, \ldots, f_q$ are holomorphic functions then the *Coleff–Herrera product* of the currents $\bar{\partial}(1/f_j)$ is defined as follows. For a test form φ of bidegree $(n, n-q)$ consider the residue integral

$$I_f^{\varphi}(\epsilon) = \int_{T(\epsilon)} \frac{\varphi}{f_1 \cdots f_q},$$

where $T(\epsilon) = \cap_1^q \{|f_j|^2 = \epsilon_j\}$. It is proved in [12] that the limit of $\epsilon \mapsto I_f^{\varphi}(\epsilon)$ exists if $\epsilon = (\epsilon_1, \ldots, \epsilon_q) \to 0$ along a path in $\mathbb{R}_+^q$ such that $\epsilon_j / \epsilon_{j+1}^k \to 0$ for all $k \in \mathbb{N}$ and

First three authors partly supported by the Swedish Research Council.

$j = 1, \dots, q-1$; such a path is said to be *admissible*. Moreover, the limit defines the action of a current, the Coleff–Herrera product

$$\bar\partial \frac{1}{f_q} \wedge \cdots \wedge \bar\partial \frac{1}{f_1} . \varphi := \text{“}\lim_{\epsilon \to 0}\text{”}\, I_f^\varphi(\epsilon), \tag{1.1}$$

where “lim” means the limit along an admissible path as above. Following Passare [19], let χ be a smooth approximation of the characteristic function $\mathbf{1}_{[1,\infty)}$ and consider the smooth form

$$\frac{\bar\partial \chi(|f_q|^2/\epsilon_q)}{f_q} \wedge \cdots \wedge \frac{\bar\partial \chi(|f_1|^2/\epsilon_1)}{f_1}. \tag{1.2}$$

It follows from [16, Theorem 2] or the proof of [19, Proposition 2] that the limit in the sense of currents of (1.2) as $\epsilon \to 0$ along an admissible path equals the Coleff–Herrera product, and moreover, that one gets the same result if one first lets $\epsilon_1 \to 0$, then lets $\epsilon_2 \to 0$ and so on. The Coleff–Herrera product is thus indeed the result of an iterative procedure. In general there are no obvious commutation properties, e.g., $\bar\partial(1/zw) \wedge \bar\partial(1/z) = 0$ whereas $\bar\partial(1/z) \wedge \bar\partial(1/zw) = \bar\partial(1/z^2) \wedge \bar\partial(1/w)$, where the last product is simply a tensor product. However, if $f = (f_1, \dots, f_q)$ defines a complete intersection, i.e., $\operatorname{codim}\{f = 0\} = q$, then the Coleff–Herrera product depends in an anticommutative way of the ordering of the tuple f; in fact by [11] the smooth form (1.2) then converges unconditionally. Moreover, also in the complete intersection case, a holomorphic function annihilates the Coleff–Herrera product if and only if it is in the ideal $\langle f_1, \dots, f_q \rangle$; this last property is called the *duality property* and it was proved independently by Dickenstein–Sessa, [13], and Passare, [18].

In this paper we consider another approach to Coleff–Herrera type products; it is based on analytic continuation and has been studied in, e.g., [6, 7, 10, 20, 27]. For $\lambda_j \in \mathbb{C}$ with $\mathfrak{Re}\, \lambda_j \gg 0$, let

$$\Gamma_f^\varphi(\lambda_1, \dots, \lambda_q) = \int \frac{\bar\partial |f_q|^{2\lambda_q} \wedge \cdots \wedge \bar\partial |f_1|^{2\lambda_1}}{f_1 \cdots f_q} \wedge \varphi,$$

where φ is a test form. It is standard to see that $\lambda_1 \mapsto \Gamma_f^\varphi(\lambda_1, \dots, \lambda_q)$ has an analytic continuation to a neighborhood of 0 and that $\Gamma_f^\varphi(0, \lambda_2, \dots, \lambda_q)$ equals

$$\frac{\bar\partial |f_q|^{2\lambda_q}}{f_q} \wedge \cdots \wedge \frac{\bar\partial |f_2|^{2\lambda_2}}{f_2} \wedge \bar\partial \frac{1}{f_1} . \varphi.$$

From [5, Proposition 2.1] it follows that $\lambda_2 \mapsto \Gamma_f^\varphi(0, \lambda_2, \dots, \lambda_q)$ is analytic at 0, that $\lambda_3 \mapsto \Gamma_f^\varphi(0, 0, \lambda_3, \dots, \lambda_q)$ is too, and so on. Once one knows that the Coleff–Herrera product is obtained by letting $\epsilon_j \to 0$ successively in (1.2) it is not that hard to see that

$$\bar\partial \frac{1}{f_q} \wedge \cdots \wedge \bar\partial \frac{1}{f_1} . \varphi = \Gamma_f^\varphi(\lambda_1, \dots, \lambda_q)|_{\lambda_1 = 0} \cdots |_{\lambda_q = 0},$$

where the expression on the right-hand side means that we first let $\lambda_1 \to 0$, then let $\lambda_2 \to 0$ etc; see, e.g., [16, Theorem 2]. However, from an algebraic point of view, cf. [8, Theorem 3.2], it is often desirable to have a current given as the value at 0 of a single one-variable analytic function; this is the motivation for this paper. From Theorem 1.2 below it follows that if $\mu_1 > \cdots > \mu_q > 0$ are integers, then $\lambda \mapsto \Gamma_f^{\varphi}(\lambda^{\mu_1}, \ldots, \lambda^{\mu_q})$, a priori defined for $\mathfrak{Re}\,\lambda \gg 0$, has an analytic continuation to a neighborhood of $[0, \infty) \subset \mathbb{C}$ and that the value at $\lambda = 0$ equals the Coleff–Herrera product (1.1). Notice that this way of letting $(\lambda_1, \ldots, \lambda_q) \to 0$ is analogous to limits along admissible paths in the sense that λ_j goes to zero much faster than λ_{j+1}, $j = 1, \ldots, q-1$.

We remark that if f defines a complete intersection then it is showed in [23] that $\Gamma_f^{\varphi}(\lambda)$ is analytic in a neighborhood of the half-space $\{\mathfrak{Re}\,\lambda_j \geq 0, j = 1, \ldots, q\}$.

Let us now consider a more general setting. Let f be a section of a Hermitian vector bundle E of rank m over a reduced complex space X of pure dimension n. In [22] and [1] were introduced currents U and R, generalizing the currents $1/f$ and $\bar{\partial}(1/f)$, respectively. These currents are based on Bochner–Martinelli type expressions. To be precise, let $f = f_1 e_1 + \cdots + f_m e_m$, where $\{e_k\}_k$ is a local holomorphic frame for E with dual frame $\{e_k^*\}_k$, and let $s = s_1 e_1^* + \cdots + s_m e_m^*$ be the section of the dual bundle E^* with pointwise minimal norm such that $f \cdot s = |f|_E^2$. For $\lambda \in \mathbb{C}$, $\mathfrak{Re}\,\lambda \gg 0$, we let

$$U^\lambda := \sum_{k=1}^{m} |f|_E^{2\lambda} \frac{s \wedge (\bar{\partial} s)^{k-1}}{|f|_E^{2k}}, \tag{1.3}$$

where $(0,1)$-forms anticommute with the e_k^*. It turns out, [1], [22], that $\lambda \mapsto U^\lambda$, considered as a current-valued map, has an analytic continuation to a neighborhood of 0. The value at $\lambda = 0$ is a current U on X that takes values in ΛE^*; U is the standard extension of $\sum_k s \wedge (\bar{\partial} s)^{k-1}/|f|_E^{2k}$ across $\{f = 0\}$. If E has rank 1, then $U = (1/f)e^*$ for any choice of metric. Let

$$R^\lambda := 1 - |f|_E^{2\lambda} + \sum_{k=1}^{m} \bar{\partial} |f|_E^{2\lambda} \wedge \frac{s \wedge (\bar{\partial} s)^{k-1}}{|f|_E^{2k}}. \tag{1.4}$$

Letting $\nabla_f := \delta_f - \bar{\partial}$, where δ_f denotes interior multiplication with f, one can check that $R^\lambda = 1 - \nabla_f U^\lambda$, see [1] for details. It follows that $\lambda \mapsto R^\lambda$ has an analytic continuation to a neighborhood of 0 and the value at $\lambda = 0$ is the current R; it is straightforward to check that R has support on $\{f = 0\}$. If E has rank 1 then $R = \bar{\partial}(1/f) \wedge e^*$ and more generally, if f defines a complete intersection then $R = \bar{\partial}(1/f_m) \wedge \cdots \wedge \bar{\partial}(1/f_1) \wedge e_1^* \wedge \cdots \wedge e_m^*$ for any choice of metric, see [1] and [22].

The value at $\lambda = 0$ of the term $1 - |f|_E^{2\lambda}$ of R^λ is the restriction $\mathbf{1}_{\{f=0\}}$ to the zero set of f, see [5]. In itself it is zero unless f vanishes identically on some components of X in which case it simply is 1 there. However, when forming products of R's the role of $\mathbf{1}_{\{f=0\}}$ is much more significant, cf. [3] and Example 1.4.

Remark 1.1. For future reference we notice that if $\pi \colon X' \to X$ is a modification such that the pullback of the ideal sheaf defined by f is principal, then one can write $\pi^* f = f^0 f'$, where f^0 is a section of the line bundle $L \to X'$ corresponding to the exceptional divisor and f' is a non-vanishing section of $L^{-1} \otimes \pi^* E$. Equipping

L with some Hermitian metric, for instance by setting $|f^0|_L := |\pi^* f|_{\pi^* E}$, we can thus write $\pi^*|f|_E = |f^0|_L |f'|_{L^{-1}\otimes \pi^* E}$. Locally on X' we can identify f^0 and f' by a holomorphic function and a non-vanishing holomorphic tuple, respectively, still denoted f^0 and f'. Hence, locally on X' we have $\pi^*|f|_E = |f^0| u$ for some smooth positive function u.

Let f_j be a section of a Hermitian vector bundle E_j of rank m_j, let U^j and R^j be the associated currents, and let $U^{j,\lambda}$ and $R^{j,\lambda}$ be the corresponding λ-regularizations. Following, e.g., [3] and [16] we define products of the currents R^j recursively as follows. Having defined $R^{k-1}\wedge\cdots\wedge R^1$, consider the current-valued function

$$\lambda \mapsto R^{k,\lambda}\wedge R^{k-1}\wedge\cdots\wedge R^1,$$

a priori defined for $\mathfrak{Re}\,\lambda \gg 0$. It turns out, see, e.g., [5] or [16], that it can be analytically continued to a neighborhood of 0, and we define $R^k\wedge\cdots\wedge R^1$ to be the value at $\lambda = 0$.

Theorem 1.2. *Let $\mu_1 > \cdots > \mu_q$ be positive integers. Then the current-valued function*

$$\lambda \mapsto R^{q,\lambda^{\mu_q}}\wedge\cdots\wedge R^{1,\lambda^{\mu_1}},$$

a priori defined for $\mathfrak{Re}\,\lambda \gg 0$, has an analytic continuation to a neighborhood of the half-axis $[0,\infty)\subset\mathbb{C}$ and the value at $\lambda = 0$ is $R^q\wedge\cdots\wedge R^1$.

To connect with Coleff–Herrera type products, let χ be the characteristic function $\mathbf{1}_{[1,\infty)}$ or a smooth regularization thereof and let

$$R^{j,\epsilon_j} := 1 - \chi(|f_j|^2_{E_j}/\epsilon_j) + \sum_{k=1}^{m_j} \bar\partial\chi(|f_j|^2_{E_j}/\epsilon_j)\wedge \frac{s_j\wedge(\bar\partial s_j)^{k-1}}{|f_j|^{2k}_{E_j}}.$$

If φ is a test form on X, then the limit of

$$\int_X R^{q,\epsilon_q}\wedge\cdots\wedge R^{1,\epsilon_1}\wedge\varphi \tag{1.5}$$

as $\epsilon\to 0$ along an admissible path exists and equals the action of $R^q\wedge\cdots\wedge R^1$ on φ, see [16].

Let us mention a version of Theorem 1.2 with connection to intersection theory. Let f be a section of E and let

$$M^\lambda := 1 - |f|^{2\lambda}_E + \sum_{k\geq 1}\bar\partial |f|^{2\lambda}_E \wedge \frac{\partial \log|f|^2_E}{2\pi i}\wedge (dd^c\log|f|^2_E)^{k-1},$$

where $dd^c = \bar\partial\partial/2\pi i$. It is shown in [3] that $\lambda\mapsto M^\lambda$ has an analytic continuation to a neighborhood of 0 and that the value at $\lambda = 0$ is a positive closed current, which we denote by M. One can give a meaning to the product $(dd^c\log|f|^2_E)^k$ for arbitrary k that extends the classical one for $k \leq \operatorname{codim}\{f=0\}$, and from [3] it follows that

$$M = \mathbf{1}_Z + \sum_{k\geq 1} \mathbf{1}_Z (dd^c \log |f|^2_E)^k,$$

where $\mathbf{1}_Z$ is the restriction to the zero set Z of f. The current M is closely connected to R. For instance, if X is smooth and D is the Chern connection on E then it follows from [2] that

$$M_k = R_k \cdot (Df/2\pi i)^k/k!,$$

where the subscript k means the component of bidegree $(*, k)$.

Let $f_1, \ldots, f_q$ be sections of Hermitian vector bundles E_j and let $M^1, \ldots, M^q$ be the associated currents. One can define products of the M^j recursively as for the R^j and we have the following analogue of Theorem 1.2.

Theorem 1.3. *Let $\mu_1 > \cdots > \mu_q$ be positive integers. Then the current-valued function*

$$\lambda \mapsto M^{q,\lambda^{\mu_q}} \wedge \cdots \wedge M^{1,\lambda^{\mu_1}},$$

a priori defined for $\mathfrak{Re}\,\lambda \gg 0$, has an analytic continuation to a neighborhood of the half-axis $[0,\infty) \subset \mathbb{C}$ and the value at $\lambda = 0$ is $M^q \wedge \cdots \wedge M^1$.

Example 1.4 (Corollary 5.6 in [3]). Let $\mathcal{J}_x \subset \mathscr{O}_{X,x}$ be an ideal and let $h_1, \ldots, h_n \in \mathcal{J}_x$ be a generic Vogel sequence of $\mathcal{J}_x$; see, e.g., [3] for the definition. By the Stückrad–Vogel procedure, [24], adapted to the local situation, [17], [25], one gets an associated Vogel cycle V^h; the multiplicities of the components of various dimensions of V^h are the Segre numbers, [14], used in excess intersection theory. By Theorem 1.3 we have that

$$\lambda \mapsto \bigwedge_{k=1}^{n} \left(1 - |h_k|^{2\lambda^{\mu_k}} + \bar\partial |h_k|^{2\lambda^{\mu_k}} \wedge \partial \log |h_k|^2/2\pi i\right)$$

is analytic at 0 and by [3] the value there is the Lelong current associated with V^h; see [3] for more details.

Remark 1.5. Assume that $\operatorname{codim} \cap_j \{f_j = 0\} = m_1 + \cdots + m_q$. Then $M^j = (dd^c \log |f_j|^2_{E_j})^{m_j} = [f_j = 0]$, where $[f_j = 0]$ is the Lelong current of the fundamental cycle of f_j, and more generally,

$$M^q \wedge \cdots \wedge M^1 = [f_q = 0] \wedge \cdots \wedge [f_1 = 0],$$

i.e., the current representing the proper intersection of the cycles $[f_j = 0]$.

In this case the current-valued function

$$(\lambda_1, \ldots, \lambda_q) \mapsto R^{q,\lambda_q} \wedge \cdots \wedge R^{1,\lambda_1}$$

has an analytic continuation to a neighborhood of the origin in $\mathbb{C}^q$, [16], and the value at $\lambda = 0$ is the R-current associated to $\oplus_j f_j$, [26]. Moreover, by [16], (1.5) depends Hölder continuously on $\epsilon \in [0,\infty)^q$ if χ is smooth. The smoothness of χ is necessary in view of the example in [21, Section 1].

2. Proof of Theorems 1.2 and 1.3

We will actually prove a slightly more general result than Theorem 1.2; we will allow mixed products of U^j and R^k. Let P^j denote either U^j or R^j and let P^{j,λ_j} be the corresponding λ-regularization, (1.3) or (1.4). One defines products of the P^j recursively as above.

Theorem 1.2′. *Let $\mu_1 > \cdots > \mu_q$ be positive integers. Then the current-valued function*

$$\lambda \mapsto P^{q,\lambda^{\mu_q}} \wedge \cdots \wedge P^{1,\lambda^{\mu_1}},$$

a priori defined for $\mathfrak{Re}\,\lambda \gg 0$, has an analytic continuation to a neighborhood of the half-axis $[0,\infty) \subset \mathbb{C}$ and the value at 0 is $P^q \wedge \cdots \wedge P^1$.

Let $\pi\colon X' \to X$ be a smooth modification of X such that $\{\pi^* f_j = 0\}$, $j = 1, \dots, q$, and $\cup_j \{\pi^* f_j = 0\}$ are normal crossings divisors. Then locally in X' we can write $\pi^* f_j = f_j^0 f_j'$, where f_j^0 is a monomial in local coordinates and f_j' is a non-vanishing holomorphic tuple. It follows that $s_j = \bar{f}_j^0 s_j'$, where s_j' is a smooth section. A straightforward computation shows that

$$\pi^* R^{j,\lambda_j} = 1 - |f_j^0|^{2\lambda_j} u_j^{2\lambda_j} + \sum_{k=1}^{m_j} \frac{\bar\partial(|f_j^0|^{2\lambda_j} u_j^{2\lambda_j})}{(f_j^0)^k} \wedge \vartheta_{jk},$$

$$\pi^* U^{j,\lambda_j} = \sum_{k=1}^{m_j} \frac{|f_j^0|^{2\lambda_j} u_j^{2\lambda_j}}{(f_j^0)^k} \wedge \vartheta_{jk},$$

where u_j is a smooth non-vanishing function and $\vartheta_{jk} = s_j' \wedge (\bar\partial s_j')^{k-1}/u_j^{2k}$ is a smooth form, cf. Remark1.1. In the same way,

$$\pi^* M^{f_j,\lambda_j} = 1 - |f_j^0|^{2\lambda_j} u_j^{2\lambda_j} + \sum_{k\ge 1} \bar\partial\big(|f_j^0|^{2\lambda_j} u_j^{2\lambda_j}\big) \wedge \partial \log(|f_j^0|^2 u_j^2) \wedge \omega_{jk},$$

where ω_{jk} is smooth, cf. [3, Section 4]. Taking the identity $\partial \log(|f_j^0|^2 u_j^2) = df_j^0/f_j^0 + 2\partial u_j/u_j$ into account, Theorems 1.2′ and 1.3 are consequences of the following quite technical lemma.

Lemma 2.1. *Let $u_1, \dots, u_r$ be smooth non-vanishing functions defined in some neighborhood of the origin in $\mathbb{C}^n$, with coordinates $x_1, \dots, x_n$.*

For $\lambda = (\lambda_1, \dots, \lambda_r) \in \mathbb{C}^r$, $\mathfrak{Re}\,\lambda_j \gg 0$, $\alpha_1, \dots, \alpha_r \in \mathbb{N}^n$, and $k_1, \dots, k_r \in \mathbb{N}$, let

$$\Gamma(\lambda) := \frac{|u_r x^{\alpha_r}|^{2\lambda_r} \cdots |u_{p+1} x^{\alpha_{p+1}}|^{2\lambda_{p+1}} \bar\partial |u_p x^{\alpha_p}|^{2\lambda_p} \wedge \cdots \wedge \bar\partial |u_1 x^{\alpha_1}|^{2\lambda_1}}{x^{k_r\alpha_r} \cdots x^{k_1\alpha_1}};$$

here $x^{k_\ell \alpha_\ell} = x_1^{k_\ell \alpha_{\ell,1}} \cdots x_n^{k_\ell \alpha_{\ell,n}}$ if $\alpha_\ell = (\alpha_{\ell,1}, \dots, \alpha_{\ell,n})$. If σ is a permutation of $\{1, \dots, r\}$, write $\Gamma^\sigma(\lambda_1, \dots, \lambda_r) := \Gamma(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})$.

Let $\mu_1, \dots, \mu_r$ be positive integers. Then $\Gamma^\sigma(\kappa^{\mu_1}, \dots, \kappa^{\mu_r})$ has an analytic continuation to a connected neighborhood of the half-axis $[0,\infty)$ in $\mathbb{C}$, and if $\mu_1 > \cdots > \mu_r$, then

$$\Gamma^\sigma(\kappa^{\mu_1}, \dots, \kappa^{\mu_r}) \,|_{\kappa=0} = \Gamma^\sigma(\lambda_1, \dots, \lambda_r) \,|_{\lambda_1=0} \cdots |_{\lambda_r=0} \,. \tag{2.1}$$

The reason for the permutation σ is that we have mixed products of U's and R's in Theorem 1.2$'$.

Proof. To begin with let us assume that all $u_j = 1$. A straightforward computation shows that

$$\Gamma(\lambda) = \lambda_1 \cdots \lambda_p \frac{\prod_{j=1}^r |x^{\alpha_j}|^{2\lambda_j}}{x^{\sum_{j=1}^r k_j \alpha_j}} \sideset{}{'}\sum_I A_I \frac{d\bar{x}_{i_1} \wedge \cdots \wedge d\bar{x}_{i_p}}{\bar{x}_{i_1} \cdots \bar{x}_{i_p}} =: \lambda_1 \cdots \lambda_p \sideset{}{'}\sum_I \Gamma_I,$$

where the sum is over all increasing multi-indices $I = \{i_1, \ldots, i_p\} \subset \{1, \ldots, n\}$ and A_I is the determinant of the matrix $(\alpha_{\ell, i_j})_{1 \le \ell \le p, 1 \le j \le p}$.

Pick a non-vanishing summand Γ_I; without loss of generality, assume that $I = \{1, \ldots, p\}$ and $A_I = 1$. With the notation $b_k(\lambda) := \sum_{\ell=1}^r \lambda_\ell \alpha_{\ell,k}$ for $1 \le k \le n$,

$$\begin{aligned}\Gamma_I &= \frac{\prod_{k=1}^n |x_k|^{2b_k(\lambda)}}{x^{\sum_{j=1}^r k_j \alpha_j}} \frac{d\bar{x}_1 \wedge \cdots \wedge d\bar{x}_p}{\bar{x}_1 \cdots \bar{x}_p} \\ &= \frac{1}{b_1(\lambda) \cdots b_p(\lambda)} \frac{\bigwedge_{k=1}^p \bar\partial |x_k|^{2b_k(\lambda)} \prod_{k=p+1}^n |x_k|^{2b_k(\lambda)}}{x^{\sum_{j=1}^r k_j \alpha_j}}.\end{aligned}$$

Now the current-valued function

$$\widetilde{\Gamma}_I : (\lambda_1, \ldots, \lambda_r) \mapsto \frac{\bigwedge_{j=1}^p \bar\partial |x_j|^{2b_j(\lambda)} \prod_{j=p+1}^n |x_j|^{2b_j(\lambda)}}{x^{\sum k_j \alpha_j}}$$

has an analytic continuation to a neighborhood of the origin in $\mathbb{C}^r$; in fact, it is a tensor product of one-variable currents. In particular, $\widetilde{\Gamma}_I(\kappa^{\mu_1}, \ldots, \kappa^{\mu_r}) \mid_{\kappa=0} = \widetilde{\Gamma}_I(\lambda) \mid_{\lambda_1=0} \cdots \mid_{\lambda_r=0}$. Let

$$\gamma(\lambda) = \frac{\lambda_1 \cdots \lambda_p}{b_1(\lambda) \cdots b_p(\lambda)}$$

and $\gamma^\sigma = \gamma(\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(r)})$. We claim that if $\mu_1 > \cdots > \mu_r$, then

$$\gamma^\sigma(\lambda) \mid_{\lambda_1=0} \cdots \mid_{\lambda_r=0} = \gamma^\sigma(\kappa^{\mu_1}, \ldots, \kappa^{\mu_r})|_{\kappa=0},$$

where it is a part of the claim that both sides make sense.

Let us prove the claim. Since $A_I = 1$, reordering the factors $b_1, \ldots, b_p$ and multiplying $\gamma(\lambda)$ by a non-zero constant, we may assume that $\alpha_{kk} = 1$, $k = 1, \ldots, p$, so that

$$\gamma(\lambda) = \frac{\lambda_1}{\lambda_1 + \alpha_{21}\lambda_2 + \cdots + \alpha_{r1}\lambda_r} \cdots \frac{\lambda_p}{\alpha_{p1}\lambda_1 + \cdots + \lambda_p + \cdots + \alpha_{rp}\lambda_r}.$$

For $j < r$ set $\tau_j := \lambda_j / \lambda_{j+1}$ and $\widetilde{\gamma}^\sigma(\tau_1, \ldots, \tau_{r-1}) := \gamma^\sigma(\lambda)$; notice that γ^σ is 0-homogeneous, so that $\widetilde{\gamma}^\sigma$ is well defined. Then $\lambda_j = \tau_j \cdots \tau_{r-1} \lambda_r$, and therefore $\widetilde{\gamma}^\sigma$ consists of p factors of the form

$$\frac{\tau_k \cdots \tau_{r-1}}{\alpha_{k1}\tau_1 \cdots \tau_{r-1} + \cdots + \tau_k \cdots \tau_{r-1} + \cdots + \alpha_{k,r-1}\tau_{r-1} + \alpha_{kr}}. \tag{2.2}$$

Observe that (2.2) is holomorphic in τ in some neighborhood of the origin. Indeed, if $\alpha_{kr} \neq 0$, then (2.2) is clearly holomorphic, whereas if $\alpha_{kr} = 0$ we can factor out τ_{r-1} from the denominator and numerator. In the latter case (2.2) is clearly holomorphic if $\alpha_{k,r-1} \neq 0$ etc; since $\alpha_{kk} = 1$ this procedure eventually stops. Hence, $\tilde{\gamma}^\sigma(\tau)$ is holomorphic in a neighborhood of 0. It follows that $\gamma^\sigma(\kappa^{\mu_1}, \ldots, \kappa^{\mu_r}) = \widetilde{\gamma}^\sigma(\kappa^{\mu_1-\mu_2}, \ldots, \kappa^{\mu_{r-1}-\mu_r})$ is holomorphic in a neighborhood of 0 and since the denominator of $\gamma^\sigma(\kappa^{\mu_1}, \ldots, \kappa^{\mu_r})$ is a polynomial in κ with non-negative coefficients it is in fact holomorphic in a neighborhood of $[0, \infty)$. Moreover, $\gamma^\sigma(\lambda_1, \ldots, \lambda_r)$ is holomorphic in $\Delta = \{|\lambda_1/\lambda_2| < \epsilon, \ldots, |\lambda_{r-1}/\lambda_r| < \epsilon\}$. Let us now fix $\lambda_2 \neq 0, \ldots, \lambda_r \neq 0$ in Δ. Then $\gamma^\sigma(\lambda)$ is holomorphic in λ_1 in a neighborhood of the origin. Next, for $\lambda_3 \neq 0, \ldots, \lambda_r \neq 0$ fixed in Δ, $\gamma^\sigma(\lambda)|_{\lambda_1=0}$ is holomorphic in λ_2 in a neighborhood of the origin, etc. It follows that

$$\gamma^\sigma(\lambda)|_{\lambda_1} \cdots |_{\lambda_r=0} = \widetilde{\gamma}^\sigma(\tau)|_{\tau=0} = \gamma^\sigma(\kappa^{\mu_1}, \ldots, \kappa^{\mu_r})|_{\kappa=0},$$

which proves the claim. Thus (2.1) follows in the case $u_j = 1$, $j = 1, \ldots, r$.

Now, consider the general case. Replace each $|u_j|^{2\lambda_j}$ in $\Gamma(\lambda)$ by $|u_j|^{2\omega_j}$, where $\omega_j \in \mathbb{C}$. Then Γ is a sum of terms of the following representative form:

$$\prod_{j=p+1}^{r} |u_j|^{2\omega_j} \prod_{j=1}^{p'} |u_j|^{2\omega_j} \bigwedge_{p'+1}^{p} \bar{\partial}|u_j|^{2\omega_j} \wedge \frac{\prod_{j=p'+1}^{r} |x^{\alpha_j}|^{2\lambda_j} \bigwedge_{j=1}^{p'} \bar{\partial}|x^{\alpha_j}|^{2\lambda_j}}{x^{k_r\alpha_r} \cdots x^{k_1\alpha_1}} \tag{2.3}$$

Fixing all λ_j and ω_j except for $\lambda_{\sigma(1)}$ and $\omega_{\sigma(1)}$, (2.3) becomes an analytic (current-valued) function $g(\lambda_{\sigma(1)}, \omega_{\sigma(1)})$ in a neighborhood of $0 \in \mathbb{C}^2$. Thus, the value at 0 of $g(\lambda_{\sigma(1)}, \lambda_{\sigma(1)})$ is the same as first letting $\omega_{\sigma(1)} = 0$ (which corresponds to setting $u_{\sigma(1)} = 1$) and then letting $\lambda_{\sigma(1)} = 0$ in $g(\lambda_{\sigma(1)}, \omega_{\sigma(1)})$. Continuing analogously for $(\lambda_{\sigma(2)}, \omega_{\sigma(2)})$ and so on, it follows that the right-hand side of (2.1) is independent of the u_j.

To see that the left-hand side of (2.1) is independent of u_j, replace each λ_j in (2.3) by $\kappa^{\mu_{\sigma(j)}}$ and denote the resulting expression by $\tilde{g}(\kappa, \omega_1, \ldots, \omega_r)$. Then $\tilde{g}$ is clearly analytic in the ω_j and by the first part of the proof it is also analytic in a neighborhood of $[0, \infty) \subset \mathbb{C}_\kappa$. Hence, $\tilde{g}$ is analytic in a neighborhood of $0 \in \mathbb{C}^{r+1}$. The left-hand side of (2.1) is obtained by evaluating $\kappa \mapsto \tilde{g}(\kappa, \kappa^{\mu_{\sigma(1)}}, \ldots, \kappa^{\mu_{\sigma(r)}})$ at $\kappa = 0$; this is thus the same as evaluating $\tilde{g}(\kappa, 0)$ (which corresponds to setting all $u_j = 1$) at $\kappa = 0$. Hence also the left-hand side of (2.1) is independent of the u_j and the lemma follows. □

References

[1] M. Andersson: *Residue currents and ideals of holomorphic functions*, Bull. Sci. math. **128** (2004), 481–512.

[2] M. Andersson: *Residues of holomorphic sections and Lelong currents*, Ark. Mat. **43** (2005), 201–219.

[3] M. Andersson, H. Samuelsson Kalm, E. Wulcan, A. Yger: *Segre numbers, a generalized King formula, and local intersections*, J. reine angew. Math., Ahead of Print. http://dx.doi.org/10.1515/crelle-2014-0109.

[4] M. Andersson, D. Eriksson, H. Samuelsson, E. Wulcan, A. Yger: *Non-proper intersections products, Segre numbers and generalized cycles*, in preparation.

[5] M. Andersson, E. Wulcan: *Decomposition of residue currents*, J. reine angew. Math. **638** (2010), 103–118.

[6] M. F. Atiyah: *Resolution of singularities and division of distributions*, Commun. Pure Appl. Math. **23** (1970), 145–150.

[7] D. Barlet: *Contribution effective de la monodromie aux développements asymptotiques*, Ann. Sci. École Norm. Sup. (4) **17(2)** (1984), 293–315.

[8] Berenstein, C. A., Yger, A.: *Analytic residue theory in the non-complete intersection case*, J. reine angew. Math. **527** (2000), 203–235.

[9] C. A. Berenstein, R. Gay, A. Vidras, A. Yger: *Residue currents and Bezout identities*, Progress in Mathematics 114. Birkhäuser Verlag, Basel, 1993.

[10] I. N. Bernstein, S. I. Gelfand: *Meromorphy of the function* P^{λ}, Funkcional. Anal. i Priložen **3(1)** (1969), 84–85.

[11] J.-E. Björk, H. Samuelsson: *Regularizations of residue currents*, J. reine angew. Math. **649** (2010), 33–54.

[12] N. R. Coleff, M. E. Herrera: *Les courants résiduels associés à une forme meromorphe*, Lecture Notes in Mathematics, 633, Springer, Berlin, 1978.

[13] A. Dickenstein, C. Sessa: *Canonical representatives in moderate cohomology*, Invent. Math. **80(3)** (1985), 417–434.

[14] T. Gaffney, R. Gassler: *Segre numbers and hypersurface singularities*, J. Algebraic Geom. 8 (1999), 695–736.

[15] M. Herrera, D. Lieberman: *Residues and principal values on complex spaces*, Math. Ann. **194** (1971), 259–294.

[16] R. Lärkäng, H. Samuelsson Kalm: *Various approaches to products of residue currents*, J. Funct. Anal. **264(1)** (2013), 118–138.

[17] D. Massey: *Lê cycles and hypersurface singularities*, Lecture Notes in Mathematics 1615, Springer-Verlag, Berlin (1995), xii+131 pp..

[18] M. Passare: *Residues, currents, and their relation to ideals of holomorphic functions*, Math. Scand. **62(1))** (1988), 75–152.

[19] M. Passare: *A calculus for meromorphic currents*, J. reine angew. Math. **392)** (1988), 37–56.

[20] M. Passare, A. Tsikh: *Residue integrals and their Mellin transforms*, Canad. J. Math. **47(5)** (1995), 1037–1050.

[21] M. Passare, A. Tsikh: *Defining the residue of a complete intersection*, *Complex analysis, harmonic analysis and applications* (*Bordeaux,* 1995), 250–267, Pitman Res. Notes Math. Ser., 347, *Longman, Harlow*, 1996.

[22] M. Passare, A. Tsikh, A. Yger: *Residue currents of the Bochner–Martinelli type*, Publ. Mat. **44** (2000), no. 1, 85–117.

[23] H. Samuelsson: *Analytic continuation of residue currents*, Ark. Mat. **47** (2009), no. 1, 127–141.

[24] J. Stückrad, W. Vogel: *An algebraic approach to the intersection theory*, Queen's Papers in Pure and Appl. Math. 61 (1982), 1–32.

[25] P. Tworzewski: *Intersection theory in complex analytic geometry*, Ann. Polon. Math. **62** (1995), 177–191.

[26] E. Wulcan: *Products of residue currents of Cauchy–Fantappiè–Leray type*, Ark. Mat. **45** (2007), 157–178.

[27] A. Yger: *Formules de division et prolongement méromorphe*, *Séminaire d'Analyse P. Lelong–P. Dolbeault–H. Skoda, Années* 1985/1986, 226–283, Lecture Notes in Math., 1295, *Springer, Berlin*, 1987.

Mats Andersson, Håkan Samuelsson Kalm, and Elizabeth Wulcan
Department of Mathematical Sciences, Division of Mathematics
University of Gothenburg and Chalmers University of Technology
SE-412 96 Göteborg, Sweden

e-mail: matsa@chalmers.se
hasam@chalmers.se
wulcan@chalmers.se

Alain Yger
Institut de Mathématiques
Université Bordeaux 1
F-33405 Talence, France

e-mail: Alain.Yger@math.u-bordeaux1.fr

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 91–109

On the Effective Membership Problem for Polynomial Ideals

Mats Andersson and Elizabeth Wulcan

Dedicated to the memory of Mikael Passare

Abstract. We discuss the possibility of representing elements in polynomial ideals in $\mathbb{C}^N$ with optimal degree bounds. Classical theorems due to Macaulay and Max Noether say that such a representation is possible under certain conditions on the variety of the associated homogeneous ideal. We present some variants of these results, as well as generalizations to subvarieties of $\mathbb{C}^N$.

1. Introduction

Let V be an algebraic subvariety of $\mathbb{C}^N$ of pure dimension n and let $F_1, \dots, F_m$ be polynomials in $\mathbb{C}^N$. We are interested in finding solutions to the polynomial division problem

$$F_1 Q_1 + \cdots + F_m Q_m = \Phi \tag{1.1}$$

on V with degree estimates, provided Φ is in the ideal (F_j) on V. By a result of Hermann, [18], if $\deg F_j \le d$, there are polynomials Q_j such that $\deg(F_j Q_j) \le \deg \Phi + C(d, N)$, where $C(d, N)$ is like $2(2d)^{2^N - 1}$ for large d and thus doubly exponential. It is shown in [24] (see also [10, Example 3.9]) that in general this estimate cannot be substantially improved.

If one imposes conditions on V and F_j one can, however, obtain much sharper estimates. The following two results in $\mathbb{C}^n$ are classical.

If $F_1, \dots, F_m$ are polynomials in $\mathbb{C}^n$ of degrees $d_1 \ge \cdots \ge d_m$ with no common zeros even at infinity and Φ is any polynomial, then one can solve (1.1) *with* $\deg(F_j Q_j) \le \max(\deg \Phi, d_1 + \cdots + d_{n+1} - n)$.

If $F_1, \dots, F_n$ are polynomials in $\mathbb{C}^n$ such that their common zero set is discrete and does not intersect the hyperplane at infinity, and Φ belongs to the ideal (F_j), then one can find polynomials Q_j such that (1.1) *holds and* $\deg(F_j Q_j) \le \deg \Phi$.

The first author was partially supported by the Swedish Research Council. The second author was partially supported by the Swedish Research Council and by the NSF.

The first theorem is due to Macaulay, [23], and the second one is Max Noether's AF+BG theorem, [25], originally stated for $n = 2$. Noether's result is clearly optimal.

In this paper we present extensions of these results to the case of more general varieties $V \subset \mathbb{C}^N$, and also generalizations in which we relax the condition on (the zero set of) the F_j. It grew out of our paper [9], in which we extended to the singular setting a framework for solving polynomial ideal membership problems with residue techniques introduced in [3] and further developed in [5, 30, 31], see below. The proofs in this paper follow the same setup. However, at least some of the results also admit algebraic proofs, see Remark 6.2.

Throughout we will let X denote the closure of V in $\mathbb{P}^N$, and $\operatorname{reg} X$ the *regularity* of X, see Section 4 for the definition. For each F_j we let f_j denote the induced section of $\mathcal{O}(\deg F_j)|_X$.

We begin with an extension of Macaulay's theorem to singular varieties; this can easily be proved by standard arguments, cf. Remark 6.2.

Theorem 1.1. *Let V be an algebraic subvariety of $\mathbb{C}^N$, with closure X in $\mathbb{P}^N$, and let $F_1, \ldots, F_m$ be polynomials in $\mathbb{C}^N$ of degrees $d_1 \geq \cdots \geq d_m$. Assume that f_j have no common zeros on X. Then for each polynomial Φ in $\mathbb{C}^N$ there are polynomials Q_j such that* (1.1) *holds and*

$$\deg(F_j Q_j) \leq \max(\deg \Phi, d_1 + \cdots + d_{n+1} - (n+1) + \operatorname{reg} X).$$

If X is smooth, then $\operatorname{reg} X \leq (n+1)(\deg X - 1) + 1$; this is Mumford's bound, see, e.g., [22, Example 1.8.48]. If X is Cohen–Macaulay in $\mathbb{P}^N$ (and N is minimal) then $\operatorname{reg} X \leq \deg X - (N - n)$, see, [17, Corollary 4.15]. In particular, if $V = \mathbb{C}^n$ so that $X = \mathbb{P}^n$, then $\operatorname{reg} X = 1$; thus we get back Macaulay's theorem. For a discussion of bounds on $\operatorname{reg} X$ for a general X, see, e.g., [10, Section 3].

Let Z^f denote the common zero set of $f_1, \ldots, f_m$ in X. Moreover, let $X_\infty := X \setminus V$. For smooth varieties we have the following version of Max Noether's theorem.

Theorem 1.2. *Let V be an algebraic subvariety of $\mathbb{C}^N$ of dimension n such that its closure X in $\mathbb{P}^N$ is smooth, and let $F_1, \ldots, F_m$ be polynomials in $\mathbb{C}^N$ of degrees $d_1 \geq \cdots \geq d_m$. Assume that $m \leq n$, that*

$$\operatorname{codim}(Z^f \cap V) \geq m, \tag{1.2}$$

and that Z^f has no irreducible component contained in X_∞. If Φ is a polynomial that belongs to the ideal (F_j) on V, then there is a representation (1.1) *with*

$$\deg(F_j Q_j) \leq \max(\deg \Phi, d_1 + \cdots + d_m - m + \operatorname{reg} X). \tag{1.3}$$

If in addition X is Cohen–Macaulay in $\mathbb{P}^N$ one can choose Q_j so that

$$\deg(F_j Q_j) \leq \deg \Phi. \tag{1.4}$$

Remark 1.3. If X is Cohen–Macaulay it suffices that V is smooth to obtain (1.4). $\square$

For $V = \mathbb{C}^n$ Theorem 1.2 appeared in [3, Theorem 1.2].

For a general X, in order to have a Max Noether theorem, we need the common zero set of the f_j not to intersect the singular locus of X too badly. To make this statement more precise we need to introduce what we call the *intrinsic BEF-varieties*

$$X^{n-1} \subset \cdots \subset X^1$$

of $X \subset \mathbb{P}^N$. These are the sets where the mappings in a locally free resolution of $\mathcal{O}^{\mathbb{P}^N}/\mathcal{J}_X$ do not have optimal rank. They are intrinsically defined subvarieties of X that are contained in $X^0 := X_{\text{sing}}$. The codimension of X^ℓ is at least $\ell + 1$, and if X is locally Cohen–Macaulay X^ℓ is empty for $\ell \geq 1$, see Sections 2.3 and 2.5.

Theorem 1.4. *Let V be an algebraic subvariety of $\mathbb{C}^N$ of dimension n, with closure X in $\mathbb{P}^N$, and let F_j be as in Theorem* 1.2. *Assume that Z^f satisfies* (1.2), *that Z^f has no irreducible component contained in X_∞, and moreover that*

$$\operatorname{codim}(Z^f \cap X^\ell) \geq m + \ell + 1, \qquad \ell \geq 0. \tag{1.5}$$

If Φ is a polynomial that belongs to the ideal (F_j) on V, then there is a representation (1.1) *such that* (1.3) *holds. If in addition X is Cohen–Macaulay in $\mathbb{P}^N$, and $m \leq n$, we can choose Q_j such that* (1.4) *holds.*

Notice that (1.5) forces that either $Z^f \cap X_{\text{sing}} = \emptyset$ or $m < n$. If X is smooth, then (1.5) is vacuous, and thus Theorem 1.2 follows immediately from Theorem 1.4. If only V is smooth but X is Cohen–Macaulay, then by the assumption on Z^f $\operatorname{codim}(Z^f \cap X_\infty) \geq m + 1$ and since $X^0 \subset X_\infty$, (1.5) is satisfied. This proves the claim in Remark 1.3.

Next we will present some generalizations of Theorem 1.4 where we relax the hypotheses on the common zero set Z^f of the f_j. First, we drop the size hypothesis (1.2) on $Z^f \cap V$. We then still get an estimate of the form (1.3) but the second entry on the right-hand side is now replaced by a constant that depends on F_j in a more involved manner. The condition that Z^f has no irreducible component at infinity should now be understood as that the ideal sheaf $\mathcal{J}_f$ over X generated by the sections $f_1, \ldots, f_m$ has no associated variety, in the sense of [28], contained in X_∞, see Section 3. This means that at each $x \in X_\infty$, $(\mathcal{J}_f)_x$ has no (varieties of) associated prime ideals contained in X_∞. Let J_f be the homogeneous ideal in $\mathbb{C}[z_0, \ldots, z_N]$ associated with $\mathcal{J}_f$, and let $\operatorname{reg} J_f$ be the *regularity* of J_f, cf. Section 4.

Theorem 1.5. *Let V be an algebraic subvariety of $\mathbb{C}^N$, with closure X in $\mathbb{P}^N$, and let $F_1, \ldots, F_m$ be polynomials in $\mathbb{C}^N$. Assume that $\mathcal{J}_f$ has no associated variety contained in X_∞. Then there is a constant $\beta = \beta(X, F_1, \ldots, F_m)$ such that if $\Phi \in (F_j)$, then there is a representation* (1.1) *on V with*

$$\deg(F_j Q_j) \leq \max(\deg \Phi, \beta). \tag{1.6}$$

If $V = \mathbb{C}^N$, one can take $\beta = \operatorname{reg} J_f$.

Conversely, if there is an associated prime of $\mathcal{J}_f$ contained in X_∞, then there is no β such that one can solve (1.1) *with* (1.6) *for all Φ in (F_j).*

In [27] Shiffman computed the regularity of a zero-dimensional homogeneous polynomial ideal J_f to be $\leq d_1 + \cdots + d_{n+1} - n$. Using this he obtained (the first part of) Theorem 1.5 for $V = \mathbb{C}^N$ and $\dim Z^f = 0$ with $\beta = \operatorname{reg} J_f = d_1 + \cdots + d_{n+1} - n$, i.e., the same bound as in Macaulay's theorem, see [27, Theorem 2(iv)]. Theorem 1.5 can thus be seen as a generalization of Shiffman's result.

The estimate (1.6) is clearly sharp if $\deg \Phi \geq \beta$. If the ideal sheaf $\mathcal{J}_f$ is locally Cohen–Macaulay, for instance locally a complete intersection, then there are no embedded primes of $\mathcal{J}_f$, and so the hypothesis that $\mathcal{J}_f$ has no associated variety at infinity just means that no irreducible component of Z^f is contained in X_∞. Thus we get back the hypothesis in Theorems 1.2 and 1.4.

Next, let us instead relax the condition that Z^f has no irreducible components at infinity. If the degrees of F_j are $\leq d$, we let $\tilde{f}_j$ denote the section of $\mathcal{O}(d)|_X$ corresponding to F_j. We let $Z^{\tilde{f}}$ be the common zero set of $\tilde{f}_1, \ldots, \tilde{f}_m$ and $\mathcal{J}_{\tilde{f}}$ the coherent analytic sheaf over X generated by the $\tilde{f}_j$. Moreover, we let c_∞ be the maximal codimension of the so-called (*Fulton–MacPherson*) *distinguished varieties* of $\mathcal{J}_{\tilde{f}}$ that are contained in X_∞, see Section 5.1. If there are no distinguished varieties contained in X_∞, then we interpret c_∞ as $-\infty$. Note that it is not sufficient that $Z^{\tilde{f}} \cap V = Z^{\tilde{f}}$, since there may be embedded distinguished varieties contained in X_∞. It is well known that the codimension of a distinguished variety cannot exceed the number m, see, e.g., Proposition 2.6 in [15], and thus $c_\infty \leq \mu$, where

$$\mu := \min(m, n).$$

Theorem 1.6. *Let V be an algebraic subvariety of $\mathbb{C}^N$, with closure X in $\mathbb{P}^N$, and let $F_1, \ldots, F_m$ be polynomials in $\mathbb{C}^N$ of degree $\leq d$. Assume that $Z^{\tilde{f}}$ satisfies*

$$\operatorname{codim}(Z^{\tilde{f}} \cap X) \geq m \tag{1.7}$$

and

$$\operatorname{codim}(Z^{\tilde{f}} \cap X^\ell) \geq m + \ell + 1, \qquad \ell \geq 0. \tag{1.8}$$

If Φ is a polynomial that belongs to (F_j) on V, then there is a representation (1.1) *on V with*

$$\deg(F_j Q_j) \leq \max(\deg \Phi + \mu d^{c_\infty} \deg X, (d-1)\min(m, n+1) + \operatorname{reg} X). \tag{1.9}$$

If in addition X is locally Cohen–Macaulay in $\mathbb{P}^N$ and $m \leq n$, then we can choose Q_j such that

$$\deg(F_j Q_j) \leq \deg \Phi + m d^{c_\infty} \deg X.$$

Note that for most choices of F_j and Φ the first entry in (1.9) is much larger than the second entry. For instance this is true for all Φ if $c_\infty \geq 2$ and d is large enough. In particular, if $X = \mathbb{P}^n$, so that $\operatorname{reg} X = 1$, and $c_\infty \geq 2$, the first entry is the largest for all d.

For $X = \mathbb{P}^n$ Theorem 1.6 is due to the first author and Götmark, [5, Theorem 1.3]. In the case when $\deg F_j = d$, so that $\tilde{f}_j = f_j$, Theorem 1.6 generalizes Theorems 1.1–1.4, see Remark 6.3.

Example 1.7. If the F_j have no common zeros on V, then Theorem 1.6 gives a solution to

$$F_1 Q_1 + \cdots + F_m Q_m = 1$$

with $\deg(F_j Q_j) \le \mu d^\mu \deg X$ if d is large enough. Except for the annoying factor μ we then get back is Jelonek's optimal effective Nullstellensatz, [20]. □

Note that the estimates of $\deg(F_j Q_j)$ in the theorems above hold for representations of *all* Φ in (F_j). If one, instead of adding conditions on V and F_j, imposes further conditions on Φ, then Hermann's degree estimate for solutions to (1.1) can also be essentially improved. Theorem 1.1 in our recent paper [9] asserts that for any $V \subset \mathbb{C}^N$ there is a number μ_0 such that if $F_1, \ldots, F_m$ are polynomials in $\mathbb{C}^N$ of degree $\le d$ and Φ is a polynomial such that $|\Phi| \le C|F|^{\mu+\mu_0}$ locally on V, where $|F|^2 = |F_1|^2 + \cdots + |F_m|^2$, then one can solve (1.1) with

$$\deg(F_j Q_j) \le \max\big(\deg \Phi + (\mu + \mu_0) d^{c_\infty} \deg X, (d-1)\min(m, n+1) + \operatorname{reg} X\big). \tag{1.10}$$

The statement that $|\Phi| \le C|F|^{\mu+\mu_0}$ implies that there is a representation (1.1) is a direct consequence of Huneke's uniform Briançon–Skoda theorem, [12, 19], and thus the degree estimate (1.10) can be seen as a global effective Briançon–Skoda–Huneke theorem.

2. Residue currents

We will briefly recall some residue theory. For more details we refer to [9] and the references therein.

2.1. Currents on a singular variety

If nothing else is mentioned X will be a reduced subvariety of $\mathbb{P}^N$ of pure dimension n. The sheaf $\mathcal{C}_{\ell,k}$ of currents of bidegree (ℓ, k) on X is by definition the dual of the sheaf $\mathcal{E}_{n-\ell,n-k}$ of smooth $(n-\ell, n-k)$-forms on X. If $i \colon X \to \mathbb{P}^N$ is an embedding of X, then $\mathcal{E}_{n-\ell,n-k}$ can be identified with the quotient sheaf $\mathcal{E}^{\mathbb{P}^N}_{n-\ell,n-k}/\operatorname{Ker} i^*$, where $\operatorname{Ker} i^*$ is the sheaf of forms ξ on $\mathbb{P}^N$ such that $i^*\xi$ vanish on X_{reg}. It follows that the currents τ in $\mathcal{C}_{\ell,k}$ can be identified with currents $\tau' = i_*\tau$ on $\mathbb{P}^N$ of bidegree $(N-n+\ell, N-n+k)$ that vanish on $\operatorname{Ker} i^*$.

Given a holomorphic function f on X, we write $1/f$ for the *principal value distribution*, defined for instance as $\lim_{\epsilon\to 0} \chi(|f|^2/\epsilon)(1/f)$, where $\chi(t)$ is the characteristic function of the interval $[1, \infty)$ or a smooth approximand of it, or as the analytic continuation of $\lambda \to |f|^{2\lambda}(1/f)$ to $\lambda = 0$. It is readily checked that $f(1/f) = 1$

as distributions and that the *residue current* $\bar\partial(1/f)$ satisfies $f\bar\partial(1/f)=0$. We will need the fact that

$$v^\lambda |f|^{2\lambda}\, \frac{1}{f}\Big|_{\lambda=0} = \frac{1}{f} \tag{2.1}$$

if v is a strictly positive smooth function; cf. [1, Lemma 2.1].

2.2. Pseudomeromorphic currents

The notion of pseudomeromorphic currents on manifolds was introduced in [8]. A slightly extended version appeared in [6]: A current on X is *pseudomeromorphic* if it is (the sum of terms that are) the pushforward under (a composition of) modifications, projections, and open inclusions of currents of the form

$$\frac{\xi}{s_1^{\alpha_1}\cdots s_{n-1}^{\alpha_{n-1}}}\wedge\bar\partial\frac{1}{s_n^{\alpha_n}},$$

where s is a local coordinate system and ξ is a smooth form with compact support, see, e.g., [6] for details.

Pseudomeromorphic currents in many respects behave like positive closed currents. For example they satisfy the *dimension principle: If τ is a pseudomeromorphic current on X of bidegree $(*,p)$ that has support on a variety of codimension $>p$, then $\tau=0$.*

Also, pseudomeromorphic currents allow for multiplication with characteristic functions of constructible sets so that ordinary computational rules hold. If τ is a pseudomeromorphic current on X and V is a subvariety of X, then the natural restriction of τ to the open set $X\setminus V$ has a canonical extension $\mathbf{1}_{X\setminus V}\tau := |h|^{2\lambda}\tau|_{\lambda=0}$, where h is any holomorphic tuple such that $\{h=0\}=V$. It follows that $\mathbf{1}_V\tau := \tau-\mathbf{1}_{X\setminus V}\tau$ is a pseudomeromorphic current with support on V. Note that if α is a smooth form, then $\mathbf{1}_V\alpha\wedge\tau=\alpha\wedge\mathbf{1}_V\tau$ and if W are W' are constructible sets, then

$$\mathbf{1}_W\mathbf{1}_{W'}\tau=\mathbf{1}_{W\cap W'}\tau. \tag{2.2}$$

Moreover, if $\pi\colon \widetilde X\to X$ is a modification, $\tilde\tau$ is a pseudomeromorphic current on $\widetilde X$, and $\tau=\pi_*\tilde\tau$, then

$$\mathbf{1}_V\tau=\pi_*\big(\mathbf{1}_{\pi^{-1}V}\tilde\tau\big) \tag{2.3}$$

for any subvariety $V\subset X$. If W is a subvariety of X and $\mathbf{1}_V\tau=0$ for all subvarieties $V\subset W$ of positive codimension we say that τ has the *the standard extension property*, SEP with respect to W, see [11].

Recall that a current is *semi-meromorphic* if it is the quotient of a smooth form and a holomorphic function. Following [6] we say that a current τ is *almost semi-meromorphic* in X if there is a modification $\pi\colon\widetilde X\to X$ and a semi-meromorphic current $\tilde\tau$ such that $\tau=\pi_*\tilde\tau$.

2.3. Residue currents associated with Hermitian complexes

Consider a complex of Hermitian holomorphic vector bundles over a complex manifold Y of dimension n,

$$0 \to E_M \xrightarrow{f^M} \cdots \xrightarrow{f^3} E_2 \xrightarrow{f^2} E_1 \xrightarrow{f^1} E_0 \to 0, \tag{2.4}$$

that is pointwise exact outside an analytic variety $Z \subset Y$ of positive codimension p. Suppose that the rank of E_0 is 1. In [2, 7] was associated to (2.4) a $\bigoplus \mathrm{Hom}\,(E_0, E_k)$-valued pseudomeromorphic current $R = R^f$; it has support on Z and in a certain sense it measures the lack of exactness of the associated sheaf complex of holomorphic sections

$$0 \to \mathcal{O}(E_M) \xrightarrow{f^M} \cdots \xrightarrow{f^3} \mathcal{O}(E_2) \xrightarrow{f^2} \mathcal{O}(E_1) \xrightarrow{f^1} \mathcal{O}(E_0). \tag{2.5}$$

Proposition 2.1. *If ϕ is a holomorphic section of E_0 such that $R\phi = 0$, then $\phi \in \mathrm{Im}\, f^1$. Moreover, if*

$$H^{k-1}(Y, \mathcal{O}(E_k)) = 0, \quad 1 \le k \le \min(M, n+1), \tag{2.6}$$

then there is a global holomorphic section q of E_1 such that $f^1 q = \phi$.

We also have the *duality principle: If* (2.5) *is exact, i.e., if it is a locally free resolution of the sheaf $\mathcal{O}(E_0)/\mathrm{Im}\, f^1$, then $R\phi = 0$ if and only if $\phi \in \mathrm{Im}\, f^1$.*

As in [9] we will refer to a (locally) free resolution (2.5) of $\mathcal{O}(E_0)/\mathcal{J}$ together with Hermitian metrics on the corresponding vector bundles as a *Hermitian (locally) free resolution.*

Let us look at the construction of R in a special case; see, e.g., [9] for more details and the general case. Let R_k denote the component of R that takes values in $\mathrm{Hom}\,(E_0, E_k)$.

Example 2.2 (The Koszul complex). Given Hermitian line bundles $S \to Y$ and $L_1, \dots, L_m \to Y$ and a tuple f of holomorphic sections $f_1, \dots, f_m$ of $L_1, \dots, L_m$, respectively, let (2.4) be the (twisted) Koszul complex of f: Let E^j be disjoint trivial line bundles with basis elements e_j, let $E = L_1^{-1} \otimes E^1 \oplus \cdots \oplus L_m^{-1} \otimes E^m$, and identify f with a section $f = \sum f_j e_j^*$ of E^*, where e_j^* are the dual basis elements. Moreover, let

$$E_0 = S, \quad E_k = S \otimes \Lambda^k E,$$

and let all f^k in (2.4) be interior multiplication δ_f by the section f.

The current associated with the Koszul complex was introduced in [1]; we will briefly recall the construction. Let σ be the section of E over $Y \setminus Z$ with pointwise minimal norm such that $f \cdot \sigma = \delta_f \sigma = 1$, i.e.,

$$\sigma = \sum_j \frac{f_j^* e_j}{|f|^2},$$

where f_j^* is the section of L_j^{-1} of minimal norm such that $f_j f_j^* = |f_j|^2_{L_j}$, and $|f|^2 = |f_1|^2_{L_1} + \cdots + |f_m|^2_{L_m}$. Then R_k equals the analytic continuation to $\lambda = 0$ of

$$R_k^\lambda = R_k^{f,\lambda} := \bar\partial |f|^{2\lambda} \wedge \sigma \wedge (\bar\partial\sigma)^{k-1}. \tag{2.7}$$

Here the exterior product is with respect to the exterior algebra over $E \oplus T^*(Y)$ so that $d\bar z_j \wedge e_\ell = -e_\ell \wedge d\bar z_j$ etc; in particular, $\bar\partial\sigma$ is a form of even degree.

If $m = 1$, then σ is just $(1/f_1)e_1$ and $R = \bar\partial(1/f_1) \wedge e_1$. In general, the coefficients of R are the Bochner–Martinelli residue currents introduced by Passare–Tsikh–Yger [26]. The sheaf complex associated with the Koszul complex is exact if and only if f is a *complete intersection*, i.e., $\operatorname{codim} Z^f = m$. In this case one can prove that (the coefficient of) $R = R_m$ coincides with the classical *Coleff–Herrera residue current* $\bar\partial(1/f_1) \wedge \cdots \wedge \bar\partial(1/f_m)$. □

Since, in light of the above example, R generalizes the classical Coleff–Herrera residue current (as well as the Bochner–Martinelli residue currents), we say that R is the *residue current* associated with the Hermitian complex (2.4).

The construction of R in general involves the minimal inverse σ_k of each f^k in (2.4); R is defined as the analytic continuation to $\lambda = 0$ of a regularization R^λ which generalizes (2.7). The component R_k is of the form $\bar\partial|f|^{2\lambda} \wedge \sigma_k \bar\partial\sigma_{k-1} \cdots \bar\partial\sigma_1|_{\lambda=0}$; see, e.g., [7] for a precise interpretation of this. It follows that outside the set Z_k where f^k does not have optimal rank,

$$R_k = \alpha_k R_{k-1}, \tag{2.8}$$

where α_k is a smooth $\operatorname{Hom}(E_{k-1}, E_k)$-valued $(0,1)$-form. If (2.5) is exact, these sets are independent of the resolution; we call them *BEF varieties* (which is an acronym for Buchsbaum–Eisenbud–Fitting, cf. [9]) and denote them $Z_k^{\rm bef} = Z_k^{\rm bef}(\mathcal{J}_f)$. The Buchsbaum–Eisenbud theorem asserts that $\operatorname{codim} Z_k^{\rm bef} \geq k$; more precisely it says that the complex (2.5) is exact if and only if the codimension of the set where f_k does not have optimal rank is $\geq k$, see, e.g., [17, Theorem 3.3]. If $\mathcal{J}_f$ has pure codimension p, then $\operatorname{codim} Z_k^{\rm bef} \geq k+1$ for $k > p$, see [16, Corollary 20.14]. Also, note that if in addition X is locally Cohen–Macaulay, then $Z_k = \emptyset$ for $k > p$. The current R_k has bidegree $(0,k)$, and thus, by the dimension principle, $R_k = 0$ for $k < p$, and for degree reasons, $R_k = 0$ for $k > n$.

If the complex (2.4) is twisted by a Hermitian line bundle, the residue current R is not affected. This follows since the σ_k are not affected by the twisting.

2.4. BEF-varieties on singular varieties

Let $i : X \to Y$ be a (local) embedding of X of dimension n into a smooth manifold Y of dimension N. Note that if $\mathcal{J}_f$ is a coherent ideal sheaf on X, then $\mathcal{J}_f + \mathcal{J}_X$ is a well-defined sheaf on Y. Indeed, locally $\mathcal{J}_f$ is the pullback $i^*\widetilde{\mathcal{J}}_f$ of an ideal sheaf on Y and the sheaf $\widetilde{\mathcal{J}}_f + \mathcal{J}_X$ is independent of the choice of $\widetilde{\mathcal{J}}_f$. We define *$k$th BEF-variety* $Z_k^{\rm bef}(\mathcal{J}_f)$ of $\mathcal{J}_f$ as $Z_{k+N-n}^{\rm bef}(\mathcal{J}_f + \mathcal{J}_X)$, which clearly is a subvariety of X.

This definition is independent of the embedding i. To see this recall that (locally) i can be factorized as $X \xrightarrow{\iota} \Omega \to \Omega \times \mathbb{C}^r = Y$, where ι is a minimal embedding. From a locally free resolution of $\mathcal{O}^\Omega/\mathcal{J}$, where $\mathcal{J}$ is a coherent ideal sheaf over Ω, it is not hard to construct a locally free resolution of $\mathcal{O}^Y/(\mathcal{J}+\mathcal{J}_\Omega)$. By relating the sets where the mappings in these resolutions do not have optimal rank one can show that the BEF-varieties of $\mathcal{J}$ are independent of i, cf. [4, Remark 4.6] and [9, Section 3].

2.5. The structure form ω on a singular variety

Now assume that X is as in Section 2.1, and let R be the residue current associated with a Hermitian free resolution $\mathcal{O}(E_\bullet), g^\bullet$ of the sheaf $\mathcal{J}_X$ of X, and let $\varOmega$ be a global nonvanishing $(\dim \mathbb{P}^N, 0)$-form with values in $\mathcal{O}(N+1)$. It was shown in [6, Proposition 3.3] that there is a (unique) almost semi-meromorphic current $\omega = \omega_0 + \cdots + \omega_{n-1}$ on X, that is smooth on $X_{\rm reg}$ and such that

$$i_*\omega = R \wedge \varOmega.$$

We say that ω is a *structure form* on X. Let E^ℓ denote the restriction of $E_{N-n+\ell}$ to X. Then the component ω_ℓ is an (n,ℓ)-form taking values in $\mathrm{Hom}\,(E^0, E^\ell)$. Moreover, let $X^0 = X_{\rm sing}$ and $X^\ell = X_{N-n+\ell}$, where X_j are the BEF-varieties of $\mathcal{J}_X$. In the language of the previous section X^ℓ is the ℓth BEF-variety of the zero sheaf. It follows from that section that the X^ℓ are independent of the embedding $i: X \to Y$ of X into a smooth manifold Y; we therefore call them the *intrinsic BEF-varieties of* X. In light of (2.8) there are almost semi-meromorphic forms α^ℓ, smooth outside X^ℓ, such that

$$\omega_\ell = \alpha^\ell \omega_{\ell-1}. \tag{2.9}$$

on X.

3. Gap sheaves and primary decomposition of sheaves

Recall that any ideal $\mathfrak{a}$ in a Noetherian ring A admits a *primary decomposition* (or *Noether–Lasker decomposition*), i.e., it can be written as $\mathfrak{a} = \bigcap \mathfrak{a}_k$, where $\mathfrak{a}_k$ is $\mathfrak{p}_k$-primary ($ab \in \mathfrak{a}_k$ implies $a \in \mathfrak{a}_k$ or $b^s \in \mathfrak{a}_k$ for some s and $\sqrt{\mathfrak{a}_k} = \mathfrak{p}_k$) for some prime ideal $\mathfrak{p}_k$. The primes in a minimal such decomposition are called the *associated primes* of $\mathfrak{a}$ and the set $\mathrm{Ass}(\mathfrak{a})$ of associated primes is independent of the primary decomposition.

Given a coherent subsheaf $\mathcal{J}$ of $\mathcal{O}^X$, Siu [28] gave a way of defining a "global" primary decomposition. Let us briefly recall his construction. First, for $p = 0, 1, \ldots, \dim X$, let $\mathcal{J}_{[p]} \supset \mathcal{J}$ be the pth *gap sheaf* (*Lückergarbe*), introduced by Thimm [29]: A germ $s \in \mathcal{O}_x$ is in $(\mathcal{J}_{[p]})_x$ if and only if there is a neighborhood U of x and a section $t \in \mathcal{J}(U)$ such that $s_x = t_x$ and $t_y \in \mathcal{J}_y$ for all $y \in U$ outside an analytic set of dimension at most p. It is not hard to see that $\mathcal{J}_{[p]}$ is a coherent sheaf, see [29], and that the set Y^p where $(\mathcal{J}_{[p]})_x \neq \mathcal{J}_x$ is an analytic variety of dimension at most p, see [28, Theorem 3]. The irreducible components of Y^p, $p = 0, 1, \ldots, \dim X$, are called the *associated (sub)varieties* of $\mathcal{J}$. A coherent sheaf $\mathcal{J}$ is said to be *primary* if it has only one associated variety Y; we then say that

$\mathcal{J}$ is Y*-primary.* Theorem 6 in [28] asserts that each coherent $\mathcal{J} \subset \mathcal{O}^X$ admits a decomposition

$$\mathcal{J} = \bigcap \mathcal{J}_i, \tag{3.1}$$

where there is one Y_i-primary intersectand $\mathcal{J}_i$ for each associated variety Y_i of $\mathcal{J}$. For a radical sheaf $\mathcal{J}_X$, the decomposition (3.1) corresponds to decomposing X into irreducible components.

By Theorem 4 in [28] if Y is an associated prime variety of $\mathcal{J}$, then at $x \in X$ the irreducible components $\mathrm{Ass}(\mathcal{J}_{Y_x})$ of Y_x are germs of varieties of associated primes of $\mathcal{J}_x$. Furthermore, if Y_x is (the variety of) an associated prime of $\mathcal{J}_x$, then Y_x is contained in Y_x^p for $p \geq \dim Y_x$. For fixed x we get that

$$\bigcup_{Y \in \mathrm{Ass}(\mathcal{J}), Y \ni x} \mathrm{Ass}(\mathcal{J}_{Y_x})$$

is a disjoint union of $\mathrm{Ass}(\mathcal{J}_x)$. Thus we have

Lemma 3.1. *The germ at x of $\mathcal{J}_{[p]}$ is precisely the intersection of the primary components of $\mathcal{J}_x$ that are of dimension $> p$.*

Given a subvariety Z of X, the *gap sheaf* $\mathcal{J}[Z] \supset \mathcal{J}$ is defined as follows: A germ $s \in \mathcal{O}_x$ is in $\mathcal{J}[Z]_x$ if and only if it extends to a section of $\mathcal{J}(U)$ for some neighborhood U of x, where $s_y \in \mathcal{J}_y$ for all $y \in U \setminus Z$. Note that $\mathcal{J}[Z]_x$ is the intersection of all components in a primary decomposition of $\mathcal{J}_x$ for which the associated varieties are not contained in Z. It is not hard to see that $\mathcal{J}[Z]$ is coherent, see [29]. Observe that $\mathcal{J}_{[p]} = \mathcal{J}[Y^p]$.

Remark 3.2. We claim that in fact

$$\mathcal{J}_{[p]} = \mathcal{J}[Z^{\mathrm{bef}}_{n-p}]. \tag{3.2}$$

To see this assume first that X is smooth. Then the (germs of) varieties of associated prime ideals of $\mathcal{J}$ of dimension $\leq p$ are precisely the (germs of) varieties of associated prime ideals that are contained in Z^{bef}_{n-p}, see, e.g., [16, Corollary 20.14]. Now (3.2) follows from Lemma 3.1.

For a general X, let $i : X \to Y$ be a local embedding of X into a manifold Y of dimension N and let $\widetilde{\mathcal{J}} = \mathcal{J} + \mathcal{J}_X$, cf. Section 2.4. It is not hard to verify that if $\mathfrak{a}$ is an ideal in $\mathcal{O}^X_x$ and $\tilde{\mathfrak{a}} := \mathfrak{a} + (\mathcal{J}_X)_x$ is the corresponding ideal in $\mathcal{O}^Y_x$ then $\mathfrak{a} = \cap \mathfrak{a}_k$ is a primary decomposition of $\mathfrak{a}$ if and only if $\tilde{\mathfrak{a}} = \cap \tilde{\mathfrak{a}}_k$ is a primary decomposition of $\tilde{\mathfrak{a}}$. Hence, in light of Lemma 3.1, $i^*\widetilde{\mathcal{J}}[V] = \mathcal{J}[V \cap X]$ and $i^*\widetilde{\mathcal{J}}_{[p]} = \mathcal{J}_{[p]}$. By the definition of BEF-varieties in Section 2.4, thus $i^*\widetilde{\mathcal{J}}[Z^{\mathrm{bef}}_{N-p}(\widetilde{\mathcal{J}})] = \mathcal{J}[Z^{\mathrm{bef}}_{N-p}(\widetilde{\mathcal{J}})] = \mathcal{J}[Z^{\mathrm{bef}}_{n-p}(\mathcal{J})]$, which proves (3.2) since $\widetilde{\mathcal{J}}_{[p]} = \widetilde{\mathcal{J}}[Z^{\mathrm{bef}}_{N-p}(\widetilde{\mathcal{J}})]$. $\square$

Given a residue current R constructed from a Hermitian locally free resolution of $\mathcal{O}^X/\mathcal{J}$ on a smooth X as in Section 2.3, in [8] we showed that the germ R_x of the current R at $x \in X$ can be written as $R_x = \sum R^{\mathfrak{p}}$, where the sum is over the associated primes of $\mathcal{J}_x$, and $R^{\mathfrak{p}}$ has support on the variety $V(\mathfrak{p})$ of $\mathfrak{p}$ and has the SEP with respect to $V(\mathfrak{p})$.

4. Resolutions of homogeneous ideals

Let $\mathcal{J}$ be a coherent ideal sheaf on $\mathbb{P}^N$. Then there is a locally free resolution $\mathcal{O}(E^f_\bullet), f^\bullet$, where E_k is a direct sum of line bundles $E_k = \bigoplus_i \mathcal{O}(-d^i_k)$ and $f^k = (f^k_{ij})$ are matrices of homogeneous forms with $\deg f^k_{ij} = d^j_k - d^i_{k-1}$, see, e.g., [22, Ch.1, Example 1.2.21]. Let J denote the homogeneous ideal in the graded ring $\mathcal{S} = \mathbb{C}[z_0, \dots, z_N]$ associated with $\mathcal{J}$, and let $\mathcal{S}(\ell)$ denote the module $\mathcal{S}$ where all degrees are shifted by ℓ. Then $\mathcal{O}(E^f_\bullet), f^\bullet$ corresponds to a free resolution

$$\cdots \to \oplus_i \mathcal{S}(-d^i_k) \to \cdots \to \oplus_i \mathcal{S}(-d^i_2) \to \oplus_i \mathcal{S}(-d^i_1) \to \mathcal{S} \tag{4.1}$$

of the module $\mathcal{S}/J$. Conversely, any such free resolution corresponds to a locally free resolution $\mathcal{O}(E_\bullet), f^\bullet$.

Recall that the *regularity* of a homogeneous module with a minimal graded free resolution (4.1) is defined as $\max_{k,i}(d^i_k - k)$, see, e.g., [17, Ch. 4]. The *regularity* $\operatorname{reg} J$ of the ideal J equals $\operatorname{reg}(\mathcal{S}/J) + 1$, cf. [17, Exercise 4.3].

If X is a subvariety of $\mathbb{P}^N$, then the *regularity* of X, $\operatorname{reg} X$, is defined as the regularity of J_X. Notice that if X has pure dimension, then the ideal J_X has pure dimension in $\mathcal{S}$; in particular the ideal associated to the origin is not an associated prime ideal. Theorem 20.14 in [16] thus implies that Z_0^{bef} is empty. Therefore the depth of $\mathcal{S}/J_X$ is at least 1, and hence a minimal free resolution of $\mathcal{S}/J_X$ has length $\leq N$. For such a resolution we thus get

$$\operatorname{reg} X = \max_{k \leq \min(M,N)} (d^i_k - k) + 1. \tag{4.2}$$

A global section of $\mathcal{O}(s)|_X \to X$ extends to a global section of $\mathcal{O}(s) \to \mathbb{P}^N$ as soon as $s \geq \operatorname{reg} X - 1$, see, e.g., [17, Chapter 4].

5. Division problems on singular varieties

Let $E^g_\bullet, g^\bullet$ be a complex that corresponds to a Hermitian free resolution of $\mathcal{O}^{\mathbb{P}^N}/\mathcal{J}_X$ as above, and let $E^f_\bullet, f^\bullet$ be an arbitrary Hermitian pointwise generically surjective complex over $\mathbb{P}^N$. Then the product current

$$R^f \wedge R^g := R^{f,\lambda} \wedge R^g|_{\lambda=0}$$

is well-defined on $\mathbb{P}^n$,

$$R^f \wedge \omega := R^{i^* f,\lambda} \wedge \omega|_{\lambda=0}$$

is a well-defined current on X, and $i_*(R^f \wedge \omega) = R^f \wedge R^g$, see [9, Section 2]. In particular, $R^f \wedge R^g$ and $R^f \wedge \omega$ only depend on the restriction of f to X, and thus these currents are well-defined even if f is only defined over X. Moreover $R^f \wedge R^g \phi = 0$ if and only if $R^f \wedge \omega i^* \phi = 0$. On X_{reg}, $R^f \wedge \omega$ is just the product of the current R^f and the smooth form ω.

The current $R^f \wedge R^g$ is related to the tensor product complex $E^h_\bullet, h^\bullet$, where

$$E^h_k = \bigoplus_{i+j=k} E^f_i \otimes E^g_j,$$

and $h = f + g$, cf. [9, Section 2.5], in a similar way as is the current R^h associated with this complex, see [4]. In particular, if ϕ is a section of $E_0^h = E_0^f \otimes E_0^g$ such that $R^f \wedge R^g \phi = 0$, one can locally solve $f^1 q + g^1 q' = \phi$. Moreover if (2.6) is satisfied for the product complex there is a global such section (q, q') of $E_1^h = E_1^f \otimes E_0^g \oplus E_0^f \otimes E_1^g$. In general, however, $R^f \wedge R^g$ does not coincide with R^h.

In fact, the definition of R^f in Section 2.3 works also when Y is singular. However, Proposition 2.1 and the duality principle do not hold in general, see, e.g., [21], and therefore R^f itself is not so well suited for division problems.

Example 5.1. Assume that $E_\bullet^f, f^\bullet$ is the Koszul complex generated by sections f_j of $L_j = \mathcal{O}(d_j)|_X$, where $X \subset \mathbb{P}^N$, twisted by $S = \mathcal{O}(\rho)$, as in Example 2.2, and that $E_\bullet^g, g^\bullet$ is a complex associated with a minimal Hermitian free resolution of $\mathcal{S}/J_X$ as in Section 4. Note that then E_ℓ^h is a direct sum of line bundles

$$\mathcal{O}(\rho - (d_{i_1} + \cdots + d_{i_\ell}) - d^i_{k-\ell}).$$

Recall that

$$H^k(\mathbb{P}^N, \mathcal{O}(\ell)) = 0 \qquad \text{if} \quad \ell \geq -N \quad \text{or} \quad k < N, \tag{5.1}$$

see, e.g., [13]. Thus (2.6) is satisfied if $\rho \geq d_{i_1} + \cdots + d_{i_\ell} + d^i_{N+1-\ell} - N$ for $\ell = 1, 2, \ldots, \min(m, n+1)$ and all choices of i and i_j. Notice that, cf. (4.2),

$$d^i_{N+1-\ell} - N = \big(d^i_{N+1-\ell} - (N+1-\ell)\big) + 1 - \ell \leq \operatorname{reg} X - \ell.$$

Hence (2.6) is satisfied if

$$\rho \geq d_1 + \cdots + d_{\min(m,n+1)} - \min(m, n+1) + \operatorname{reg} X. \tag{5.2}$$

Summing up we have:

If ρ satisfies (5.2) and ϕ is a section of $\mathcal{O}(\rho)$ on $\mathbb{P}^N$ such that $R^f \wedge R^g \phi = 0$ (or equivalently $R^f \wedge R^g i^* \phi = 0$) then there are global sections q_j of $\mathcal{O}(\rho - d_j)$ such that $f_1 q_1 + \cdots + f_m q_m = \phi$ on X.

If X is Cohen–Macaulay we may assume that $E_\bullet^g, g^\bullet$ ends at level $N - n$. If moreover $m \leq n$, then $E_\bullet^h, h^\bullet$ ends at level $\leq N$ and thus (2.6) is satisfied for any ρ. $\square$

Example 5.2. Let F_j be polynomials in $\mathbb{C}^N$, let $\hat{f}_j$ be the sections of $\mathcal{O}(\deg F_j) \to \mathbb{P}^N$ corresponding to F_j, and let $\mathcal{J}_{\hat{f}}$ be the ideal sheaf on $\mathbb{P}^N$ generated by the $\hat{f}_j$. Moreover, let $E_\bullet^f, f^\bullet$ and $E_\bullet^g, g^\bullet$ be complexes associated with minimal free resolutions of $\mathcal{J}_{\hat{f}}$ and $\mathcal{J}_X$ as in Section 4, where X is a subvariety of $\mathbb{P}^N$; say $E_k^f = \bigoplus \mathcal{O}(\delta_k^i)$ and $E_k^g = \bigoplus \mathcal{O}(d_k^i)$. Then E_k^h is a direct sum of line bundles $\mathcal{O}(-\delta_\ell^i - d_{k-\ell}^j)$, and thus (2.6) is satisfied if $\rho \geq \delta_\ell^i + d^j_{N+1-\ell} - N$ for all i, j, ℓ, cf. Example 5.1. Notice that, in light of Section 4,

$$\delta_\ell^i + d^j_{N+1-\ell} - N = (\delta_\ell^i - \ell) + (d^j_{N+1-\ell} - (N+1-\ell)) + 1 \leq \operatorname{reg} J_{\hat{f}} + \operatorname{reg} X - 1,$$

where $J_{\hat{f}}$ is the homogeneous ideal associated with $\mathcal{J}_{\hat{f}}$. Thus (2.6) is satisfied if $\rho \geq \operatorname{reg} J_{\hat{f}} + \operatorname{reg} X - 1$.

Let $Z_k^{\hat f}$ and Z_ℓ^g be the BEF-varieties of $\mathcal{J}_{\hat f}$ and $\mathcal{J}_X$, respectively. Theorem 4.2 in [4] asserts that if

$$\operatorname{codim}(Z_k^{\hat f} \cap Z_\ell^g) \geq k+\ell, \tag{5.3}$$

then $R^f \wedge R^g\phi = 0$ if and only if $\phi \in \mathcal{J}_{\hat f} + \mathcal{J}_X = \mathcal{J}_f + \mathcal{J}_X$, where $\mathcal{J}_f$ is the sheaf on X generated by the restrictions f_j of $\hat f_j$, cf. Section 2.4. If moreover $\mathcal{J}_{\hat f}$ and $\mathcal{J}_X$ are both Cohen–Macaulay and the resolutions $\mathcal{O}(E^f_\bullet), f^\bullet$ and $\mathcal{O}(E^g_\bullet), g^\bullet$ have minimal length, then $R^f \wedge R^g = R^h$, see [4, Theorem 4.2]. □

5.1. Distinguished varieties

Let X be a subvariety of $\mathbb{P}^N$ and let $\tilde f_j$ be sections of $L = \mathcal{O}(d)|_X$. Moreover, let $\nu\colon X_+ \to X$ be the normalization of the blow-up of X along $\mathcal{J}_{\tilde f}$, and let $W = \sum r_j W_j$ be the exceptional divisor; here W_j are irreducible Cartier divisors. The images $Z_j := \nu(W_j)$ are called the (*Fulton–MacPherson*) *distinguished varieties* associated with $\mathcal{J}_{\tilde f}$, see, e.g., [22]. If we consider $\tilde f = (\tilde f_1, \dots, \tilde f_m)$ as a section of $E^* := \oplus_1^m \mathcal{O}(-d)$, then $\nu^* \tilde f = \tilde f^0 \tilde f'$, where $\tilde f^0$ is a section of the line bundle $\mathcal{O}(-W)$ and $\tilde f' = (\tilde f_1', \dots, \tilde f_m')$ is a nonvanishing section of $\nu^* E^* \otimes \mathcal{O}(W)$, where $\mathcal{O}(W) = \mathcal{O}(-W)^{-1}$. Furthermore, $\omega_{\tilde f} := dd^c \log|\tilde f'|^2$ is a smooth first Chern form for $\nu^* L \otimes \mathcal{O}(W)$. We will use the geometric estimate

$$\sum r_j \deg_L Z_j \leq \deg_L X \tag{5.4}$$

from [15, Proposition 3.1], see also [22, (5.20)].

Let $R^{\tilde f}$ be the residue current associated with the Koszul complex of the $\tilde f_j$ as in Example 2.2 and consider the regularization (2.7) of $R^{\tilde f}$. Using the notation in Example 2.2, $\nu^*\sigma = (1/\tilde f^0)\sigma'$, where $1/\tilde f^0$ is a meromorphic section of $\mathcal{O}(W)$ and σ' is a smooth section of $\nu^* E \otimes \mathcal{O}(-W)$. It follows that

$$\nu^*(\sigma \wedge (\bar\partial\sigma)^{k-1}) = \frac{1}{(\tilde f^0)^k}\sigma' \wedge (\bar\partial\sigma')^{k-1},$$

and hence

$$\nu^* R_k^\lambda = \bar\partial|\tilde f^0 \tilde f'|^{2\lambda} \wedge \frac{1}{(\tilde f^0)^k}\sigma' \wedge (\bar\partial\sigma')^{k-1} \text{ for } \operatorname{Re}\lambda >> 0,$$

when $k \geq 1$. Since $\tilde f'$ is nonvanishing, by (2.1) the value at $\lambda = 0$ is precisely

$$R_k^+ := \bar\partial \frac{1}{(\tilde f^0)^k} \wedge \sigma' \wedge (\bar\partial\sigma')^{k-1}. \tag{5.5}$$

Thus

$$\nu_* R_k^+ = R_k^{\tilde f}.$$

6. Proofs

Proof of Theorem 1.5. For $j = 1, \dots, m$, let $\hat{f}_j$ be the $\deg F_j$-homogenization of the polynomial F_j, considered as a section of $\mathcal{O}(\deg F_j) \to \mathbb{P}^N$. Moreover let $g_1, \dots, g_r$ be global generators of the ideal sheaf $\mathcal{J}_X$; assume they are sections of $\mathcal{O}(d_1), \dots, \mathcal{O}(d_r)$, respectively. Let $\mathcal{J} = \mathcal{J}_{\hat{f}} + \mathcal{J}_X = \mathcal{J}_f + \mathcal{J}_X$. Then there is a locally free resolution $\mathcal{O}(E^h_\bullet), h^\bullet$ of $\mathcal{O}/\mathcal{J}$, where each E^h_k is a direct sum of line bundles $E_k = \bigoplus_i \mathcal{O}(-d^i_k)$ and in particular $E^1 = \bigoplus_1^m \mathcal{O}(-\deg F_j) \oplus_1^r \bigoplus \mathcal{O}(-d_k)$ and $h^1 = (f_1, \dots, f_m, g_1, \dots, g_r) =: f + g$, cf. Section 4. Let $R = R^h$ be the residue current associated with $E^h_\bullet, h^\bullet$.

Recall from Section 3 that for fixed $x \in X$, $R_x = \sum R^{\mathfrak{p}}$, where the sum is over $\mathrm{Ass}(\mathcal{J}_x)$ and where $R^{\mathfrak{p}}$ has the SEP with respect to $V(\mathfrak{p})$; in particular, $\mathbf{1}_{H_\infty} R^{\mathfrak{p}} = R^{\mathfrak{p}}$ if $V(\mathfrak{p}) \subset H_\infty$ and $\mathbf{1}_{H_\infty} R^{\mathfrak{p}} = 0$ otherwise. Thus

$$\mathbf{1}_{H_\infty} R_x = \sum_{\mathfrak{p} \in \mathrm{Ass}(\mathcal{J}_x), V(\mathfrak{p}) \subset H_\infty} R^{\mathfrak{p}}. \tag{6.1}$$

In Remark 3.2 we saw that $\mathfrak{a} = \cap \mathfrak{a}_k$ is a primary decomposition of the ideal $\mathfrak{a}$ in $\mathcal{O}^X_x$ if and only if $\tilde{\mathfrak{a}} = \cap \tilde{\mathfrak{a}}_k$ is a primary decomposition of the ideal $\tilde{\mathfrak{a}} = \mathfrak{a} + (\mathcal{J}_X)_x$ in $\mathcal{O}^Y_x$. Thus, that $\mathcal{J}_f$ has no associated varieties contained in X_∞ implies that, for a fixed $x \in X$, $\mathcal{J}_x$ has no (varieties of) associated primes contained in the hyperplane H_∞ at infinity in $\mathbb{P}^N$. We conclude, in light of (6.1), that $\mathbf{1}_{H_\infty} R = 0$. If ϕ is any homogenization of Φ then $\mathbf{1}_{\mathbb{C}^N} R\phi = 0$ because of the duality principle and hence $R\phi = \mathbf{1}_{H_\infty} R\phi + \mathbf{1}_{\mathbb{C}^N} R\phi = 0$.

Assume that the complex $E^h_\bullet, h^\bullet$ ends at level M (by Hilbert's syzygy theorem we may assume that $M \le N + 1$) and let

$$\beta := \max_i d^i_{N+1} - N \text{ if } M = N + 1 \quad \text{ and } \beta := 0 \text{ otherwise.} \tag{6.2}$$

If $\rho \ge \beta$ then (2.6) is satisfied for $E^h_\bullet, h^\bullet$ twisted by $\mathcal{O}(\rho)$ in light of (5.1) and thus by Proposition 2.1 there are global holomorphic sections $q = (q_j)$ of $\bigoplus \mathcal{O}(\rho - \deg F_j)$ and $q' = (q'_k)$ of $\bigoplus \mathcal{O}(\rho - d_k)$ over $\mathbb{P}^N$ such that $\hat{f} q + g q' = \phi$. Indeed, recall from the end of Section 2.3 that R is also the residue current associated with the twisted complex. Dehomogenizing gives polynomials Q_j, Q'_j, and G_j in $\mathbb{C}^N$ such that

$$\sum F_j Q_j + \sum G_j Q'_j = \Phi$$

and where $\deg(F_j Q_j) \le \rho$. Since the G_j vanish on V we get the desired solution to (1.1) on V, and thus the first part of Theorem 1.5 follows with β as in (6.2).

If $V = \mathbb{C}^N$, $\mathcal{O}_X$ should be interpreted as the zero sheaf. Then $E^h_\bullet, h^\bullet$ is a locally free resolution of $\mathcal{O}/\mathcal{J}_f$ and $\beta \le \operatorname{reg} J_f$, cf. Section 4.

For the second part of Theorem 1.5, assume that $\mathcal{J}_f$ has an associated variety contained in X_∞. We are to prove that for arbitrarily large ℓ there is a polynomial $\Phi = \Phi_\ell$ of degree $\ge \ell$ in (F_j) on V for which one cannot solve (1.1) with $\deg(F_j Q_j) \le \deg \Phi_\ell$.

Let $L = \mathcal{O}(1)|_X$. The hypothesis on $\mathcal{J}_f$ then means that $\mathcal{J}_f[X_\infty]$ is strictly larger than $\mathcal{J}_f$. Therefore, since L is ample, for some large enough s_0 there is a global section ψ_0 of $L^{\otimes s_0} \to X$ such that ψ_0 is in $\mathcal{J}_f[X_\infty]$ but not in $\mathcal{J}_f$. Moreover we can find a global section ψ of $L^{\otimes s}$ for some $s \geq 1$ such that ψ does not vanish identically on any of the associated varieties of $\mathcal{J}_f$ that are contained in X_∞. We may assume that $s_0, s \geq \operatorname{reg} X - 1$, so that ψ_0 and ψ extend to global sections $\hat{\psi}_0$ and $\hat{\psi}$ of $\mathcal{O}(s_0)$ and $\mathcal{O}(s)$, respectively. Let Ψ_0 and Ψ be the corresponding dehomogenized polynomials in $\mathbb{C}^N$. For $\ell \geq 0$, let $\phi_\ell = \psi_0\psi^\ell$ and $\Phi_\ell = \Psi_0\Psi^\ell$. Since $\mathcal{J}_f[X_\infty]_x = (\mathcal{J}_f)_x$ for all $x \in V$, Φ_ℓ is in the ideal (F_j) on V, and thus we can solve (1.1) for $\Phi = \Phi_\ell$ on V. Assume that there is a solution to (1.1) with $\deg(F_jQ_j) \leq \rho_\ell$. Then there are sections q_j of $L^{\rho_\ell - \deg F_j}$ such that

$$\sum f_j q_j = z_0^{\rho_\ell - (s_0 + s\ell)} \phi_\ell$$

on X. Since ϕ_ℓ is not in $\mathcal{J}_f$ it follows that $\rho_\ell - (s_0 + s\ell) \geq 1$ and thus $\rho_\ell \geq 1 + (s_0 + s\ell) \geq 1 + \deg \Phi_\ell$. Since $\hat{\psi}$ does not vanish identically at X_∞, $\deg \Psi \geq 1$ and hence $\deg \Phi_\ell \geq \ell$. Hence we have found Φ_ℓ with the desired properties and the second part of Theorem 1.5 follows. $\square$

Remark 6.1. If $\mathcal{J}_{\hat{f}}$ and $\mathcal{J}_X$ are Cohen–Macaulay and the BEF-varieties of $\mathcal{J}_{\hat{f}}$ and $\mathcal{J}_X$ satisfy (5.3), then we can choose the complex $E^h_\bullet, h^\bullet$ in the above proof to be the tensor product of the complexes $E^f_\bullet, f^\bullet$ and $E^g_\bullet, g^\bullet$ corresponding to minimal resolutions of $\mathcal{J}_{\hat{f}}$ and $\mathcal{J}_X$, see Example 5.2. In this case, by Example 5.2, we get Theorem 1.5 for $\beta = \operatorname{reg} J_{\hat{f}} + \operatorname{reg} X - 1$. $\square$

The residue current technique in the preceding proof is convenient and makes it possible to carry out the proof within our general framework, but it is not crucial.

Remark 6.2 (The algebraic approach). Let us first sketch an algebraic proof of the first part of Theorem 1.5. We use the notation from the proof above. To begin with we have to prove that ϕ is in $\mathcal{J}$, which of course precisely corresponds to proving that $R\phi = 0$. Since (the restriction to V of) ϕ is in $\mathcal{J}_f$ on V it follows that $\phi_{x'}$ is in $\mathcal{J}$ outside H_∞. Since moreover $\mathcal{J} = \mathcal{O}^{\mathbb{P}^N}$ outside X, we have to prove that $\phi_x \in \mathcal{J}_x$ for each $x \in X_\infty$. At such a point x we have a minimal primary decomposition $\mathcal{J}_x = \cap_\ell \mathcal{J}^\ell_x$. Since $\mathcal{J}$ is coherent, $\mathcal{J} \subset \mathcal{J}^\ell$ in a neighborhood $\mathcal{U}$ of x, where $\mathcal{J}^\ell$ is the coherent sheaf defined by $\mathcal{J}^\ell_x$. Let Z^ℓ be the zero-set of $\mathcal{J}^\ell$. Since $\phi_{x'}$ is in $\mathcal{J}_{x'}$ for x' outside H_∞ it follows that $\phi_{x'}$ is in $\mathcal{J}^\ell_{x'}$ for $x' \in Z^\ell \setminus H_\infty$. Hence $\mathcal{F} := (\mathcal{J}^\ell + (\phi))/\mathcal{J}^\ell$ is a coherent sheaf in $\mathcal{U}$ with support on $Z^\ell \cap H_\infty$. Since by assumption $\mathcal{J}_f$ has no associated varieties contained in X_∞ it follows that $Z^\ell \cap H_\infty$ has positive codimension in Z^ℓ, cf. the proof of Theorem 1.5 above. Therefore, by the Nullstellensatz there is a holomorphic function h, not vanishing identically on Z^ℓ such that $h\mathcal{F} = 0$. In particular, $h_x\phi_x \in \mathcal{J}^\ell_x$. Since h_x is not in the radical of $\mathcal{J}^\ell_x$ and $\mathcal{J}^\ell_x$ is primary it follows that $\phi_x \in \mathcal{J}^\ell_x$. We conclude that $\phi_x \in \mathcal{J}_x$. Notice that the last arguments above can be thought of as an algebraic version of the SEP-argument in the proof of Theorem 1.5 above.

Next we would like to use that $\phi \in \mathcal{J}$ to conclude that there is a global holomorphic solution to $hq = \phi$. By a partition of unity, using that $E^h_\bullet, h^\bullet$ is exact, one can glue local such solutions together to obtain a global smooth solution to $(h - \bar\partial)\psi = \phi$, cf. [9, Section 4]. By solving a certain sequence of $\bar\partial$-equations in $\mathbb{P}^N$ we can modify ψ to a global holomorphic solution q to $hq = \phi$. These $\bar\partial$-equations are solvable if $\rho \geq \beta$ defined by (6.2). Alternatively, one can directly refer to the well-known result that there is a solution to $hq = \phi$ if $\rho \geq \operatorname{reg} J$, where J is the homogeneous ideal corresponding to $\mathcal{J}$, see, e.g., [17, Proposition 4.16].

In the same way Theorem 1.1 and 1.2 follow without any reference to residues. Probably one can also find give an algebraic proof of Theorem 1.4. $\square$

In the next proof the residue technique plays a more decisive role.

Proof of Theorem 1.6. Let

$$\rho = \max(\deg \Phi + \mu d^{c_\infty} \deg X, (d-1)\min(m, n+1) + \operatorname{reg} X),$$

or if X is Cohen–Macaulay and $m \leq n$ let $\rho = \deg \Phi + m d^{c_\infty} \deg X$, and let ϕ be the ρ-homogenization of Φ considered as a section of $\mathcal{O}(\rho)|_X$. Note that then $\phi = z_0^{\rho - \deg \Phi}\tilde\phi$, where $\tilde\phi$ is the $\deg \Phi$-homogenization of Φ. Moreover, let $R^{\tilde f} \wedge \omega$ be the residue current associated with the (twisted) Koszul complex $E^{\tilde f}_\bullet, \tilde f^\bullet$ of the sections $\tilde f_j$ of $\mathcal{O}(d)|_X$ associated with F_j, and a complex $E^g_\bullet, g^\bullet$ associated with a minimal resolution of $\mathcal{O}/\mathcal{J}_X$ as in Example 5.1 (with $d_j = d$ for all j).

Claim: $R^{\tilde f} \wedge \omega_0 \phi$ has support on $Z^{\tilde f} \cap X^0$.

To prove the claim, since ω is smooth on $X_{\rm reg}$, it is enough to show that $R^{\tilde f}\phi = 0$ on $X_{\rm reg}$. First, since $\operatorname{codim} Z^{\tilde f} \cap V \geq m$, the duality principle for a complete intersection, cf. Example 2.2, implies that $R^{\tilde f}\phi = 0$ on $V_{\rm re}$.

Next, to prove that $\mathbf{1}_{X_\infty \setminus X^0} R^{\tilde f}\phi = 0$ we consider the normalization of the blow-up $\nu\colon X_+ \to X$, and let $R^+ := \sum R^+_k$ be as in Section 5.1. Let W' be the union of the irreducible components of $W = \nu^{-1} Z^{\tilde f}$ that are contained in $\nu^{-1}X_\infty$. We claim that

$$\mathbf{1}_{X_\infty} R^{\tilde f} = \nu_*\big(\mathbf{1}_{W'} R^+\big). \tag{6.3}$$

In fact, by (2.3),

$$\mathbf{1}_{X_\infty} R^{\tilde f} = \nu_*\big(\mathbf{1}_{\nu^{-1}X_\infty} R^+\big) = \nu_*\big(\mathbf{1}_{\nu^{-1}X_\infty}(\mathbf{1}_{W'} + \mathbf{1}_{W\setminus W'})R^+\big). \tag{6.4}$$

By, (2.2), $\mathbf{1}_{\nu^{-1}X_\infty}\mathbf{1}_{W'}R^+ = \mathbf{1}_{W'}R^+$. Moreover,

$$\mathbf{1}_{\nu^{-1}X_\infty}\mathbf{1}_{W\setminus W'}\bar\partial\frac{1}{(\tilde f^0)^k} = \mathbf{1}_{\nu^{-1}X_\infty \cap (W\setminus W')}\bar\partial\frac{1}{(\tilde f^0)^k} = 0$$

by (2.2) and the dimension principle, since $\nu^{-1}X_\infty \cap (W \setminus W')$ has codimension at least 2 in X_+. In view of (5.5) we conclude that $\mathbf{1}_{\nu^{-1}X_\infty}\mathbf{1}_{W\setminus W'}R^+ = 0$, and thus (6.3) follows from (6.4).

It follows from (6.3) that $\mathbf{1}_{X_\infty \setminus X^0} R^{\tilde f}\phi = 0$ if $\mathbf{1}_{W'}R^+\nu^*\phi = 0$. To show that $\mathbf{1}_{W'}R^+\nu^*\phi$ vanishes first note that it is sufficient to show that it vanishes in a neighborhood of each point x on W' where W is smooth. Indeed, since $W_{\rm sing}$ has codimension at least 2 in W, $\mathbf{1}_{W_{\rm sing}}\bar\partial(1/(\tilde f^0)^k)=0$ by the dimension principle. Hence, using (5.5) and (2.2) we get that

$$\mathbf{1}_{W'}R^+=\mathbf{1}_{W'}(\mathbf{1}_{W_{\rm reg}}+\mathbf{1}_{W_{\rm sing}})R^+=\mathbf{1}_{W'\cap W_{\rm reg}}R^+.$$

Consider now $x\in 1_{W'\cap W_{\rm reg}}$; say x is contained in the irreducible component W_j of W'. In a neighborhood of x we have that $\tilde f^0=s^{r_j}v$, where s is a local coordinate function and v is nonvanishing and r_j is as in Section 5.1. Since $\phi=z_0^{\rho-\deg\Phi}\tilde\phi$, by the choice of ρ, $\nu^*\phi$ vanishes to order (at least) $\mu d^{c_\infty}\deg X$ on W'.

If Ω is a first Chern form for $\mathcal{O}(1)|_X$, e.g., $\Omega=dd^c\log|z|^2$, then $d\Omega$ is a first Chern form for $L=\mathcal{O}(d)|_X$ on X (notice that d denotes the degree and not the differential). By (5.4) we therefore have that

$$r_j\int_{Z_j}(d\Omega)^{\dim Z_j}\le\int_X(d\Omega)^n,$$

which implies that

$$r_j\le d^{\operatorname{codim} Z_j}\deg X.$$

It follows that $\nu^*\phi$ vanishes (at least) to order μr_j on W_j and hence it has a factor $s^{\mu r_j}$. In a neighborhood of x,

$$\bar\partial\frac{1}{(\tilde f^0)^k}=\bar\partial\frac{1}{s^{kr_j}}\wedge smooth$$

and thus, in light of (5.5), $R_k^+\nu^*\phi=0$ for $k\le\mu$ there. Hence $\mathbf{1}_{W'\cap W_{\rm reg}}R_k^+\nu^*\phi=0$ for $k\le\mu$ and $\mathbf{1}_{X_\infty\setminus X^0}R^{\tilde f}\phi=0$. We conclude that $\mathbf{1}_{X\setminus X^0}R^{\tilde f}\phi=\mathbf{1}_{V_{\rm reg}}\ R^{\tilde f}\phi+\mathbf{1}_{X_\infty\setminus X^0}R^{\tilde f}\phi=0$, which proves the claim that $R^{\tilde f}\wedge\omega_0\phi$ has support on $Z^{\tilde f}\cap X^0$.

By (1.8) and the dimension principle we conclude that $R^{\tilde f}\wedge\omega_0\phi$ vanishes identically, since the bidegree of $R^{\tilde f}$ is at most $(0,m)$ and ω_0 has bidegree $(n,0)$. Thus $R^{\tilde f}\wedge\omega_1\phi=R^{\tilde f}\wedge\alpha^1\omega_0\phi$, see (2.9), vanishes outside X^1. By (1.8) and the dimension principle, it vanishes identically since the bidegree of $R^{\tilde f}\wedge\omega_1$ is at most $(n,m+1)$. By induction, it follows that $R^{\tilde f}\wedge\omega_\ell\phi=0$ for each ℓ. We conclude that $R^{\tilde f}\wedge\omega\phi=0$.

Since ρ satisfies (5.2) (with $d_j=d$) and $R^{\tilde f}\wedge\omega\phi=0$, by Example 5.1 there is a global section $q=(q_j)$ of $\sum_1^m\mathcal{O}(\rho-d)$ such that $fq=\phi$ on X. Dehomogenizing gives polynomials Q_j such that (1.1) holds on V and $\deg(F_jQ_j)\le\rho$. $\square$

Proof of Theorems 1.1 *and* 1.4. Let

$$\rho=\max(\deg\Phi,d_1+\cdots+d_{\min(m,n+1)}-\min(m,n+1)+\operatorname{reg}X),$$

or if X is Cohen–Macaulay and $m\le n$ let $\rho=\deg\Phi$. Moreover let ϕ be the ρ-homogenization of Φ and let $R^f\wedge\omega$ be the residue current associated with the twisted Koszul complex $E_\bullet^f,f^\bullet$ of the $\deg F_j$-homogenizations f_j of F_j and a minimal resolution of $\mathcal{O}/\mathcal{J}_X$ as in Example 5.1.

We claim that under the hypotheses of both theorems $R^f \wedge \omega_0\phi$ has support on $Z^f \cap X^0$. Since ω is smooth outside X^0 it is enough to show that $R^f\phi = 0$ there. First in the case of Theorem 1.1, R^f vanishes for trivial reasons, since Z^f is empty. In the case of Theorem 1.4, first $R^f\phi$ vanishes on $V_{\rm reg}$ by the duality principle. Next, since by assumption (1.2) holds and Z^f has no irreducible components in X_∞, it holds that $\operatorname{codim}(X_\infty \cap Z^f) > m$. Since the components of R^f have bidegree at most $(0, m)$, we conclude that $\mathbf{1}_{X_\infty \setminus X_0} R^f = 0$ by the dimension principle. This proves that $R^f \wedge \omega\phi$ has support on $Z^f \cap X^0$.

Now arguing as in the end of the proof of Theorem 1.6, we get that $R^f \wedge \omega\phi = 0$, and the results follow from Example 5.1. $\square$

Remark 6.3. If $\deg F_j = d$, then Theorems 1.1 and 1.4 follow directly from Theorem 1.6. First, notice that Theorem 1.1 follows if we apply Theorem 1.6 to F_j with no common zeros on X. Indeed, since Z^f is empty, $\operatorname{codim}(Z^f \cap X) = \infty$ and thus (1.7) and (1.8) are satisfied, and moreover $c_\infty = -\infty$.

Next, assume that F_j satisfy the hypothesis of Theorem 1.4. Since the codimension of a distinguished variety is at most m the condition that Z^f satisfies (1.2) and has no irreducible component contained in X_∞ means that (1.7) is satisfied and no distinguished varieties can be contained in X_∞. Thus $c_\infty = -\infty$ and $d^{c_\infty} = 0$ and Theorem 1.4 follows from Theorem 1.6. $\square$

Acknowledgment. We thank Richard Lärkäng for helpful discussions.

References

[1] M. Andersson: *Residue currents and ideals of holomorphic functions*, Bull. Sci. Math. **128**, (2004), 481–512.

[2] M. Andersson: *Integral representation with weights. II. Division and interpolation*, Math. Z. **254** (2006), no. 2, 315–332.

[3] M. Andersson: *The membership problem for polynomial ideals in terms of residue currents*, Ann. Inst. Fourier **56** (2006), 101–119.

[4] M. Andersson: *A residue criterion for strong holomorphicity*, Arkiv Mat. **48** (2010), 1–15.

[5] M. Andersson & E. Götmark: *Explicit representation of membership of polynomial ideals*, Math. Ann. **349** (2011), 345–365.

[6] M. Andersson & H. Samuelsson: *A Dolbeault–Grothendieck lemma on a complex space via Koppelman formulas*, Invent. Math. **190** (2012), 261–297.

[7] M. Andersson, E. Wulcan: *Residue currents with prescribed annihilator ideals*, Ann. Sci. École Norm. Sup. **40** (2007), 985–1007.

[8] M. Andersson & E. Wulcan: *Decomposition of residue currents*, J. Reine Angew. Math. **638** (2010), 103–118.

[9] M. Andersson & E. Wulcan: *Global effective versions of the Briançon–Skoda–Huneke theorem*, Invent. Math. **200** (2015), 607–651.

[10] D. Bayer & D. Mumford: *What can be computed in algebraic geometry?*, Computational algebraic geometry and commutative algebra (Cortona, 1991), 1–48, Sympos. Math., XXXIV, Cambridge Univ. Press, Cambridge, 1993.

[11] J.-B. Björk: *Residues and $\mathcal{D}$-modules*, The legacy of Niels Henrik Abel, 605–651, Springer, Berlin, 2004.

[12] J. Briançon & H. Skoda: *Sur la clôture intégrale d'un idéal de germes de fonctions holomorphes en un point de $\mathbb{C}^n$*, C. R. Acad. Sci. Paris Sér. A 278 (1974), 949–951.

[13] J-P Demailly: *Complex Analytic and Differential Geometry*, Monograph Grenoble (1997).

[14] A. Dickenstein & C. Sessa: *Canonical representatives in moderate cohomology*, Invent. Math. **80** (1985), 417–434.

[15] L. Ein & R. Lazarsfeld: *A geometric effective Nullstellensatz*, Invent. Math. **135** (1999), 427–448.

[16] D. Eisenbud: *Commutative algebra. With a view toward algebraic geometry*, Graduate Texts in Mathematics, 160. Springer-Verlag, New York, 1995.

[17] D. Eisenbud: *The geometry of syzygies. A second course in commutative algebra and algebraic geometry*, Graduate Texts in Mathematics, 229. Springer-Verlag, New York, 2005.

[18] G. Hermann: *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann., **95** (1926), 736–788.

[19] C. Huneke: *Uniform bounds in Noetherian rings*, Invent. Math. **107** (1992) 203–223.

[20] Z. Jelonek: *On the effective Nullstellensatz*, Invent. Math. **162** 1–17 (2005).

[21] R. Lärkäng: *On the duality theorem on an analytic variety*, Math. Ann. **355** (2013), 215–234.

[22] R. Lazarsfeld: *Positivity in algebraic geometry I and II*, Springer-Verlag 2004.

[23] F. S. Macaulay: *The algebraic theory of modular systems*, Cambridge Univ. Press, Cambridge 1916.

[24] E. Mayr & A. Mayer: *The complexity of the word problem for commutative semigroups and polynomial ideals*, Adv. in math. **46** (1982), 305–329.

[25] M. Nöther: *Über einen Satz aus der Theorie der algebraischen Functionen*, Math. Ann. (1873), 351–359.

[26] M. Passare & A. Tsikh & A. Yger: *Residue currents of the Bochner–Martinelli type*, Publicacions Mat. **44** (2000), 85–117.

[27] B. Shiffman: *Degree bounds for the division problem in polynomial ideals*, Michigan Math. J. **36** (1989), no. 2, 163–171.

[28] Y.-T. Siu: *Noether–Lasker decomposition of coherent analytic subsheaves*, Trans. Amer. Math. Soc. **135** 1969 375–385.

[29] W. Thimm: *Lückengarben von kohärenten analytischen Modulgarben*, Math. Ann. **148** 1962 372–394.

[30] E. Wulcan: *Sparse effective membership problems via residue currents*, Math. Ann. **350** (2011), 661–682.

[31] E. Wulcan: *Some variants of Macaulay's and Max Noether's theorems*, J. Commut. Algebra **2** (2010), no. 4, 567–580.

Department of Mathematics
Chalmers University of Technology and the University of Gothenburg
S-412 96 Göteborg, Sweden
e-mail: matsa@chalmers.se
wulcan@chalmers.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 111–120

On the Optimal Regularity of Weak Geodesics in the Space of Metrics on a Polarized Manifold

Robert J. Berman

To the memory of Mikael Passare

Abstract. Let (X, L) be a polarized compact manifold, i.e., L is an ample line bundle over X and denote by $\mathcal{H}$ the infinite-dimensional space of all positively curved Hermitian metrics on L equipped with the Mabuchi metric. In this short note we show, using Bedford–Taylor type envelope techniques developed in the authors previous work [3], that Chen's weak geodesic connecting any two elements in $\mathcal{H}$ are $C^{1,1}$-smooth, i.e., the real Hessian is bounded, for any fixed time t, thus improving the original bound on the Laplacians due to Chen. This also gives a partial generalization of Blocki's refinement of Chen's regularity result. More generally, a regularity result for complex Monge–Ampère equations over $X \times D$, for D a pseudoconvex domain in $\mathbb{C}^n$ is given.

1. Introduction

Let X be an n-dimensional compact complex manifold equipped with a Kähler form ω and denote by $[\omega]$ the corresponding cohomology class in $H^2(X, \mathbb{R})$. The space of all Kähler metrics in $[\omega]$ may be identified with the space $\mathcal{H}(X, \omega)$ of all Kähler potentials, modulo constants, i.e., the space of all functions u on X such that

$$\omega_u := \omega + dd^c u, \quad \left(dd^c := \frac{i}{2\pi} \partial \bar{\partial} \right)$$

is positive, i.e., defines a Kähler form on X. Mabuchi introduced a natural Riemannian metric on $\mathcal{H}(X, \omega)$ [21], where the squared norm of a tangent vector $v \in C^\infty(X)$ at u is defined by

$$g_{|u}(v, v) := \int_X v^2 \omega_u^n \tag{1.1}$$

The main case of geometric interest is when the cohomology class $[\omega]$ is integral, which equivalently means that it can be realized as the first Chern class

$c_1(L)$ of an ample line bundle L over the projective algebraic manifold X. Then the space $\mathcal{H}(X, \omega)$ may be identified with the space $\mathcal{H}(L)$ of all positively curved metrics ϕ on the line bundle L and as pointed by Donaldson [15] the space $\mathcal{H}(L)$ may then be interpreted as the symmetric space dual of the group $\mathrm{Ham}(X, \omega)$ of Hamiltonian diffeomorphisms of (X, ω). Under this (formal) correspondence the geodesics in $\mathcal{H}(X, \omega)$ correspond to one-parameter subgroups in the (formal) complexification of $\mathrm{Ham}(X, \omega)$ and this motivated Donaldson's conjecture concerning the existence of geodesics in $\mathcal{H}(X, \omega)$, connecting any two given elements.

However, Donaldson's existence problem has turned out to be quite subtle. In fact, according to the recent counter-examples in [20, 11] the existence of bona fide geodesic segments fails in general. On the other hand, there always exists a (unique) *weak* geodesic u_t connecting given points u_0 and u_1 in $\mathcal{H}(X, \omega)$ defined as follows. First recall that, by an important observation of Semmes [23] and Donaldson [15], after a complexification of the variable t, the geodesic equation for u_t on $X \times [0, 1]$ may be written as the following complex Monge–Ampère equation on a domain $M := X \times D$ in $X \times \mathbb{C}$ for the function $U(x, t) := u_t(x)$:

$$(\pi^*\omega + dd^c U)^{n+1} = 0. \tag{1.2}$$

As shown by Chen [9], with complements by Blocki [8], for any smoothly bounded domain D in $\mathbb{C}$ the corresponding boundary value problem on M admits a unique solution U such that $\pi^*\omega + dd^c U$ is a positive current with coefficients in L^∞, satisfying the equation 1.2 almost everywhere. In particular, when D is an annulus in $\mathbb{C}$ this construction gives rise to the notion of a weak geodesic curve u_t in the space of all functions u such that ω_u is a positive current with coefficients in L^∞ (the latter regularity equivalently means that the Laplacian of u is in L^∞). In particular, by standard linear elliptic estimates, U is "almost $C^{1,1}$" in the sense that U is in the Hölder class $C^{1,\alpha}$ for any $\alpha < 1$. As shown by Blocki [8], in the case when X admits a Kähler metric with non-negative holomorphic bisectional curvature Chen's regularity result can be improved to give that U is $C^{1,1}$-smooth. However, the assumption on X appearing in Blocki's result is very strong and essentially implies that X is a homogeneous manifold. In this short note we point out that, in the case when the given Kähler class $[\omega]$ is an integral the function u_t on X is in general, for any fixed t, in $C^{1,1}(X)$, i.e., its first derivatives are Lipschitz continuous. More precisely, the real Hessian of u_t has bounded coefficients with a bound which is independent of t :

Theorem 1.1. *For any integral Kähler class $[\omega]$ the weak geodesic u_t connecting any two points u_0 and u_1 in the space $\mathcal{H}(X, \omega)$ of ω-Kähler potentials has the property that, for any fixed t, the function u_t is in $C^{1,1}(X)$. More precisely, the upper bound on the sup norm on X of the real Hessian of u_t only depending on an upper bound of sup norms of the real Hessians of u_0 and u_1.*

This regularity result should be compared with recent results of Darvas–Lempert [11] showing that the solution $U(x, t) := u_t(x)$ is not, in general, C^2-smooth up to the boundary of M in (more precisely $dd^c U$ is not represented by a continuous form). However, the argument in [11], which is inspired by a similar argument in the case of $M = D$ for a pseudoconvex domain D in $\mathbb{C}^2$ due to Bedford–Fornaess [1], does not seem to exclude the possibility that U be C^2-smooth in the *interior* of M. Anyway, the latter scenario appears to be highly unlikely in view of the explicit counter-example of Gamelin–Sibony [17] to interior C^2-regularity for the case when D is the unit-ball in $\mathbb{C}^2$. Note also that, since the bounds on the real Hessian of u_t are controlled by the Hessians of u_0 and u_1 the previous theorem shows that $PSH(X, \omega) \cap C^{1,1}(X)$ is closed with respect to weak geodesics. By the very recent work of Darvas [10] and Guedj [18] this the latter property equivalently means that $PSH(X, \omega) \cap C^{1,1}(X)$ defines a geodesic subspace of the metric completion of the space $\mathcal{H}$ equipped with the Mabuchi metric.

The starting point of the proof of Theorem 1.1 is the well-known Perron type envelope representation of the solution to the Dirichlet problem for the complex Monge–Ampère operator. The proof, which is inspired by Bedford–Taylor's approach in their seminal paper [2], proceeds by a straightforward generalization of the technique used in [3] to establish the corresponding regularity result for certain envelopes of positively curved metrics in a line bundle $L \to X$ (which can be viewed as solutions to a free boundary value problem for the complex Monge–Ampère equation on X). In fact, the situation here is considerably simpler than the one in [3] which covers the case when the line bundle L is merely big (the $C^{1,1}$-regularity then holds on the ample locus of L in X) and one of the motivations for the present note is to highlight the simplicity of the approach in [3] in the present situation (see also [22] for other generalizations of [3]). But it should be stressed that, just as in [3], the results can be generalized to more general line bundles. For example, by passing to a smooth resolution, Theorem 1.1 be generalized to show that the weak geodesic connecting any two smooth metrics with non-negative curvature current on an ample line bundle L over a singular compact normal complex variety X is $C^{1,1}$-smooth on the regular locus of X (for a fixed "time").

As it turns out one can formulate a general result (Theorem 2.1 below) which contains both Theorem 1.1 and the corresponding regularity result in [3]. In particular, the latter result covers the case when the domain D is the unit disc (or more generally, the unit ball in $\mathbb{C}^n$, where the following more precise regularity result holds:

Theorem 1.2. *For any integral Kähler class* $[\omega]$ *on a compact complex manifold* X *the solution* U *to the Dirichlet problem for the complex Monge–Ampère equation* 1.2 *with* C^2*-boundary data,* ω*-psh along the slices* $\{t\} \times X$*, is* $C^{1,1}$*-smooth in the interior of* $X \times D$*, if* D *is the unit disc in* $\mathbb{C}$.

As pointed out by Donaldson [15] the boundary value problem appearing in the previous theorem can be viewed as an infinite-dimensional analog of a standard boundary value problem for holomorphic discs in the complexification of a compact Lie group G or more precisely the classical factorization theorem for loops in G (recall that the role of G in the present infinite-dimensional setting is played by the group $\mathrm{Ham}(X,\omega)$ of Hamiltonian diffeomorphisms). As shown by Donaldson [16] the solution U is in general not smooth and Donaldson raised the problem of studying the singularities of Chen's weak solution; the paper can thus be seen as one step in this direction.

One potentially useful consequence of the regularity results in Theorems 1.1, 1.2 is that, for a fixed "time" t the differential of u_t (which geometrically represents the connection one form of the corresponding metric on the line bundle L) is Lipschitz continuous and in particular differentiable on $X-E$, where the exceptional set E is a null set for the Lebesgue measure. For example, it then follows from the results in [3] that the corresponding scaled Bergman kernel $B_k(x,x)/k^n$, attached to high tensor powers $L^{\otimes k}$, converges when $k\to\infty$ point-wise on $X-E$ to the density of $\omega_{u_t}^n$. By a circle of ideas going back to Yau such Bergman kernels can be used to approximate differential geometric objects in Kähler geometry. Accordingly, the precise $C^{1,1}$-regularity established in the present paper will hopefully find applications in Kähler geometry in the future. In fact, one of the initial motivations for writing the present note came from a very recent joint work with Bo Berndtsson [5] where Bergman kernel asymptotics are used to establish the convexity of Mabuchi's K-energy along weak geodesics and where the precise $C^{1,1}$-regularity was needed at an early stage of the work. Eventually it turned that Chen's regularity, or more precisely the fact that u_t has a bounded Laplacian, is sufficient to get the point-wise convergence of B_k/k^n for some *subsequence* away from some (non-explicit) null set E (see Theorem 2.1 in [5]) which is enough to run the approximation argument. But with a bit of imagination one could envisage future situations where the more precise $C^{1,1}$-regularity would be needed.

Let us finally point out that in a very recent article Darvas and Rubinstein [12] consider psh-envelopes of functions of the form $f=\min\{f_1,f_2,\dots,f_m\}$. Such envelopes appear in the Legendre transform type formula for weak geodesics introduced in [12] which has remarkable applications to the study of the completion of the Mabuchi metric space [10]. The same technique from [3] we describe here implies $C^{1,1}$-regularity of such envelopes in the case the Kähler class is integral (see the first point in Section 2.3). In [12] the authors give a different proof of this result (still using [3]) and also prove a Laplacian bound in the case of a general Kähler class.

2. $C^{1,1}$-regularity of solutions to complex Monge–Ampère equations over products

2.1. Notation: quasi-psh functions vs metrics on line bundles

Here we will briefly recall the notion for (quasi-) psh functions and metrics on line bundles that we will use. Let (X,ω_0) be a compact complex manifold of dimension n equipped with a fixed Kähler form ω_0, i.e., a smooth real positive closed $(1,1)$-form on X. Denote by $PSH(X,\omega_0)$ be the space of all ω_0-psh functions u on X, i.e., $u \in L^1(X)$ and u is (strongly) upper-semicontinuous (usc) and

$$\omega_u := \omega_0 + \frac{i}{2\pi}\partial\bar{\partial}u := \omega_0 + dd^c u \geq 0,$$

holds in the sense of currents.

We will write $\mathcal{H}(X, \omega_0)$ for the interior of $PSH(X, \omega_0) \cap \mathcal{C}^\infty(X)$, i.e., the space of all Kähler potentials (w.r.t ω_0). In the *integral case*, i.e., when $[\omega] = c_1(L)$ for a holomorphic line bundle $L \to X$, the space $PSH(X, \omega_0)$ may be identified with the space $\mathcal{H}_L$ of (singular) Hermitian metrics on L with positive curvature current. We will use additive notion for metrics on L, i.e., we identify a Hermitian metric $\|\cdot\|$ on L with its "weight" ϕ. Given a covering (U_i, s_i) of X with local trivializing sections s_i of $L_{|U_i}$ the object ϕ is defined by the collection of open functions $\phi_{|U_i}$ defined by

$$\|s_i\|^2 = e^{-\phi_{|U_i}}.$$

The (normalized) curvature ω of the metric $\|\cdot\|$ is the globally well-defined $(1,1)$-current defined by the following local expression:

$$\omega = dd^c \phi_{|U_i}.$$

The identification between $\mathcal{H}_L$ and $PSH(X, \omega_0)$ referred to above is obtained by fixing ϕ_0 and identifying ϕ with the function $u := \phi - \phi_0$, so that $dd^c\phi = \omega_u$.

2.2. The $C^{1,1}$-regularity of weak geodesics

Let (X, ω) be a compact Kähler manifold and D a domain in $\mathbb{C}^n$. Set $M := X \times D$ and denote by π the natural projection from M to X. Given a continuous function f on $\partial M (= X \times \partial D)$ we define the following point-wise Perron type upper envelope on the interior of M :

$$U := P(f) := \sup\{V : \; V \in \mathcal{F}\}, \tag{2.1}$$

where $\mathcal{F}$ denotes the set of all $V \in PSH(M, \pi^*\omega)$ such that $V_{|\partial M} \leq f$ on the boundary ∂M (in a point-wise limiting sense). In the case when D is a smoothly bounded pseudoconvex domain and f is ω-psh in the "X-directions", i.e., $f(\cdot, t) \in PSH(X, \omega)$ it was shown in [4] that $P(f)$ is continuous up to the boundary of M and U then coincides with the unique solution of the Dirichlet problem for the corresponding complex Monge–Ampère operator with boundary data f, in the weak sense of pluripotential theory [2]. Here we will establish the following higher-order regularity result for the envelope $P(f)$:

Theorem 2.1. *Let (X, ω_0) be an n-dimensional integral compact Kähler manifold manifold and D a bounded domain in $\mathbb{C}^m$ and set $M := X \times D$. Then, given f a function on ∂M such that $f(\cdot, \tau)$ is in $C^{1,1}(X)$, with a uniform bound on the corresponding real Hessians, the function $u_\tau := P(f)_{|X \times \{\tau\}}$ is in $C^{1,1}(X)$ and satisfies*

$$\sup_X |\nabla^2 u_\tau|_{\omega_0} \leq C,$$

where $|\nabla^2 u_\tau|_{\omega_0}$ denotes the point-wise norm of the real Hessian matrix of the function u_τ on X defined with respect to the Kähler metric ω_0. Moreover, the constant C only depends on an upper bound on the sup norm of the real Hessians of f_τ for $\tau \in \partial D$. In the case when D is the unit ball the function $U(x,\tau)$ is in $C^{1,1}_{\text{loc}}$ in the interior of M.

2.2.1. Proof of Theorem 2.1. In the course of the proof of the theorem we will identify an $\pi^*\omega$-psh function U on M with a positively curved metric Φ on the line bundle $\pi^*L \to M$. The case when D is a point is the content of Theorem 1.1 in [3] and as will be next explained the general case can be proved in completely analogous manner. First recall that the argument in [3] is modelled on Bedford–Taylor's proof of the case when X is a point and D is the unit-ball [2] (see also Demailly's simplifications [13]). The latter proof uses that B is a homogeneous domain. In order to explain the idea of the proof of Theorem 2.1 first consider the case when (X, L) is *homogeneous*, i.e., the group Aut (X, L) of all biholomorphic automorphisms of X lifting to L acts transitively on X. In particular, there exists a family F_λ in Aut (X, L) parametrized by $\lambda \in \mathbb{C}^n$ such that, for any fixed point $x \in X$, the map $\lambda \mapsto F_\lambda(x)$ is a biholomorphism (onto its image) from a sufficiently small ball centered at the origin in $\mathbb{C}^n$. Given a metric ϕ on L we set $\phi^\lambda := F_\lambda^*\phi$. Similarly, given a metric $\Phi(= \Phi(x,\tau))$ on π^*L we set

$$\Phi^\lambda := (F_\lambda \times I)^*\Phi.$$

Since F_λ is holomorphic the metric Φ^λ has positive curvature iff Φ has positive curvature. Now to first prove a Lipschitz bound on $P\Phi_f$, where Φ_f is the metric on $L \to \partial M$ corresponding to the given boundary data f, we take any candidate Ψ for the sup defining $P\Phi_f$ and note that, on ∂M, i.e., for $\tau \in \partial D$:

$$\Psi^\lambda \leq \Phi_f^\lambda \leq \Phi_f + C_1|\lambda|, \tag{2.2}$$

where C_1 only depends on the Lipschitz bounds in the "X-direction" of the given function f on $X \times \partial D$. But this means that $\Psi^\lambda - C_1|\lambda|$ is also a candidate for sup defining $P\Phi_f$ and hence $\Psi^\lambda - C_1|\lambda| \leq P\Phi_f$ on all of $X \times D$. Finally, taking the sup over all candidates Ψ gives, on $X \times D$, that

$$(P\Phi_f)^\lambda \leq (P\Phi_f) + C_1|\lambda|.$$

Since this holds for any λ and in particular for $-\lambda$ this concludes the proof of the desired Lipschitz bound on $P\Phi_f$. Next, to prove the bound on the real Hessian one first replaces Ψ^λ in the previous argument with $\frac{1}{2}(\Psi^\lambda + \Psi^{-\lambda})$ and deduces, precisely as before, that

$$\frac{1}{2}\left((P\Phi_f)^\lambda + (P\Phi_f)^{-\lambda}\right) \leq (P\Phi_f) + C_2|\lambda|^2,$$

where now C_2 depends on the upper bound in the "X-direction" of the real Hessian of the function f on $X \times \partial D$. The previous inequality implies an upper bound on the real Hessians of the local regularizations Ψ_ϵ of $P\Phi_f$ defined by local convolutions. Moreover, since $dd^c\Psi_\epsilon \geq 0$ it follows from basic linear algebra that a lower bound on the real Hessians also holds. Hence, letting $\epsilon \to 0$ shows that $P\Phi_f$ is in $C^{1,1}_{\mathrm{loc}}$ in the "X-direction" with a uniform upper bound on the real Hessians (compare [2, 13]).

Of course, a general polarized manifold (X, L) may not admit even a single (non-trivial) holomorphic vector field. But as shown in [3] this problem can be circumvented by passing to the total space Y of the dual line bundle $L^* \to X$, which does admit an abundance of holomorphic vector fields. The starting point is the standard correspondence between positively curved metrics ϕ on L and psh "log-homogeneous" functions χ on Y induced by the following formula:

$$\chi(z, w) = \phi(z) + \log |w|^2,$$

where z denotes a vector of local coordinates on X and (z, w) denote the corresponding local coordinates on Y induced by a local trivialization of L. Accordingly, the envelope $P\Phi_f$ on X corresponds to an envelope construction on Y, defined w.r.t the class of psh log-homogenous functions on Y. Fixing a metric ϕ_0 on L we denote by K the compact set in Y defined by the corresponding unit circle bundle. By homogeneity any function χ as above is uniquely determined by its restriction to K. Now, for any fixed point y_0 in K there exists an $(n+1)$-tuple of global holomorphic vector fields V_i on Y defining a frame in a neighborhood of y_0:

Lemma 2.2. *Given any point y_0 in the space Y^* defined as the complement of the zero-section in the total space of L^* there exist holomorphic vector fields $V_1, \ldots,$ V_{n+1} on Y^* which are linearly independent close to y_0.*

Proof. This follows from Lemma 3.7 in [3]. For completeness and since we do not need the explicit estimates furnished by Lemma 3.7 in [3] we give a short direct proof here. Set $Z := \mathbb{P}(L^* \oplus \mathbb{C})$, viewed as the fiber-wise $\mathbb{P}^1$-compactification of Y. Denote by π the natural projection from Z to X and by $\mathcal{O}(1)$ the relative (fiber-wise) hyper plane line bundle on Z. As is well known, for any sufficiently positive integer the line bundle $L_m := (\pi^* L) \otimes \mathcal{O}(1)^{\otimes m}$ on Z is ample and holomorphically trivial on Y^*. As a consequence, the rank $n+1$-vector bundle $E := TZ \otimes L_m^{\otimes k}$ is globally generated for k sufficiently large, i.e., any point z_0 in Z there exists global holomorphic sections $S_1, \ldots, S_{n+1}$ spanning $E_{|z_0}$. Since, L_m is holomorphically trivial on $Y^* \subset Z$ this concludes the proof. $\square$

Now, integrating the (short-time) flow of the holomorphic vector field $V(\lambda) := \sum \lambda_i V_i$ gives a family of holomorphic maps $F_\lambda(y)$ defined for $y \in K$ and λ in a sufficiently small ball B centered at the origin in $\mathbb{C}^{n+1}$ such that $\lambda \mapsto F_\lambda(y_0)$ is a biholomorphism. However, the problem is that the corresponding function $\chi^\lambda := F_\lambda^* \chi$ is only defined in a neighborhood of K in Y (and not log-homogeneous). But this issue can be bypassed by replacing χ^λ with a new function that we will denote by $T(\chi^\lambda)$, where $T(f)$, for f a function on K, is obtained by first taking the sup of f over the orbits of the standard S^1-action on Y to get an S^1-invariant function $g := \hat{f}$ and then replacing g with its log-homogeneous extension $\tilde{g}$, i.e.,

$$T(f) := \widetilde{\left(\widehat{\chi^\lambda}\right)}.$$

The following lemma follows from basic properties of plurisubharmonic functions (see [3] for a proof):

Lemma 2.3. *If f is the restriction to the unit circle bundle $K \subset Y$ of a psh function, then $T(f)$ is a psh log-homogeneous function on Y*

Now performing the previous constructions for any fixed $\tau \in D$ and identifying a candidate Ψ with a function χ on $Y \times D$, as above, gives

$$\chi^{\lambda}(y_0) \leq \widehat{\chi^{\lambda}}(y_0) = \widetilde{\left(\widehat{\chi^{\lambda}}\right)}(y_0) := T(\chi^{\lambda})(y_0). \tag{2.3}$$

But, by construction, for $\tau \in \partial D$ we have $T(\chi^{\lambda}) \leq T(\chi^{\lambda}_{\Phi_f})$ and since f_τ is assumed Lipschitz for $\tau \in \partial D$ we also have that

$$T(\chi^{\lambda}_{\Phi_f}) \leq T(\chi_{\Phi_f}) + C_1|\lambda| = \chi_{\Phi_f} + C_1|\lambda|.$$

But this means that $T(\chi^{\lambda}) - C_1|\lambda|$ is a candidate for the sup in question and hence bounded from above by $\chi_{P\Phi_f}$, which combined with the inequality 2.3 gives

$$\chi^{\lambda}(y_0) - C_1|\lambda| \leq \chi_{\Phi_f}(y_0).$$

Taking the sup over all candidates χ and replacing λ with $-\lambda$ hence gives the desired Lipschitz bound on $P\Phi_f$ at the given point y_0 and hence, by compactness, for any point in K. The estimate on the Hessian then proceeds precisely as above.

Finally, in the case when B is the unit ball one can exploit that B is homogeneous (under the action of the Möbius group), replacing the holomorphic maps $(x, \tau) \mapsto (F_\lambda(x), \tau)$ used above with $(x, \tau) \mapsto (F_\lambda(x), G_a(\tau))$, where G_a is a suitable family of Möbius transformations (the case when X is point is precisely the original situation in [2]). Then the proof proceeds precisely as before.

2.3. Further remarks

- The proof of the previous theorem also applies in the more general situation where f may be written as $f = \inf_{\alpha \in A} f_\alpha$ for a given family of functions f_α, as long as the Hessians of $f_\alpha(\tau, \cdot)$ are uniformly bounded on X (by a constant C independent of τ and α)) and similarly for the Lipschitz bound. Indeed, then equation 2.2 holds with f replaced by f_α for any $\alpha \in A$ with the same constant C. For D equal to a point this result has been obtained in [12] using a different proof.
- As shown in [4] (using a different pluripotential method), in the case of a general, possibly non-integral, Kähler class $[\omega]$ a bounded Laplacian in the X-directions of the boundary data f results in a bounded Laplacian of the corresponding envelope. In the case of geodesics this result has also recently been obtained in [19] by refining Chen's proof.
- By the proof of the previous theorem, the Lipschitz norm $\|u_t\|_{C^{0,1}(X)}$ of a weak geodesic u_t only depends on an upper bound on the Lipschitz norms of u_0 and u_1. Since the Lipschitz norm in the t-variable is controlled by the C^0-norm of $u_0 - u_1$ [6] it follows that the Lipschitz norm $\|U\|_{C^{0,1}(X \times A)}$ of the

corresponding solution U on $X \times A$ is controlled by the Lipschitz norms of u_0 and u_1 and the C^0-norm of $u_0 - u_1$. For a general Kähler class this result also follows from Blocki's gradient estimate [7, 8].

Acknowledgment

I am grateful to Jean-Pierre Demailly for many stimulating and illuminating discussions on the topic of the present note. Also thanks to Slawomir Dinew for comments and Yanir Rubinstein for discussions related to [12]. This work has been supported by grants from the Swedish and European Research Councils and the Wallenberg Foundation.

References

[1] Bedford, E.; Fornæss, J.-E.: Counterexamples to regularity for the complex Monge–Ampère equation. Invent. Math. 50 (1978/79), no. 2, 129–134

[2] Bedford, Eric; Taylor, B. A.: The Dirichlet problem for a complex Monge–Ampère equation. Invent. Math. 37 (1976), no. 1, 1–44

[3] Berman, R. J.: Bergman kernels and equilibrium measures for line bundles over projective manifolds. The Amer. J. of Math., Vol. 131, Nr. 5, October 2009

[4] Berman, R. J.; Demailly, J.-P.: Regularity of plurisubharmonic upper envelopes in big cohomology classes. In "Perspectives in Analysis, Geometry, and Topology", *Springer-Verlag*

[5] Berman, R. J.; Berndtsson, B.: Convexity of the K-energy on the space of Kähler metrics and uniqueness of extremal metrics. arXiv:1405.0401

[6] Berndtsson, B.: A Brunn–Minkowski type inequality for Fano manifolds and some uniqueness theorems in Kähler geometry. arXiv:1303.4975

[7] Blocki, Z.: A gradient estimate in the Calabi–Yau theorem, Math. Ann. 344 (2009), 317–327

[8] Blocki, Z.: On geodesics in the space of Kähler metrics, Proceedings of the "Conference in Geometry" dedicated to Shing-Tung Yau (Warsaw, April 2009), in "Advances in Geometric Analysis", eds. S. Janeczko, J. Li, D. Phong, Advanced Lectures in Mathematics 21, pp. 3–20, International Press, 2012

[9] Chen, X. X.: The space of Kähler metrics, J. Diff. Geom. 56 (2000), 189–234

[10] Darvas, T.: The Mabuchi geometry of finite energy classes. Adv. Math. 285 (2015), 182–219

[11] Darvas, T.; Lempert, L.: Weak geodesics in the space of Kähler metrics. Math. Res. Lett. 19 (2012), no. 5, 1127–1135

[12] Darvas, T.; Rubinstein, Y. A.: Kiselman's principle, the Dirichlet problem for the Monge–Ampère equation, and rooftop obstacle problems. J. of the Math. Soc. of Japan 68, no. 2 (2016): 773–796

[13] Demailly, J.-P.: Potential Theory in Several Complex Variables. Notes from the Trenot conference in 1992.
http://www-fourier.ujf-grenoble.fr/~demailly/documents.html

[14] Donaldson, S. K.: Remarks on gauge theory, complex geometry and 4-manifold topology. Fields Medallists' lectures, 384–403, World Sci. Ser. 20th Century Math., 5, World Sci. Publ., River Edge, NJ, 1997

[15] Donaldson, S. K.: Symmetric spaces, Kähler geometry and Hamiltonian dynamics. In Northern California Symplectic Geometry Seminar, volume 196 of Amer. Math. Soc. Transl. Ser. 2, pages 13–33. Amer. Math. Soc., Providence, RI, 1999

[16] Donaldson, S. K.: Holomorphic discs and the complex Monge–Ampère equation. J. Symplectic Geom. 1 (2002), no. 2, 171–196

[17] Gamelin, W.; Sibony, N.: Subharmonicity for uniform algebras, J. of Funct. Analysis 35(1980), 64–108

[18] Guedj, V.: The metric completion of the Riemannian space of Kähler metrics. arXiv:1401.7857

[19] He, W.: On the space of Kähler potentials. Comm. Pure Appl. Math. 68 (2015), no. 2, 332–343

[20] Lempert, L.; Vivas, L.: Geodesics in the space of Kähler metrics. Duke Math. J. Volume 162, Number 7 (2013), 1369–138

[21] Mabuchi, T.: Some symplectic geometry on compact Kähler manifolds. I. Osaka J. Math., 24(2):227–252, 1987

[22] Ross, J.: Witt Nystrom, D.: Envelopes of positive metrics with prescribed singularities. Preprint

[23] Semmes, S.: Complex Monge–Ampère and symplectic manifolds. Amer. J. Math., 114(3):495–550, 1992

Robert J. Berman
Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
412 96 Gothenburg, Sweden
e-mail: robertb@chalmers.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 121–126

A Comparison Principle for Bergman Kernels

Bo Berndtsson

Dedicated to the memory of Mikael Passare

Abstract. We give a version of the comparison principle from pluripotential theory where the Monge–Ampère measure is replaced by the Bergman kernel and use it to derive a maximum principle.

1. Introduction

Let ϕ and ψ be two plurisubharmonic functions in a complex manifold X, and let Ω be a relatively compact subdomain in X. Assume that on the boundary of Ω, $\phi \leq \psi$, and that inside the domain the Monge–Ampère measures of ϕ and ψ satisfy

$$(dd^c\phi)^n \geq (dd^c\psi)^n.$$

Then the maximum principle for the Monge–Ampère equation asserts that the inequality $\phi \leq \psi$ holds inside the domain Ω too. (Here of course both the inequality between ϕ and ψ on the boundary and the Monge–Ampère equation have to be given a precise meaning.) The maximum principle is easy to prove if the functions are sufficiently smooth, e.g., of class $\mathcal{C}^2$. For non-regular functions the maximum principle can be derived from the so-called comparison principle (see [2]) of Bedford and Taylor, which also serves as a substitute for the maximum principle in some cases. The comparison principle states (again omitting precise assumptions) that

$$\int_{\{\psi<\phi\}} (dd^c\phi)^n \leq \int_{\{\psi<\phi\}} (dd^c\psi)^n.$$

On the other hand it is well known that Monge–Ampère measures often can be approximated by measures defined by *Bergman functions*. Suppose that we have given on our manifold X a positive measure, μ, and consider the L^2-space of holomorphic functions

$$A^2 = A^2(X,\mu,\phi) = \left\{ h \in H(X); \int |h|^2 e^{-\phi} d\mu < \infty \right\},$$

or its closure in $L^2(X,\mu,\phi)$.

We denote by $K_\phi(z,\zeta)$ the Bergman kernel for A^2 and let

$$B_\phi(z) = K_\phi(z,z)e^{-\phi}$$

be the Bergman function, also known as the *density of states function*. It is a consequence of the asymptotic expansion formula of Tian–Catlin–Zelditch (see [3]) that we have

$$\lim_{k\to\infty} k^{-n} B_{k\phi} d\mu = c_n (dd^c\phi)^n$$

if ϕ is plurisubharmonic and ϕ and μ are sufficiently regular. We can therefore think of $B_\phi d\mu$ as an approximation, or perhaps quantization, of the Monge–Ampère measure of ϕ.

The main observation in this note is that a version of the comparison principle holds if we replace the Monge–Ampère operator by the density of states function, so that

$$\int_{\psi<\phi} B_\phi d\mu \leq \int_{\psi<\phi} B_\psi d\mu.$$

As it turns out, this is an almost completely formal phenomenon, and it holds under very (but not completely) general circumstances. In particular, the plurisubharmonicity of ϕ and ψ plays no role at all, and even the holomorphicity of functions in A^2 enters only in a very weak form, so similar results also hold in many other situations when we have a well-behaved Bergman kernel, and also if we consider sections of line bundles instead of scalar-valued functions. However, the setting of plurisubharmonic weights and holomorphic functions allows a slightly stronger statement with strict inequality, and in that context our main theorem is as follows.

Theorem 1.1. *Let L be a holomorphic line bundle over a complex manifold X, and let ϕ and ψ be two, possibly singular, metrics on L. Suppose that $dd^c\phi \geq -\omega$ and $dd^c\psi \geq -\omega$ for some smooth Hermitian $(1,1)$-form ω. Assume also that for some constant C, $\phi \leq \psi + C$ and that μ is given by a strictly positive continuous volume form. Then*

$$\int_{\psi<\phi} B_\phi d\mu \leq \int_{\psi<\phi} B_\psi d\mu. \tag{1.1}$$

Moreover, if $\emptyset \neq \{\psi < \phi\} \neq X$ and if B_ψ is not identically equal to 0, then strict inequality holds if the left integral is finite.

A few remarks are in order. The strict inequality is of less importance when we deal with Monge–Ampère measures, since one can often arrange that by an ad hoc small perturbation. For Bergman kernels this is less clear and that is the reason why we mention the (very weak) conditions for strict inequality. The condition that $\phi \leq \psi + C$ is sometimes phrased as 'ψ is less singular than ϕ', and some condition like that is necessary. Indeed, if $\psi < \phi$ everywhere and we assume X compact, the two integrals equal the dimensions of the space of sections of L that are square integrable with respect to the respective metrics. If ψ is more singular than ϕ it may well happen that the space of sections that have finite norm measured by ψ is smaller than the space of sections that have finite norm measured by ϕ, so the inequality cannot hold.

2. The abstract setting

We will first deal with the abstract setting of general L^2-spaces with a Bergman kernel. Let (X, μ) be a measure space, let $e^{-\phi}$ be a measurable weight function on X, and let $\mathcal{H}_\phi$ be a Hilbert subspace of $L^2(e^{-\phi}d\mu)$. We assume that for any $z \in X$, point evaluation at z is a bounded linear functional on $\mathcal{H}_\phi$. Then $\mathcal{H}_\phi$ has a Bergman kernel, $K_\phi(z, \zeta)$ and we denote $B_\phi(z) = K_\phi(z, z)e^{-\phi}$.

Theorem 2.1 (Comparison principle for Bergman spaces). *Let ϕ and ψ be two weight functions on X such that for some constant C, $\phi \leq \psi + C$. Then*

$$\int_{\psi<\phi} B_\phi d\mu \leq \int_{\psi<\phi} B_\psi d\mu. \tag{2.1}$$

To prove the comparison principle we need a, basically standard, lemma on derivatives of Bergman kernels.

Proposition 2.2. *Let ϕ_t be a differentiable family of weight functions with uniformly bounded derivative with respect to t. Put $K_t = K_{\phi_t}$. Then K_t is differentiable with respect to t. Let $\dot{K}_t$ and $\dot{\phi}_t$ be the derivatives of K_t and ϕ_t with respect to t. Then for z and ζ fixed*

$$\dot{K}_t(z, \zeta) = \int_X \dot{\phi}_t K_t(z, w)K_t(w, \zeta)e^{-\phi_t(w)}d\mu(w). \tag{2.2}$$

Moreover, for the difference quotients we have, if $|\tau| \leq 1$,

$$|(K_{t+\tau}(z, z) - K_t(z, z))/\tau| \leq AK_t(z, z) \tag{2.3}$$

for some constant A depending on the sup-norm of $\dot{\phi}$.

Proof. Note first that since $\dot{\phi}$ is bounded, $\phi_t - \phi_{t+\tau}$ is bounded for $|\tau| \leq 1$. Let $\Delta(t, \tau) = e^{-\phi_t} - e^{-\phi_{t+\tau}}$. Since

$$\Delta(t, \tau) = \int_0^\tau \dot{\phi}_s e^{-\phi_{t+s}}ds,$$

$|\Delta(t, \tau)| \leq A|\tau|e^{-\phi_t}$ if $|\tau| \leq 1$. Next note that by the reproducing property of Bergman kernels

$$(K_{t+\tau} - K_t)(z, \zeta) = \int_X K_t(z, w)K_{t+\tau}(w, \zeta)(e^{-\phi_t} - e^{-\phi_{t+\tau}})d\mu(w). \tag{2.4}$$

Hence for $|\tau| \leq 1$

$$|(K_{t+\tau}(z, z) - K_t(z, z))/\tau| \leq A \int_X |K_t(z, w)K_{t+\tau}(w, z)|e^{-\phi_t}d\mu(w).$$

Since $\phi_t - \phi_{t+\tau}$ is bounded this is less than

$$A\left(\int_X |K_t(z, w)|^2 e^{-\phi_t}d\mu(w) + \int_X |K_{t+\tau}(z, w)|^2 e^{-\phi_{t+\tau}}d\mu(w)\right) \leq A'K_t(z, z),$$

so we have proved (2.3). To prove (2.2) is very easy formally, just differentiating under the integral sign, but to prove that this is legitimate we have to work a bit more. We first multiply (2.4) by its conjugate and integrate with respect to ζ.

Letting $f(z,\zeta) := (K_{t+\tau} - K_t)(z,\zeta)$ we get

$$\int |f(z,\zeta|^2 e^{-\phi_{t+\tau}} d\mu(\zeta) = \int K_t(z,w)K_t(w',z)\Delta(t,\tau)(w)\Delta(t,\tau)(w') \\ \times \int K_{t+\tau}(\zeta,w')K_{t+\tau}(w,\zeta)e^{-\phi_{t+\tau}(\zeta)}d\mu(\zeta)d\mu(w)d\mu(w').$$

Using the reproducing property of Bergman kernels in the inner integral this is

$$\int K_t(z,w)K_t(w',z)K_{t+\tau}(w,w')\Delta(t,\tau)(w)\Delta(t,\tau)(w')d\mu(w)d\mu(w').$$

Next we apply (2.4) to the integral with respect to w' and get

$$\int f(w,z)K_t(z,w)\Delta(t,\tau)(w)d\mu(w).$$

Then use that $|\Delta(t,\tau)| \leq |\tau|e^{-\phi_t}$ and apply Cauchy's inequality to get

$$\int |f(z,\zeta)|^2 e^{-\phi_t} d\mu(\zeta) \leq A|\tau|K_t(z,z). \tag{2.5}$$

We are now finally ready to prove (2.2). By (2.4)

$$(K_{t+\tau} - K_t)(z,\zeta)/\tau = \int_X K_t(z,w)K_{t+\tau}(w,\zeta)(e^{-\phi_t} - e^{-\phi_{t+\tau}})/\tau d\mu(w).$$

By (2.5) we may replace $K_{t+\tau}$ by K_t in the integral. After that we let τ tend to zero and get (2.2) by dominated convergence. □

We now turn to the proof of the comparison principle Theorem 2.1. We first claim that we may assume that $\phi - \psi$ is bounded. To see this, let $u := \psi - \phi$ so that $u \geq -C$. Put $u_0 := \min(u,0)$, $\psi_0 = \phi + u_0$. Then $\psi_0 \leq \psi$ and $\psi_0 - \phi$ is bounded. By the extremal characterization of Bergman kernels $K_{\psi_0}(z,z) \leq K_\psi(z,z)$. On the other hand, where $\psi < \phi$, $u < 0$ so $u_0 = u$. Hence $\psi_0 = \psi$ and $B_{\psi_0} \leq B_\psi$. Moreover $\psi < \phi$ if and only if $u < 0$ which is equivalent to $u_0 < 0$, so $\psi < \phi$ if and only if $\psi_0 < \phi$. Hence it suffices to prove the theorem for ψ_0 since then

$$\int_{\psi<\phi} B_\phi d\mu \leq \int_{\psi_0<\phi} B_{\psi_0} d\mu \leq \int_{\psi<\phi} B_\psi d\mu.$$

From now on we assume that $\phi - \psi$ is bounded and let ρ be a measurable function on X such that

$$\int_X \rho(z)K_\phi(z,z)e^{-\phi}d\mu(z) < \infty.$$

The same integral with ϕ replaced by ψ is then also bounded. Let $\phi_t = \phi + tu$, so that $\phi_0 = \phi$ and $\phi_1 = \psi$. Then we claim that by Proposition 2.2, if

$$G(t) := \int_X \rho(z)B_{\phi_t} d\mu$$

then

$$G'(t) = \int_X -\rho(z)\dot{\phi}_t(z)K_t(z,z)e^{-\phi_t}d\mu \\ + \int_X\int_X \rho(z)\dot{\phi}_t(w)K_t(z,w)K_t(w,z)e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w). \tag{2.6}$$

Again, this follows formally by the proposition and to justify the limit process we write

$$(G(t+\tau)-G(t)) = \int_X \rho K_t(e^{-\phi_{t+\tau}} - e^{-\phi_t})d\mu + \int_X \rho(K_{t+\tau}-K_t)e^{-\phi_{t+\tau}}d\mu.$$

When we divide by τ and let $\tau \to 0$ we see that the first term converges to the first term of (2.6) by dominated convergence. For the second term we use (2.3) to conclude that we have dominated convergence in that integral as well.

In the first integral on the right-hand side we insert the reproducing formula for the Bergman kernel

$$K_t(z,z) = \int_X K_t(z,w)K_t(w,z)e^{-\phi_t(w)}d\mu(w)$$

which changes the right-hand side to

$$\int_X\int_X \rho(z)(\dot{\phi}_t(w)-\dot{\phi}_t(z))K_t(z,w)K_t(w,z)e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w).$$

We can write this more symmetrically as

$$(1/2)\int_X\int_X (\rho(z)-\rho(w))(\dot{\phi}_t(w)-\dot{\phi}_t(z))|K_t(z,w)|^2e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w). \tag{2.7}$$

Now recall that $\phi_t = \phi + tu$ so $\dot{\phi}_t = u$. Let ρ be the characteristic function of the set where $\psi - \phi = u < 0$. Then (2.7) becomes

$$(1/2)\int\int_{\{u(z)<0<u(w)\}} (u(w)-u(z))|K_t(z,w)|^2e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w) \\ -(1/2)\int\int_{\{u(w)<0<u(z)\}} (u(w)-u(z))|K_t(z,w)|^2e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w).$$

Again using symmetry we get

$$\frac{d}{dt}\int_{u<0} B_{\phi_t}d\mu \\ = \int\int_{\{u(z)<0<u(w)\}} (u(w)-u(z))|K_t(z,w)|^2e^{-\phi_t(z)-\phi_t(w)}d\mu(z)d\mu(w). \tag{2.8}$$

Clearly this expression is non-negative, so we have proved Theorem 2.1.

Remark. Since the Bergman function $B_\phi(z) = K_\phi(z,z)e^{-\phi}$ does not change if we add a constant to ϕ, we also have that $\int_{\psi<\phi+c} B_\phi d\mu \leq \int_{\psi<\phi+c} B_\psi d\mu$ for any constant c, as soon as the sets $\{\psi < \phi + c\}$ and $\{\psi > \phi + c\}$ are both nonempty.

3. The proof of Theorem 1.1

It is now an easy matter to deduce Theorem 1.1 from Theorem 2.1. First we note that the setting of line bundles instead of scalar-valued functions causes no extra difficulties. Indeed the proof goes through in the same way with only nominal changes. Alternatively, we could use that any line bundle has a discontinuous trivializing section and since continuity played no role in the proof, the line bundle case follows. It remains to prove that we have strict inequality if B_ψ is non trivial and $\emptyset \neq \{\psi < \phi\} \neq X$. For this it suffices to show that the right-hand side of (2.8) is strictly positive. But

$$V := \{(z,w); u(w) < 0 < u(z)\}$$

is by assumption nonempty. Moreover, this set is open for the plurifine topology and therefore has positive Lebesgue measure, [1]. Hence it has positive μ-measure if μ is given by a strictly positive continuous density. From this it follows that $K_t(z,w)$ is different from zero almost everywhere on V, since it is holomorphic with respect to z and w (this is the only place we use holomorphicity), so it follows that the derivative of G is strictly positive.

Finally we give a 'maximum principle' for Bergman kernels which follows from Theorem 1.1 in the same way that the Monge–Ampère maximum principle follows from the classical comparison principle.

Theorem 3.1. *Under the same assumptions as in Theorem* 1.1*, let $\Omega \neq X$ be a subset of X such that $B_\phi \geq B_\psi$ on Ω. Assume that $\phi \leq \psi$ on $X \setminus \Omega$. Then $\phi \leq \psi$ everywhere.*

Proof. Assume the set U where $\psi < \phi$ is nonempty. Then U is a subset of Ω, and Ω is not equal to X. This contradicts Theorem 1.1. □

References

[1] Bedford E. and Taylor, B. A.: Fine topology, Shilov boundary and $(dd^c)^n$. J. Func. Anal. 72 (1987), pp. 225–251

[2] Kolodziej, S.: The Complex Monge–Ampère equation and pluripotential theory. Mem. Am. Math. Soc., 178 (2005)

[3] Zelditch, S.: Szegö kernels and a theorem of Tian. Internat. Math. Res. Notices 1998, no. 6, pp. 317–331

Bo Berndtsson
Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg, Sweden
e-mail: bob@chalmers.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 127–133

Suita Conjecture from the One-dimensional Viewpoint

Zbigniew Błocki

Dedicated to the memory of Mikael Passare

Abstract. The Suita conjecture predicted the optimal lower bound for the Bergman kernel of a domain on the plane in terms of logarithmic capacity. It was recently proved as a special case of the optimal version of the Ohsawa–Takegoshi extension theorem. We present here a purely one-dimensional approach that should be suited to readers not interested in several complex variables.

Introduction

For a domain Ω in $\mathbb{C}$ by $A^2(\Omega)$ we denote the space of holomorphic square integrable functions in Ω. The Bergman kernel K_Ω is defined by the reproducing property

$$f(w) = \int_\Omega f\overline{K_\Omega(\cdot, w)}d\lambda, \quad f \in A^2(\Omega),\ w \in \Omega.$$

On the diagonal we have

$$K_\Omega(w, w) = ||K_\Omega(\cdot, w)||^2 = \sup\{|f(w)|^2 \colon f \in A^2(\Omega),\ ||f|| \leq 1\}$$

where $||\cdot||$ denotes the L^2-norm. Suita [17] conjectured that

$$c_\Omega(w)^2 \leq \pi K_\Omega(w, w) \tag{1}$$

where

$$c_\Omega(w) = \exp \lim_{z \to w} (G_\Omega(z, w) - \log|z - w|)$$

is the logarithmic capacity of the complement of Ω with respect to w. Here G_Ω is the Green function, it is the maximal negative function such that $G_\Omega(\cdot, w) - \log|\cdot - w|$ is harmonic in Ω (or $\equiv -\infty$).

Partially supported by the Ideas Plus grant 0001/ID3/2014/63 of the Polish Ministry of Science and Higher Education.

Ohsawa [15] was the first to notice that the right approach is to treat it an L^2-extension problem: one has to construct holomorphic f in Ω such that $f(w) = 1$ and $||f||^2 \leq \pi/c_\Omega(w)^2$. Using the methods of the original proof of the Ohsawa–Takegoshi extension theorem [16] he managed to show the estimate

$$c_\Omega(w)^2 \leq C\pi K_\Omega(w, w)$$

with $C = 750$. This was improved to $C = 2$ in [3] and to $C = 1.95388\ldots$ by Guan–Zhou–Zhu [11] who proved the extension theorem with this constant using an ODE with one unknown (see also [4]).

The estimate with $C = 1$ was established in [5] where also the optimal version of the Ohsawa–Takegoshi theorem was obtained. The main tool was the Hörmander L^2-estimate [12] for the $\bar\partial$-equation as well as some ideas of Chen [8] who was the first to show that the extension theorem (without an optimal constant) can be deduced directly from this estimate. One of the key steps was a solution of an ODE with two unknowns. Guan–Zhou [9, 10] later proved some generalizations of the extension theorem with optimal constant but similarly as in [5] the key was essentially the same ODE with two unknowns.

Two other proofs of the Suita conjecture were found afterwards. Both of them gave the estimate

$$K_\Omega(w, w) \geq \frac{1}{e^{-2t}\lambda(\{G_\Omega(\cdot, w) < t\})}, \tag{2}$$

where $t \leq 0$, from which (1) easily follows when $t \to -\infty$. The first from [6] used the tensor-power trick and thus effectively needed an arbitrarily high dimension in order to obtain this one-dimensional result. The second was due to Lempert [13] who noticed that (2) can be deduced from subharmonicity property of the Bergman kernel for sections of a pseudoconvex domain in $\mathbb{C}^2$, see [14] and [2]. This way one had to use two dimensions to get the Suita conjecture. One can add that it was shown in [7] using the isoperimetric inequality that the right-hand side of (2) is monotone in t.

Using some ideas of Berndtsson [1] and essentially following the approach of Guan–Zhou [9] we will give a self-contained one-dimensional proof of the Suita conjecture (1). We will obtain the same ODE as in [5]. It would be interesting to find such a one-dimensional proof of (2). As a by-product in Section 2 we present a new formula for the Bergman kernel on the diagonal as an extremal for a family of test functions.

The author is grateful to Bo Berndtsson for finding an error in the first version of the paper.

1. Proof of the Suita conjecture

It is well known that the Bergman kernel, Green function and logarithmic capacity converge locally uniformly as Ω_j is an increasing sequence of domains whose union is Ω. Without loss of generality we may therefore assume that Ω is bounded and has smooth boundary.

We will use the notation

$$\partial = \frac{\partial}{\partial z}, \quad \bar\partial = \frac{\partial}{\partial \bar z}.$$

The following description of the Bergman kernel as a solution of the Dirichlet problem is well known, we present the proof for the sake of completeness:

Proposition 1. *For $w \in \Omega$ where Ω is a bounded domain in $\mathbb{C}$ with C^1 boundary, let v be the complex-valued harmonic function in Ω such that $v(z) = 1/(\pi\overline{(z-w)})$ for $z \in \partial\Omega$. Then $K_\Omega(\cdot, w) = \partial v$.*

Proof. It is clear that ∂v is holomorphic and we have to show that the reproducing formula is satisfied. Take $f \in A^2(\Omega)$, by the approximating property we may assume that f is defined in a neighbourhood of $\bar\Omega$. By the Cauchy–Green formula

$$f(w) = \frac{1}{2\pi i}\int_{\partial\Omega}\frac{f(z)}{z-w}dz = -\frac{i}{2}\int_{\partial\Omega} f\bar v dz = \int_\Omega f\overline{\partial v}d\lambda$$

and the result follows. □

For a real-valued $\varphi \in C^1(\Omega)$ we consider the weighted scalar product

$$\langle\alpha,\beta\rangle_\varphi = \int_\Omega \alpha\bar\beta e^{-\varphi}d\lambda$$

and the adjoint operator

$$\bar\partial^*_\varphi\alpha = -e^\varphi\partial(\alpha e^{-\varphi}) = -\partial\alpha + \alpha\partial\varphi,$$

so that

$$\langle\bar\partial^*_\varphi\alpha,\beta\rangle_\varphi = \langle\alpha,\bar\partial\beta\rangle_\varphi,$$

provided that $\varphi,\alpha,\beta \in C^1(\bar\Omega)$ are such that on $\partial\Omega$ either $\alpha = 0$ or $\beta = 0$. We have

$$\bar\partial\bar\partial^*_\varphi\alpha = \bar\partial^*_\varphi\bar\partial\alpha + \alpha\partial\bar\partial\varphi. \tag{3}$$

To prove (1) assume for simplicity that $w = 0$ and set

$$\alpha := e^\varphi(1 - \pi\bar z v), \tag{4}$$

where v is as in Proposition 1 and φ will be determined later. We have $\alpha = 0$ on $\partial\Omega$, $\alpha(0) = e^{\varphi(0)}$ and

$$\bar\partial^*_\varphi\alpha = \pi\bar z K_\Omega(\cdot,0)e^\varphi.$$

Then

$$K_\Omega(0,0) = \frac{1}{\pi^2}\int_\Omega |\bar\partial^*_\varphi\alpha|^2\frac{e^{-2\varphi}}{|z|^2}d\lambda. \tag{5}$$

We will need the following:

Proposition 2. *Assume that $\mu,\varphi \in C^2(\bar\Omega)$ are real-valued and $\alpha \in C^1(\bar\Omega)$ is such that $\alpha = 0$ on $\partial\Omega$. Then*

$$\int_\Omega \mu|\bar\partial^*_\varphi\alpha|^2 e^{-\varphi}d\lambda = \int_\Omega \left[\mu|\bar\partial\alpha|^2 + |\alpha|^2\big(\mu\partial\bar\partial\varphi - \partial\bar\partial\mu\big) + 2\Re\big(\bar\alpha\bar\partial\mu\,\bar\partial^*_\varphi\alpha\big)\right]e^{-\varphi}d\lambda.$$

Proof. We have

$$\langle \mu\bar\partial^*_\varphi\alpha, \bar\partial^*_\varphi\alpha\rangle_\varphi = \langle \bar\partial\mu\, \bar\partial^*_\varphi\alpha, \alpha\rangle_\varphi + \langle \mu\bar\partial\bar\partial^*_\varphi\alpha, \alpha\rangle_\varphi$$

and by (3)

$$\langle \mu\bar\partial\bar\partial^*_\varphi\alpha, \alpha\rangle_\varphi = \langle \bar\partial^*_\varphi\bar\partial\alpha, \mu\alpha\rangle_\varphi + \langle \alpha\partial\bar\partial\varphi, \mu\alpha\rangle_\varphi.$$

Further,

$$\langle \bar\partial^*_\varphi\bar\partial\alpha, \mu\alpha\rangle_\varphi = \langle \bar\partial\alpha, \mu\bar\partial\alpha\rangle_\varphi + \langle \bar\partial\alpha, \alpha\bar\partial\mu\rangle_\varphi$$

and

$$\langle \bar\partial\alpha, \alpha\bar\partial\mu\rangle_\varphi = \langle \alpha, \bar\partial\mu\bar\partial^*_\varphi\alpha\rangle_\varphi - \langle \alpha, \alpha\partial\bar\partial\mu\rangle_\varphi. \qquad \square$$

We obtain the following version of the Nakano inequality:

Corollary 3. *Let α and φ be as in Proposition 2. Assume that both $\mu_1 \in C^2(\bar\Omega)$ and integrable μ_2 are positive. Then*

$$\int_\Omega (\mu_1+\mu_2)|\bar\partial^*_\varphi\alpha|^2 e^{-\varphi} d\lambda \geq \int_\Omega |\alpha|^2 \left(\mu_1\partial\bar\partial\varphi - \partial\bar\partial\mu_1 - \frac{|\partial\mu_1|^2}{\mu_2}\right) e^{-\varphi} d\lambda.$$

Proof. By Proposition 2 and the Cauchy–Schwarz inequality

$$\int_\Omega \mu_1|\bar\partial^*_\varphi\alpha|^2 e^{-\varphi} d\lambda \geq \int_\Omega \left[|\alpha|^2\left(\mu_1\partial\bar\partial\varphi - \partial\bar\partial\mu_1\right) - \frac{|\alpha\bar\partial\mu_1|^2}{\mu_2} - \mu_2|\bar\partial^*_\varphi\alpha|^2\right] e^{-\varphi} d\lambda. \ \square$$

By (5) we see that we should use Corollary 3 with $\mu_1+\mu_2 = e^{-\varphi}/|z|^2$. Denote $G = G_\Omega(\cdot, 0)$ and set $\psi := 2G - \log|z|^2$, so that ψ is harmonic in Ω and $c_\Omega(0)^2 = e^{\psi(0)}$. We will look for

$$\varphi = \psi + \chi(-2G), \quad \mu_1 = e^{-\gamma(-2G)},$$

where $\chi(t)$ and $\gamma(t)$ defined for $t = -2G \geq 0$ will be determined later. Note that

$$\mu_2 = \frac{e^{-\varphi}}{|z|^2} - \mu_1 = e^{t-\chi} - e^{-\gamma}.$$

Using the fact that

$$\partial\bar\partial G = \frac{\pi}{2}\delta_0$$

we will obtain

$$\begin{aligned}
&\mu_1\partial\bar\partial\varphi - \partial\bar\partial\mu_1 - \frac{|\partial\mu_1|^2}{\mu_2} \\
&\quad = -\pi(\chi'+\gamma')e^{-\gamma}\delta_0 + 4\left(\chi''+\gamma'' - \frac{(\gamma')^2}{1-e^{\chi-\gamma-t}}\right)e^{-\gamma}|\partial G|^2 \\
&\quad = -\pi\eta' e^{-\gamma}\delta_0 + 4\left(\eta'' - \frac{(\gamma')^2}{1-e^{\eta-2\gamma-t}}\right)e^{-\gamma}|\partial G|^2,
\end{aligned}$$

where $\eta = \chi+\gamma$. It is convenient to choose γ and η satisfying $-\eta' e^{-\gamma} = 1$ and

$$\eta'' - \frac{(\gamma')^2}{1-e^{\eta-2\gamma-t}} = 0$$

which is the same equation as in [5]. We can take the solutions obtained there:

$$\eta = -\log(t + e^{-t} - 1)$$
$$\gamma = -\log(t + e^{-t} - 1) + \log(1 - e^{-t}),$$

so that

$$\chi = -\log(1 - e^{-t})$$

and

$$\varphi = \psi - \log(1 - e^{2G}) = \psi - \log(1 - |z|^2 e^{\psi}).$$

By Corollary 3, since $\alpha(0) = e^{\varphi(0)} = c_\Omega(0)^2$,

$$\int_\Omega |\bar\partial^*_\varphi \alpha|^2 \frac{e^{-2\varphi}}{|z|^2} d\lambda \geq \pi c_\Omega(0)^2$$

and it is enough to use (5) to obtain (1). (Although we have used Corollary 3 for μ_1 which is not C^2 at the origin – in fact it is of the form $\mu_1 = -2G + \rho$ where ρ is smooth – by approximation it is clear that it holds also for such a function.)

2. A Formula for the Bergman kernel

Using similar methods as before we will prove the following result:

Theorem 4. *For a domain Ω in $\mathbb{C}$ and $w \in \Omega$ one has*

$$K_\Omega(w,w) = \frac{1}{\pi^2} \inf \left\{ \int_\Omega \frac{|\partial\alpha(z)|^2}{|z-w|^2} d\lambda(z) \colon \alpha \in C_0^\infty(\Omega),\ \alpha(w) = 1 \right\}. \tag{6}$$

Proof. We may assume that $w = 0$. Take $\alpha \in C_0^\infty(\Omega)$ and $f \in A^2(\Omega)$ with $\alpha(0) = f(0) = 1$. Then $u := f/(\pi z)$ solves $\bar\partial u = \delta_0$ and

$$1 = |\alpha(0)|^2 = \left| \int_\Omega \bar\alpha\, \bar\partial u \right|^2 = \left| - \int_\Omega u\, \overline{\partial\alpha}\, d\lambda \right|^2 \leq \frac{1}{\pi^2} ||f||^2 \int_\Omega \frac{|\partial\alpha|^2}{|z|^2} d\lambda.$$

This gives $\leq$ in (6). To prove $\geq$ we first assume that Ω is bounded and has smooth boundary. Let v be harmonic in Ω and such that $v = 1/(\pi\bar z)$ on $\partial\Omega$. Then $\alpha := 1 - \pi\bar z v$ is smooth up to the boundary, vanishes there and $\alpha(0) = 1$. By Proposition 1 we have $\partial\alpha = -\pi\bar z K_\Omega(\cdot, 0)$ and it is enough to show that α can be well approximated by test forms. Let ρ be a defining function for Ω, so that $\Omega = \{\rho > 0\}$, and let $\chi \in C^\infty(\mathbb{R})$ be such that $\chi(t) = 0$ for $t \leq 1$ and $\chi(t) = 1$ for $t \geq 2$. One can easily show that for the test forms $\alpha_j := \chi(j\rho)\alpha$ one has

$$\int_\Omega \frac{|\partial\alpha_j|^2}{|z|^2} d\lambda \longrightarrow \int_\Omega \frac{|\partial\alpha|^2}{|z|^2} d\lambda$$

as $j \to \infty$. If Ω is arbitrary and $K_\Omega(0,0) < a$ then we can find $\Omega' \Subset \Omega$ with smooth boundary such that $K_{\Omega'}(0,0) < a$. By the previous part there exists $\alpha \in C_0^\infty(\Omega')$

such that $\alpha(0) = 1$ and

$$\frac{1}{\pi^2}\int_{\Omega'}\frac{|\partial\alpha|^2}{|z|^2}d\lambda < a.$$

This finishes the proof. $\square$

Similarly, for any $\varphi \in C^1(\Omega)$ one can show

$$K_\Omega(0,0) = \frac{1}{\pi^2}\inf\left\{\int_\Omega |\bar\partial^*_\varphi\alpha|^2\frac{e^{-2\varphi}}{|z|^2}d\lambda\colon \alpha \in C^\infty_0(\Omega),\ \alpha(0) = e^{\varphi(0)}\right\}.$$

If Ω is bounded with smooth boundary and $\varphi \in C^1(\bar\Omega)$ then instead of test forms we can take $\alpha \in C^1(\bar\Omega)$ with $\alpha = 0$ on $\partial\Omega$ and α given by (4) realizes the infimum.

References

[1] B. Berndtsson, *Integral formulas and the Ohsawa–Takegoshi extension theorem*, Science in China Ser. A Math. 48 suppl. (2005), 61–73

[2] B. Berndtsson, *Subharmonicity properties of the Bergman kernel and some other functions associated to pseudoconvex domains*, Ann. Inst. Fourier 56 (2006), 1633–1662

[3] Z. Błocki, *Some estimates for the Bergman kernel and metric in terms of logarithmic capacity*, Nagoya Math. J. 185 (2007), 143–150

[4] Z. Błocki, *On the Ohsawa–Takegoshi extension theorem*, Univ. Iag. Acta Math. 50 (2012), 53–61

[5] Z. Błocki, *Suita conjecture and the Ohsawa–Takegoshi extension theorem*, Invent. Math. 193 (2013), 149–158

[6] Z. Błocki, *A lower bound for the Bergman kernel and the Bourgain–Milman inequality*, Geometric Aspects of Functional Analysis, Israel Seminar (GAFA) 2011–2013, eds. B. Klartag, E. Milman, Lect. Notes in Math. 2116, pp. 53–63, Springer, 2014

[7] Z. Błocki, W. Zwonek, *Estimates for the Bergman kernel and the multidimensional Suita conjecture*, New York J. Math. 21 (2015) 151–161

[8] B.-Y. Chen, *A simple proof of the Ohsawa–Takegoshi extension theorem*, arXiv: 1105.2430

[9] Q. A. Guan, X. Y. Zhou, *Optimal constant problem in the L^2 extension theorem*, C. R. Acad. Sci. Paris, Ser. I 350 (2012) 753–756

[10] Q. A. Guan, X. Y. Zhou, *A solution of an L^2 extension problem with optimal estimate and applications*, Ann. of Math. 181 (2015), 1139–1208

[11] Q. A. Guan, X. Y. Zhou, L. F. Zhu, *On the Ohsawa–Takegoshi L^2 extension theorem and the Bochner–Kodaira identity with non-smooth twist factor*, J. Math. Pures Appl. 97 (2012), 579–601

[12] L. Hörmander, *L^2 estimates and existence theorems for the $\bar\partial$ operator*, Acta Math. 113 (1965), 89–152

[13] L. Lempert, private communication, October 2013

[14] F. Maitani, H. Yamaguchi, *Variation of Bergman metrics on Riemann surfaces*, Math. Ann, 330 (2004), 477–489

[15] T. Ohsawa, *Addendum to "On the Bergman kernel of hyperconvex domains"*, Nagoya Math. J. 137 (1995), 145–148

[16] T. Ohsawa, K. Takegoshi, *On the extension of L^2 holomorphic functions*, Math. Z. 195 (1987), 197–204

[17] N. Suita, *Capacities and kernels on Riemann surfaces*, Arch. Ration. Mech. Anal. 46 (1972), 212–217

Zbigniew Błocki
Uniwersytet Jagielloński
Instytut Matematyki
Łojasiewicza 6
PL-30-348 Kraków, Poland
e-mail: Zbigniew.Blocki@im.uj.edu.pl
umblocki@cyf-kr.edu.pl

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 135–145

Siciak's Theorem on Separate Analyticity

Jan Boman

Dedicated to the memory of Mikael Passare

Abstract. We give a simple proof of an important special case of the famous theorem of Jósef Siciak on separate analyticity.

1. Introduction

The well-known theorem of Hartogs states that a function $u(z_1, z_2)$ of two complex variables which is separately analytic must be analytic. By separately analytic we mean here that $z_1 \mapsto u(z_1, z_2)$ is analytic for each fixed z_2 and vice versa. Similar statements are also true if the "fixed" variables are restricted to sets of real dimension 1, or even to arbitrary sets of positive capacity. An important theorem of that kind was proved by Siciak in 1969, [11]. In the same paper Siciak gave a precise description of the maximal domain in $\mathbf{C}^2$ to which the function can be analytically continued. Many sharpenings and extensions of Siciak's theorem have been given later, for instance in [15], [13], [14], and [8], and a couple of years ago Jarnicki and Pflug wrote a whole book on the subject, [4]. Surveys of results related to separate analyticity can be found in [10] and [5]. The purpose of this note is to give a short proof of the most important special case of Siciak's theorem (Theorem 1′ in Section 3) using only very well-known tools. More exactly, we shall treat the special case when the "fixed" variables range over bounded intervals on the real line. We will treat only the case of functions of two variables; the extension to functions on $\mathbf{C}^n \times \mathbf{C}^m$ is straightforward.

The well-known example $x_1x_2/(x_1^2+x_2^2)$ shows that a separately real analytic function need not be real analytic. Let us say that a function $u(x_1, x_2)$ is *uniformly* separately real analytic in the domain $D \Subset \mathbf{R}^2$, if the functions $x_1 \mapsto u(x_1, x_2^0)$ and $x_2 \mapsto u(x_1^0, x_2)$ are analytically continuable to complex disks with radius $r(x_1^0, x_2^0)$ around x_1^0 and x_2^0, respectively, where r is some positive continuous function on D. It is a corollary of Siciak's theorem that a uniformly separately real analytic function is real analytic. If the additional assumption is made that the continued function u is locally bounded, then this statement is easy to prove and very well known, but in its general form we think the theorem is not as well known as it deserves to be. Here we give a proof of that theorem (Proposition $1'$) using an important lemma of Lelong (Lemma 2), which is a sharpening of the well-known Hartogs lemma.

The notation and terminology used here follows that of Siciak's papers. Let E be an open bounded interval on the real axis $\mathbf{R}$ and let G be a simply connected bounded domain in $\mathbf{C}$ which contains $\overline{E}$, the closure of E. If G is regular with respect to the Dirichlet problem we define the function $h_{G,E}(z)$ on $\overline{G}$ as the solution to the Dirichlet problem in $G \setminus \overline{E}$ with boundary values 0 on $\overline{E}$ and 1 on ∂G, the boundary of G. Note that the domain $G \setminus \overline{E}$ is also regular, since E is an interval. We remark that a bounded domain in $\mathbf{C}$ must be regular if it has finite connectivity and ∂G contains no isolated point.

A first version of this paper was written in 1994 while my student Ozan Öktem was writing the paper [9] and we both were struggling to understand Siciak's proof from his original paper [11]. Having learnt from Christer Kiselman about Lelong's lemma (Lemma 2) and its relevance in this context I wrote a new version containing Theorem $1'$ in 2004. In submitting the paper I made a new revision and added references to the book [4] and a couple of articles that have appeared after 2004.

I am indebted to two referees for a number of very valuable comments leading to a considerably improved article.

2. Separate analyticity with boundedness assumption

We shall begin by making the simplifying assumption that the original function is locally bounded. In Section 3 we shall discuss the case when no boundedness assumptions are made.

Theorem 1. *Let E_1 and E_2 be open bounded intervals on $\mathbf{R}$, and let G_1 and G_2 be simply connected bounded regular domains in $\mathbf{C}$ such that $G_j \supset \overline{E}_j$ for $j = 1, 2$. Let u be defined in the set*

$$X = (E_1 \times G_2) \cup (G_1 \times E_2) \subset \mathbf{C}^2,$$

and assume that u is separately analytic in X; by definition this means that for every $x_2 \in E_2$ the function $z_1 \mapsto u(z_1, x_2)$ is analytic in G_1, and for every $x_1 \in E_1$ the function $z_2 \mapsto u(x_1, z_2)$ is analytic in G_2. Assume furthermore that u is locally bounded on X. Then u can be continued analytically to the set

$$\widetilde{X} = \{z \in G_1 \times G_2;\ h_{G_1,E_1}(z_1) + h_{G_2,E_2}(z_2) < 1\}. \tag{1}$$

Our proof consists of three steps. The first step is to use the above-mentioned fact that a uniformly separately real analytic function is real analytic, in the easy special case when the function is assumed bounded (Proposition 1). It follows that u must be real analytic on $E_1 \times E_2$, which means by definition that u can be continued to an analytic function on some neighborhood of $E_1 \times E_2$ in $\mathbf{C}^2$.

In the second step (Proposition 2) we construct an analytic extension of u to an open neighborhood Σ in $\mathbf{C}^2$ of the set X. The third step is to prove that any function which is analytic in Σ can be extended to an analytic function on $\widetilde{X}$ (Proposition 3).

Proposition 1. *Let E_1 and E_2 be open bounded intervals on $\mathbf{R}$ with open complex neighborhoods V_1 and V_2, respectively, and assume that $z_1 \mapsto u(z_1, x_2)$ is analytic on V_1 for each $x_2 \in E_2$ and that $z_2 \mapsto u(x_1, z_2)$ is analytic on V_2 for each $x_1 \in E_1$. Assume furthermore that u is bounded on $(E_1 \times V_2) \cup (V_1 \times E_2)$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $E_1 \times E_2$ in $\mathbf{C}^2$ that agrees with u on $E_1 \times E_2$.*

This theorem is well known; for the proof we refer for instance to [6].

Using the notion of analytic wave front set one can give a short proof of Proposition 1 as follows.[1] Let u be a compactly supported integrable function or distribution in $\mathbf{R}^n$. It is known that $(x^0, \xi^0) \in T^*(\mathbf{R}^n)$, $\xi^0 \neq 0$, belongs to the complement of the analytic wave front set of u, $WF_A(u)$, if and only if the so-called FBI transform of u,

$$F_u(x, \xi) = \int_{\mathbf{R}^n} u(y) e^{-|\xi||y-x|^2} e^{-iy\cdot\xi} dy,$$

decays exponentially as $|\xi|$ tends to infinity for ξ in a conic neighborhood of ξ^0 uniformly for x in some neighborhood of x^0. It follows immediately from the definition that the analytic wave front set is conic in the second variable, i.e., that $(x^0, \xi^0) \in WF_A(u)$ if and only if $(x^0, \lambda\xi^0) \in WF_A(u)$ and $\lambda > 0$. It is a basic fact that a function (distribution) is real analytic in some neighborhood of x^0 if and only if $(x^0, \xi^0) \notin WF_A(u)$ for every $\xi^0 \neq 0$. Assume now that $u \in L^1(\mathbf{R}^2)$ is uniformly separately real analytic in some neighborhood of x^0. Write $F_u(x, \xi)$ as a repeated integral with inner integral

$$\int_{\mathbf{R}} u(y_1, y_2) e^{-|\xi|(y_1 - x_1)^2} e^{-iy_1\xi_1} dy_1. \tag{2}$$

By the assumption of real analyticity with respect to x_1 we can use Cauchy's theorem to deform the path of integration a little bit into the complex near $y_1 = x_1^0$ and thereby prove that the integral (2) tends to zero exponentially as $|\xi_1|$ tends to infinity for x_1 close to x_1^0, and hence the same is true of $F_u(x, \xi)$. Similarly, real analyticity with respect to x_2 implies that $F_u(x, \xi)$ is exponentially decreasing as

[1]The analytic wave front set for distributions was introduced by Hörmander 1970 in connection with a new proof of Holmgren's uniqueness theorem for partial differential equations with real analytic coefficients. A parallel theory was developed independently by M. Sato. There the so-called singular support of a hyperfunction was defined in terms of the possibility to represent the (hyper-)function as a sum of boundary values of analytic functions in regions $\{x + iy;\, x \in U \subset \mathbf{R}^n,\, y \in \Gamma_k,\, |y| < \varepsilon\}$, where U is open and Γ_k are certain cones in $\mathbf{R}^n$; see [1], ch. 9. The fact that the concepts were equivalent for distributions was proved a few years later. The third equivalent definition used here was given by Bros and Iagolnizer in 1975; see [1], Theorem 9.6.3.

$|\xi_2| \to \infty$ for x_2 close to x_2^0. Hence $(x^0, \xi^0) \notin WF_A(u)$ for every $\xi^0 \neq 0$, so u is real analytic in a neighborhood of x^0.

The second step consists in using the assumption of separate analyticity to extend the region of joint analyticity in one direction at the time.

Proposition 2. *Let G be a simply connected bounded domain in $\mathbf{C}$, U an open disk with $\overline{U} \subset G$, and F a compact interval in $\mathbf{R} \subset \mathbf{C}$. Let u be an analytic function in a complex neighborhood of $U \times F$. Assume that the function $U \ni z_1 \mapsto u(z_1, x_2)$ can be extended to an analytic function in G for every $x_2 \in F$ and that the extended function $u(z_1, x_2)$ is locally bounded in $G \times F$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $G \times F$ that agrees with u on $U \times F$.*

As a preparation for the proof of Proposition 2 we shall first consider the case when G is the open disk $U_R = \{\zeta \in \mathbf{C}; |\zeta| < R\}$ and U is a smaller disk containing the origin. This lemma is part of the standard proof of Hartogs' theorem on separate analyticity (see, e.g., [2], Lemma 2.2.11), but we include it here for the sake of completeness and in order to facilitate the discussion in Section 3.

Lemma 1. *Let $U_\varepsilon = \{\zeta \in \mathbf{C}; |\zeta| < \varepsilon\}$ and let F be a compact interval in $\mathbf{R} \subset \mathbf{C}$. Let u be analytic in some complex neighborhood of $U_\varepsilon \times F$ and assume that $U_\varepsilon \ni z_1 \mapsto u(z_1, x_2)$ can be extended to an analytic function in U_R for every $x_2 \in F$. Assume moreover that the function $u(z_1, x_2)$ is bounded for $(z_1, x_2) \in U_R \times F$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $U_R \times F$ that agrees with u on $U_\varepsilon \times F$.*

Proof. By the first assumption u can be expanded in a Taylor series with respect to z_1

$$u(z_1, z_2) = \sum_{k=0}^{\infty} a_k(z_2) z_1^k, \tag{3}$$

where $a_k(z_2)$ are analytic in some neighborhood of F. Denoting by $V_\delta(F)$ the complex δ-neighborhood of $F \subset \mathbf{C}$ we choose $\delta > 0$ and $\varepsilon > 0$ so that u is analytic and bounded in $U_\varepsilon \times V_\delta(F)$. By Cauchy's inequality we then obtain

$$|a_k(z_2)| \leq C_1 \varepsilon^{-k}, \quad z_2 \in V_\delta(F), \quad k = 0, 1, \ldots. \tag{4}$$

By the second assumption we also have the estimates

$$|a_k(x_2)| \leq C_0 R^{-k}, \quad x_2 \in F, \quad k = 0, 1, \ldots. \tag{5}$$

Let $g(w)$ be the solution to the Dirichlet problem in $V_\delta(F) \setminus F$ with boundary values 0 on F and 1 on the boundary of $V_\delta(F)$. The function $\log |a_k(w)|$ is subharmonic in $V_\delta(F)$ and $\leq A_0$ in F and $\leq A_1$ in $V_\delta(F)$, where $A_0 = \log(C_0 R^{-k})$ and $A_1 = \log(C_1 \varepsilon^{-k})$. Hence $\log |a_k(w)| \leq A_0 + (A_1 - A_0) g(w)$ in $V_\delta(F)$, or

$$|a_k(w)| \leq C R_0(w)^{-k},$$

where $R_0(w) = R^{1-g(w)} \varepsilon^{g(w)}$ and $C = \max(C_0, C_1)$. But $R_0(w)$ is continuous and $R_0(w) = R$ for $w \in F$, hence for any given $r < R$ there exists a neighborhood W_r

of F such that $R_0(w) > r$ for $w \in W_r$. This proves that the series (3) converges in $U_r \times W_r$ for every $r < R$, which completes the proof of the lemma. □

Proof of Proposition 2. Since G is simply connected it is easy to see that one can construct a locally finite covering $G = \cup_{k=0}^{\infty} G_k$ of G by open disks $G_k \subset G$ with $\overline{G}_k \subset G$ such that $G_0 = U$ and for every k the union $\cup_{j=0}^{k} G_j$ is simply connected and contains the center of G_{k+1}. Set $H_k = \cup_{j=0}^{k} G_j$ for all k. We claim that

$$\text{for every } k \text{ there exists a complex neighborhood } V_k \text{ of } F \text{ and an analytic function } \widetilde{u}_k \text{ in } H_k \times V_k \text{ that agrees with } u \text{ on } U \times F. \tag{P_k}$$

To prove this we use induction over k. For $k = 0$ there is nothing to prove. Assume that the statement (P_k) is true. Let ζ_0 be the center of G_{k+1}. Since $\zeta_0 \in G_k$ we can choose $\varepsilon > 0$ so that $U_\varepsilon(\zeta_0) = \{\zeta \in \mathbf{C};\ |\zeta - \zeta_0| < \varepsilon\}$ is contained in $G_k \cap G_{k+1}$. By the induction assumption $\widetilde{u}_k$ is then analytic in $U_\varepsilon(\zeta_0) \times V_k$. Applying Lemma 1 with $U_R = U_R(\zeta_0)$ and $U_\varepsilon = U_\varepsilon(\zeta_0)$ and R chosen so that $\overline{G}_{k+1} \subset U_R(\zeta_0) \subset G$ we can find a complex neighborhood V_{k+1} of F and a function $\widetilde{u}_{k+1}$ that is analytic in $G_{k+1} \times V_{k+1}$ and agrees with $\widetilde{u}_k$ on $U_\varepsilon(\zeta_0) \times F$. Shrinking V_{k+1}, if necessary, we may assume that $V_{k+1} \subset V_k$. Extending $\widetilde{u}_{k+1}$ suitably we therefore get an analytic function on $H_{k+1} \times V_{k+1}$, that we also denote by $\widetilde{u}_{k+1}$. Since H_{k+1} is simply connected, analytic continuation from G_0 to G_{k+1} along a different chain of disks would give the same values in $G_{k+1} \times V_{k+1}$. This proves (P_{k+1}) and hence shows that the statement (P_k) is true for all k.

To finish the proof of the proposition we observe that the union W of all $H_k \times V_k$ is a complex neighborhood of $G \times F$ and that all the $\widetilde{u}_k$ agree on their common domains, which shows that they define an analytic function on W. This completes the proof. □

Proposition 3. *Let G_1 and G_2 be simply connected bounded regular domains in $\mathbf{C}$ and let E_1 and E_2 be open bounded intervals on the real axis such that $\overline{E}_j \subset G_j$ for $j = 1, 2$. Let X and $\widetilde{X}$ be defined as in Theorem 1, and let u be analytic in some open neighborhood Σ of X. Then there exists an analytic function $\widetilde{u}$ on $\widetilde{X}$ that agrees with u on X.*

Proof. Set $h(z) = h_{G_1,E_1}(z_1) + h_{G_2,E_2}(z_2)$, and for $\varepsilon > 0$ and $0 < t < 1$ define the region $\widetilde{X}_\varepsilon(t)$ by

$$\widetilde{X}_\varepsilon(t) = \{z \in G_1 \times G_2;\ h(z) < \min(1 - \varepsilon, t + \varepsilon |z|^2\}.$$

Choose $M > 1$ so that $|z|^2 \le M$ in $G_1 \times G_2$. We shall prove that, for every sufficiently small $\varepsilon > 0$, there exists an analytic function $\widetilde{u}_\varepsilon$ on

$$\widetilde{X}_\varepsilon(1 - 2\varepsilon M) \tag{6}$$

that agrees with u on X. Since the union of all the regions (6) is equal to $\widetilde{X}$ and all the functions $\widetilde{u}_\varepsilon$ agree on their common domains of definition, this proves the assertion of the proposition. We first claim that

$$\widetilde{X}_\varepsilon(t) \subset \Sigma$$

if ε and t are sufficiently small. To prove this we observe that the continuous function $h(z)$ is positive on the compact set $(\overline{G}_1 \times \overline{G}_2) \setminus \Sigma$, hence $h(z) \geq \delta$ on that set for some $\delta > 0$. It follows that $\widetilde{X}_\varepsilon(t) \subset \Sigma$ if ε and t are so small that $t + \varepsilon M < \delta$. Fix an arbitrary $\varepsilon > 0$ with $2\varepsilon M < \delta < 1$ and set

$$t_0 = \sup\{t < 1 - 2\varepsilon M; \text{there exists a function } \widetilde{u}_{\varepsilon,t} \text{ that is analytic in } \widetilde{X}_\varepsilon(t) \\ \text{and is equal to } u \text{ on } X\}.$$

Assuming that $t_0 < 1 - 2\varepsilon M$ we shall obtain a contradiction. It is clear that all the functions $\widetilde{u}_{\varepsilon,t}$ with $t < t_0$ define a function $\widetilde{u}_{\varepsilon,t_0}$ that is analytic on $\widetilde{X}_\varepsilon(t_0)$. On the other hand, by the definition of t_0 there must exist $z^0 \in \partial\widetilde{X}_\varepsilon(t_0)$ such that $\widetilde{u}_{\varepsilon,t_0}$ cannot be continued to any neighborhood of z^0. We claim that

$$\partial\widetilde{X}_\varepsilon(t) \subset \{z \in \mathbf{C}^2;\, h(z) = t + \varepsilon|z|^2\}, \quad \text{if } t < 1 - 2\varepsilon M. \tag{7}$$

Indeed, if $h(z) < t+\varepsilon|z|^2$ and $t < 1-2\varepsilon M$, then $h(z) < 1-2\varepsilon M+\varepsilon M = 1-\varepsilon M < 1-\varepsilon$, so z cannot belong to the boundary of $\widetilde{X}_\varepsilon(t)$, which proves (7).

The functions h_{G_j,E_j} are harmonic in $G_j \setminus \overline{E}_j$, hence $h(z) - \varepsilon|z|^2$ is *strictly plurisuperharmonic*, so (7) implies that the domain $\widetilde{X}_\varepsilon(t_0)$ is *strictly pseudoconcave*. This implies that u must be continuable to an analytic function in some neighborhood of z^0. This is a contradiction and hence completes the proof of the proposition. □

Proof of Theorem 1. By Proposition 1 there exists an analytic function $\widetilde{u}_0$ in some open neighborhood W_0 of $E_1 \times E_2$ that agrees with u on $E_1 \times E_2$. We may assume that W_0 is connected, and then it is clear that $\widetilde{u}_0$ agrees with the given function u on $X \cap W_0$. Applying Proposition 2 to $G = G_1$, an arbitrary closed subinterval $F \subset E_2$, and an open disk $U \subset G_1$ such that $U \times F \subset W_0$ we can then find $\widetilde{u}_1$ that is analytic in some complex neighborhood of $G_1 \times F$ and agrees with u on $U \times F$, hence agrees with u on $G_1 \times F$. Varying $F \subset E_2$ we get a function $\widetilde{u}_1$ that is analytic in some complex neighborhood W_1 of $G_1 \times E_2$ and agrees with u on $G_1 \times E_2$. Similarly we can find $\widetilde{u}_2$ that is analytic in some complex neighborhood W_2 of $E_1 \times G_2$ and agrees with u on $E_1 \times G_2$. Since $\widetilde{u}_1$ and $\widetilde{u}_2$ agree on an open set, it is clear that they together define an analytic function $\widetilde{u}$ in a complex neighborhood Σ of X. An application of Proposition 3 now completes the proof of the theorem. □

3. The general case

We shall now discuss the situation when no boundedness assumption is made in Theorem 1. We shall use the convention that a primed theorem, proposition etc. is the analogue without boundedness assumption of the unprimed theorem (proposition etc.) with the same number.

Theorem 1′. *The statement of Theorem* 1 *is true without the assumption that u is locally bounded.*

The proof of this theorem consists of three steps, analogous to those of the proof of Theorem 1. Only the first two steps need to be modified. The following lemma of Lelong is essential for both those steps ([7], Théorème 10; see also [11], Theorem 2.1). It is an important extension of the well-known Hartogs lemma.

Lemma 2. *Let F be a compact interval on* $\mathbf{R}$ *and let* $G \subset \mathbf{C}$ *be an open set containing F. Let $\varphi_k(z)$, $k = 1, 2, \ldots$, be a sequence of subharmonic functions in G satisfying*

$$\varphi_k(z) \leq B, \quad z \in G, \quad k = 1, 2, \ldots, \tag{8}$$

and

$$\overline{\lim_{k \to \infty}}\, \varphi_k(x) \leq A, \quad x \in F.$$

Then for every $\eta > 0$ there exists a complex neighborhood U of F and a number k_0 such that

$$\varphi_k(z) < A + \eta, \quad z \in U, \quad k \geq k_0. \tag{9}$$

The neighborhood U depends only on the numbers η, B, A, and on the sets F and G (not on the sequence φ_k).

We shall sketch a proof of this lemma using facts from [3]. Let us first make a couple of remarks. If we knew that the function $\varphi(z) = \overline{\lim}\, \varphi_k(z)$ were subharmonic, then φ would be majorized in G by the function $h(z) = A + (B - A)h_{G,F}$, the solution to the Dirichlet problem in $G \setminus F$ with boundary values A on F and B on ∂G. Then (9) could be proved just as the classical Hartogs lemma (use the mean value property of φ_k and Fatou's lemma to find k_0 independent of z such that $\varphi_k(z) < \varphi(z) + \varepsilon$ for $k > k_0$). But the limes superior of a sequence of subharmonic functions is not always subharmonic. (What is true is that it must be subharmonic if it is upper semicontinuous; more generally, the upper semicontinuous regularization $\overline{\varphi}$ of $\overline{\lim}\, \varphi_k$ is subharmonic, but we do not know that $\overline{\varphi} \leq A$ on F.) Lelong proves Lemma 2 by establishing the majorization just mentioned for a class of functions which includes upper limits of sequences of subharmonic functions.

Sketch of proof of Lemma 2. As was indicated above it is sufficient to prove the estimate

$$\overline{\lim_{k \to \infty}}\, \varphi_k(z) \leq h(z) = A + (B - A)h_{G,F} \tag{10}$$

for $z \in G$. It is clearly sufficient to prove (10) for $z \in G \setminus F$. Assume (10) is false at some point $z \in G \setminus F$. Then there exists a number c such that

$$\varphi_k(z_0) > c > h(z_0) \tag{11}$$

for infinitely many k. Since φ_k is bounded from above, we can take a subsequence $\widetilde{\varphi}_\nu = \varphi_{k_\nu}$ satisfying (11) and converging in $\mathcal{D}'(G)$ to some subharmonic function ψ (Theorem 3.2.12 in [3]). Then $\overline{\lim}\, \widetilde{\varphi}_\nu \leq \psi$ (Theorem 3.2.13 in [3]), and according to Theorem 3.4.14 in [3] the set

$$M = \{z \in G; \overline{\lim}\, \widetilde{\varphi}_\nu(z) < \psi(z)\}$$

is polar (a polar set is by definition any set on which a subharmonic function can be equal to $-\infty$ without being identically $-\infty$). But $\overline{\lim}\,\widetilde{\varphi}_\nu \leq \overline{\lim}\,\varphi_k \leq A$ on F by assumption, hence ψ must be $\leq A$ on $F \setminus M$, and it is clear that $\psi \leq B$ on all of G. Since M is polar, and ψ is subharmonic in G this implies in fact that $\psi \leq h$ in G (M is so small that the boundary values on M do not influence the solution to the Dirichlet problem). On the other hand

$$\psi(z_0) \geq \overline{\lim}\,\widetilde{\varphi}_\nu(z_0) \geq c > h(z_0).$$

Thus we have obtained a contradiction and (10) is proved. □

Using Lemma 2 it is easy to prove the analogues of Lemma 1 and Proposition 2 without boundedness assumptions:

Lemma 1′. *Let $U_\varepsilon = \{\zeta \in \mathbf{C};\ |\zeta| < \varepsilon\}$ and let F be a compact interval in $\mathbf{R} \subset \mathbf{C}$. Let u be analytic in some complex neighborhood of $U_\varepsilon \times F$ and assume that $U_\varepsilon \ni z_1 \mapsto u(z_1, x_2)$ can be extended to an analytic function in U_R for every $x_2 \in F$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $U_R \times F$ that agrees with u on $U_\varepsilon \times F$.*

Proof. Let φ_k be the subharmonic function

$$\varphi_k(w) = \frac{1}{k} \log |a_k(w)|,$$

where $a_k(\cdot)$ is defined by (3). By the first assumption the sequence $\varphi_k(w)$ is uniformly bounded from above in $V_\delta(F)$ for some $\delta > 0$. By the second assumption

$$\overline{\lim_{k\to\infty}}\, \varphi_k(x_2) \leq \log(1/R), \quad \text{for all} \quad x_2 \in F.$$

According to Lemma 2 there must then exist for any $r < R$ a number k_0 such that

$$\varphi_k(x_2) < \log(1/r), \quad \text{if} \quad k > k_0,\ x_2 \in F,$$

or equivalently

$$|a_k(x_2)| \leq r^{-k}, \quad \text{if} \quad k > k_0,\ x_2 \in F.$$

Thus we have estimates corresponding to (4) and (5), and the proof can now be finished exactly in the same way as the proof of Lemma 1. □

We can now prove Proposition 2 without boundedness assumption:

Proposition 2′. *Let G be a simply connected bounded domain in $\mathbf{C}$, U an open disk with $\overline{U} \subset G$, and F a compact interval in $\mathbf{R} \subset \mathbf{C}$. Let u be an analytic function in a complex neighborhood of $U \times F$. Assume that the function $U \ni z_1 \mapsto u(z_1, x_2)$ can be extended to an analytic function in G for every $x_2 \in F$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $G \times F$ that agrees with u on $U \times F$.*

Proof. This statement is proved using Lemma 1′ in exactly the same way as Proposition 2 was proved using Lemma 1. □

We are now ready to prove the theorem on separate real analyticity without boundedness assumptions.

Proposition 1′. *Let E_1 and E_2 be bounded open intervals on* $\mathbf{R}$ *with complex neighborhoods V_1 and V_2, respectively, and assume that $z_1 \mapsto u(z_1, x_2)$ is analytic on V_1 for each $x_2 \in E_2$ and that $z_2 \mapsto u(x_1, z_2)$ is analytic on V_2 for each $x_1 \in E_1$. Then there exists an analytic function $\widetilde{u}$ on some complex neighborhood of $E_1 \times E_2$ in* $\mathbf{C}^2$ *that agrees with u on $E_1 \times E_2$.*

Proof. It is enough to prove the assertion for arbitrary closed subintervals $F_1 \subset E_1$ and $F_2 \subset E_2$. Shrinking V_1 and V_2, if necessary, we may also assume that V_1 and V_2 are simply connected and that $z_1 \mapsto u(z_1, x_2)$ is bounded on V_1 for each $x_2 \in F_2$ and that $z_2 \mapsto u(x_1, z_2)$ is bounded on V_2 for each $x_1 \in F_1$. For any natural number N define the set

$$K_N = \{x_1 \in F_1;\ |u(x_1, z_2)| \le N \text{ for all } z_2 \in V_2\}.$$

We claim that K_N is closed for each N. In fact, let $x_1^\nu \in K_N$ for $\nu = 1, 2, \dots$ and $\lim_{\nu\to\infty} x_1^\nu = x_1^0$. We have to prove that $x_1^0 \in K_N$. Since the family of analytic functions $w_\nu(z_2) = u(x_1^\nu, z_2)$ is uniformly bounded, there exists a subsequence of x_1^ν such that $w_\nu(z_2)$ converges to an analytic function $w(z_2)$ on V_2 with $|w(z_2)| \le N$. Since $E_1 \ni x_1 \mapsto u(x_1, x_2)$ must be continuous for each $x_2 \in E_2$, we must have $w(x_2) = u(x_1^0, x_2)$ for each $x_2 \in F_2 \subset E_2$. But this implies that $w(z_2) = u(x_1^0, z_2)$ for all $z_2 \in V_2$, and hence proves our claim that K_N is closed. Since $V_2 \ni z_2 \mapsto u(x_1, z_2)$ is bounded for each $x_1 \in F_1$, the union of all K_N must be equal to all of F_1. By Baire's theorem K_N must have an interior point for some N, in other words, we can choose N_1, $x_1^0 \in F_1$ and $\delta_1 > 0$ such that $\{x_1;\ |x_1 - x_1^0| < \delta_1\} \subset F_1$ and

$$|u(x_1, z_2)| \le N_1 \quad \text{whenever } |x_1 - x_1^0| < \delta_1 \text{ and } z_2 \in V_2. \tag{12}$$

Set $I_{\delta_1} = \{x_1;\ |x_1 - x_1^0| < \delta_1\}$. Applying the same argument to the function u on $(I_{\delta_1} \times V_2) \cup (V_1 \times F_2)$ with the variables interchanged we can find a number $N \ge N_1$, $x_2^0 \in F_2$, and $\delta_2 > 0$ such that $\{x_2;\ |x_2 - x_2^0| < \delta_2\} \subset F_2$ and, in addition to (12),

$$|u(z_1, x_2)| \le N \quad \text{whenever } |x_2 - x_2^0| < \delta_2 \text{ and } z_1 \in V_1.$$

Set $J_{\delta_2} = \{x_2;\ |x_2 - x_2^0| < \delta_2\}$. Now we can apply Proposition 1 to conclude that u must be real analytic on $I_{\delta_1} \times J_{\delta_2}$. By definition this implies that there exist complex neighborhoods U_1 of I_{δ_1} and U_2 of J_{δ_2} and an analytic function $\widetilde{u}_0$ in $U_1 \times U_2$ that agrees with u on $I_{\delta_1} \times J_{\delta_2}$. Applying Proposition 2′ (Proposition 2 would actually suffice here) with $G = V_1$, $F = F_2$ equal to a closed subinterval of J_{δ_2}, and a disk $U \subset U_1$, we can find an analytic function $\widetilde{u}_1$ in some neighborhood of $V_1 \times F_2$ that agrees with $\widetilde{u}_0$ on $U \times F_2$, hence agrees with u on $(U \cap E_1) \times F_2$. Since $E_1 \ni x_1 \mapsto u(x_1, x_2)$ is real analytic for each x_2, $\widetilde{u}_1$ must agree with u on $E_1 \times F_2$. Thus for an arbitrary closed subinterval $F_1 \subset E_1$ we can now choose a disk $U \subset U_2$ such that $\widetilde{u}_1$ is analytic in a complex neighborhood of $F_1 \times U$. Then we can apply Proposition 2′ with those choices of F_1 and U and $G = V_2$ to conclude that there exists an analytic function $\widetilde{u}_2$ that is analytic in a complex neighborhood of $F_1 \times V_2$ that agrees with $\widetilde{u}_1$ on $F_1 \times U$, hence agrees with u on $F_1 \times E_2$. Since F_1 was an arbitrary closed subinterval of E_1 this gives an analytic function in a complex neighborhood of $E_1 \times E_2$ that is equal to u on $E_1 \times E_2$. The proof is complete. □

Proof of Theorem $1'$. We argue in the same way as in the proof of Theorem 1. By Proposition $1'$ there exists an analytic function $\widetilde{u}_0$ in some neighborhood W_0 of $E_1 \times E_2$ that agrees with u on $E_1 \times E_2$. Applying Proposition $2'$ to $G = G_1$, an arbitrary closed subinterval $F \subset E_2$, and open disks $U \subset G_1$ such that $U \times F \subset W_0$ we find $\widetilde{u}_1$ that is analytic in some complex neighborhood W_1 of $G_1 \times E_2$ and agrees with u on $G_1 \times E_2$. Similarly we can find $\widetilde{u}_2$ that is analytic in a complex neighborhood W_2 of $E_1 \times G_2$ and agrees with u on $E_1 \times G_2$. It is clear that $\widetilde{u}_1$ and $\widetilde{u}_2$ together define an analytic function $\widetilde{u}$ in a complex neighborhood Σ of X. The proof is completed by means of Proposition 3 exactly in the same way as before. □

In [11] Siciak treats also the case when E_1 and E_2 are allowed to be general compact subsets of G_1 and G_2, respectively, not necessarily subsets of the real line. It is assumed that the boundaries of E_1 and E_2 are regular for the Dirichlet problem, which implies in particular that E_1 and E_2 are not too small. An analogous statement in n dimensions where $u(x_1, \ldots, x_n)$ is assumed to be separately analytic in each variable is also proved in [11]. Extension to the case when G_1 and G_2 may be higher-dimensional manifolds is given in [15]. In [13] Siciak gave a new proof of his main result in [11], based on his theory of so-called extremal plurisubharmonic functions. A theorem analogous to Theorem 1 where u is allowed to have singularities on an algebraic curve Γ in $\mathbf{C}^2$ was given in [14]; this proved a conjecture by Öktem, who treated the special case when Γ is a complex line, [9].

References

[1] Hörmander, L., *The analysis of linear partial differential operators, vol.* 1, Springer-Verlag, Berlin 1983.

[2] Hörmander, L., *An introduction to complex analysis in several variables*, Van Nostrand, Princeton 1966.

[3] Hörmander, L., *Notions of convexity*, Birkhäuser, Boston 1994.

[4] Jarnicki, M. and Pflug, P., *Separately analytic functions*, European Mathematical Society, Zürich 2011.

[5] Jarnicki, M. and Pflug, P., *Directional regularity vs. joint regularity*, Notices of the AMS, **58** (2011), 896–904.

[6] Krantz, S., and Parks, H., *A primer of real analytic functions*, Birkhäuser, Basel 1992.

[7] Lelong, P., *Fonctions plurisousharmoniques et fonctions analytiques de variables réelles*, Ann. Inst. Fourier (Grenobles), **11** (1961), 515–562.

[8] Nguyen Than Van and Zeriahi, A., *Une extension du théorème de Hartogs sur les fonctions séparément analytiques*, in Analyse complexe multivariable, Récents développements et applications, A. Meril (ed.), EditEl, Rende 1991, 183–194.

[9] Öktem, O., *Extension of separately analytic functions and applications to range characterization of the exponential Radon transform*, Ann. Polon. Math. **70** (1998), 195–213.

[10] Pflug, P., *Extension of separately holomorphic functions – a survey* 1899–2001, Ann. Polon. Math. **80** (2003), 21–36.

[11] Siciak, J., *Separately analytic functions and envelopes of holomorphy of some lower dimensional subsets of* $\mathbf{C}^n$, Ann. Polon. Math. **22** (1969), 145–171.

[12] Siciak, J., *Analyticity and separate analyticity of functions defined on lower dimensional subsets of* $\mathbf{C}^n$, Zeszyty Nauk. Uniw. Jagiello. Prace Mat. **13** (1969), 53–70.

[13] Siciak, J., *Extremal plurisubharmonic functions in* $\mathbf{C}^n$, Ann. Polon. Math. **39** (1981), 175–211.

[14] Siciak, J., *Holomorphic functions with singularites on algebraic sets*, Univ. Iagel. Acta Math. **39** (2001), 9–16.

[15] Zaharjuta, V. P., *Separately analytic functions, generalizations of Hartogs' theorem, and envelopes of holomorphy*, Mat. Sbornik **101** (1976). English translation: Math. USSR Sb. **30** (1976), 51–76.

Jan Boman
Department of Mathematics
Stockholm University
106 91 Stockholm, Sweden
e-mail: jabo@math.su.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 147–156

Mikael Passare, a Jaunt in Approximation Theory

Jean-Paul Calvi

Abstract. A personal and informal presentation of the contributions of Mikael Passare to the field of approximation theory.

Mathematics Subject Classification (2010). Primary 41A17; Secondary 32E30.

Keywords. Kergin interpolation, Hermite interpolation, mean value interpolation, simplex functional, $\mathbb{C}$-convexity.

The axe bites into the tree,
But the snail
Is calm and serene.
Baishitsu

1. In the late eighties, Mikael Passare was a frequent visitor of the (late) *Laboratoire d'Analyse complexe* at the *University Paul Sabatier* in Toulouse. His expertise in pluri-complex integration was much appreciated by a few colleagues who worked on the $\bar{\partial}$-equation, then a most popular topic. He arrived one spring just after Thomas Bloom returned to Toronto. In these years, Tom was working on a certain multivariate polynomial interpolation method that had been introduced a little earlier by his Ph.D. Student Paul Kergin [13] and was already known to many approximation theorists as *Kergin interpolation* [9]. He had just given a lecture on the subject and some of his documents were forgotten somewhere in the department (Tom can be abstracted) and, by chance, say, came into the hands of Mikael. I was not a direct witness of this encounter since I became a member of the *laboratoire* only a few months later. Yet, I was told the story by different colleagues with reasonable variations and predictable ornaments – should Mikael have found Tom's papers in a waste-paper basket? – in a way which ultimately convinced me of its truth. The way the subject came to Mikael Passare was typical of his relation to mathematics, as I understood it, and I believe that the story of his contribution shed a beautiful light on the life of mathematics, perhaps of a certain old-fashioned form of mathematics.

2. Back to Hermite

One of the many beginnings of polynomial interpolation is an extraordinary paper written by Charles Hermite in 1878. Hermite considered the problem of finding a (univariate) complex polynomial p of degree $n-1$ such that $p^{(j)}(a_i) = f^{(j)}(a_i)$ where $i = 1, \ldots, k$, $j = 0, \ldots, \alpha_i - 1$, $\sum \alpha_i = n$ and f was a certain analytic function whose regularity was to be made precise. In those days, a mix of statements and proofs was not uncommon. Although he applied his procedure to derive some quadrature formulas, Hermite did not offer any motivation apart from that of generalizing Lagrange interpolation and (subsequently) connecting it to the Taylor polynomial. As far as I know, he was not supported, either financially or morally, by a scientific prospective committee or by any other sort of bureaucratic creature. He was soberly introducing *Hermite interpolation*, a future fundamental tool of applied mathematics. The Hermite problem, of course, is readily solved by nowadays elementary linear algebra. You consider the vector space $\mathcal{P}_{n-1}$ of polynomials of degree at most $n-1$ and the linear map

$$\mathcal{P}_{n-1} \ni p \mapsto (p(a_1), \ldots, p^{(\alpha_1-1)}(a_1), \ldots, p(a_k), \ldots, p^{(\alpha_k-1)}(a_k)) \in \mathbb{C}^n, \qquad (1)$$

and you prove the existence and uniqueness of the searched polynomial by merely checking that the kernel of this map is trivial. In fact, if p is in the kernel, it has a zero of order at least α_i at a_i, which gives at least n roots taking multiplicity into account, and this is too much for a *non zero* polynomial of degree at most $n-1$. In 1878, the required abstract linear algebra formalism was still to come (Hermite only observed that the problem was well posed) and, in lack of a plain argument, he had to produce a subtle one. He observed that if f is analytic in a region S containing the points a_i and bounded by a contour S then the function

$$x \mapsto \int_S \frac{f(z)\Phi(x)}{(z-x)\Phi(z)} dz, \quad \Phi(z) = \prod_{i=1}^{k} (z - a_i)^{\alpha_i}, \qquad (2)$$

differs from f by a polynomial of degree at most $n-1$ which satisfies the required properties. As the experienced mathematician easily guesses, it might very well be that Hermite made the observation first and then understood the potential interest of such a polynomial. This would not have been a less estimable way of obtaining his result. The above integral formula is now refereed to as the *Hermite remainder formula*. It is the basic tool for the theory of interpolation of univariate analytic functions whose main results, including its connection to plane potential theory, were established in the first half of the twentieth century[1]. Hermite was pleased to find a way of seeing Lagrange interpolation and Taylor approximation as opposite sides of a same question (n points, each with a minimal interpolation condition, for Lagrange and one point with a maximal interpolation condition for Taylor). Something, however, worried him after he finished the first draft of his paper in July 1877. Observing that Taylor polynomials and Lagrange polynomials

[1] The classical treatments are [16, 15].

are well defined in elementary analysis, he felt that it should be possible to derive his result without using complex integration. Again, the modern mathematician will soon recognize that Hermite's problem is purely algebraic and, if you decide to match $y_i^j \in K$ instead of $f^{(j)}(a_i)$, you may even solve it, with the same argument as above, with a polynomial whose coefficients belongs to any field K you like.

Two months later, Hermite was ready for adding a remarkable *post scriptum*. Due to an exceptional computational dexterity, he discovered, admittedly, with some surprise, that complex integration could be replaced by multivariate real integration. In fact, removing the non integrated term $\Phi(x)$ and setting $x = a_0$, $\Pi(z) = (z - a_0)\Phi(z)$ and $u = (a_0 - a_1)t_1 + (a_1 - a_2)t_2 + \cdots + (a_n - a_{n-1})t_n$,, the integral in formula (2) could be rewritten as

$$\int_S \frac{f(z)}{\Pi(z)} dz = \int_0^1 dt_n \int_0^{t_n} dt_{n-1} \int_0^{t_{n-2}} dt_{n-2} \dots \int_0^{t_1} f^{(n)}(u) dt_1. \tag{3}$$

This formula was valid when all α_i equal 1. Hermite then derived a similar formula in the general case by differentiating both sides with respect to the a_i. He did not mention the case of real (differentiable) functions. Nearly at the same period, the Italian analyst and number theorist Angelo Genocchi found out a similar formula in the case of Lagrange interpolation, in a more natural way, by expressing as an integral the coefficients, which are called *divided differences*, of the Newton formula for Lagrange interpolation. This would be recalled as the Hermite–Genocchi formula for divided differences.

3. From Hermite to Kergin

Although Hermite achieved more generality, his formula was not satisfying in the case where the α_i were greater than one. It occulted several fundamental properties, the most important being a simple dependency as a function of the interpolation points. The last step required a slight formal modification. Instead of choosing points a_i with multiplicity α_i, you start from a 'set' $X = \{x_1, \dots, x_n\}$ in which a same point may be repeated. There is an immediate one-to-one correspondence between Hermite data and sets with repetition: just take $x_i = a_1$ for $i = 1, \dots, \alpha_1$, $x_{\alpha_1 + i} = a_2$ for $i = 1, \dots, \alpha_2$ and so on. In that case, setting $x_0 = x$, the right-hand side of (3) can be expressed (one needs to use a certain linear change of variables) as

$$\int_{\Delta_n} f^{(n)} \left(\sum_{i=0}^n t_i x_i \right) dm(t), \tag{4}$$

where m is the Lebesgue measure on the simplex

$$\Delta_n = \left\{ (t_0, \dots, t_n) \in [0,1]^{n+1} \ : \ \sum_{i=0}^n t_i = 1 \right\}. \tag{5}$$

From this, one can show that the Hermite interpolation polynomial $H(X, f)$ of degree n corresponding to the data $X = \{x_0, \dots, x_n\}$ and the function f is given by the formula

$$H(X,f)(x)=\sum_{j=0}^{n}(x-x_0)\cdots(x-x_{j-1})\int_{\Delta_j} f^{(j)}\left(\sum_{i=0}^{j}t_ix_i\right)dm(t), \tag{6}$$

where the empty product (when $j=0$) is taken as 1 and the corresponding integral term reduces to $f(x_0)$. The above is certainly not a standard presentation of Hermite interpolation. In fact, if all mathematicians have heard about Hermite interpolation, the knowledge of the Hermite–Genocchi formula is reserved for specialists in approximation theory. However, once we have it in mind, the definition of Kergin interpolation becomes obvious: it is the very natural multivariate counterpart of (6). It suffices to observe that

$$(x-x_0)\cdots(x-x_{j-1})f^{(j)}\left(\sum_{i=0}^{j}t_ix_i\right)=D^jf\left(\sum_{i=0}^{j}t_ix_i\right)\big(x-x_0,\dots,x-x_{j-1}\big). \tag{7}$$

In fact, the Kergin interpolation polynomial of a multivariate function f at the (non necessarily distinct) point of $X=\{x_0,\dots,x_n\}\subset\mathbb{R}^N$ is given by

$$K(X,f)(x)=\sum_{j=0}^{n}\int_{\Delta_j}D^jf\left(\sum_{i=0}^{j}t_ix_i\right)\big(x-x_0,\dots,x-x_{j-1}\big)dm(t), \tag{8}$$

where D^jf denotes the jth total (Fréchet) derivative of f (which is a symmetric j-linear form). This was observed by Micchelli and Milman [14]; Kergin himself arrived to its procedure in a different manner. Formula (8) defines an operator $f\mapsto K(X,f)$ which possesses remarkable properties to which I will come back later.

For now, let me concentrate on the definition. I did not specify any assumption on the points or the function f. If, as is natural, we want to compute $K(X,f)$ for any choice of points X in $\Omega\subset\mathbb{R}^N$ then any convex combination $\sum_{i=0}^{j}t_jx_j$ of points of Ω should be included in Ω which therefore needs to be convex while f will be required to be, say, n times continuously differentiable on Ω. The same reasoning works in the complex case.

In the real case, the assumption on Ω cannot be weakened, it definitely has to be convex, and nothing more general was known in the complex case until Mikael Passare entered the game. Just like Hermite felt the necessity of eliminating complex integration in formula (2), Mikael felt the necessity of putting back complex integration in formula (8).

4. Complex Kergin interpolation

Would it be possible to define Kergin interpolation for a holomorphic function defined on a domain Ω that is not convex? Mikael studied the problem with his long-time collaborator Mats Andersson. The starting point was clear (in complex integration a simplex can be deformed) but the actual resolution was to lead them to investigate quite a few involved notions of complex convexity – first considered a few decades before by A. Martineau and further studied by several USSR analysts, and that they would conclude, jointly with Sigurdsson, with their monograph *Complex convexity and analytic functionals* [3].

In [1] Andersson and Passare showed that complex Kergin interpolation naturally lives on $\mathbb{C}$-convex domains. A domain in $\mathbb{C}^N$ is $\mathbb{C}$-*convex* when any of its intersections with a complex line is connected and simply connected (or empty). The point, of course, is to extend the integral terms in (8), that is the so-called *simplex functional*

$$f \mapsto \int_{\Delta^n} f\left(\sum_{i=0}^{n} t_i x_i\right) dm(t). \tag{9}$$

This functional actually plays a fundamental role in several problems on multivariate polynomial approximation. The description of the general extension is rather technical but it is easy to understand in the case of two points for which we have to extend the functional

$$\int_0^1 f\big(x_0 + t(x_1 - x_0)\big)dt \quad \text{where} \quad x_0,\ x_1 \in \Omega \subset \mathbb{C}^N, \tag{10}$$

and this is sufficient for understanding how $\mathbb{C}$-convexity intervenes. Indeed, since Ω is supposed to be $\mathbb{C}$-convex, the complex line $L = x_0 + \mathbb{C}(x_1 - x_0)$ intersects Ω in a simply connected domain Ω_L of L. You consider the affine map Φ from $\mathbb{C}$ to L defined by $\phi(z) = x_0 + z(x_1 - x_0)$. Then $\phi^{-1}(\Omega_L)$ is a simply connected domain in the ordinary complex plane (which contains 0 and 1) and

$$\int_\gamma (f \circ \phi)(z)dz = \int_\gamma f\big(x_0 + z(x_1 - x_0)\big)dz \tag{11}$$

does not depend on the regular path γ joining 0 to 1; this provides the correct extension of (10). Andersson and Passare also showed that the assumption of $\mathbb{C}$-convexity cannot be relaxed.

The same paper contains a second, essentially independent part, in which the authors derives a remainder formula for Kergin interpolation based on a Cauchy–Fantappiè representation formula that enabled them to extend a convergence result for entire functions due to Bloom [4]. I believe that the consideration of such an error formula directed them toward a formally simpler presentation of Kergin interpolation on $\mathbb{C}$-convex domain. To understand this second approach, it is necessary to briefly turn back to the main properties of Kergin's map.

5. Kergin and Fantappiè

The fundamental property of real Kergin interpolation (on convex sets) is that it provides a *lifting* of Hermite interpolation in the sense that, if f is a *ridge function*, that is, a *univariate* function h composed with a *linear form* ℓ, $g = h \circ \ell$, then

$$K(X, f) = H\big(\ell(X), h\big) \circ \ell, \tag{12}$$

thus, the Kergin interpolation polynomial of a ridge function is a *ridge polynomial* obtained as a composition of a Hermite interpolation polynomial. Such an invariance property implies many other interesting properties, most notably, the fact that $K(X, f)$ interpolates f at the points of X in the Hermite sense: if a point w is repeated α times in X then the first $\alpha - 1$ total derivatives of $K(X, f)$ at x equal the corresponding total derivatives of f. In particular, of course, if all points in X coincides then $K(X, f)$ is a multivariate Taylor polynomial. In fact, Formula (8) clearly presents Kergin interpolation polynomials as de-centred Taylor polynomials.

The space generated by ridge functions is dense in every space where polynomials are but, in general, there is no simple linear operator realizing such approximation. Things are fundamentally different in the complex setting, for many representation formulas can precisely be stated as relations of the following form

$$f(z) = \int f(w)\, k\Big(w\,,\, s_w^\star(z)\Big) d\mu(w), \quad u^\star(z) := \langle z, u\rangle = \sum_{i=1}^{n} z_i u_i, \tag{13}$$

where $k(\cdot,\cdot)$ is a certain kernel, $s : w \mapsto s_w \in \mathbb{C}^N$ and μ a measure. In such an expression, the connection to ridge functions becomes obvious. This observation has deep consequences.

The first one is that if Kergin interpolation is well defined, in a reasonable sense, then one may permute K and $\int$ and, in view of (12) and (13), we will have

$$K(X, f)(z) = \int f(w)\, H\big(s_w^\star(X)\,,\, k\,(w, \cdot)\,\big)\big(s_w^\star(z)\big)\; d\mu(w). \tag{14}$$

The trick enables transforming a problem on Kergin interpolation (which is a multivariate problem) into a (univariate) Hermite interpolation problem for which the data, the function and the point of evaluation all depend on a parameter w. One may then expect to use the available one-variable machinery to study approximation by Kergin interpolation polynomials. All convergence results on approximation of holomorphic functions by Kergin interpolants are based on this principle. It also explains why we cannot expect as rich a theory for Kergin interpolation as for (univariate) Hermite interpolation: roughly speaking, we need a sort of w-uniform dependency of the interpolation points $s_w^\star(x)$, the functions $k(w, \cdot)$ and the evaluation points $s_w^\star(z)$. Optimal results can be achieved only in quite particular cases[2].

The second consequence is due to Andersson and Passare: roughly, they guessed that a necessary and sufficient condition for defining complex Kergin interpolation should be the existence of a relation of the form (13). This is the content of their second paper on Kergin interpolation [2]. They first observe that a formally better setting naturally asks for a use of projective spaces on which the notion of $\mathbb{C}$-convexity easily extends. Given a set Ω in $\mathbb{P}^N$, one defines $\Omega^\star$ as the set of hyperplanes in $\mathbb{P}^N$ which do not intersect Ω: an element of $\Omega^\star$ is given by

[2]For convergence results on Kergin interpolants of holomorphic functions, see [5, 6]

$[\xi]$ where $\langle \xi, z\rangle \neq 0$ for every z in Ω. When Ω is open $\Omega^\star$ is compact. When Ω is a domain in $\mathbb{P}^N$, given two elements $\eta^\star \in \Omega^\star$ and $z, \eta \in \Omega$, the function

$$R_z : w \in \Omega^\star \mapsto \frac{\langle z, \eta^\star\rangle\langle \eta, w\rangle}{\langle z, w\rangle} \in \mathcal{H}(\Omega^\star), \tag{15}$$

that is, defines a holomorphic function on $\Omega^\star$. One may regard R_z as an example of the projective version of a ridge function. If $\mu \in \mathcal{H}'(\Omega^\star)$, i.e., μ is a continuous linear form on $\mathcal{H}(\Omega^\star)$, we may compute μ on R_z. This gives the Fantappiè transform,

$$\mathcal{F}\mu(z) = \mu(R_z), \tag{16}$$

and, as a function of z, $\mathcal{F}\mu$ will be a holomorphic function on Ω. Andersson and Passare used the remarkable fact that when the domain is $\mathbb{C}$-convex then the Fantappiè transform $\mathcal{F} : \mathcal{H}'(\Omega^\star) \to \mathcal{H}(\Omega)$ is a topological isomorphism (for the usual topologies). Thus every $f \in \mathcal{H}(\Omega)$ can be written as $f(z) = \mu_f(R_z)$ and the definition of the (projective) Kergin interpolation polynomial now becomes

$$PK(X, f)(z) = \mu_f \left\{ H\left(\frac{\langle X, \cdot\rangle}{\langle X, \eta^\star\rangle\langle \eta, \cdot\rangle}\ ,\ \frac{1}{\cdot} \right) \left(\frac{\langle z, \cdot\rangle}{\langle z, \eta^\star\rangle\langle \eta, \cdot\rangle} \right) \right\}, \tag{17}$$

where the letter P in front of K means that we use a certain projective version of Kergin polynomials.

This was only the door through which they entered their study of $\mathbb{C}$-convexity and analytic functionals. Passare and Andersson decided to clarify the notions and began to write some surveys that would reconstruct the whole edifice of $\mathbb{C}$-convexity in a way they judged satisfactory. I had these preprints on my work table for years and, every time I met with Mikael, I did not forget to ask him about the next chapter. When Ragnar Sigurdsson entered the team, the project finally ended in the publication of the beautiful monograph [3], a dedicated copy of which I am proud to have on my bookshelf.

6. Mean value interpolation

I first encountered Mikael at one session of the *Journées complexes du sud* in the early nineties. He gave a beautiful talk about Kergin interpolation both on his results and on those of his student Xing Yang who had just obtained a certain extension of the famous Müntz theorem on the approximation of continuous functions by the linear span of monomials with real exponents [17] as a by-product of Kergin interpolation. He accepted to be reviewer for my Ph.D. thesis; I found him a *revue technique* for a certain old French car. One evening, when we were waiting for his train, having a drink at the Matabiau station buffet, we understood that we stood firmly on the opposite sides of the fundamental dichotomy: one believed in Chance, the other in Will. This irreconcilable disagreement left us with solely frivolous topics of conversation such as, for instance, mathematics; and I visited him a couple of times in Stockholm. There, one afternoon, I mentioned a generalization of (real) Kergin interpolation that had been considered by Micchelli and his collaborators and which they called *Mean value interpolation* [8]. I pointed out that it should be possible to construct some *complex mean value interpolation* following the same lines as in his papers. He immediately showed his interest. I did not understand. Early, the next day, Mikael, myself and a few other colleagues met in the morning train to Kiselman's seminar in Uppsala. During the journey, he repeated his interest in mean value interpolation and patiently guided me to understand it as follows. The simplest univariate (real) mean value $M(X)$ operator is related to Hermite interpolation by the relation

$$M(X,f) = \mathcal{D}\Big(H(X\ ,\ \mathcal{D}^{-1}f)\Big); \tag{18}$$

where f is a continuous function, $\mathcal{D}$ indicates derivation and $\mathcal{D}^{-1}$ anti-derivation. It is readily seen that $M(X,f)$ does not depend on the anti-derivative that we choose so that $M(X,f)$ is correctly defined. If the definition is formally elegant, the utility of such an operator in approximation theory is perhaps not evident. Assume, for simplicity, that all the points in $X = \{x_0,\dots,x_n\}$ are pairwise distinct so that $H(X,\cdot)$ reduces to a Lagrange interpolation operator. Then,

$$\int_{x_0}^{x_j} M(X,f)(t)dt = \Big[H(X\ ,\ \mathcal{D}^{-1}f)\Big]_{x_0}^{x_j} = \Big[\mathcal{D}^{-1}f\Big]_{x_0}^{x_j} = \int_{x_0}^{x_j} f(t)dt. \tag{19}$$

Thus, in that case, $M(X,f)$ is the unique polynomial P of degree at most $n-1$ satisfying $\int_{x_i}^{x_j} P(t)dt = \int_{x_i}^{x_j} f(t)dt$ and such interpolation conditions naturally occur in approximation theory. Now, looking at the complex version, either from (18) or by considering the translation of the interpolation conditions that must be $\int_\gamma f(z)dz$ where γ is a regular path joining x_i to x_j, it is clear that simple connectedness is required. In other words, $\mathbb{C}$-convexity appears even in the univariate case unlike the case of Kergin interpolation, making it a still more natural condition. That was what Mikael understood the day before. He encouraged me to work jointly with Yang on this question but both of us were already busy with other problems and the project failed. The problem was finally, and excellently, handled by Lars Filipsson, another student of Mikael [12, 10][3]. Before Lars began to work, Mikael asked me whether the problem was free; I were still to receive further evidence of his uncommon elegance.

7. I still remember the last time we met in a seminar. He had moved to other mathematical subjects, years before. As usual, he gave an excellent talk in his peculiar style: he used to come with a few slides, parsimoniously filled with a few formulas. He spoke and you left the room with something new in you mind. He was able to arouse the interest of every open-minded mathematician. This day, someone who was not – and whose tone I remember – asked: "Why, why should we study such things?" It was a standard Tartuffe tone made of a mix of contempt, compassion and envy. Mikael answered as we do in similar cases, with

[3]Lars' thesis [11] also contains further results on Kergin interpolation, see also [12, 7].

an imperceptible smile on his face, he pointed out the elegance of the theory and mentioned applications. I think I know what he could not say: “Sorry, I don’t choose my mathematics, I am chosen by mathematics.”

References

[1] Andersson, M., and Passare, M. Complex Kergin interpolation. *J. Approx. Theory* 64, 2 (1991), 214–225.

[2] Andersson, M., and Passare, M. Complex Kergin interpolation and the Fantappiè transform. *Math. Z.* 208, 2 (1991), 257–271.

[3] Andersson, M., Passare, M., and Sigurdsson, R. *Complex convexity and analytic functionals*, vol. 225 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2004.

[4] Bloom, T. Kergin interpolation of entire functions on $\mathbf{C}^n$. *Duke Math. J.* 48, 1 (1981), 69–83.

[5] Bloom, T., and Calvi, J.-P. Kergin interpolants of holomorphic functions. *Constr. Approx.* 13, 4 (1997), 569–583.

[6] Bloom, T., and Calvi, J.-P. The distribution of extremal points for Kergin interpolation: real case. *Ann. Inst. Fourier* (*Grenoble*) 48, 1 (1998), 205–222.

[7] Calvi, J.-P., and Filipsson, L. The polynomial projectors that preserve homogeneous differential relations: a new characterization of Kergin interpolation. *East J. Approx.* 10, 4 (2004), 441–454.

[8] Cavaretta, Jr., A. S., Goodman, T. N. T., Micchelli, C. A., and Sharma, A. Multivariate interpolation and the Radon transform. III. Lagrange representation. In *Second Edmonton conference on approximation theory* (*Edmonton, Alta.,* 1982), vol. 3 of *CMS Conf. Proc.* Amer. Math. Soc., Providence, RI, 1983, pp. 37–50.

[9] de Boor, C. Polynomial interpolation. In *Proceedings of the International Congress of Mathematicians* (*Helsinki,* 1978) (1980), Acad. Sci. Fennica, Helsinki, pp. 917–922.

[10] Filipsson, L. Complex mean-value interpolation and approximation of holomorphic functions. *J. Approx. Theory* 91, 2 (1997), 244–278.

[11] Filipsson, L. *On polynomial interpolation and complex convexity.* ProQuest LLC, Ann Arbor, MI, 1999. Thesis (Dr. Techn.) – Kungliga Tekniska Hogskolan (Sweden).

[12] Filipsson, L. Kergin interpolation in Banach spaces. *J. Approx. Theory* 127, 1 (2004), 108–123.

[13] Kergin, P. G. *Interpolation of C^k functions.* ProQuest LLC, Ann Arbor, MI, 1978. Thesis (Ph.D.) – University of Toronto (Canada).

[14] Micchelli, C. A., and Milman, P. A formula for Kergin interpolation in $\mathbf{R}^k$. *J. Approx. Theory* 29, 4 (1980), 294–296.

[15] Smirnov, V. I., and Lebedev, N. A. *Functions of a complex variable: Constructive theory.* Translated from the Russian by Scripta Technica Ltd. The M.I.T. Press, Cambridge, Mass., 1968.

[16] Walsh, J. L. *Interpolation and approximation by rational functions in the complex domain.* Third edition. American Mathematical Society Colloquium Publications, Vol. XX. American Mathematical Society, Providence, R.I., 1960.

[17] Yang, X. Une généralisation à plusieurs variables du théorème de Müntz–Szász. *C. R. Acad. Sci. Paris Sér. I Math.* 312, 8 (1991), 575–578.

Jean-Paul Calvi
Institut de Mathématiques
Université de Toulouse III and CNRS (UMR 5219)
F-31062, Toulouse Cedex 9, France
e-mail: jean-paul.calvi@math.univ-toulouse.fr

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 157–190

Amoebas and their Tropicalizations – a Survey

Timo de Wolff

Dedicated to Mikael Passare (1959–2011)

Abstract. Let $\mathcal{V}(f)$ be the complex hypersurface of a Laurent polynomial f. The amoeba $\mathcal{A}(f)$ is the projection of $\mathcal{V}(f)$ under the Log-absolute map. Amoebas have countless applications and, in particular, they form a key connection between "classical" algebraic geometry and tropical geometry. There exist multiple different tropical hypersurfaces related to amoebas. In this survey, we introduce the most important of these tropical hypersurfaces and compare their relations to amoebas. Moreover, we discuss related open problems in amoeba theory.

As a new contribution we provide an example of an amoeba in $\mathbb{R}^2$ which has a component in the complement with an order not contained in the support of the defining polynomial. As a consequence, we conclude that an amoeba and its corresponding complement induced tropical hypersurface are not homotopy equivalent in general. Similarly, we prove that Archimedean amoebas and non-Archimedean amoebas are not homotopy equivalent in general.

Mathematics Subject Classification (2010). 12K10, 14M25, 14Q10, 14T05, 32Axx, 52B20, 55P10.

Keywords. Amoeba, complement induced tropicalization, Maslov dequantization, non-Archimedean amoeba, ArchTrop, spine, tropical geometry, tropicalization, valuation.

1. Introduction

In 1993/94 Gelfand, Kapranov and Zelevinsky introduced a new mathematical object in their book "Discriminants, Resultants and Multidimensional Determinants"[14], which they named "amoeba".

Specifically, let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}] := \mathbb{C}[z_1^{\pm 1}, \ldots, z_n^{\pm 1}]$ be a Laurent polynomial defining a *hypersurface* $\mathcal{V}(f) \subset (\mathbb{C}^*)^n := (\mathbb{C} \setminus \{0\})^n$. The *amoeba* $\mathcal{A}(f)$ (also referred to as *Archimedean amoeba*) of f is the image of $\mathcal{V}(f)$ under the Log-absolute map given by

$$\mathrm{Log}\,|\cdot|: \ (\mathbb{C}^*)^n \to \mathbb{R}^n, \quad (z_1,\ldots,z_n) \mapsto (\log|z_1|,\ldots,\log|z_n|)\,. \tag{1.1}$$

The *Newton polytope* $\mathrm{New}(f)$ of a Laurent polynomial f is the convex hull of the exponents of f. Gelfand, Kapranov and Zelevinsky motivated the definition of the amoeba $\mathcal{A}(f)$ by the relations between the hypersurface $\mathcal{V}(f)$ and the Newton polytope $\mathrm{New}(f)$ which are displayed by $\mathcal{A}(f)$; see [14, Section 6.1.B, Page 194], see also Section 2.1. Nowadays, we know that amoebas are objects of high interest themselves. They have amazing structural properties and are related to various mathematical subjects. These include dynamical systems [9], complex analysis [11, 33], nonnegativity of real polynomials [16], dimers and crystal shapes [19], the topology of real algebraic curves [23], statistical thermodynamics [30], the theory of stability preservers [32], discriminants [34], hyperbolicity and stability of polynomials [37], and the geometry of polynomials [47].

Most importantly, amoebas form a *bridge* between classical algebraic geometry and *tropical geometry*. Tropical geometry is a rising mathematical subject, which has been intensively discussed during roundabout the last 15 years. It investigates objects defined over the *tropical semiring* $(\mathbb{R}\cup\{-\infty\},\oplus,\odot)$. In this semiring the *tropical addition* $\oplus$ represents the usual maximum and the *tropical multiplication* $\odot$ represents the usual addition of two numbers. For a given support set $A \subset \mathbb{N}$ one defines a counterpart to “classical” polynomials in tropical geometry. A *tropical polynomial* is given by

$$f(\mathbf{x}) \ := \ \bigoplus_{\alpha\in A} b_\alpha \odot x_1^{\alpha_1} \odot \cdots \odot x_n^{\alpha_n}$$

with variables $x_1,\ldots,x_n \in \mathbb{R}$ and a coefficients $b_\alpha \in \mathbb{R}$ for all $\alpha \in A$. Note that for $1 \le j \le n$ we have

$$x_j^{\alpha_j} \ := \ \underbrace{x_j \odot \cdots \odot x_j}_{\alpha_j \text{ times}} \ = \ \alpha_j \cdot x_j.$$

The *tropical hypersurface* $\mathcal{T}(f)$, the zero set of a tropical polynomial f, is the piecewise linear subset of $\mathbb{R}^n$ where the maximum is attained at least by two tropical monomials of f. We give a more detailed description of tropical geometry in Section 2.2. For an overview about tropical geometry see, e.g., [5, 13, 17, 21, 39].

Tropical objects are due to their combinatorial nature much easier to handle than algebraic ones, but they still contain surprisingly much information of their algebraic counterparts. This makes them a powerful tool. In the abstract of their recent book Maclagan and Sturmfels summarize this fact as follows [21]:

“*Tropical geometry is a combinatorial shadow of algebraic geometry.*”

Amoebas live in between these two worlds. On the one hand, amoebas are projections of algebraic varieties. On the other hand, certain tropical hypersurfaces are deformation retracts of amoebas. In this sense amoebas can be described themselves by tropical polynomials and their hypersurfaces. Hence, from the perspective of amoeba theory, we might vary the previous quote to:

> "*Amoebas are shadows of algebraic varieties and tropical varieties are their combinatorial spirit.*"

Nevertheless, one important fact is often overseen in this context: There does not exist a unique tropical polynomial or a unique tropical hypersurface which can be associated to a given amoeba in a meaningful way. There exist several such tropical polynomials and tropical hypersurfaces with different properties. In what follows we refer to these tropical hypersurfaces as *tropicalizations* of amoebas. For example, the following two claims are common misconceptions about key objects in amoeba theory:

Misconception 1. The *non-Archimedean amoeba* (see Section 3) and the *spine* (see Section 6) of an amoeba are identical (or at least homotopy equivalent) tropical hypersurfaces.

Misconception 2. The spine is both a tropical hypersurface and a deformation retract of its corresponding amoeba; see for example [33, Theorem 1] and [35]. Since tropical hypersurfaces are piecewise linear, they can be computed easily via methods from linear algebra. Thus, the problem about the existence of components of the complement of amoebas originally stated by Gelfand, Kapranov and Zelevinsky (see [14, Remark 1.10, Page 198] and also Problem 2.6) can also be solved easily.

Both statements are wrong, as we conclude at the end of this survey in Corollary 7.1 and Remark 7.2.

Compared to tropical geometry, there exist few books and surveys about amoeba theory. Particularly, there exist the surveys [24, 35], the article [8] the mini survey [49], and the recently finished book [51] (in Russian). None of them, however, focuses on comparing the different relationships of amoebas with their tropicalizations. This fact combined with the existence of such common misunderstandings was the first motivation to write this survey.

The second motivation is that many contributions about amoebas and their tropicalizations were made by Mikael Passare, who tragically passed away in 2011, and by his students, particularly by Hans Rullgård in his seminal thesis [42]. Examples are [11, 12, 28, 29, 30, 31, 33, 34, 35, 41, 42]. I hope that this survey helps to popularize their work to a broader audience.

In this survey we focus on four key tropicalizations of amoebas:

1. the *non-Archimedean amoeba*, see Sections 3 and 4,
2. the *Archimedean tropical hypersurface*, see Section 5,
3. the *complement induced tropical hypersurface*, see Section 5, and
4. the *spine*, see Section 6.

Roughly speaking, these tropicalizations yield more accurate descriptions of the original amoeba and have more of its original properties in the order of appearance. In the same order it becomes more complicated to determine their defining tropical polynomial; see the comparison in Section 7 at the end of the article. In summary, *all* of these tropicalizations are useful and it depends on the situation which one is the best choice.

Some of the tropicalizations (1)–(4) have been discussed in other surveys, too. The non-Archimedean amoeba can be found in most books and surveys about tropical geometry like [21], mostly from the valuation point of view. It is also covered in Mikhalkin's article [25] and his survey about amoebas [24] and Viro's surveys like [48] (mostly) about his patchworking technique. The spine was discussed in the survey [35] by Passare and Tsikh. The tropicalizations (2) and (3) were, to the best of my knowledge, not part of a survey so far. However, (3) was mentioned in Rullgård's thesis, see [42, Remark on Page 33 and proof of Theorem 12], and was also discussed in my thesis; see [6, Section 4.1.1]. However, there has not been a detailed *comparison* of the different relationships of these tropicalizations with amoebas (a partial, brief one is given in [21, Section 1.4]). The main purpose of this survey is to introduce these four tropicalizations, explain how they are constructed, and to point out their properties and their differences with respect to amoebas. Moreover, I want to emphasize some open problems in amoeba theory, which are related to the different tropicalizations.

This survey also contains some new contributions. Using a statement from Rullgård's thesis [42, Theorem 11, Part 2, Page 36] we are able to conclude that the complement induced tropical hypersurface is not a deformation retract of its corresponding amoeba in general; see Corollary 5.14. We provide an explicit counterexample in Example 5.13. This result and further parts of this survey were also covered in my thesis; see [6, Example 4.44, Corollary 4.45]. We show furthermore that amoebas and their corresponding non-Archimedean amoebas are not homotopy equivalent in general using a class of polynomials introduced by Passare and Rullgård; see Theorem 3.3.

2. Preliminaries

In this section we introduce the notation needed for amoebas and tropical geometry. Moreover, we recall some fundamental statements about amoebas, which partially also motivate our notation.

2.1. Amoebas

For a broader introduction to amoebas I recommend the reader to also consult the following sources: The book [14] by Gelfand, Kapranov and Zelevinsky, the surveys [24] by Mikhalkin and [35] by Passare and Tsikh, and the theses [6] by myself and [42] by Rullgård.

The first statements about amoebas were proven by Gelfand, Kapranov and Zelevinsky together with their initial definition of amoebas. We start with some topological facts.

Theorem 2.1 (Gelfand, Kapranov, Zelevinsky, [14]). *For $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ with $f(\mathbf{z}) \not\equiv 0$ the amoeba $\mathcal{A}(f)$ is a closed set (with respect to the standard topology) with non-empty complement.*

Let $\mathcal{A}(f)^c$ denote the *complement* of the amoeba $\mathcal{A}(f)$. We say $E \subseteq \mathcal{A}(f)^c$ with $E \neq \emptyset$ is a *component* of the complement of $\mathcal{A}(f)$ if E and $\mathcal{A}(f)^c \setminus E$ are not connected. Gelfand, Kapranov and Zelevinsky knew already in 1993/94 that these components of the complement contain crucial information about the original Laurent polynomial f and its hypersurface $\mathcal{V}(f)$; see [14, Cor. 1.6, Page 195]

Theorem 2.2 (Gelfand, Kapranov, Zelevinsky, [14]). *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$. Every component of the complement of the amoeba $\mathcal{A}(f)$ is convex. The set of all components of the complement of $\mathcal{A}(f)$ corresponds bijectively to the set of all Laurent expansions of $1/f$ centered at the origin.*

We call the set of all exponents of a Laurent polynomial f the *support* of f. Recall that the *Newton polytope* $\mathrm{New}(f)$ of a Laurent polynomial f is the lattice polytope given by the convex hull of the support of f, see, e.g., [14, 53]. The components of the complement of $\mathcal{A}(f)$ have an important combinatorial relation to the Newton polytope $\mathrm{New}(f)$ of f. This was partially known by Gelfand, Kapranov, and Zelevinsky (see the part about normal fans below) and discovered in depth by Forsberg, Passare, and Tsikh with their introduction of the order map in [11]. Every component of the complement of a given amoeba $\mathcal{A}(f)$ corresponds to a unique lattice point in the Newton polytope $\mathrm{New}(f)$ of f via the *order map*:

$$\begin{aligned} \mathrm{ord}: \mathbb{R}^n \setminus \mathcal{A}(f) \quad &\to \quad \mathrm{New}(f) \cap \mathbb{Z}^n, \quad \mathbf{w} \mapsto (u_1, \ldots, u_n) \ \text{ with} \\ u_j \quad &:= \quad \frac{1}{(2\pi i)^n} \int_{\mathrm{Log}\,|\mathbf{z}| = \mathbf{w}} \frac{z_j \partial_j f(\mathbf{z})}{f(\mathbf{z})} \frac{dz_1 \cdots dz_n}{z_1 \cdots z_n} \quad \text{for all } \ 1 \leq j \leq n\,. \end{aligned} \tag{2.1}$$

The order map can be understood as a multivariate analogue of the classical *argument principle* from complex analysis. It states that for a function f in z, which is meromorphic in a domain $\Omega \subseteq \mathbb{C}$, the following statement holds: Let γ be a closed path in Ω such that f has no zeros and poles on γ. Then

$$\frac{1}{2\pi i} \cdot \int_\gamma \frac{f'(z)}{f(z)} dz = \#\text{ zeros of } f - \#\text{ poles of } f \text{ in the region bounded by } \gamma.$$

Particularly, the image of the order map (2.1) is a set of integer vectors, the order map is constant on each component of the complement of $\mathcal{A}(f)$, and it is injective on the set of all components; see [11, Prop. 2.5, Theorem 2.8]

Theorem 2.3 (Forsberg, Passare, Tsikh). *The image of the order map is contained in $\mathrm{New}(f) \cap \mathbb{Z}^n$. Let $\mathbf{w}, \mathbf{w}' \in \mathcal{A}(f)^c$. Then $\mathbf{w}$ and $\mathbf{w}'$ belong to the same component of the complement of $\mathcal{A}(f)$ if and only if $\mathrm{ord}(\mathbf{w}) = \mathrm{ord}(\mathbf{w}')$.*

We define for each $\alpha \in \mathrm{New}(f) \cap \mathbb{Z}^n$ the set

$$E_\alpha(f) := \{\mathbf{w} \in \mathbb{R}^n \setminus \mathcal{A}(f) \ : \ \mathrm{ord}(\mathbf{w}) = \alpha\}. \tag{2.2}$$

As a consequence of Theorem 2.3 each $E_\alpha(f)$ equals the connected component in the complement of the amoeba $\mathcal{A}(f)$ which has the order α. Hence, we call $E_\alpha(f)$ *the component of order* α of the complement of $\mathcal{A}(f)$. For a given support set $A \subset \mathbb{Z}^n$ and $f \in (\mathbb{C}^*)^A$ we define the set:

$$\mathrm{Comp}(f) := \{\alpha \in \mathrm{conv}(A) \cap \mathbb{Z}^n \ : \ E_\alpha(f) \neq \emptyset\}. \tag{2.3}$$

By the definition of the sets $E_\alpha(f)$ this means $\mathrm{Comp}(f)$ contains all lattice points in $\mathrm{conv}(A)$ which are associated to existing components of the complement of $\mathcal{A}(f)$ via the order map.

For a given Newton polytope $\mathrm{New}(f)$ we denote its *normal fan* by $\mathrm{NF}(f)$. A definition of the normal fan can be found in [53, Page 193]. If S is a face of $\mathrm{New}(f)$, then let $\mathrm{NF}_S(f)$ denote the corresponding *dual cone* in $\mathrm{NF}(f)$. It was already shown by Gelfand, Kapranov and Zelevinsky [Prop. 1.7., Page 195][14] that the amoeba $\mathcal{A}(f)$ is related to the normal fan $\mathrm{NF}(f)$. This relation is an example for the connection between hypersurfaces and their Newton polytopes, which is displayed in the amoeba and which was mentioned in the introduction.

Theorem 2.4 (Gelfand, Kapranov, Zelevinsky, [14]). *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ with support set $A \subset \mathbb{Z}^n$. If $\alpha \in A$ is a vertex of* $\mathrm{New}(f)$*, then $E_\alpha(f) \neq \emptyset$. Furthermore, $E_\alpha(f)$ contains an affine translation of the full-dimensional cone* $\mathrm{NF}_{\{\alpha\}}(f)$ *in* $\mathrm{NF}(f)$*, which is dual to the vertex α in* $\mathrm{New}(f)$*.*

For a given polytope $P \subset \mathbb{R}^n$ we denote the set of its vertices by $\mathrm{vert}(P)$. In summary, we can conclude the following at this point.

Corollary 2.5. *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$. Then we have*

$$\mathrm{vert}(\mathrm{New}(f)) \ \subseteq \ \mathrm{Comp}(f) \ \subseteq \ \mathrm{New}(f) \cap \mathbb{Z}^n.$$

Thus, the number of vertices of $\mathrm{New}(f)$ *is a lower bound and the number of lattice points in* $\mathrm{New}(f)$ *is an upper bound for the number of components of the complement of an amoeba.*

Later, we will see that there is a deeper connection between hypersurfaces, amoebas, Newton polytopes and their normals fans based on tropical geometry.

We define the *parameter space* or *configuration space* of coefficients for a fixed (finite) support set $A \subset \mathbb{Z}^n$ as

$$(\mathbb{C}^*)^A \quad := \quad \left\{ \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \in \mathbb{C}[\mathbf{z}^{\pm 1}] \ : \ b_\alpha \in \mathbb{C}^* \right\}.$$

That means, $(\mathbb{C}^*)^A$ is the set of all Laurent polynomials which have exactly the set A as support set. The investigation of such spaces $(\mathbb{C}^*)^A$ for a fixed support set A was excessively used by Gelfand, Kapranov and Zelevinsky [14] and later by Rullgård [42] and others and has proven to be very powerful. It is also referred to as *A-philosophy*; see [14, Chapter 5, Section 1, Part A]. In $(\mathbb{C}^*)^A$ we can identify every polynomial with its coefficient vector. Thus, we can identify the parameter space $(\mathbb{C}^*)^A$ with a $(\mathbb{C}^*)^d$ space, where $d := \#A$. Note that it is also common to investigate a space $\mathbb{C}^A$ instead of $(\mathbb{C}^*)^A$ where coefficients are allowed to be zero. Note furthermore that for every $f \in (\mathbb{C}^*)^A$ it holds that $\mathrm{New}(f) = \mathrm{conv}(A)$. For the remainder of the article we always assume that $\mathrm{conv}(A)$ is a full-dimensional polytope in $\mathbb{R}^n$ and hence $d \geq n+1$.

One key problem in amoeba theory is to understand the sets

$$U_\alpha^A \quad := \quad \{f \in (\mathbb{C}^*)^A \ : \ E_\alpha(f) \neq \emptyset\},$$

i.e., the set of all polynomials in $(\mathbb{C}^*)^A$, whose amoebas have a component of order α in the complement.

Problem 2.6. *Find an algebraic and/or a topological description of the sets U_α^A. Furthermore, determine for a fixed $A \subset \mathbb{Z}^n$ for which $\alpha \in \mathrm{conv}(A) \cap \mathbb{Z}^n$ it holds that $U_\alpha^A = \emptyset$.*

The problem to find an algebraic and topological description of the sets U_α^A was already stated (in a simpler way) by Gelfand, Kapranov and Zelevinsky; see [14, Remark 1.10, Page 198]. The sets U_α^A were first studied systematically by Rullgård. He showed that every U_α^A is an open, semi-algebraic set and its complement $(U_\alpha^A)^c$ is connected. Note that the semi-algebraicity is non-explicit and follows via a Tarski–Seidenberg argument. Unless α is a vertex of $\mathrm{conv}(A)$ the set U_α^A is a strict subset of $(\mathbb{C}^*)^A$, which is possibly empty if $\alpha \notin A$; see [42, Theorem 10, Corollary 5, Theorem 14], see also [41]. For background information about semi-algebraic sets see for example [2].

Example 2.7. Let $f(z_1, z_2) := 1 + z_1^3 + z_2^3 + 3z_1z_2^2 + 3z_1 + 10z_1z_2$. The Newton polytope $\mathrm{New}(f)$ is the simplex given by $\mathrm{conv}\{(0,3),(3,0),(0,0)\}$. Thus, the complement of the amoeba $\mathcal{A}(f)$ has three unbounded components corresponding to the vertices $(0,3),(3,0)$, and $(0,0)$. Additionally, by Corollary 2.5, the complement of an amoeba of a polynomial with support set $A = \{(0,3),\ (1,0),\ (1,1),\ (1,2),\ (3,0),\ (0,0)\}$ can have at most seven non-vertex components. More precisely, it can have six unbounded non-vertex components and a single bounded one, corresponding to the non-vertex lattice points in $\mathrm{New}(f)$. The existence of these components depends on the choice of the coefficients of f. For the given choice of coefficients, the bounded and two of the unbounded components exist in addition to the three mandatory ones; see Figure 1.

For some of the following statements it is necessary to have a brief look at the *fiber* of a point $\mathbf{w} \in \mathbb{R}^n$. We define the fiber $\mathbb{F}_{\mathbf{w}}$ of $\mathbf{w} \in \mathbb{R}^n$ with respect to the $\mathrm{Log}\,|\cdot|$ map as

$$\mathbb{F}_{\mathbf{w}} := \mathrm{Log}^{-1}|\mathbf{w}| = \{\mathbf{z} \in (\mathbb{C}^*)^n \ : \ \mathrm{Log}\,|\mathbf{z}| = \mathbf{w}\}.$$

Recall that a branch of the holomorphic logarithm is defined as

$$\log_{\mathbb{C}} : \mathbb{C}^* \to \mathbb{C}, \quad z \mapsto \log|z| + i\arg(z),$$

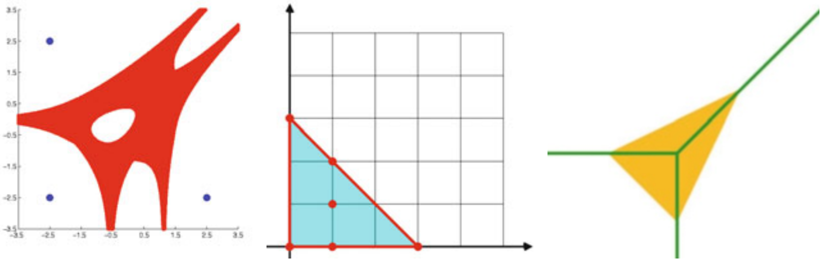


FIGURE 1. The left picture contains an approximation of the amoeba of $f(z_1, z_2) := 1 + z_1^3 + z_2^3 + 3z_1z_2^2 + 3z_1 + 10z_1z_2$. Its complement has six components. The three components corresponding to the vertices of $\mathrm{New}(f)$ via the order map contain a blue dot. One can see that each of them contains an affine translation of a cone of the normal fan of $\mathrm{New}(f)$, which is shown in the right picture. The three remaining components can vanish for another choice of the coefficients of f. The middle picture contains the Newton polytope of f.

where $\arg(z)$ denotes the argument of the complex number z. This means that the log absolute map $\log|\cdot|$ equals the real part of the complex logarithm. The multivariate case works componentwise like the univariate case. As a consequence, the holomorphic logarithm $\mathrm{Log}_{\mathbb{C}}$ yields a fiber bundle structure $(S^1)^n \to (\mathbb{C}^*)^n \to \mathbb{R}^n$ for the map $\mathrm{Log}\,|\cdot|$ such that the following diagram commutes; see [6, 23, 24, 44] for further details:

$$\begin{tikzcd}
(\mathbb{C}^*)^n \arrow[rr, "\mathrm{Log}_{\mathbb{C}}"] \arrow[dr, "\mathrm{Log}\,|\cdot|"'] & & \mathbb{R}^n \times (S^1)^n. \arrow[dl, "\mathrm{Re}"] \\
& \mathbb{R}^n &
\end{tikzcd}$$

Hence, every fiber $\mathbb{F}_{\mathbf{w}}$ is homeomorphic to a torus $(S^1)^n$. For background literature about fibrations see for example [15]. In particular, every fiber $\mathbb{F}_{\mathbf{w}}$ is compact and a point $\mathbf{w} \in \mathbb{R}^n$ is contained in the amoeba $\mathcal{A}(f)$ if and only if $f(\mathbf{z}) = 0$ for some $\mathbf{z} \in \mathbb{F}_{\mathbf{w}}$, see Figure 2.

2.2. The tropical semi-ring

Tropical geometry has been an emerging topic in mathematics within the last roundabout 15 years. It investigates the geometrical properties of the *tropical semi-ring* $(\mathbb{R} \cup \{-\infty\}, \oplus, \odot)$. Recall from the introduction that

$$a \oplus b \;:=\; \max\{a, b\}, \quad \text{and} \quad a \odot b \;:=\; a + b.$$

Thus, the neutral element for the tropical addition is $-\infty$ and the neutral element for tropical multiplication is 0. It is not a ring since we do not have inverse elements with respect to addition. Note that some authors prefer the minimum together with

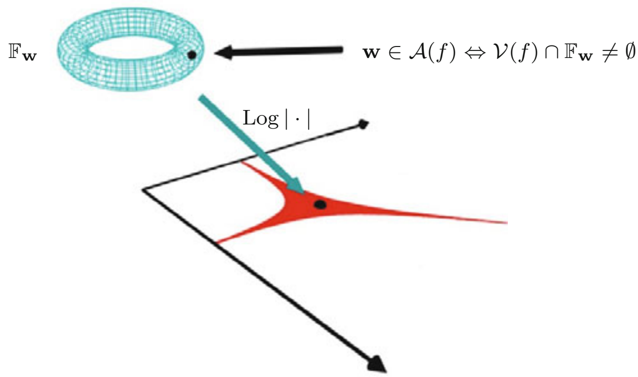


FIGURE 2. An amoeba $\mathcal{A}(f)$ with a fiber of a point $\mathbf{w} \in \mathbb{R}^n$ with respect to the $\mathrm{Log}\,|\cdot|$ map.

$+\infty$ instead of the maximum as tropical addition. For a general introduction to tropical geometry see [5, 13, 17, 21, 39].

Recall from the introduction that a *tropical polynomial* with support set $A \subset \mathbb{N}^n$ is a finite tropical sum of tropical monomials, i.e., it is a function

$$\mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto \bigoplus_{\alpha \in A} b_\alpha \odot \mathbf{x}^\alpha \;=\; \max_{\alpha \in A}\{b_\alpha + \langle \mathbf{x}, \alpha \rangle\}.$$

with $b_\alpha \in \mathbb{R}$. Note that a tropical monomial $b_\alpha \odot \mathbf{x}^\alpha$ does not vanish if $b_\alpha = 0$, but it vanishes if $b_\alpha = -\infty$. For formal reasons we have to allow in what follows to add redundant terms $-\infty \odot \mathbf{x}^\alpha$ as they can appear as a result of computations in certain cases. As for classical polynomials we do, however, disregard such redundant terms regarding the support, the Newton polytope, or the tropical hypersurface of a tropical polynomial. Thus, the support set $A \subset \mathbb{N}^n$ is, analog to classical polynomials, the set of non-vanishing terms of the tropical polynomial.

We see that a tropical monomial in terms of classical operations is the affine linear form $b_\alpha + \langle \mathbf{x}, \alpha \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product.

For a tropical polynomial h, the *tropical hypersurface* $\mathcal{T}(h)$ is defined as the set of points in $\mathbb{R}^n$ where *the maximum is attained at least by two tropical monomials*. In several contexts $\mathcal{T}(h)$ is defined over $(\mathbb{R} \cup \{-\infty\})^n$ instead, see, e.g., [21, 27] for further details. In this survey, we restrict ourselves to $\mathbb{R}^n$ for convenience of the reader and to avoid certain technicalities.

A tropical hypersurface $\mathcal{T}(h)$ is a polyhedral complex, which is dual to a subdivision of the Newton polytope $\mathrm{New}(h)$ of h, where $\mathrm{New}(h)$ is defined as the convex hull of the set of exponents of h, analogously to usual Newton polytopes. For background information about polyhedral complexes, (lower) convex hulls, and related objects see [53].

The subdivision of the Newton polytope is induced in the following way: Every lattice point $\alpha \in A$ is lifted to $(\alpha, -b_\alpha) \in \mathbb{R}^{n+1}$ where we ignore possible redundant terms with $b_\alpha = -\infty$, which can never attain a maximum. We refer to the convex hull of these lifted points as the *lifted Newton polytope*. Projecting the lower convex hull of the lifted Newton polytope down to $\mathbb{R}^n$ yields the desired subdivision of $\mathrm{New}(h)$. Such a subdivision is called a *regular* subdivision [14, 20]. See Figure 3 for some examples of tropical hypersurfaces.

In the dual picture, every tropical monomial represents an affine linear hyperplane. Therefore, the set of all monomials in a tropical polynomial corresponds to a hyperplane arrangement. The intersection of the positive half-spaces of the hyperplanes in the arrangement defines a polyhedron P. Taking the maximum of the tropical monomials corresponds geometrically to taking the lower convex hull of the intersection of the positive half-spaces of the arrangement. That is, it corresponds to taking the lower convex hull of P; see [21, 39]. Since a tropical hypersurface is given by the points where the maximum is attained twice, it corresponds to the projection of the non-smooth points of this lower convex hull to $\mathbb{R}^n$. This is the projection of the points where at least two facets belonging to the lower convex hull of P intersect. See Figure 4 for an example of the polyhedron given by the intersection of positive half-spaces of the hyperplane arrangement and the corresponding tropical hypersurface. See [21] for further details.

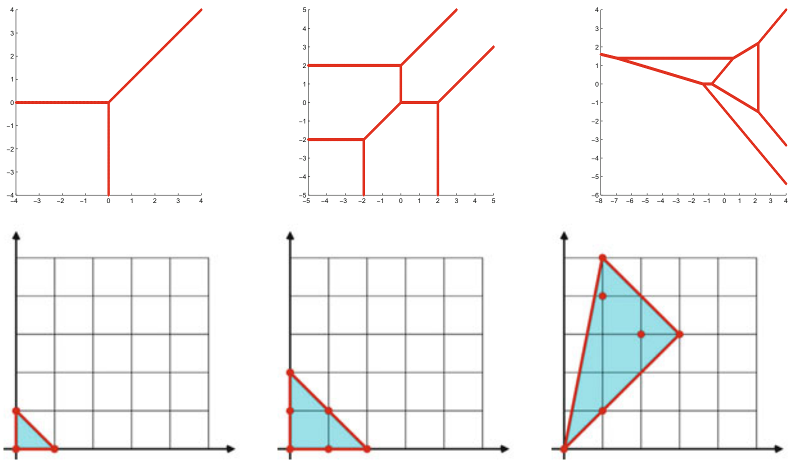


FIGURE 3. The tropical hypersurfaces of the tropical polynomials $f_1(x_1, x_2) := 1x_1 \oplus 1x_2 \oplus 1$, $f_2(x_1, x_2) := 1x_1^2 \oplus 3x_1 \oplus 1x_2^2 \oplus 3x_2 \oplus 3x_1x_2 \oplus 1$ and $f_3(x_1, x_2) := 0x_1^3x_2^3 \oplus \log|9|x_1^2x_2^3 \oplus 0x_1x_2^5 \oplus \log|4|x_1x_2^4 \oplus \log|4|x_1x_2 \oplus 0$ and their corresponding Newton polytopes. All multiplications of coefficients and variables are meant to be tropical. That is, e.g., $1x_1^2$ means $1 \odot x_1 \odot x_1$.

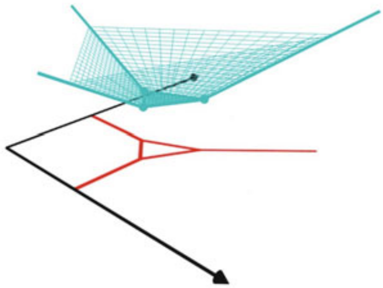


FIGURE 4. A lower convex hull of an intersection of positive half-spaces of a hyperplane arrangement and its corresponding tropical hypersurface.

3. Valuations and the non-Archimedean amoeba

The probably most common way to *tropicalize* a hypersurface, i.e., to associate a tropical hypersurface to an algebraic hypersurface, is to use *valuations*. The resulting tropical hypersurface is commonly called the *non-Archimedean amoeba*. In this section we describe valuations, tropicalizations, and we explain the origin of the term "non-Archimedean amoeba". Furthermore, we demonstrate why valuation maps are a proper choice for the connection between the classical and the tropical world, but also why they are often not sufficient to tackle key questions in amoeba theory itself.

Note that not only hypersurfaces can be tropicalized, but also general varieties, which are not given by principal ideals. Here, we restrict ourselves to the hypersurface case since amoebas of ideals are merely understood so far. For the more general case we refer the reader to the literature mentioned in Section 2.2, particularly [21]. For background on valuations the reader may study literature about commutative algebra like [3]. For more information about the non-Archimedean amoeba the reader may consult literature about tropical geometry like [5, 13, 17, 21, 39].

For a ring $(R, \tilde{+}, \tilde{\cdot})$ and a totally ordered commutative group $(G, +)$ a *valuation* on R with values in G is a map $\nu : R \to G \cup \{\infty\}$, which satisfies the following axioms, see [3, Page 386]:

$$\begin{aligned} \nu(x \,\tilde{\cdot}\, y) &= \nu(x) + \nu(y) \quad \text{for every} \quad x, y \in R, \\ \nu(x \,\tilde{+}\, y) &\geq \inf\{\nu(x), \nu(y)\} \quad \text{for every} \quad x, y \in R, \\ \nu(1) &= 0 \quad \text{and} \quad \nu(0) \ = \ \infty. \end{aligned}$$

If $G = \mathbb{R}$, then we say that R is *real valuated*. In this case the image of ν forms an additive subgroup of $\mathbb{R}$, [21]. On real valuated fields F the valuation map

$\nu : F \to \mathbb{R} \cup \{\infty\}$ induces a norm on F; [3, p. 428 et seq.], see also [21]. This norm is given by

$$|\cdot|_\nu : F \to \mathbb{R}, \quad z \mapsto e^{-\nu(z)}.$$

The norm $|\cdot|_\nu$ is *non-Archimedean*, i.e., it satisfies $|x+y|_\nu \leq \max\{|x|_\nu, |y|_\nu\}$.

The *field of Puiseux series* $\mathbb{K}$ is the set of all formal power sums $\sum_{q \in Q} b_q t^q$ such that all $b_q \in \mathbb{C}$, and Q is a well-ordered subset of the rational numbers, such that all $q \in Q$ share a common denominator [8, 21, 39]. On the field of Puiseux series $\mathbb{K}$ there exists a real valuation map $\mathrm{val} : \mathbb{K} \to \mathbb{R} \cup \{\infty\}$, which is given by

$$\mathrm{val}\left(\sum_{q \in Q} b_q t^q\right) \quad := \quad \min\{q \,:\, b_q \neq 0\}. \tag{3.1}$$

Note that the minimum always exists since the support set Q of every element in $\mathbb{K}$ is well ordered. Hence, $\mathbb{K}$ is a real valuated field.

Let $f \in \mathbb{K}[\mathbf{z}^{\pm 1}]$ be a Laurent polynomial over the field of Puiseux series. Kapranov [8, 18], see also [24, 26, 39] defined for a given algebraic variety $\mathcal{V}(f) \subset (\mathbb{K}^*)^n$ its *non-Archimedean amoeba* $\mathcal{A}_\mathbb{K}(f)$ by $\mathcal{A}_\mathbb{K}(f) := \overline{\mathrm{Log}_\mathbb{K}|\mathcal{V}(f)|}$, where $\mathrm{Log}_\mathbb{K}|\cdot|$ is given by

$$\mathrm{Log}_\mathbb{K}|\cdot| : (\mathbb{K}^*)^n \to \mathbb{R}^n, \ (z_1, \ldots, z_n) \mapsto (\log|z_1|_{\mathrm{val}}, \ldots, \log|z_n|_{\mathrm{val}}),$$

where $|\cdot|_{\mathrm{val}}$ denotes the norm on $\mathbb{K}$ induced by the valuation val. In other words, we have $\log|z_j|_{\mathrm{val}} = -\mathrm{val}(z_j)$ for every $1 \leq j \leq n$. Note that $\mathrm{Log}_\mathbb{K}|\cdot|$ is well defined here, since we assumed that $\mathcal{V}(f) \subset (\mathbb{K}^*)^n$. The analogy between the maps $\mathrm{Log}|\cdot|$, see (1.1), and $\mathrm{Log}_\mathbb{K}|\cdot|$ explains the name *non-Archimedean amoeba.* In fact, the non-Archimedean amoeba is a tropical hypersurface as the following statement, known as *Kapranov's theorem*, shows.

Theorem 3.1 (Kapranov [18]; see also [21, 24, 39]). *Let $A \subset \mathbb{Z}^n$ finite, $f(\mathbf{z}) := \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \in \mathbb{K}[\mathbf{z}^{\pm 1}]$ with $\mathcal{V}(f) \subset (\mathbb{K}^*)^n$, and let $h(\mathbf{x}) := \bigoplus_{\alpha \in A} -\mathrm{val}(b_\alpha) \odot \mathbf{x}^\alpha$. Then the non-Archimedean amoeba $\mathcal{A}_\mathbb{K}(f)$ equals the tropical hypersurface $\mathcal{T}(h)$.*

The generalization of Kapranov's theorem for ideals is referred to as *Fundamental Theorem of Tropical Geometry.* It was formulated by Speyer and Sturmfels in [45]; see [21, Theorem 3.2.5].

Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ be a Laurent polynomial with a hypersurface $\mathcal{V}(f) \subseteq (\mathbb{C}^*)^n$. Since $\mathbb{C}^*$ is isomorphic to a subring of $\mathbb{K}^*$ one can investigate $\mathrm{Log}_\mathbb{K}|\mathcal{V}(f)|$. This is, however, not helpful, since

$$\mathrm{Log}_\mathbb{K}|(\mathbb{C}^*)^n| \ = \ \mathbf{0} \ \text{ and thus } \ \mathrm{Log}_\mathbb{K}|\mathcal{V}(f)| \ = \ \mathbf{0}.$$

Since $\mathbb{C}[\mathbf{z}^{\pm 1}]$ is isomorphic to a subring of $\mathbb{K}[\mathbf{z}^{\pm 1}]$ we can instead interpret the Laurent polynomial f as a polynomial $f_\mathbb{K}$ over field of Puiseux series $\mathbb{K}[\mathbf{z}^{\pm 1}]$ via the following embedding:

$$\Psi : \mathbb{C}[\mathbf{z}^{\pm 1}] \to \mathbb{K}[\mathbf{z}^{\pm 1}], \quad \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \mapsto \sum_{\alpha \in A} \left(\sum_{q \in \{0\}} b_{\alpha,q} \cdot t^0 \right) \mathbf{z}^\alpha = \sum_{\alpha \in A} \left(b_\alpha \cdot t^0 \right) \mathbf{z}^\alpha.$$

That means, we interpret every complex coefficient b_α of f as a Puiseux series coefficient given by a series with a single term $b_\alpha t^0$. Hence, we also obtain a hypersurface $\mathcal{V}_\mathbb{K}(f_\mathbb{K}) \subseteq (\mathbb{K}^*)^n$ which contains the hypersurface $\mathcal{V}(f) \subseteq (\mathbb{C}^*)^n \subset (\mathbb{K}^*)^n$ as a proper subset. Therefore, we also gain a canonical map between Archimedean amoebas and non-Archimedean amoebas for every f:

$$\Psi_\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n, \quad \mathcal{A}(f) \mapsto \mathcal{A}_\mathbb{K}(f_\mathbb{K}) = \mathcal{A}_\mathbb{K}(\Psi(f))$$

such that we obtain the following diagram:

$$\begin{tikzcd} f \arrow[r, hook, "\Psi(f)"] \arrow[d, "\mathrm{Log}\,|\mathcal{V}(f)|"'] & f_\mathbb{K} \arrow[d, "\mathrm{Log}_\mathbb{K}\,|\mathcal{V}_\mathbb{K}(f)|"] \\ \mathcal{A}(f) \arrow[r, "\Psi_\mathcal{A}(\mathcal{A}(f))"] & \mathcal{A}_\mathbb{K}(f_\mathbb{K}) \end{tikzcd} \tag{3.2}$$

Example 3.2. Let $f(z_1, z_2) := z_1 + z_2 - 1$ with an amoeba $\mathcal{A}(f)$. Over the field of Puiseux series f is of the form $f_\mathbb{K}(z_1, z_2) = t^0 z_1 + t^0 z_2 - t^0$. Thus, by Kapranov's theorem, the corresponding non-Archimedean amoeba $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ is the tropical hypersurface given by the tropical polynomial $h(x_1, x_2) = \mathrm{val}(0) \odot x_1 \oplus \mathrm{val}(0) \odot x_1 \oplus \mathrm{val}(0) = 0 \odot x_1 \oplus 0 \odot x_2 \oplus 0$.

We have, however, not developed an explanation in this article so far, how the amoeba $\mathcal{A}(f)$ of a Laurent polynomial $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ is related to the non-Archimedean amoeba $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ geometrically. More precisely, so far, we cannot apply the map $\Psi_\mathcal{A}$ on its own, since we cannot determine f from $\mathcal{A}(f)$. But even if both f and $\mathcal{A}(f)$ are given, then we have to compute $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ via applying the valuation map. From the algebraic and the computational perspective valuations are a powerful tool, but from the geometric perspective they are an "algebraic black box". A valuation map takes a classical polynomial as input and provides a tropical polynomial as output without any geometry involved. Later, in Section 4, we solve this issue and also provide the missing geometric relation between the objects $\mathcal{A}(f)$ and $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ by investigating Maslov dequantizations.

Before we get there we make an observation, which gives a first idea why for some Laurent polynomial $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$, from the viewpoint of amoeba theory, the non-Archimedean amoeba $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ is not sufficient to understand the amoeba $\mathcal{A}(f)$.

Theorem 3.3. *For $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ the non-Archimedean amoeba $\mathcal{A}_\mathbb{K}(f_\mathbb{K})$ is not homotopy equivalent to $\mathcal{A}(f)$ in general.*

This theorem is a consequence of a crucial class of polynomials given by Passare and Rullgård. To my awareness, Theorem 3.3 is very well-known in the amoeba community. However, I am neither aware of a reference nor of a published proof. Therefore, we provide a proof for Theorem 3.3 in this section; see also Figure 6. The class of polynomials given by Passare and Rullgård and its key properties are described in the following theorem; see [33, Proposition 2].

Theorem 3.4 (The Passare–Rullgård Polynomials). *Let* $f(\mathbf{z}) := 1 + \sum_{j=1}^{n} z_j^{n+1} + a \cdot z_1 \cdots z_n$ *with* $a \in \mathbb{C}$. *Then the complement of* $\mathcal{A}(f)$ *has at most one bounded component and the following statements are equivalent:*

1. *The complement of* $\mathcal{A}(f)$ *has a bounded component of order* $(1, \ldots, 1)$.
2. $\mathcal{A}(f)$ *does not contain the origin.*
3. $a \notin \{-\sum_{j=0}^{n} e^{i \cdot 2\pi\phi_j} \;:\; \phi_1, \ldots, \phi_n \in [0, 2\pi)$ *and* $\sum_{j=0}^{n} \phi_j = 2\pi\}$.

This example class is very important since it provides counterexamples for various properties one might conjecture amoebas to have. Although this class of polynomials is easy to construct and to understand, it gives a good idea, why many questions about amoebas are hard. I recommend the reader to keep this class in mind.

The Passare–Rullgård Polynomials can be generalized strongly. Theobald and I showed that the equivalence described in Theorem 3.4 holds for a more general class: polynomials supported on circuits such that the circuits satisfy some additional barycentric condition; see [46, Theorem 6.1]. A support set A is called a *circuit* if A is an affine dependent set, but all proper subset of A are affinely independent. For the purpose of this survey, however, it will be sufficient to restrict ourselves to the Passare–Rullgård Polynomials. In the generalized version of Theorem 3.4, the set in (3) is a region in the complex plane containing the origin, which is bounded by a rotated *hypocycloid*. A hypocycloid is a particular plane algebraic curve belonging to the family of *roulette curves*; see, e.g., [4]. For further details see [6, 46]. For the special case of Theorem 3.4 and $n = 2$ this hypocycloid coincides with the *Steiner Curve* $(x_1^2 + x_2^2)^2 - 8x_1(3x_2^2 - x_1^2) + 18(x_1^2 + x_2^2) - 27$, see [33, 35] and Figure 5.

We provide a proof for Theorem 3.3.

Proof of Theorem 3.3. Consider the Laurent polynomial $f(z_1, z_2) := 1 + z_1^3 + z_2^3 + 2z_1z_2$. The corresponding amoeba $\mathcal{A}(f)$ has a bounded component by Theorem 3.4. The corresponding non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ is the tropical hypersurface given by the tropical polynomial $h(x_1, x_2) := -\operatorname{val}(1 \cdot t^0) \oplus -\operatorname{val}(1 \cdot t^0)x_1^3 \oplus -\operatorname{val}(1 \cdot t^0)x_2^3 \oplus -\operatorname{val}(2 \cdot t^0)x_1x_2$. Since the valuation of every complex number interpreted as a Puiseux series equals zero by (3.1) we have $h(x_1, x_2) = 0 \oplus 0 \odot x_1^3 \oplus 0 \odot x_2^3 \oplus 0 \odot x_1x_2$. The corresponding tropical hypersurface $\mathcal{T}(h) = \mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ is a genus zero tropical curve with a single vertex at the origin. Thus, $\mathcal{A}(f)$ and $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ are not homotopy equivalent in general and Theorem 3.3 follows. See also Figure 6. $\square$

We remark that the non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ of a Laurent polynomial $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ embedded into $\mathbb{K}[\mathbf{z}^{\pm 1}]$ does not depend on the coefficients of f. Thus, an alternative proof for Theorem 3.3 is given by the fact that the image

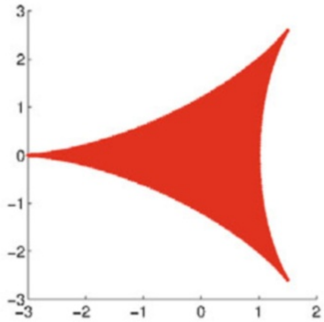


FIGURE 5. The region in the complex plane which is bounded by the Steiner Curve. By Theorem 3.4 the amoeba of $f(\mathbf{z}) := 1 + \sum_{j=1}^{n} z_j^{n+1} + a \cdot z_1 \cdots z_n$ has a bounded component whenever the coefficient $a \in \mathbb{C}$ is not contained in this region.

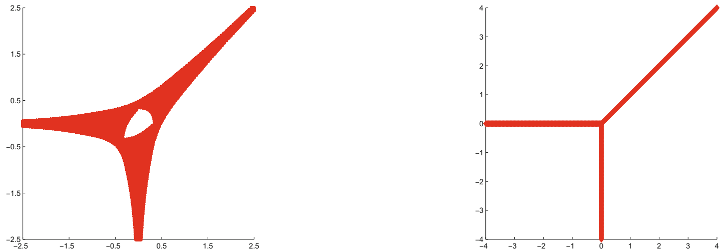


FIGURE 6. Let $f(z_1, z_2) := 1 + z_1^3 + z_2^3 + 2z_1z_2$. The amoeba $\mathcal{A}(f)$ has a bounded component as stated by Theorem 3.4 (left picture), but the non-Archimedean amoeba has not (right picture).

of the order map depends non-trivially on the coefficients of f. The latter is well known; see, e.g., [42].

4. Maslov dequantization

In this section we consider a family of semi-rings together with semi-ring isomorphisms and a particular limit process known as *Maslov dequantization*. This allows us to provide a geometric explanation how amoebas and non-Archimedean amoebas are related to each other.

The amount of background literature specifically about Maslov dequantization is limited. Unfortunately, Maslov dequantization is not explained in full detail in many surveys about tropical geometry. From the tropical point of view, valuations are often a sufficient tool to tackle the particular problems of interest. Hence, one should additionally consult literature about Viro's patchworking, which is closely related to Maslov dequantization. I refer the reader particularly to Mikhalkin's survey about amoebas [24], his article [25], the survey [5] by Brugallé and Shaw about tropical geometry, Viro's articles/surveys [48, 50, 52], and the books [17, 27]. Some details of the terminology differ slightly from source to source.

Let $(R_0, \oplus_0, \odot_0)$ be a semi-ring. A *quantization* of $(R_0, \oplus_0, \odot_0)$ is a family of semi-rings $(R_s, \oplus_s, \odot_s)$, $s \geq 0$ such that two rings R_s and $R_{s'}$ are isomorphic for every $s, s' > 0$ but no R_s with $s > 0$ is isomorphic to R_0. The operations $\oplus_s$ and $\odot_s$ hereby depend continuously on s. The ring R_0 is referred to as the *classical* object, while the rings R_s with $s > 0$ are *quantum* objects. One calls R_s with $s > 0$ a *quantized version* of R_0. Similarly, for a given family R_s, $s \geq 0$ of such semi-rings, the process of starting at a fixed $s > 0$ and taking a limit $\lim_{s \to 0} R_s$ is called a *dequantization* of R_s, [24, 25, 48, 50]. Indeed, these terms are motivated by quantum mechanics, see [50], particularly the "Litvinov–Maslov Correspondence Principle".

The term *Maslov dequantization* refers to a dequantization in a particular family of semi-rings. In our modern terminology, Maslov observed [22] that the standard semi-ring $(\mathbb{R}_{>0}, +, \cdot)$ is a quantized version of the tropical $(\max, +)$ semi-ring in the following way. For all parameters $s \geq 0$ let $t := e^{1/s} \in (1, \infty]$ and let $(R_t)_{t \in (1,\infty]}$ denote a family of semi-rings such that each $R_t := (\mathbb{R}, \oplus_t, \odot_t)$ satisfies for every $x, y \in \mathbb{R}$

$$\begin{aligned} x \oplus_t y \ &:= \ \begin{cases} \log_t(t^x + t^y) & \text{for } 1 < t < \infty \\ \max\{x, y\} & \text{for } t = \infty, \text{ and} \end{cases} \\ x \odot_t y \ &:= \ \log_t(t^{x+y}) \ = \ x + y \end{aligned}$$

Note that for every $t \in (1, \infty)$ there exists a semi-ring isomorphism D_t from the semi-ring $\mathbb{R}_+ := (\mathbb{R}_{>0}, +, \cdot)$ to $R_t = (\mathbb{R}, \oplus_t, \odot_t)$ given by $D_t : \mathbb{R}_+ \to R_t, x \mapsto \log_t |x|$ satisfying for every $x, y \in \mathbb{R}_{>0}$

$$D_t(x + y) \ = \ D_t(x) \oplus_t D_t(y) \ \text{ and } \ D_t(x \cdot y) \ = \ D_t(x) \odot_t D_t(y).$$

Thus, all R_t with $t \in (1, \infty)$ are isomorphic, but there is exists no isomorphism between some R_t with $t \in (1, \infty)$ and R_∞. Namely, the maximum is idempotent, while every $\oplus_t$ for $t \in (1, \infty)$ is not, [25].

Let $\operatorname{Log}_t |\cdot| : (\mathbb{C}^*)^n \to \mathbb{R}, \ \mathbf{z} \mapsto (\log_t |z_1|, \ldots, \log_t |z_n|) = (x_1, \ldots, x_n)$ and let $A \subset \mathbb{Z}^n$ be a support set. Maslov and Viro [22, 48], see also [25, 24], showed that for every $t \in (1, \infty)$ and every polynomial

$$g_t(\mathbf{x}) \quad := \quad \bigoplus_{\alpha \in A} {}_t \ b_\alpha \odot_t \langle \alpha, \mathbf{x} \rangle,$$

mapping from $(R_t)^n$ to R_t, the function $f_t = \log_t^{-1} \circ g_t \circ \operatorname{Log}_t |\cdot|$ is a classical polynomial with standard operators $+$ and $\cdot$ mapping from $(\mathbb{C}^*)^n$ to $\mathbb{C}$. Namely, we have

$$\log_t^{-1}(g_t) = \log_t^{-1}\left(\bigoplus_{\alpha\in A}{}_t\; b_\alpha \odot_t \langle\alpha,\mathbf{x}\rangle\right) \;=\; \log_t^{-1}\left(\log_t\left(\sum_{\alpha\in A} t^{b_\alpha+\langle\alpha,\mathbf{x}\rangle}\right)\right)$$
$$= \sum_{\alpha\in A} t^{b_\alpha+\langle\alpha,\mathrm{Log}_t|\mathbf{z}|\rangle} \;=\; \sum_{\alpha\in A} t^{b_\alpha}\cdot\mathbf{z}^\alpha \;=:\; f_t.$$

Thus, for every $t \in (1,\infty)$ the polynomial f_t has a variety $\mathcal{V}(f_t) \subset (\mathbb{C}^*)^n$ and a corresponding amoeba $\mathcal{A}_t(f_t) \subset \mathbb{R}^n$ given by $\mathrm{Log}_t|\mathcal{V}(f_t)|$. On the one hand, Maslov dequantization yields a tropical polynomial $\lim_{t\to\infty} f_t$ with a tropical hypersurface as a limit out of a classical polynomial f_t with classical hypersurface. Hence, we might expect that taking the amoeba and deforming it via changing the log-basis yields a tropical hypersurface in its limit $\lim_{t\to\infty}\mathcal{A}(f_t)$. On the other hand, we can interpret an entire family $f_t(\mathbf{z}) = \sum_{\alpha\in A} b_\alpha t^{c_\alpha}\mathbf{z}^\alpha$ with $b_\alpha \in \mathbb{C}, c_\alpha \in \mathbb{R}\cup\{-\infty\}$, $t \in (1,\infty)$ of polynomials as a *single* polynomial with coefficients in the *field of Puiseux series with real exponents* $\mathbb{K}_\mathbb{R}[\mathbf{z}^{\pm 1}]$. This field $\mathbb{K}_\mathbb{R}$ is literally defined as the usual Puiseux series $\mathbb{K}$ except that the exponents in the power series can also be non-rational real numbers. The maps val and $\mathrm{Log}_{\mathbb{K}_\mathbb{R}}|\cdot|$ work analogously as over $\mathbb{K}$; see, e.g., [24]. Note that such polynomials f_t are referred to as *patchwork polynomials* due to their role in Viro's patchworking; see, e.g., [25, 48, 52] Let $\mathcal{V}_{\mathbb{K}_\mathbb{R}}(f_t)$ denote the corresponding hypersurface in $(\mathbb{K}_\mathbb{R}^*)^n$ and let $\mathcal{A}_{\mathbb{K}_\mathbb{R}}(f_t)$ denote the non-Archimedean amoeba given by $\mathrm{Log}_{\mathbb{K}_\mathbb{R}}|\mathcal{V}_{\mathbb{K}_\mathbb{R}}(f_t)|$. It turns out that $\lim_{t\to\infty}\mathcal{A}(f_t)$ coincides with the tropical hypersurfaces $\mathrm{Log}_{\mathbb{K}_\mathbb{R}}|\mathcal{V}_{\mathbb{K}_\mathbb{R}}(f_t)|$.

More precisely, let A, B be closed subsets of $\mathbb{R}^n$ and $d(\cdot,\cdot)$ denote the Euclidean metric on $\mathbb{R}^n$. For $\mathbf{x} \in \mathbb{R}^n$ let $d(\mathbf{x}, A) := \inf_{a\in A} d(\mathbf{x}, a)$. The *Hausdorff distance* of two closed subsets A, B of $\mathbb{R}^n$ is defined as

$$\mathrm{Hausd}(A,B) \quad := \quad \max\{\sup_{a\in A} d(a,B), \sup_{b\in B} d(A,b)\}.$$

Note that if one of the sets A, B is unbounded, then the Hausdorff distance is possibly infinite. Let $\mathcal{A}_t(f_t)$ and $\mathcal{A}_{\mathbb{K}_\mathbb{R}}(f_t)$ as before. Then the following theorem holds; see [25, Corollary 6.4], [42, Theorem 9]; see also [24].

Theorem 4.1 (Mikhalkin, Rullgård). *Let $f_t(\mathbf{z}) = \sum_{\alpha\in A} b_\alpha t^{c_\alpha}\mathbf{z}^\alpha$ with $b_\alpha \in \mathbb{C}^*, c_\alpha \in \mathbb{R}, t \in (1,\infty)$ be a family of polynomials in $\mathbb{C}[\mathbf{z}^{\pm 1}]$ and simultaneously a single polynomial in $\mathbb{K}_\mathbb{R}[\mathbf{z}^{\pm 1}]$. Then for $t \to \infty$ the amoebas $\mathcal{A}_t(f_t)$ converge to the non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}_\mathbb{R}}(f_t)$ in Hausdorff distance.*

In Section 3 we pointed out that we can interpret every Laurent polynomial $f(\mathbf{z}) = \sum_{\alpha\in A} b_\alpha\mathbf{z}^\alpha$ as a patchwork polynomial $f_t(\mathbf{z}) = \sum_{\alpha\in A} b_\alpha t^0\mathbf{z}^\alpha$. Since 0 is the only exponent of the parameter t, we do not have to distinguish between $\mathbb{K}$ and $\mathbb{K}_\mathbb{R}$. Since we obtain $\mathcal{A}_e(f_e) = \mathcal{A}(f)$ for $t = e$, Theorem 4.1 yields the geometric interpretation for the relation between $\mathcal{A}(f)$ and $\mathcal{A}_\mathbb{K}(f_\mathbb{K}) = \mathrm{Log}_\mathbb{K}|\mathcal{V}_\mathbb{K}(\Psi(f))|$ we were looking for in Section 3. Indeed, we have $\mathcal{V}_t(f_t) = \mathcal{V}(f)$ for all $t \in (1,\infty)$, but the corresponding amoebas $\mathcal{A}_t(f_t) = \mathrm{Log}_t|\mathcal{V}(f)|$ converge to $\mathcal{A}_{\mathbb{K}_\mathbb{R}}(f_t) = \mathcal{A}_\mathbb{K}(\Psi(f)) =$

$\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$. Theorem 4.1 also allows us to modify our diagram (3.2) in the following way:

$$\begin{tikzcd} f \arrow[r, hook, "\Psi(f)"] \arrow[d, "\mathrm{Log}\,|\mathcal{V}(f)|"'] & f_{\mathbb{K}} \arrow[d, "\mathrm{Log}_{\mathbb{K}}\,|\mathcal{V}_{\mathbb{K}}(f)|"] \\ \mathcal{A}(f) \arrow[r, "\text{converg. via Log}", "\text{basis change}"'] & \mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}}) \end{tikzcd} \tag{4.1}$$

Example 4.2. Let $f(z_1, z_2) := z_1 + z_2 - 1$. Its non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ is given by the tropical hypersurface of $0 \oplus 0 \odot x_1 \oplus 0 \odot x_2$ due to Kapranov's Theorem 3.1; see also Example 3.2. If we (formally) consider the family $f_t(z_1, z_2) := t^0 z_1 + t^0 z_2 - 1t^0$, then Theorem 4.1 yields that $\mathrm{Log}_t\,|\mathcal{V}(f_t)| = \mathrm{Log}_t\,|\mathcal{V}(f)|$ converges to $\mathcal{A}_{\mathbb{K}_{\mathbb{R}}}(f_t) = \mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ for $t \to \infty$.

It is a well-known result by Forsberg, Passare, and Tsikh [11] that for this particular f (and similarly for other linear polynomials) the boundary of $\mathcal{A}(f)$ is given by the $\mathrm{Log}\,|\cdot|$ image of the real locus of $\mathcal{V}(f)$, i.e., the zeros of the line $l(x_1, x_2) := x_1 + x_2 - 1$ for $x_1, x_2 \in \mathbb{R}^2$. We visualize the convergence of $\mathcal{A}(f)$ to $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ by drawing the $\mathrm{Log}_t\,|\cdot|$ image of the real zero set $\mathcal{V}_{\mathbb{R}}(l)$, see Figure 7.

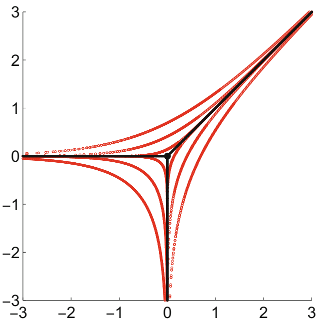


FIGURE 7. Maslov dequantization: In red (light) we see the $\mathrm{Log}_t\,|\cdot|$ image of the real zeros of $l(x_1, x_2) := x_1 + x_2 - 1$, $x_1, x_2 \in \mathbb{R}^2$ for $t := e, t := 2e$ and $t := 50e$. The non-Archimedean amoeba of $z_1 + z_2 - 1$ is the black (dark) tropical curve.

5. The Archimedean tropical hypersurface and the complement induced tropical hypersurface

In the two previous Sections 3 and 4, we have seen that Maslov dequantizations, valuations, and as a consequence the non-Archimedean amoeba are powerful tools. For questions regarding amoebas themselves, however, they are often not suitable, since we only have convergence in Hausdorff distance, and, in particular, since the resulting tropical objects are not homotopy equivalent to the original amoeba in general; see Theorem 3.3, Figure 6. Hence, we would like to find another tropical hypersurface, which is more sensitive about the amoeba structure. In this section we discuss two such tropical hypersurfaces: The *Archimedean tropical hypersurface* and the *complement induced tropical hypersurface*.

Let $A \subset \mathbb{Z}^n$ finite and let $f(\mathbf{z}) := \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha$ with $b_\alpha \in \mathbb{C}^*$. A canonical guess for constructing an improved tropicalization of the amoeba $\mathcal{A}(f)$ is to consider $\log|b_\alpha|$ as coefficients for the tropicalization instead of the all 0's given by $-\operatorname{val}(b_\alpha)$. Following this idea we define the *Archimedean tropical polynomial*

$$\operatorname{ArchTrop}(f) \quad := \quad \bigoplus_{\alpha \in A} \log|b_\alpha| \oplus \mathbf{x}^\alpha,$$

and the corresponding *Archimedean tropical hypersurface*

$$\mathcal{AR}(f) \quad := \quad \mathcal{T}(\operatorname{ArchTrop}(f)).$$

The tropical polynomial $\operatorname{ArchTrop}(f)$ was already mentioned by Passare and Tsikh in their survey [35], where it is just called *tropicalization* of f, and also already in Mikhalkin's early articles about tropical geometry, for example [26]. In a recent work by Avendaño, Kogan, Nisse, and Rojas [1], the tropical hypersurface $\mathcal{AR}(f)$ was called *Archimedean tropical variety*. The motivation for the name is that this tropical hypersurface can be seen as an Archimedean counterpart to the non-Archimedean amoeba, since the complex absolute value, as a norm, is Archimedean – in contrast to the non-Archimedean norm $|\cdot|_\nu$ on the field of Puiseux series. I keep the name Archimedean tropical polynomial here (note that in [1] $\operatorname{ArchTrop}(f)$ denotes the hypersurface, not the polynomial). In the same article [1, Theorem 1.11], the authors show the following theorem. In what follows, we define for every $X \subseteq \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}_{>0}$

$$X_\varepsilon \quad := \quad \{\mathbf{x} \in \mathbb{R}^n \ : \ |\mathbf{x} - \mathbf{y}| < \varepsilon \text{ for some } \mathbf{y} \in X\}.$$

Theorem 5.1 (Avendaño, Kogan, Nisse, Rojas, [1]). *Let $A \subset \mathbb{Z}^n$ with $\#A = k \in \mathbb{N}^*$. Let $f \in (\mathbb{C}^*)^A \subset \mathbb{C}[\mathbf{z}^{\pm 1}]$. Then*

1. $\mathcal{A}(f) \subseteq \mathcal{AR}(f)_{\log(k-1)}$.
2. $\mathcal{AR}(f) \subseteq \mathcal{A}(f)_\varepsilon$ *with* $\varepsilon := \sqrt{n}\lceil \frac{1}{4}k(k-1)\rceil(\log(9)k - \log(\frac{81}{2}))$ *for* $n \geq 3$

Note that there are better bounds for the ε in Part (2) of the theorem for the cases $n = 1$ and $n = 2$. Moreover, the authors provide generalizations for the case that the Newton polytope of the defining polynomial is not full-dimensional. See [1, Theorem 1.11] for further details.

Theorem 5.1 guarantees that $\mathcal{AR}(f)$ always yields at least a rough approximation of an amoeba, which is easy to compute. This is useful especially for practical problems, since it allows for example to certify easily that certain points are not contained in the amoeba. A weaker version of the bound of Part (1) in Theorem 5.1 for dimension two was given by Mikhalkin already in 2005 in [26].

An obvious follow-up question to Theorem 5.1 is whether $\mathcal{AR}(f)$ is homotopy equivalent or even a deformation retract of $\mathcal{A}(f)$. Unfortunately, this is not the case in general. Once again, we can use the Passare–Rullgård Polynomials given in Theorem 3.4. Note in this context that an amoeba is called *solid* if it has the minimal possible number of components of the complement. In other words, the order of every component of the complement corresponds to a vertex in the corresponding Newton polytope.

Example 5.2. Let $f(z_1, z_2) := 1+z_1^3+z_2^3-2z_1z_2$. By Theorem 3.4 $\mathcal{A}(f)$ is solid, but $\mathcal{AR}(f)$, the tropical hypersurface of $\mathrm{ArchTrop}(f) = 0 \oplus 0 \odot x_1^3 \oplus 0 \odot x_2^3 \oplus \log|2| \odot x_1x_2 = \max\{0, 3\cdot x_1, 3\cdot x_2, \log|2| + x_1 + x_2\}$, has a bounded component. This is easy to see since for example at the origin $\log|2| \odot x_1x_2$ is the only dominating term. See also Figure 8.

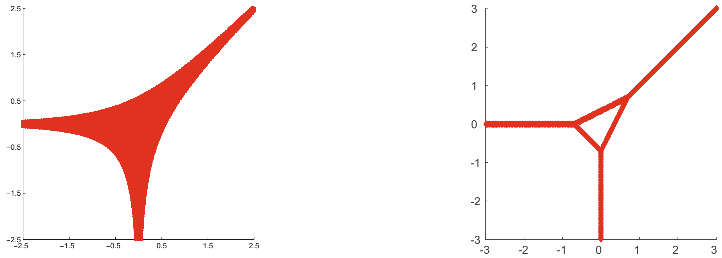


FIGURE 8. Let $f(z_1, z_2) := 1 + z_1^3 + z_2^3 - 2z_1z_2$. The left figure shows the amoeba $\mathcal{A}(f)$ and the right one the tropical hypersurface $\mathcal{AR}(f)$. Obviously, $\mathcal{A}(f)$ and $\mathcal{AR}(f)$ are not homotopy equivalent.

Another counterexample can for instance be found in [1, Example 1.6.]. In general, it is an open problem to give an exact characterization of the cases, when $\mathcal{AR}(f)$ and $\mathcal{A}(f)$ are homotopy equivalent.

Problem 5.3. *Determine for which support sets $A \subset \mathbb{Z}^n$ it holds for every $f \in (\mathbb{C}^*)^A$ that $\mathcal{A}(f)$ is homotopy equivalent to $\mathcal{AR}(f)$.*

Already more than 10 years ago people were aware that $\mathrm{ArchTrop}(f)$ is not sufficient to describe the topology of amoebas. Hence, Passare and Rullgård tried an entirely new approach, namely to *change the support* of the tropical polynomial. This was a crucial idea, which finally lead to the *spine*, which we discuss in Section 6. However, it is also a hard way to go. Namely, all tropical representations of amoebas, which arise along this way, *cannot be computed straightforwardly from f* in general, as we see in this and the following section. See also Remark 7.2 at the end of Section 7.

Recall from (2.3) that $\mathrm{Comp}(f)$ contains all lattice points in $\mathrm{conv}(A)$ which are associated to existing components of the complement of $\mathcal{A}(f)$ via the order map. We define the *complement induced tropicalization* as

$$\mathrm{CompTrop}(f) \quad := \quad \bigoplus_{\alpha \in \mathrm{Comp}(f)} \log|b_\alpha| \oplus \mathbf{x}^\alpha, \tag{5.1}$$

and the corresponding *complement induced tropical hypersurface* as

$$\mathcal{C}(f) \quad := \quad \mathcal{T}(\mathrm{CompTrop}(f)).$$

Note that for $\alpha \in \mathrm{Comp}(A) \setminus A$ we set $b_\alpha = 0$ and $\log|0| = -\infty$, which makes sense in the tropical world.

To the best of my knowledge these terms were first used by Theobald and myself in [46]. However, the objects were already investigated earlier, at least by Rullgård in his thesis [42]. In particular, he proved the following statement, see [42, Theorem 12, p. 36].

Theorem 5.4 (Rullgård, [42]). *Let $A \subset \mathbb{Z}^n$ with $\#A \leq 2n$ such that for all $1 \leq k \leq \#A$ it holds: No $k+2$ elemental subset of A lies in a k-dimensional affine subspace of $\mathbb{R}^n$. Then for all $f \in (\mathbb{C}^*)^A$ it holds: $\mathcal{C}(f)$ is a deformation retract of $\mathcal{A}(f)$.*

Note that for $n \geq 2$ these conditions are particularly satisfied if A is a circuit. Polynomials supported on circuits and their amoebas have especially nice properties, see [14, 16, 46, 47]. Another case when the relation between $\mathcal{A}(f)$ and $\mathcal{C}(f)$ is particularly easy is the following one. It is a consequence from Corollary 6.4, which we introduce in Section 6.

Theorem 5.5. *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ such that $\mathcal{A}(f)$ is solid, i.e., $E_\alpha(f) \neq \emptyset$ if and only if α is a vertex of $\mathrm{New}(f)$. Then $\mathcal{C}(f)$ is a deformation retract of $\mathcal{A}(f)$. Moreover, $\mathcal{C}(f)$ is a tree.*

Motivated by these statements we formulate the following question:

Question 5.6. *Is $\mathcal{C}(f)$ always a deformation retract of $\mathcal{A}(f)$?*

Later, we show that, unfortunately, the answer to this problem is "no". More specifically: It is not surprising that the answer is "no" in the univariate case and it is not very hard to construct counterexamples. However, the answer is also "no" in the multivariate case; see Corollary 5.14.

Question 5.6 has a partially unclear history. Forsberg mentioned a relation between amoebas and what is today known to be a tropical hypersurfaces already in his thesis [10] in 1998. Passare and Rullgård developed the *spine* of an amoeba, which *is* a deformation retract of $\mathcal{A}(f)$, already in 2004 in [33]; see also Section 6, particularly Theorem 6.2. A few years ago it seemed to me that it was folklore that $\mathcal{C}(f)$ is *not* a deformation retract of $\mathcal{A}(f)$, but (seemingly) nobody had a explicit counterexample (at least for the multivariate case). Although I would expect that Passare knew a counterexample, he did, to the best of my knowledge, not write about whether $\mathcal{C}(f)$ is a deformation retract of $\mathcal{A}(f)$. Similarly, according to the remark after [42, Theorem 8, Page 33] and [42, Theorem 12, Page 36], I am not entirely sure whether Rullgård knew the answer to Question 5.6 in 2003/04. If he, however, did not, then this would be very surprising to me since he already did almost the entire work to negate Question 5.6 as we show in the following pages. Hence, I suggest that he knew the answer already in 2003/04 and hence I believe that Passare did as well, since he was Rullgård's advisor. However, to the best of my knowledge, no explicit counterexample for the multivariate case was given elsewhere before.

In general, it seems to me that many people did not pay very much attention to Question 5.6 for a long time. Probably, since it was convenient to refer to the spine instead. In my opinion, $\mathcal{C}(f)$ deserves attention since its construction is simpler than the spine in the sense that it does not need the Ronkin function, and, as a consequence, its coefficients are easier to compute; see also Section 6.

Before we answer Question 5.6, we investigate two other questions about $\mathrm{Comp}(f)$, which "are en route", namely:

Question 5.7. *Let $A \subset \mathbb{Z}^n$ be a finite set.*

1. *Does for every $\alpha \in A$ exist some $f \in (\mathbb{C}^*)^A$ with $\alpha \in \mathrm{Comp}(f)$? In other words, is $U_\alpha^A \neq \emptyset$ for every $\alpha \in A$?*
2. *Is for every $f \in (\mathbb{C}^*)^A$ the set $\mathrm{Comp}(f)$ always a subset of A?*

This first part of Question 5.7 is well-known to have an affirmative answer. It is a consequence of the following theorem initially proven by Forsberg, Passare and Tsikh; see [11, Prop. 2.7]

Theorem 5.8 (Forsberg, Passare, Tsikh). *Let $A \subset \mathbb{Z}^n$ and $f(\mathbf{z}) = \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ with $b_\alpha \in \mathbb{C}^*$ and $\mathcal{V}(f) \subset (\mathbb{C}^*)^n$. Assume that there exists an $\alpha' \in A$ and a $\mathbf{w} \in \mathcal{A}(f)^c$ such that for all $\mathbf{z} \in (\mathbb{C}^*)^n$ with $\mathrm{Log}\,|\mathbf{z}| = \mathbf{w}$ it holds that $|b_{\alpha'} \mathbf{z}^{\alpha'}| > |\sum_{\alpha \in A \setminus \{\alpha'\}} b_\alpha \mathbf{z}^\alpha|$. Then $\mathrm{ord}(\mathbf{w}) = \alpha'$, i.e., $\mathbf{w} \in E_{\alpha'}(f)$.*

The condition of this theorem is in particular always satisfied if for some $\mathbf{z} \in (\mathbb{C}^*)^n$ it holds that

$$|b_{\alpha'} \mathbf{z}^{\alpha'}| > \sum_{\alpha \in A \setminus \{\alpha'\}} |b_\alpha \mathbf{z}^\alpha|. \tag{5.2}$$

This is a consequence of the triangle inequality. Condition (5.2) was excessively used by Purbhoo in [38] to approximate amoebas. Following Purbhoo we say f is *lopsided* at $|\mathbf{z}|$ if (5.2) is satisfied. Obviously, (5.2) will always hold for $|b_{\alpha'}|$ sufficiently large while keeping the other coefficients constant. This shows that $U_{\alpha'}^A \neq \emptyset$ for $\alpha' \in A$ and hence provides a positive answer to the first part of Question 5.7. See [38] for further details. Note that Rullgård proved an even stronger version of Theorem 5.8, see [41, 42]. Note furthermore that the univariate version of condition (5.2) and the corresponding statement is a classical result by Pellet [36]. Hence, some authors refer to lopsidedness also as *generalized Pellet condition.*

The second part of Question 5.7 is trickier. The question is: Can there exist a component in the complement of $\mathcal{A}(f)$, which is mapped to a lattice point in $\mathrm{conv}(A)$, which is not in the support A itself? Rullgård showed a very strong statement regarding this question in his thesis. Unfortunately, to my believe, the statement is almost unknown so far, since, to the best of my knowledge, it was not published elsewhere later. For a finite set $A \subset \mathbb{Z}^n$ we denote the lattice generated by A by $\mathcal{L}_A$, i.e.,

$$\mathcal{L}_A \quad := \quad \left\{ \sum_{\alpha \in A} \lambda_\alpha \cdot \alpha \ : \ \lambda_\alpha \in \mathbb{Z} \right\} \subseteq \mathbb{Z}^n.$$

Then the following theorem holds; see see [42, Theorem 11].

Theorem 5.9 (Rullgård, [42]). *Let $A \subset \mathbb{Z}^n$ be a support set and l be a line in $\mathbb{R}^n$.*

1. *If $\alpha \in \mathbb{Z}^n$ and $U_\alpha^A \neq \emptyset$, then $\alpha \in \mathrm{conv}(A) \cap \mathcal{L}_A$.*
2. *If $\alpha \in \mathrm{conv}(A \cap l) \cap \mathcal{L}_{A \cap l}$, then $U_\alpha^A \neq \emptyset$.*

The first part of the theorem shows that the upper bound from Corollary 2.5 for the number of components, which the complement of an amoeba can have, can be improved. In particular, if the support set A is sparse, then, roughly speaking, the number of components "remains low" even if the Newton polytope given by $\mathrm{conv}(A)$ is "very large" since the lattice $\mathcal{L}_A$ will be sparse, too. The statement is a consequence of amoebas and their complement under lattice transformations, see [42, Theorem 7, Page 31]. Rullgård speaks of *functorial properties of amoebas* in this context.

The second part shows in particular that the answer to Part (2) of Question 5.7 is "no". In general, there exist components in the complement of the amoeba, which have an image under the order map outside of the support set A. We write down this fact as a statement on its own.

Corollary 5.10. *Let $A \subset \mathbb{Z}^n$ and $f \in (\mathbb{C}^*)^A$. Then the image of the order map* $\mathrm{Comp}(f)$ *of the components of $\mathcal{A}(f)^c$ is contained in $\mathcal{L}_A \cap \mathbb{Z}^n$, but* $\mathrm{Comp}(f)$ *is not contained in A in general.*

In simple words, this observation *makes the world of amoebas a lot more complicated!* In particular, we have the following problem remaining open, which is a part of the general Problem 2.6 mentioned in the introduction.

Problem 5.11. *Let $A \subset \mathbb{Z}^n$ finite and $f \in (\mathbb{C}^*)^A$. Determine* $\mathrm{Comp}(f)$*.*

In what follows we construct a class of multivariate Laurent polynomials which has an amoeba with components of the complement of order α, although α is not contained in the support set $A \subset \mathbb{Z}^n$ of the corresponding Laurent polynomial. The construction is adapted from Rullgård's original proof of Theorem 5.9, see [42, Theorem 11].

Proposition 5.12. *Let $l = \{(z_1, 0, \ldots, 0) : z_1 \in \mathbb{Z}\}$ and $A \subset \mathbb{Z}^n$ be a support set with*

$$A \cap l = \{(0, \ldots, 0), (\alpha(1), 0, \ldots, 0), (\alpha(1) + \alpha(2), 0, \ldots, 0)\},$$

where $\alpha(1), \alpha(2) \in \mathbb{N}$ *are coprime. Consider the polynomials*

$$f(\mathbf{z}) := z_1^{\alpha(1)+\alpha(2)} + e^{i\cdot\pi\cdot\rho} \cdot z_1^{\alpha(2)} + 1$$
$$g(\mathbf{z}) := f(\mathbf{z}) + \sum_{\alpha \in A \setminus l} \varepsilon\, \mathbf{z}^\alpha \in (\mathbb{C}^*)^A$$

with $\varepsilon > 0$ *and* $\rho\cdot(\alpha(1)+\alpha(2)) \notin \mathbb{Z}$. *Then* $E_\alpha(g) \neq \emptyset$ *for all* $\alpha \in \operatorname{conv}(A\cap l)\cap\mathcal{L}_{A\cap l}$ *if* ε *is sufficiently small.*

Recall from Section 2.1 that $\mathbb{F}_{\mathbf{w}} = \operatorname{Log}^{-1}|\mathbf{w}|$ denotes the fiber of a point $\mathbf{w} \in \mathbb{R}^n$ with respect to the $\operatorname{Log}|\cdot|$ map and that $\mathbb{F}_{\mathbf{w}}$ is homeomorphic to a torus $(S^1)^n$. Note that the proposition remains true if one multiplies g with an arbitrary monomial $\mathbf{z}^\alpha$ with $\alpha \in \mathbb{Z}^n$.

Proof. Let $\alpha = (\alpha_1, 0, \ldots, 0) \in (\operatorname{conv}(A) \cap l) \setminus A$. $f(\mathbf{z})$ is a trinomial in z_1, which can be interpreted as a polynomial in $\mathbf{z}$, which does not depend on $z_2, \ldots, z_n$. It is a consequence of [47, Theorem 4.4] that $E_\alpha(f) \neq \emptyset$ for all $\alpha \in \operatorname{conv}(A) \cap l$ if $\rho \cdot (\alpha(1) + \alpha(2)) \notin \mathbb{Z}$.

Let $\mathbf{w} \in E_\alpha(f)$. We have to show that $E_\alpha(g) \neq \emptyset$ to complete the proof. We define $\delta := \min_{\mathbf{z}\in\mathbb{F}_{\mathbf{w}}} |f(\mathbf{z})| > 0$ as the minimal value attained by f on the fiber $\mathbb{F}_{\mathbf{w}} = \operatorname{Log}^{-1}|\mathbf{w}|$ over $\mathbf{w}$. Note that this minimum δ is greater than 0 since $\mathbf{w} \notin \mathcal{A}(f)$ and the minimum exists since the fiber $\mathbb{F}_{\mathbf{w}}$ is compact. We define $\kappa := \max_{\alpha \in A\setminus(A\cap l)} \max_{\mathbf{z}\in\mathbb{F}_{\mathbf{w}}} |\mathbf{z}^\alpha|$ as the maximal absolute value that a monomial with exponent in A but not belonging to f attains at the fiber $\mathbb{F}_{\mathbf{w}}$. Let furthermore $d := \#A - 3$.

We choose $\varepsilon := \delta/(2d\kappa)$. If we evaluate g at an arbitrary point $\mathbf{v} \in \mathbb{F}_{\mathbf{w}}$, then we obtain

$$|g(\mathbf{v})| \geq |f(\mathbf{v})| - \left|\sum_{\alpha\in A\setminus l} \varepsilon\, \mathbf{v}^\alpha\right| \geq |f(\mathbf{v})| - \sum_{\alpha\in A\setminus l} \varepsilon\, |\mathbf{v}^\alpha| \geq \delta - d \cdot \frac{\delta}{2d\kappa} \cdot \kappa = \frac{\delta}{2} > 0.$$

Thus, $\mathbf{w} \notin \mathcal{A}(g)$.

We compute the order of $\mathbf{w}$ with respect to g. Let, for each $j \in \{1, \ldots, n\}$, f_j and g_j be the univariate polynomials in z_j obtained from f and g by setting $z_i = v_i$ for all $i \neq j$. Since $|f_1(z_1)| = |f(z_1, v_2, \ldots, v_n)| \geq \delta$ for all z_1 with $|z_1| = |v_1|$ and $|g_1(z_1)|$ differs from $|f_1(z_1)|$ by at most $\delta/2$, we can conclude $w_1 \in E_{\alpha_1}(g_1)$, since we know $w_1 \in E_{\alpha_1}(f_1)$ and polynomials are continuous in their coefficients.

For every $j \neq 1$ we know that $f_j(z_j)$ equals the constant given by $f_1(v_1)$. Thus, $|f(\mathbf{v})| = \delta$ contributes to the coefficient which is the constant term of g_j and $\sum_{\alpha\in A\setminus l} \varepsilon|\mathbf{v}^\alpha| \leq \delta/2$ for all $v_j \in \mathbb{F}_{w_j}$. Thus, g_j is lopsided, see (5.2), in v_j with the constant term being the dominating term. Hence, $w_j \in E_0(g_j)$ for all $1 < j \leq n$ by Theorem 5.8. Therefore, we have $\operatorname{ord}(\mathbf{w}) = \alpha$ with respect to g and thus $E_\alpha(g) \neq \emptyset$. $\square$

We provide an explicit example.

Example 5.13 (de Wolff, [6]). Let $f(z_1, z_2) := z_2 + e^{i\cdot\pi/5} \cdot z_1 z_2 + z_1^3 z_2 + \varepsilon \cdot (z_1 + z_1^2 + z_1 z_2^2 + z_1^2 z_2^2)$ with $\varepsilon \in \mathbb{R}_{>0}$. If ε is sufficiently small, then $E_{(1,2)}(f) \neq \emptyset$ by Proposition 5.12; see Figure 9.

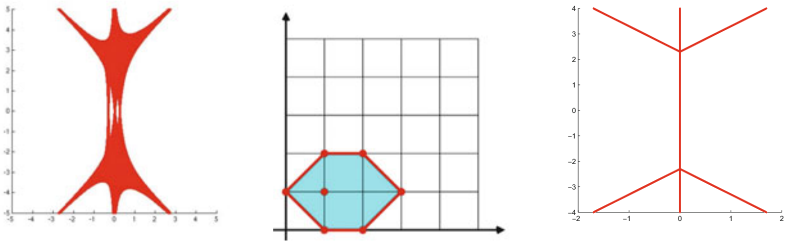


FIGURE 9. Left picture: The amoeba of $f(z_1, z_2) := z_2 + e^{i\cdot\pi/5} \cdot z_1 z_2 + z_1^3 z_2 + \varepsilon \cdot (z_1 + z_1^2 + z_1 z_2^2 + z_1^2 z_2^2)$ for $\varepsilon := 1/10$; middle picture: The Newton polytope of f; right picture: the complement induced tropical hypersurface $\mathcal{C}(f)$.

A consequence of this example or, analogously, of Theorem 5.9 is that the complement induced tropical hypersurface $\mathcal{C}(f)$ of a polynomial $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ is not homotopy equivalent to its amoeba $\mathcal{A}(f)$ in general. This answers Question 5.6. See also Figure 9 and [6, Corollary 4.45].

Corollary 5.14. *For general $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ $\mathcal{A}(f)$ and $\mathcal{C}(f)$ are not homotopy equivalent.*

Proof. We consider the polynomial in Example 5.13, i.e., $f(z_1, z_2) := z_2 + e^{i\cdot\pi/5} \cdot z_1 z_2 + z_1^3 z_2 + \varepsilon \cdot (z_1 + z_1^2 + z_1 z_2^2 + z_1^2 z_2^2)$ with a small ε. Since f does not have a term $b_{(2,1)} z_1^2 z_2$, we can formally add such a monomial with coefficient $b_{(2,1)} = 0$. We know by Example 5.13 that $\mathrm{Comp}(f) = \mathrm{conv}(A) \cap \mathbb{Z}^n$ and hence the complement of $\mathcal{A}(f)$ has eight connected components. We investigate the corresponding tropical polynomial $\mathrm{CompTrop}(f)$ in the tropical semi-ring given by

$$0 \odot x_2 \oplus 0 \odot x_1 x_2 \oplus -\infty \odot x_1^2 x_2 \oplus 0 \odot x_1^3 x_2 \oplus \log|\varepsilon| \odot (x_1 \oplus x_1^2 \oplus x_1 x_2^2 \oplus x_1^2 x_2^2).$$

The number of connected components of the complement of tropical hypersurfaces is limited by the number of terms of their defining polynomial, since every connected component is given by the points in $\mathbb{R}^n$ where a single term uniquely attains the maximum. $\mathrm{CompTrop}(f)$ has eight terms. Since $\log|b_{(2,1)}| = -\infty$, there exists no point where $-\infty \odot x_1^2 x_2$ attains the maximum in $\mathrm{CompTrop}(f)$. Thus, the complement of $\mathcal{C}(f) = \mathcal{T}(\mathrm{CompTrop}(f))$ has at most seven components. Therefore, the complement induced tropical hypersurface $\mathcal{C}(f)$ is not homotopy equivalent to $\mathcal{A}(f)$. The example can be generalized to arbitrary dimensions in the obvious way by symmetrically adding additional terms with positive and negative exponents in new variables which all have ε-coefficients. $\square$

Finally, let us point out that also Rullgård's Theorem 5.9 does not completely characterize which of the sets U_α^A are non-empty for a given $A \subset \mathbb{Z}^n$. Nevertheless, he gave a conjecture (a "rather wild guess" in his words) in his thesis [42, Problem 1, Page 60], which is still open.

Conjecture 5.15 (Rullgård, [42]). *Let $A \subset \mathbb{Z}^n$ finite. Then the sufficient condition for $U_\alpha^A \neq \emptyset$ given in Theorem* 5.9*, Part* (2)*, is also necessary.*

6. The spine

In this section we solve the issues from Section 5, which culminated in Corollary 5.14. For a given Laurent polynomial f, we introduce a tropical polynomial with a tropical hypersurface, which *is* a deformation retract of the corresponding amoeba $\mathcal{A}(f)$. This tropical hypersurface is the *spine* of an amoeba. It was introduced by Passare and Rullgård in 2004 in [33] and it was also part of Rullgård's thesis [42]. Its construction, mainly relying on the *Ronkin function* [40], is a beautiful piece of mathematics, as we will see in what follows.

Let f be a Laurent polynomial with support set $A \subset \mathbb{Z}^n$. The general idea of how to construct a tropical hypersurface associated to f, which is a deformation retract of the amoeba $\mathcal{A}(f)$, is

1. to consider $\mathrm{Comp}(f)$ instead of A as support for the tropical polynomial as we did for the complement induced tropicalization $\mathrm{CompTrop}(f)$, but
2. to replace $\log|b_\alpha|$ as coefficients by new coefficients given by the Ronkin function.

First, we define the *Ronkin function* [40]; see also [33]. Let Ω be a convex open set in $\mathbb{R}^n$ and let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ (we might also generalize f to an arbitrary holomorphic function which is defined in $\mathrm{Log}^{-1}|\Omega|$). The *Ronkin function N_f* is defined by the integral

$$N_f : \Omega \to \mathbb{R}, \quad \mathbf{x} \mapsto \frac{1}{(2\pi i)^n} \int_{\mathrm{Log}^{-1}|\mathbf{x}|} \frac{\log|f(\mathbf{z})| dz_1 \dots dz_n}{z_1 \dots z_n}.$$

Let us investigate this function N_f. Here, $\mathrm{Log}^{-1}|\mathbf{x}|$ denotes the fiber with respect to the $\mathrm{Log}|\cdot|$ map over the point $\mathbf{x} \in \mathbb{R}^n$. In Section 2.1, we have seen that this fiber is homeomorphic to a torus $(S^1)^n$; see particularly Figure 2. This means we can think about the Ronkin function as a map which sends a point $\mathbf{x}$ in the ambient $\mathbb{R}^n$ space of an amoeba $\mathcal{A}(f)$ to a mean $\log|\cdot|$-value of f on the fiber over $\mathbf{x}$; see also [35]. The Ronkin function can alternatively be defined via integrating $\log|\cdot|$ against the Haar measure of the complex torus $(\mathbb{C}^*)^n$; see [42, Page 17/18]. Note that Purbhoo's results [38, Section 4.3] provide a "discrete analogue" of the Ronkin function. More precisely, his iterated resultants provide a Riemann sum approximating the Ronkin function. This approach might be more accessible to readers with a non-analytic background.

The Ronkin function has a couple of useful properties. In particular, the following theorem holds; see [33, Theorem in Section 2]:

Theorem 6.1 (Ronkin [40] / Passare, Rullgård, [33]). *Let f be a holomorphic function as above. Then N_f is a convex function. If $U \subset \Omega$ is a connected, open set, then the restriction of N_f to U is affine linear if and only if U does not intersect the amoeba of f. If $\mathbf{x}$ is in the complement of the amoeba, then grad $N_f(\mathbf{x})$ is equal to the order of the component of the complement of the amoeba $\mathcal{A}(f)$ containing $\mathbf{x}$.*

The idea for the construction of the spine is to use these properties, particularly the affine linearity on each component of the complement, to define a hyperplane arrangement. This hyperplane arrangement yields a tropical hypersurface, which is a deformation retract of the amoeba. Passare and Rullgård proceed in the following way. Let $\langle \cdot, \cdot \rangle$ denote the usual scalar product. For every component $E_\alpha(f) \neq \emptyset$ of the complement of $\mathcal{A}(f)$, i.e., $\alpha \in \mathrm{Comp}(f)$, we define the *Ronkin coefficient*

$$r_\alpha \quad := \quad N_f(\mathbf{x}) - \langle \alpha, \mathbf{x} \rangle \quad \text{for every } \mathbf{x} \in E_\alpha(f). \tag{6.1}$$

Indeed, r_α is well defined. By Theorem 6.1 for all $\mathbf{x} \in E_\alpha(f)$ the Ronkin function is affine linear, i.e., it is for the form $r_\alpha + \langle \alpha', \mathbf{x} \rangle$ for some $r_\alpha \in \mathbb{R}$ and $\alpha' \in \mathbb{Z}^n$. Theorem 6.1 also states that the gradient of $N_f(\mathbf{x})$ equals $\mathrm{ord}(\mathbf{x})$ if $\mathbf{x} \in \mathcal{A}(f)^c$. Thus, for $\mathbf{x} \in E_\alpha(f)$ we have $\mathrm{grad}\, N_f(\mathbf{x}) = \mathrm{ord}(\mathbf{x}) = \alpha$ and therefore $\alpha' = \alpha$. This implies that (6.1) is well defined. We define a new tropical polynomial

$$\mathrm{SpineTrop}(f) \quad := \quad \bigoplus_{\alpha \in \mathrm{Comp}(f)} r_\alpha \odot \mathbf{x}^\alpha, \tag{6.2}$$

and we define the *spine* $\mathcal{S}(f)$ as

$$\mathcal{S}(f) \quad := \quad \mathcal{T}(\mathrm{SpineTrop}(f)).$$

We give a visualization of these objects in Figure 10.

As announced, comparing (5.1) and (6.2) we can see that the only difference between $\mathrm{CompTrop}(f)$ and $\mathrm{SpineTrop}(f)$ is the replacement of $\log |b_\alpha|$ by r_α for every $\alpha \in \mathrm{Comp}(f)$. This difference, however, is crucial; see [33, Theorem 1].

Theorem 6.2 (Passare, Rullgård, [33]). *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$. The spine $\mathcal{S}(f)$ is a deformation retract of $\mathcal{A}(f)$.*

Proof. (Rough idea) Let $\alpha \in \mathrm{conv}(A) \cap \mathbb{Z}^n$ such that $E_\alpha(f) \neq \emptyset$ and let $F_\alpha(f) := \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathrm{SpineTrop}(f)(\mathbf{x}) = r_\alpha + \langle \alpha, \mathbf{x} \rangle\}$, i.e., the subset of $\mathbb{R}^n$ where $r_\alpha \odot \mathbf{x}^\alpha$ is the dominating term. By the convexity of the Ronkin function and the construction of $\mathrm{SpineTrop}(f)$ it follows that $E_\alpha(f) \subseteq F_\alpha(f)$. Since moreover every $E_\alpha(f)$ is open we have $E_\alpha(f) \subset F_\alpha(f)$ and hence $\mathcal{S}(f) \subset \mathcal{A}(f)$. Since every component of the complement of $\mathcal{S}(f)$ corresponds to a different set $F_\alpha(f)$, we can deformation retract $\mathcal{A}(f)$ to $\mathcal{S}(f)$. □

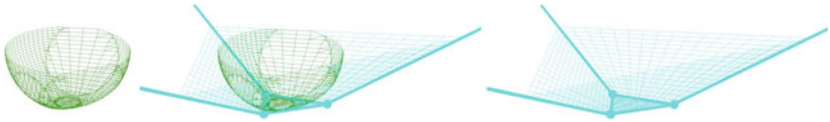


FIGURE 10. A visualization of the construction of the spine: For a given polynomial f the Ronkin function is convex and affine linear on the complement of the amoeba $\mathcal{A}(f)$ (left picture). One hyperplane is taken for every component of the complement (middle picture). The resulting hyperplane arrangement SpineTrop(f) is kept, while the Ronkin function itself is not needed anymore (right picture).

Let us assume that we know Comp(f) for some given $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$. Since the Ronkin function is defined by an integral, an immediate question for applied purposes is whether the Ronkin coefficients can be computed or at least approximated efficiently. The answer is that it depends on the particular Laurent polynomial f. But there is something that we can do. Namely, r_α is given by the real part of a function Φ_α, which almost satisfies a Gelfand–Kapranov–Zelevinsky hypergeometric system, see [33, Theorem 3]. For additional background on hypergeometric systems and related topics see [14, 43]. This fact allows to represent Φ_α via a power series representation and therefore to approximate the Ronkin coefficient r_α. This representation has some important consequences.

Theorem 6.3 (Passare, Rullgård, [33]). *Let $A \subset \mathbb{Z}^n$ a finite set and $f(\mathbf{z}) := \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha$ be a Laurent polynomial in $(\mathbb{C}^*)^A$ such that $E_{\alpha^*}(f) \neq 0$ for some fixed $\alpha^* \in \operatorname{conv}(A)$. Let Γ be a face of the Newton polytope* New$(f) = \operatorname{conv}(A)$ *of f. If $\alpha^* \in \Gamma$, then $\Phi_{\alpha^*}(f)$ and thus also r_{α^*} only depends on the terms of f with exponents in the face Γ. In particular, if α^* is a vertex of the Newton polytope of f, then $r_{\alpha^*} = \log|b_{\alpha^*}|$.*

An immediate consequence is the following corollary, which also implies Theorem 5.5.

Corollary 6.4. *Let $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ such that $\mathcal{A}(f)$ is a solid amoeba. Then $\mathcal{C}(f) = \mathcal{S}(f)$.*

Even if the power series representation of r_α is known, then it is often not obvious, how much a Ronkin coefficient r_α differs from the corresponding term $\log|b_\alpha|$. Using a result by Duistermaat and van der Kallen [7] Passare and Tsikh proved the following statement, which shows, roughly speaking, that for coefficients b_α with "large" absolute value the coefficients r_α and $\log|b_\alpha|$ "do not differ very much".

Theorem 6.5 (Passare, Tsikh, [35]). *Let $f(\mathbf{z}) := \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ be a Laurent polynomial, let $\alpha^* \in \operatorname{conv}(A) \cap \mathbb{Z}^n$ be fixed and r_{α^*} the corresponding Ronkin coefficient. Then the function $\mathbb{C} \to \mathbb{R}, b_{\alpha^*} \mapsto r_{\alpha^*} - \log|b_{\alpha^*}|$ is harmonic in a neighborhood of infinity. It has a zero of finite order at $b_{\alpha^*} = \infty$ and its power series expansion at infinity has finite radius of convergence.*

Note that in the previous theorem r_{α^*} depends on the choice of b_{α^*}. An example for the difference between $\mathcal{C}(f)$ and $\mathcal{S}(f)$ is given in Figure 11; see also [6, Figure 4.2] and [46, Figure 1].

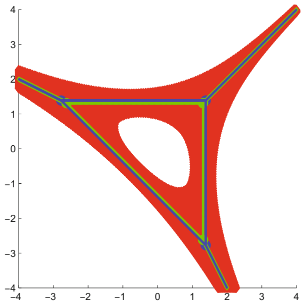


FIGURE 11. Let $f(z_1, z_2) := 1 + z_1^2 z_2 + z_1 z_2^2 - 4 z_1 z_2$. The picture shows the amoeba $\mathcal{A}(f)$ (red) with the spine $\mathcal{S}(f)$ (green, light) and the complement-induced tropical hypersurface $\mathcal{C}(f)$ (blue, dark). Note that here $\mathcal{S}(f)$ and $\mathcal{C}(f)$ coincide on the outer tentacles of $\mathcal{A}(f)$ but not on the triangles in the middle.

7. Summary and open problems

Finally, let us summarize the content of this survey and particularly compare the four different tropicalizations of amoebas that we have investigated. Moreover, let us emphasize on the open problems that we have encountered on the way.

First let us return to the two misconceptions which were stated in the introduction. We can now conclude that they are wrong and why they are wrong. For the first one we have the following corollary.

Corollary 7.1. *Let $A \subset \mathbb{Z}^n$ finite and $f \in (\mathbb{C}^*)^A$ with amoeba $\mathcal{A}(f)$. The spine $\mathcal{S}(f)$ and the non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ are not homotopy equivalent in general since the spine is a deformation retract of $\mathcal{A}(f)$ and $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ is not in general.*

Proof. Follows immediately from Theorem 3.3 and Theorem 6.2. $\square$

For the second misconception we formulate the following remark.

Remark 7.2. *For a given $f \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ it is in general not possible to compute the spine $\mathcal{S}(f)$ without knowing which components of the complement of $\mathcal{A}(f)$ exist.*

The tropical polynomials $\mathrm{SpineTrop}(f)$ and also $\mathrm{CompTrop}(f)$ require knowledge of $\mathrm{Comp}(f)$ in order to be defined, since $\mathrm{Comp}(f)$ is their support; see (5.1) and (6.2). Thus, we cannot use $\mathrm{SpineTrop}(f)$ or its tropical hypersurface, the spine $\mathcal{S}(f)$, to determine this set. The approach is circular, since one needs to know the desired set in advance. Hence, the question about the existence of particular components of the complement of an amoeba is cannot be solved by investigating the spine. This was the second misconception from the introduction. We provide an example to show this problem explicitly.

Example 7.3. Let $f(z_1, z_2) := 1 + z_1^2 z_2 + z_1 z_2^2 - a z_1 z_2$ with $a \in \mathbb{R}$. According to the generalization of Theorem 3.4 in [46] $\mathcal{A}(f)$ is solid if and only if $a \in [-3, 1]$; see Figure 11 for $a = -4$. However, one can check, e.g., using Purbhoo's results in [38] that the Ronkin coefficient $r_{(1,1)} > 1$ for $a = -3 + \varepsilon$ for $\varepsilon > 0$ sufficiently small. Thus, for $a = -3 + \varepsilon$ the tropical polynomial

$$0 \oplus 0 \odot x_1^2 x_2 \oplus 0 \odot x_1 x_2^2 \oplus r_{(1,1)} \odot x_1 x_2$$

has a tropical hypersurface of genus one, similar to the one depicted in Figure 11. To construct the correct spine $\mathcal{S}(f)$ one has to *know* that $\mathcal{A}(f)$ is solid and thus $(1, 1) \notin \mathrm{Comp}(f)$ and hence the term $r_{(1,1)} \odot x_1 x_2$ does belong to $\mathrm{SpineTrop}(f)$. The spine $\mathcal{S}(f)$ is indeed the tropical hypersurface of genus zero given by

$$\mathrm{SpineTrop}(f) \quad = \quad 0 \oplus 0 \odot x_1^2 x_2 \oplus 0 \odot x_1 x_2^2.$$

Second, we summarize the properties of the different tropicalizations that we have encountered during this survey. Let $f(\mathbf{z}) = \sum_{\alpha \in A} b_\alpha \mathbf{z}^\alpha \in \mathbb{C}[\mathbf{z}^{\pm 1}]$ be a Laurent polynomial with amoeba $\mathcal{A}(f)$. Recall that $\mathrm{Comp}(f) \subseteq \mathrm{conv}(A) \cap \mathbb{Z}^n$ is the image of the order map; see also (2.3).

The non-Archimedean amoeba

- The non-Archimedean amoeba $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}}) = \mathcal{A}_{\mathbb{K}}(\Psi(f))$ is the tropical hypersurface given by the tropical polynomial $\bigoplus_{\alpha \in A} - \mathrm{val}(\alpha) \odot \mathbf{x}^\alpha = \bigoplus_{\alpha \in A} 0 \odot \mathbf{x}^\alpha$; see Kapranov's Theorem 3.1. Since the support is given by A and all coefficients are equal, $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ is particularly very easy to compute.
- The amoeba $\mathcal{A}(f)$ can be deformed such that it converges to $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$ in Hausdorff metric. This process is based on a change of the basis of the $\mathrm{Log}\,|\cdot|$-map; see Section 4, particularly Theorem 4.1.
- $\mathcal{A}(f)$ and $\mathcal{A}(f_{\mathbb{K}})_{\mathbb{K}}$ are not homotopy equivalent in general; see Theorems 3.3 and 3.4. Also, nothing is known about the Hausdorff distance of the initial $\mathcal{A}(f)$ (that is, before applying any deformation) and $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$.

The Archimedean tropical hypersurface

- The Archimedean tropical hypersurface $\mathcal{AR}(f)$ is given by the tropical polynomial $\bigoplus_{\alpha \in A} \log |b_\alpha| \odot \mathbf{x}^\alpha$. The computation of $\mathcal{AR}(f)$ is easy in theory and can be expected to work well in practice, too.
- The distance between $\mathcal{A}(f)$ and $\mathcal{AR}(f)$ is bounded; see Theorem 5.1 for details. Thus, $\mathcal{AR}(f)$ provides a rough but quick approximation of $\mathcal{A}(f)$. See also [1] for further details.
- $\mathcal{A}(f)$ and $\mathcal{AR}(f)$ are not homotopy equivalent in general; see, e.g., Theorem 3.3 and Figure 6.

The complement induced tropical hypersurface

- The complement induced tropical hypersurface $\mathcal{C}(f)$ is given by the tropical polynomial $\bigoplus_{\alpha \in \mathrm{Comp}(A)} \log|b_\alpha| \odot \mathbf{x}^\alpha$. The computation of $\mathcal{C}(f)$ is difficult, since it is non-trivial to compute the set $\mathrm{Comp}(f)$ in general, which $\mathcal{C}(f)$ depends on.
- $\mathcal{C}(f)$ and $\mathcal{A}(f)$ are not homotopy equivalent in general; see Example 5.13 and Corollary 5.14.
- However, $\mathcal{C}(f)$ is a deformation retract of $\mathcal{A}(f)$ in certain special cases, e.g., if the amoeba is solid or if the support A of f is a circuit with $n \geq 2$; see Theorems 5.4 and 5.5. Also for polynomials supported on circuits bounds are known for the coefficients which allow to compute $\mathrm{Comp}(f)$ immediately; see [46] for further details.

The spine

- The spine $\mathcal{S}(f)$ is the tropical hypersurface given by the tropical polynomial $\bigoplus_{\alpha \in \mathrm{Comp}(A)} r_\alpha \odot \mathbf{x}^\alpha$ where r_α denotes the Ronkin coefficient of α; see (6.1). The computation of $\mathcal{S}(f)$ is very difficult, since first, as $\mathcal{C}(f)$, the spine depends on the set $\mathrm{Comp}(f)$. Moreover, the computation of the Ronkin coefficients r_α is a priori also non-trivial.
- The spine $\mathcal{S}(f)$ is a deformation retract of the amoeba $\mathcal{A}(f)$; see Theorem 6.2.

Finally, I would like to emphasize on three general problems, which are related to the content of this survey, that currently remain open.

1. The most general open problem which was discussed in this survey was stated in Problem 2.6: Find an algebraic and/or a topological description of the sets U_α^A. Furthermore, determine for a fixed $A \subset \mathbb{Z}^n$ for which $\alpha \in \mathrm{conv}(A) \cap \mathbb{Z}^n$ it holds that $U_\alpha^A = \emptyset$. Except of some special cases, e.g., certain polynomials supported on circuits, this problem is open. It is, however, likely also too general to be answered completely. Note that Problem 5.11 asking to determine $\mathrm{Comp}(f)$ for a given support set A (in dependence of the choice of coefficients) is essentially the same problem in different words. Note furthermore that if Rullgård's Conjecture 5.15 is true, then the question when $U_\alpha^A = \emptyset$ would be solved.
2. We have seen in this survey that $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$, $\mathcal{AR}(f)$, and $\mathcal{C}(f)$ are not homotopy equivalent to $\mathcal{A}(f)$ in general; see Theorem 3.3, Example 5.13, and Corollary 5.14. It would be interesting to determine classes of polynomials for which $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$, $\mathcal{AR}(f)$, and $\mathcal{C}(f)$ are homotopy equivalent to $\mathcal{A}(f)$ or even a deformation retract of $\mathcal{A}(f)$.
3. Every one of the tropical hypersurfaces $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$, $\mathcal{AR}(f)$, and $\mathcal{C}(f)$ have a canonical equivalent to the order map since every connected component of the complement of a tropical hypersurface corresponds to a unique dominating term in the defining tropical polynomial. Hence, we can consider the exponent of this dominating term as the order of the particular component. Thus, it is

a priori possible that an amoeba $\mathcal{A}(f)$ and one of the tropical hypersurfaces $\mathcal{A}_{\mathbb{K}}(f_{\mathbb{K}})$, $\mathcal{AR}(f)$ are homotopy equivalent, but the images of their order maps are not identical. It would be interesting whether this can indeed happen and, if it can happen, for which classes of polynomials and for which tropical hypersurfaces it can happen.

Acknowledgement

I cordially thank Hannah Markwig for her encouragement to write this survey, for her support, and for her various helpful comments. Furthermore, I thank Maurice Rojas, Franziska Schroeter, Thorsten Theobald and in particular Jens Forsgård for their helpful comments. I especially thank the anonymous referee for providing countless helpful comments and suggestions which improved the article incredibly. Particularly, the Problems 5.3 and Problem (3) on Page 188 are open problems that were suggested by the referee.

References

[1] M. Avendaño, R. Kogan, M. Nisse, and J. M. Rojas, *Metric estimates and membership complexity for archimedean amoebae and tropical hypersurfaces*, 2013, Preprint, arXiv:1307.3681.

[2] J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 36, Springer-Verlag, Berlin, 1998.

[3] N. Bourbaki, *Elements of mathematics. Commutative algebra*, Hermann, Paris, 1972.

[4] E. Brieskorn and H. Knörrer, *Plane algebraic curves*, Birkhäuser/Springer Basel AG, Basel, 1986.

[5] E. Brugallé and K. Shaw, *A bit of tropical geometry*, Amer. Math. Monthly **121** (2014), no. 7, 563–589.

[6] T. de Wolff, *On the Geometry, Topology and Approximation of Amoebas*, Ph.D. thesis, Goethe University, Frankfurt am Main, 2013.

[7] J. J. Duistermaat and W. van der Kallen, *Constant terms in powers of a Laurent polynomial*, Indag. Math. (N.S.) **9** (1998), no. 2, 221–231.

[8] M. Einsiedler, M. M. Kapranov, and D. Lind, *Non-Archimedean amoebas and tropical varieties*, J. Reine Angew. Math. **601** (2006), 139–157.

[9] M. Einsiedler, D. Lind, R. Miles, and T. Ward, *Expansive subdynamics for algebraic $\mathbb{Z}^d$-actions*, Ergodic Theory Dynam. Systems **21** (2001), no. 6, 1695–1729.

[10] M. Forsberg, *Amoebas and laurent series*, Ph.D. thesis, KTH Stockholm, 1998.

[11] M. Forsberg, M. Passare, and A. Tsikh, *Laurent determinants and arrangements of hyperplane amoebas*, Adv. Math. **151** (2000), 45–70.

[12] J. Forsgård, *Tropical aspects of real polynomials and hypergeometric functions*, Ph.D. thesis, Stockholm University, 2015.

[13] A. Gathmann, *Tropical algebraic geometry*, Jahresber. Deutsch. Math.-Verein. **108** (2006), no. 1, 3–32.

[14] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[15] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2001.

[16] S. Iliman and T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Res. Math. Sci. **3** (2016), 3:9.

[17] I. Itenberg, G. Mikhalkin, and E. Shustin, *Tropical algebraic geometry*, second ed., Oberwolfach Seminars, vol. 35, Birkhäuser Verlag, Basel, 2009.

[18] M. M. Kapranov, *Amoebas over non-archimedian fields*, preprint, 2000.

[19] R. Kenyon, A. Okounkov, and S. Sheffield, *Dimers and amoebae*, Ann. of Math. (2) **163** (2006), no. 3, 1019–1056.

[20] J. A. De Loera, J. Rambau, and F. Santos, *Triangulations*, Algorithms and Computation in Mathematics, vol. 25, Springer-Verlag, Berlin, 2010, Structures for algorithms and applications.

[21] D. Maclagan and B. Sturmfels, *Introduction to Tropical Geometry*, Amer. Math. Soc., Providence, R.I., 2015.

[22] V. P. Maslov, *On a new superposition principle for optimization problem*, Séminaire sur les équations aux dérivées partielles, 1985–1986, École Polytech., Palaiseau, 1986, pp. Exp. No. XXIV, 14.

[23] G. Mikhalkin, *Real algebraic curves, the moment map and amoebas*, Ann. Math. **151** (2000), 309–326.

[24] ———, *Amoebas of algebraic varieties and tropical geometry*, Different Faces of Geometry (S. K. Donaldson, Y. Eliashberg, and M. Gromov, eds.), Kluwer, New York, 2004, pp. 257–300.

[25] ———, *Decomposition into pairs-of-pants for complex algebraic hypersurfaces*, Topology **43** (2004), no. 5, 1035–1065.

[26] ———, *Enumerative tropical algebraic geometry in* $\mathbb{R}^2$, J. Amer. Math. Soc. **18** (2005), no. 2, 313–377.

[27] G. Mikhalkin and J. Rau, *Tropical geometry*, book manuscript, 2014.

[28] G. Mikhalkin and H. Rullgård, *Amoebas of maximal area*, Internat. Math. Res. Notices (2001), no. 9, 441–451.

[29] M. Passare, *How to compute* $\sum 1/n^2$ *by solving triangles*, Amer. Math. Monthly **115** (2008), no. 8, 745–752.

[30] M. Passare, D. Pochekutov, and A. Tskih, *Amoebas of complex hypersurfaces in statistical thermodynamics*, Math. Phys. Anal. Geom. **15** (2012).

[31] M. Passare and J. J. Risler, *On the curvature of the real amoeba*, Proceedings of the Gökova Geometry-Topology Conference 2010, Int. Press, Somerville, MA, 2011, pp. 129–134.

[32] M. Passare, J. M. Rojas, and B. Shapiro, *New multiplier sequences via discriminant amoebae*, Mosc. Math. J. **11** (2011), no. 3, 547–560, 631.

[33] M. Passare and H. Rullgård, *Amoebas, Monge–Ampère measures and triangulations of the Newton polytope*, Duke Math. J. **121** (2004), no. 3, 481–507.

[34] M. Passare, T. Sadykov, and A. Tsikh, *Singularities of hypergeometric functions in several variables*, Compositio Math. **141** (2005), 787–810.

[35] M. Passare and A. Tsikh, *Amoebas: their spines and their contours*, Idempotent Mathematics and Mathematical Physics, Contemp. Math., vol. 377, Amer. Math. Soc., Providence, RI, 2005, pp. 275–288.

[36] A. E. Pellet, *Sur un mode de séparation des racines des équations et la formule de Lagrange*, Darboux Bull. (2) **5** (1881), 393–395, in French.

[37] R. Pemantle, *Hyperbolicity and stable polynomials in combinatorics and probability*, Current developments in mathematics, 2011, Int. Press, Somerville, MA, 2012, pp. 57–123.

[38] K. Purbhoo, *A Nullstellensatz for amoebas*, Duke Math. J. **141** (2008), no. 3, 407–445.

[39] J. Richter-Gebert, B. Sturmfels, and T. Theobald, *First steps in tropical geometry*, Idempotent Mathematics and Mathematical Physics, Contemp. Math., vol. 377, Amer. Math. Soc., Providence, RI, 2005, pp. 289–317.

[40] L. I. Ronkin, *On zeros of almost periodic functions generated by functions holomorphic in a multicircular domain*, Complex Analysis in Modern Mathematics (Russian), FAZIS, Moscow, 2001, pp. 239–251.

[41] H. Rullgård, *Stratification des espaces de polynômes de Laurent et la structure de leurs amibes*, C. R. Acad. Sci. Paris, Sèrie I **331** (2000), 355–358, in French.

[42] ———, *Topics in Geometry, Analysis and Inverse Problems*, Ph.D. thesis, Stockholm University, 2003.

[43] M. Saito, B. Sturmfels, and N. Takayama, *Gröbner deformations of hypergeometric differential equations*, Algorithms and Computation in Mathematics, vol. 6, Springer-Verlag, Berlin, 2000.

[44] F. Schroeter and T. de Wolff, *The boundary of amoebas*, 2013, preprint, arXiv:1310.7363.

[45] D. Speyer and B. Sturmfels, *The tropical Grassmannian*, Adv. Geom. **4** (2004), no. 3, 389–411.

[46] T. Theobald and T. de Wolff, *Amoebas of genus at most one*, Adv. Math. **239** (2013), 190–213.

[47] T. Theobald and T. de Wolff, *Norms of roots of trinomials*, Math. Ann. **366** (2016), no. 1, 219–247.

[48] O. Viro, *Dequantization of real algebraic geometry on logarithmic paper*, European Congress of Mathematics, Vol. I (Barcelona, 2000), Progr. Math., vol. 201, Birkhäuser, Basel, 2001, pp. 135–146.

[49] ———, *What is an amoeba?*, Notices AMS **49** (2002), no. 8, 916–917.

[50] ———, *From the sixteenth Hilbert problem to tropical geometry*, Jpn. J. Math. **3** (2008), no. 2, 185–214.

[51] ———, *Multivariate hypergeometric and algebraic functions*, 2015, Book draft; in Russian.

[52] ———, *Patchworking real algebraic varieties*, 206, Translation of Dissertation, see arXiv:math/0611382.

[53] G. M. Ziegler, *Lectures on polytopes*, Springer-Verlag, 2007.

Timo de Wolff
Texas A&M University
Department of Mathematics
College Station, TX 77843-3368, USA
e-mail: dewolff@math.tamu.edu

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 191–212

Coamoebas of Polynomials Supported on Circuits

Jens Forsgård

Abstract. We study coamoebas of polynomials supported on circuits. Our results include an explicit description of the space of coamoebas, a relation between connected components of the coamoeba complement and critical points of the polynomial, an upper bound on the area of a planar coamoeba, and a recovered bound on the number of positive solutions of a fewnomial system.

1. Introduction

A possibly degenerate circuit is a point configuration $A \subset \mathbb{Z}^n$ of cardinality $n+2$ which span a sublattice $\mathbb{Z}A$ of rank n. That is, such that the Newton polytope $\mathcal{N}_A = \operatorname{Conv}(A)$ is of full dimension. A polynomial system $f(z) = 0$ is said to be supported on a circuit A if each polynomial occurring in $f(z)$ is supported on A. Polynomial systems supported on circuits have recently been been studied in the context of, e.g., real algebraic geometry [4, 6], complexity theory [5], and amoeba theory [18]. The name "circuit" originate from matroid theory; see [17] and [20] for further background.

The aim of this article is to describe geometrical and topological properties of coamoebas of polynomials supported on circuits. Such an investigation is motivated not only by the vast number of applications of circuits in different areas of geometry, but also since circuits provide an ideal testing ground for open problems in coamoeba theory.

This paper is organized as follows. In Section 2 we will give a brief overview of coamoeba theory. In Section 3 we will discuss the relation between real polynomials and the coamoeba of the A-discriminant. The main results of this paper are contained in Sections 4–7, each of which can be read as a standalone text.

In Section 4 we will give a complete description of the space of coamoebas. That is, we will describe how the topology of the coamoeba $\mathcal{C}_f$ depends on the coefficients of f. Describing the space of amoebas is the topic of the articles [18] and [19], and to fully appreciate our result one should consider these spaces simultaneously, see, e.g., Figure 2. The geometry of the space of coamoebas is closely related to the A-discriminantal variety, see Theorems 4.1 and 4.2.

In Section 5 we will prove that the area of a planar circuit coamoeba is bounded from above by $2\pi^2$. That is, a planar circuit coamoeba covers at most half of the torus $\mathbf{T}^2$. Furthermore, we will prove that a circuit admits a coamoeba of maximal area if and only if it admits an equimodular triangulation. Note that we calculate area without multiplicities, in contrast to [11]. However, the relation between (co)amoebas of maximal area and Harnack curves is made visible also in this setting.

In Section 6 we will prove that, under certain assumptions on A, the critical points of $f(z)$ are projected by the componentwise argument mapping into distinct connected component of the complement of the coamoeba $\mathcal{C}_f$. Furthermore, this projection gives a bijective relation between the set of critical points and the set of connected components of the complement of the closed coamoeba. This settles a conjecture used in [10] when computing monodromy in the context of dimer models and mirror symmetry.

In Section 7 we will consider bivariate systems supported on a circuit. If such a system is real, then it admits at most three roots in $\mathbb{R}^2_+$. The main contribution of this section is that we offer a new approach to fewnomial theory. Using our method, we will prove that if $\mathcal{N}_A$ is a simplex, then, for each $\theta \in \mathbf{T}^2$, a complex bivariate system supported on A has at most two roots in the *sector* $\operatorname{Arg}^{-1}(\theta)$.

2. Coamoebas and lopsidedness

Let A denote a point configuration $A = \{\mathbf{a}_0, \ldots, \mathbf{a}_{N-1}\} \subset \mathbb{Z}^n$, where $N = \#A$. By abuse of notation, we identify A with the $(1+n) \times N$-matrix

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ \mathbf{a}_0 & \cdots & \mathbf{a}_{N-1} \end{pmatrix}. \tag{1}$$

The *codimension* of A is the integer $m = N-1-n$. A *circuit* is a point configuration of codimension one. A circuit is said to be nondegenerate if it is not a pyramid over a circuit of smaller dimension. That is, if all maximal minors of the matrix A are nonvanishing. We will partition the set of circuits into two classes; *simplex circuits*, for which $\mathcal{N}_A$ is a simplex, and *vertex circuits*, for which $A = \operatorname{Vert}(\mathcal{N}_A)$.

We associate to A the family $\mathbb{C}^A_*$ consisting of all polynomials

$$f(z) = \sum_{k=0}^{N-1} f_k\, z^{\mathbf{a}_k},$$

where $f(z)$ is identified with the point $f = (f_0, \ldots, f_{N-1}) \in \mathbb{C}^A_*$. By slight abuse of notation, we will denote by $f_k(z)$ the monomial function $z \mapsto f_k\, z^{\mathbf{a}_k}$. We denote the algebraic set defined by f by $Z(f) \subset \mathbb{C}^n_*$. The *coamoeba* $\mathcal{C}_f$ is the image of $Z(f)$ under the componentwise argument mapping $\operatorname{Arg}\colon \mathbb{C}^n_* \to \mathbf{T}^n$ defined by

$$\operatorname{Arg}(z) = (\arg(z_1), \ldots, \arg(z_n)),$$

where $\mathbf{T}^n$ denotes the real n-torus. It is sometimes beneficial to consider the multivalued argument mapping, which gives the coamoeba as a multiply periodic subset of $\mathbb{R}^n$. Coameobas were introduced by Passare and Tsikh as a dual object, in an imprecise sense, of the amoeba $\mathcal{A}_f$.

We will say that a point $z \in \mathbb{C}_*^n$ is a *critical point* of f if it solves the system

$$\partial_1 f(z) = \cdots = \partial_n f(z) = 0. \tag{2}$$

If in addition $z \in Z(f)$ then z will be called a *singular point* of f. The A-discriminant $\Delta(f) = \Delta_A(f)$ is an irreducible polynomial with domain $\mathbb{C}_*^A$ which vanishes if and only if f has a singular point in $\mathbb{C}_*^n$ [9].

A *Gale dual* of A is an integer matrix B whose columns span the right $\mathbb{Z}$-kernel of A. That is, B is an integer $N \times m$-matrix, of full rank, such that its maximal minors are relatively prime. A Gale dual is unique up to the action of $\mathrm{SL}_m(\mathbb{Z})$. The rows $\mathbf{b}$ of B are indexed by the points $\mathbf{a}_k \in A$. To each Gale dual we associate a zonotope

$$\mathcal{Z}_B = \left\{ \frac{\pi}{2} \sum_{k=0}^{N-1} \lambda_k \mathbf{b}_k \,\middle|\, |\lambda_k| \leq 1 \right\} \subset \mathbb{R}^m.$$

We will say that a triangulation T of $\mathcal{N}_A$ is a triangulation of A if $\mathrm{Vert}(T) \subset A$. Such a triangulation is said to be *equimodular* if all maximal simplices has equal volume.

Let h be a *height function* $h\colon A \to \mathbb{R}$. The function h induces a triangulation T_h of A in the following manner. Let $\mathcal{N}_h$ denote the polytope in $\mathbb{R}^{n+1}$ with vertices $(\mathbf{a}, h(\mathbf{a}))$. The lower facets of $\mathcal{N}_h$ are the facets whose outward normal vector has negative last coordinate. Then, T_h is the triangulation of A whose maximal simplices are the images of the lower facets of $\mathcal{N}_h$ under the projection onto the first n coordinates. A triangulation T of A is said to be *coherent* if there exists a height function h such that $T = T_h$.

If A is a circuit then B is a column vector, unique up to sign. Hence, the zonotope $\mathcal{Z}_B$ is an interval. Let $A_k = A \setminus \{\mathbf{a}_k\}$, with associated matrix A_k, and let $V_k = \mathrm{Vol}(A_k)$. If A is a nondegenerate circuit, so that $V_k > 0$ for all k, then $\mathcal{N}_A$ admits exactly two coherent triangulations with vertices in A [9]. Denote these two triangulations by T_δ for $\delta \in \{\pm 1\}$. Each simplex $\mathcal{N}_{A_k}$ occurs in exactly one of the triangulations T_δ. Hence, there is a well-defined assignment of signs $k \mapsto \delta_k$, where $\delta_k \in \{\pm 1\}$, such that

$$T_\delta = \left\{ \mathcal{N}_{A_k} \right\}_{\delta_k = \delta}, \quad \delta = \pm 1.$$

Here, we have identified a triangulation with its set of maximal simplices. As shown in [9, Ch. 7 and Ch. 9] and [7, Sec. 5], a Gale dual of A is given by

$$\mathbf{b}_k = (-1)^k |A_k| = \delta_k V_k. \tag{3}$$

Thus, the zonotope $\mathcal{Z}_B$ is an interval of length $2\pi\, \mathrm{Vol}(A)$.

The A-discriminant Δ has $n+1$ homogeneities, one for each row of the matrix A. Each Gale dual correspond to a dehomogenization of Δ. To be specific, introducing the variables

$$\xi_j = \prod_{k=0}^{N-1} f_k^{\mathbf{b}_{kj}}, \quad j = 1, \ldots, m, \tag{4}$$

there is a Laurent monomial $M(c)$ and a polynomial $\Delta_B(\xi)$ such that

$$\Delta_B(\xi) = M(f)\Delta_A(f).$$

We will say that Δ_B is the *reduced form* of Δ. Such a reduction yields a projection $\mathrm{pr}_B : \mathbb{C}_*^A \to \mathbb{C}_*^m$, and we will say that $\mathbb{C}_*^m$ is the reduced family associated to A, and that $\mathrm{pr}_B(f)$ is the reduced form of f.

Example 2.1. Let $A = \{0, 1, 2\}$, so that $\mathbb{C}_*^A$ is the family of quadratic univariate polynomials

$$f(z) = f_0 + f_1 z + f_2 z^2.$$

Consider the Gale dual $B = (1, -2, 1)^t$, and introduce the variable $\xi = f_0 f_1^{-2} f_2$. In this case the A-discriminant Δ_A is well known, and we find that

$$f_1^{-2}\Delta_A(f) = f_1^{-2}\left(f_1^2 - 4f_0 f_2\right) = 1 - 4\xi = \Delta_B(\xi).$$

The projection pr_B correspond to performing the change of variables $z \mapsto f_0 f_1^{-1} z$, and multiplying $f(z)$ by f_0^{-1}, after with we obtain the reduced family consisting of all polynomials of the form

$$f(z) = 1 + z + \xi z^2.$$

Let S denote a subset of A. The *truncated polynomial* f_S is the image of f under the projection $\mathrm{pr}_S : \mathbb{C}_*^A \to \mathbb{C}_*^S$. Of particular interest is the case when $S = \Gamma \cap A$ for some face Γ of the Newton polytope $\mathcal{N}_A$ (denoted by $\Gamma \prec \mathcal{N}_A$). We will write $f_\Gamma = f_{\Gamma \cap A}$. It was shown in [14] that

$$\overline{\mathcal{C}}_f = \bigcup_{\Gamma \prec \mathcal{N}_A} \mathcal{C}_{f_\Gamma},$$

Let $\mathcal{E}$ denote the set of edges of $\mathcal{N}_A$, then the *shell* of the coamoeba is defined by

$$\mathcal{H}_f = \bigcup_{\Gamma \in \mathcal{E}} \mathcal{C}_{f_\Gamma}.$$

As an edge Γ is one-dimensional, the shell $\mathcal{H}_f$ is a hyperplane arrangement. Its importance can be seen in that each full-dimensional cell of $\mathcal{H}_f$ contain at most one connected component of the complement of $\overline{\mathcal{C}}_f$, see [7].

Example 2.2. The coamoeba of $f(z) = 1 + z_1 + z_2$, as described in [7] and [14], can be seen in Figure 1, where it is drawn in the fundamental domains $[-\pi, \pi]^2$ and $[0, 2\pi]^2$. The shell $\mathcal{H}_f$ consist of the hyperplane arrangement drawn in black. In this case, it is equal to the boundary of $\mathcal{C}_f$. The Newton polytope $\mathcal{N}_A$ and its outward normal vectors are drawn in the rightmost picture. If $\mathcal{H}_f$ is given orientations in accordance with the outward normal vectors of $\mathcal{N}_A$, then the interior of the coamoeba consist of the oriented cells.

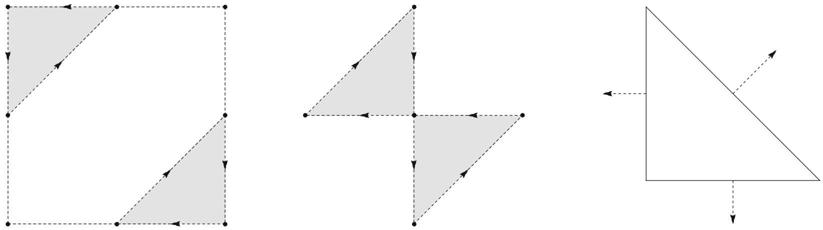


FIGURE 1. The coamoeba of $f(z) = 1 + z_1 + z_2$ in two fundamental regions, and the Newton polytope $\mathcal{N}_A$.

Acting on A by an *integer affine transformation* is equivalent to performing a monomial change of variables and multiplying f by a Laurent monomial. Such an action induces a linear transformation of the coamoeba $\mathcal{C}_f$, when viewed in $\mathbb{R}^n$ [7]. We will repeatedly use this fact to impose assumptions on A, e.g., that it contains the origin.

The polynomial f is said to be *colopsided* at a point $\theta \in \mathbb{R}^n$ if there exist a phase φ such that

$$\Re\left(e^{i\varphi} f_k(e^{i\theta})\right) \geq 0, \quad k = 0, \ldots, N-1, \tag{5}$$

with at least one of the inequalities (5) being strict. The motivation for this definition is as follows. If f is colopsided at θ, then

$$\Re\left(e^{i\varphi} f(re^{i\theta})\right) = \sum_{k=0}^{N-1} r^{\mathbf{a}_k}\, \Re\left(e^{i\varphi} f_k(e^{i\theta})\right) > 0, \quad \forall\, r \in \mathbb{R}_+^n,$$

since at least one term of the sum is strictly positive. Hence, colopsidedness at θ implies that $\theta \in \mathbf{T}^n \setminus \mathcal{C}_f$. The colopsided coamoeba, denoted $\mathcal{L}_f$, is defined as the set of all θ such that (5) does not hold for any phase φ [7]. Hence, $\mathcal{C}_f \subset \mathcal{L}_f$.

Each monomial $f_k(z)$ defines an affinity (i.e., a group homomorphism composed with a translation) $f_k \colon \mathbb{C}_*^n \to \mathbb{C}_*$ by $z \mapsto f_k\, z^{\mathbf{a}_k}$. We thus obtain unique affinities $|f_k|$ and $\hat{f}_k$ such that the following diagram of short exact sequences commutes:

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{R}_+^n \arrow[r] \arrow[d, "|f_k|"'] & \mathbb{C}_*^n \arrow[r] \arrow[d, "f_k"'] & (S^1)^n \arrow[r] \arrow[d, "\hat{f}_k"'] & 0 \\
0 \arrow[r] & \mathbb{R}_+ \arrow[r] & \mathbb{C}_* \arrow[r] & S^1 \arrow[r] & 0.
\end{tikzcd}$$

Notice that $\mathbf{T} \simeq S^1 \subset \mathbb{C}$. We denote by $\hat{f}(\theta) \subset (S^1)^A \subset \mathbb{C}_*^A$ the vector with components $\hat{f}_k(\theta)$. Assume that f contains the constant monomial 1, and consider the map $\operatorname{ord}_B(f) \colon \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$\operatorname{ord}_B(f)(\theta) = \operatorname{Arg}_\pi\left(\hat{f}(\theta)\right) \cdot B, \tag{6}$$

where Arg_π denotes the componentwise principal argument map. It was shown in [7] that the map $\mathrm{ord}_B(f)$ induces a map

$$\mathrm{ord}_B(f)\colon \mathbf{T}^n \setminus \overline{\mathcal{L}}_f \to \{\mathrm{Arg}_\pi(f)B + 2\pi\mathbb{Z}^m\} \cap \operatorname{int} \mathcal{Z}_B. \tag{7}$$

which in turn induces a bijection between the set of connected components of the complement of $\overline{\mathcal{L}}_f$ and the finite set in the right-hand side of (7). The map $\mathrm{ord}_B(f)$ is known as the *order map* of the lopsided coamoeba.

Remark 2.3. The requirement that f contains the monomial 1 is related to the choice of branch cut of the function Arg; in order to obtain a well-defined map, we need the right-hand side of (6) to be discontinuous only for θ such that two components of $\hat{f}(\theta)$ are antipodal, see [7]. If f does not contain the constant monomial 1, then one should fix a point $\mathbf{a}_k \in A$ and multiply the vector $\hat{f}(\theta)$ by the scalar $\hat{f}_k(\theta)^{-1}$ before taking principal arguments. It is shown in [7, Thm. 4.3] that the obtained map is independent of the choice of $\mathbf{a}_k$.

If $\theta \in \mathbf{T}^n \setminus \overline{\mathcal{L}}_f$, then we can choose φ such that

$$\Re\left(e^{i\varphi} f_k(e^{i\theta})\right) > 0, \quad k = 0, \dots, N-1.$$

That is, the boundary of $\mathcal{L}_f$ is contained in the hyperplane arrangement consisting of all θ such that two components of $\hat{f}(\theta)$ are antipodal.

It has been conjectured that the number of connected components of the complement of $\overline{\mathcal{C}}_f$ is at most $\mathrm{Vol}(A)$.[1] A proof in arbitrary dimension has been proposed by Nisse in [13], and an independent proof in the case $n = 2$ was given in [8]. That the number of connected components of the complement of $\overline{\mathcal{L}}_f$ is at most $\mathrm{Vol}(A)$ follows from the theory of Mellin–Barnes integral representations of A-hypergeometric functions, see [2] and [3].

A finite set $\mathcal{I} \subset \mathbf{T}^n$ which is in a bijective correspondence with the set of connected components of the complement of $\overline{\mathcal{C}}_f$ by inclusion, will be said to be an *index set* of the coamoeba complement. This notation will be slightly abused; a set $\mathcal{I}$ of cardinality $\mathrm{Vol}(A)$ will be said to be an index set of the coamoeba if each connected component of its complement contains exactly one element of $\mathcal{I}$.

The term "lopsided" was first used by Purbhoo in [15], denoting the corresponding condition to (5) for amoebas: the polynomial f is said to be *lopsided* at a point $x \in \mathbb{R}^n$ if there is a $\mathbf{a}_k \in A$ such that the moduli $|f_k|(x)$ is greater than the sum of the remaining modulis. As a comparison, note that the polynomial f is colopsided at $\theta \in \mathbf{T}$ if and only if the greatest intermediate angle of the components of $\hat{f}(\theta)$ is greater than the sum of the remaining intermediate angles.

[1] This conjecture has commonly been attributed to Mikael Passare, however, it seems to originate from a talk given by Mounir Nisse at Stockholm University in 2007.

3. Real points and the coameoba of the A-discriminant

We will say that f is *real at* θ, if there is a real subvector space $\ell \subset \mathbb{C}$ such that $\hat{f}_k(\theta) \in \ell$ for all $k = 0, \ldots, N-1$. If such a θ exist then f is real, that is, after a change of variables and multiplication with a Laurent monomial $f \in \mathbb{R}_*^A$. In this section, we will study the function $\hat{f}$ from the viewpoint of real polynomials. Our main result is the following characterization of the coamoeba of the A-discriminant of a circuit.

Proposition 3.1. *Let A be a nondegenerate circuit, and let δ_k be as in* (3)*. Then,* $\operatorname{Arg}(f) \in \mathcal{C}_\Delta$ *if and only if after possibly multiplying f with a constant, there is a $\theta \in \mathbb{R}^n$ such that $\hat{f}_k(\theta) = \delta_k$ for all k.*

If A is a circuit and B is a Gale dual of A then the Horn–Kapranov parametrization of the reduced discriminant Δ_B can be lifted to a parametrization of the discriminant surface Δ as

$$z \mapsto \left(\mathbf{b}_0 z^{\mathbf{a}_0}, \ldots, \mathbf{b}_{N-1} z^{\mathbf{a}_{N-1}}\right).$$

Taking componentwise arguments, we obtain a simple proof Proposition 3.1. In particular, the proposition can be interpreted as a coamoeba version of the Horn–Kapranov parametrization valid for circuits. Our proof of Proposition 3.1 will be more involved, however, for our purposes the lemmas contained in this section are of equal importance.

Lemma 3.2. *Assume that the polynomial f is real at $\theta_0 \in \mathbb{R}^n$. Then, f is real at $\theta \in \mathbb{R}^n$ if and only if $\theta \in \theta_0 + \pi L$, where L is the dual lattice of $\mathbb{Z}A$.*

Proof. After translating θ and multiplying f with a Laurent monomial, we can assume that $\theta_0 = 0$, that $\ell_0 = \mathbb{R}$, and that f contains the monomial 1. That is, all coefficients of f are real, in particular proving *if*-part of the statement. To show the *only if*-part, notice first that $\hat{f}(\theta) \subset \ell$ implies that ℓ contains both the origin and 1. That is, $\ell = \mathbb{R}$. Furthermore, $\hat{f}(\theta) \subset \mathbb{R}$ only if for each $\mathbf{a} \in A$ there is a $k \in \mathbb{Z}$ such that $\langle \mathbf{a}, \theta \rangle = \pi k$, which concludes the proof. □

The A-discriminant Δ related to a circuit has been described in [9, Chp. 9, Pro. 1.8] where the formula

$$\Delta(f) = \prod_{\delta_k=1} \mathbf{b}_k^{\mathbf{b}_k} \prod_{\delta_k=-1} f_k^{-\mathbf{b}_k} - \prod_{\delta_k=-1} \mathbf{b}_k^{-\mathbf{b}_k} \prod_{\delta_k=1} f_k^{\mathbf{b}_k} \tag{8}$$

was obtained. In particular, Δ is a binomial. As the zonotope $\mathcal{Z}_B$ is a symmetric interval of length $2\pi \operatorname{Vol}(A)$, the image of the map $\operatorname{ord}_B(f)$ is of cardinality $\operatorname{Vol}(A)$ unless

$$\operatorname{Arg}_\pi(f) B \equiv 2\pi \operatorname{Vol}(A) \mod 2\pi. \tag{9}$$

In particular, the complement of $\overline{\mathcal{C}}_f$ has the maximal number of connected components (i.e., $\operatorname{Vol}(A)$-many) unless the equivalence (9) holds.

Lemma 3.3. *For each* $\kappa = 0, 1, \dots, n+1$*, there are exactly* $\mathrm{Vol}(A_\kappa)$*-many points* $\theta \in \mathbf{T}$ *such that*

$$\hat{f}_k(\theta) = \delta_k, \quad \forall\, k \neq \kappa. \tag{10}$$

Proof. By applying an integer affine transformation, the statement follows from the case when A_κ consist of the vertices of the standard simplex. □

Lemma 3.4. *Fix* $\kappa \in \{0, \dots, n+1\}$*. For each* θ *fulfilling* (10)*, let* $\varphi_\theta \in \mathbf{T}$ *be defined by the condition that if* $\arg_\pi(f_\kappa) = \varphi_\theta$ *then*

$$\hat{f}_\kappa(\theta) = \delta_\kappa. \tag{11}$$

Assume that $\mathbb{Z}A = \mathbb{Z}^n$*. Then, the numbers* φ_θ *are distinct.*

Proof. We can assume that $\mathbf{a}_0 = \mathbf{0}$ and that $f_0 = 1$. Assume that $\varphi_{\theta_1} = \varphi_{\theta_2}$. Then,

$$\langle \mathbf{a}, \theta_2 \rangle = \langle \mathbf{a}, \theta_1 \rangle + 2\pi r, \quad \forall \mathbf{a} \in A.$$

By translating, we can assume that $\theta_1 = 0$, and hence, since 1 is a monomial of f, that all coefficients are real. Consider the lattice L consisting of all points $\theta \in \mathbb{R}^n$ such that f is real at θ. Since $\mathbb{Z}A = \mathbb{Z}^n$, Lemma 3.2 shows that $L = \pi\mathbb{Z}^n$. However, we find that

$$\left\langle \mathbf{a}, \frac{\theta_2}{2} \right\rangle = \pi r,$$

and hence $\frac{\theta_2}{2} \in L$. This implies that $\theta_2 \in 2\pi\mathbb{Z}^n$, and hence $\theta_2 = 0$ in $\mathbf{T}^n$. □

Proof of Proposition 3.1. Assume first that there is a θ as in the statement of the proposition, where we can assume that $\theta = 0$. Then, $\arg(f_k) = \arg(\delta_k)$. It follows that the monomials

$$\prod_{\delta_k = 1} \mathbf{b}_k^{\mathbf{b}_k} \prod_{\delta_k = -1} f_k^{-\mathbf{b}_k} \quad \text{and} \quad \prod_{\delta_k = -1} \mathbf{b}_k^{\mathbf{b}_k} \prod_{\delta_k = 1} f_k^{-\mathbf{b}_k}$$

have equal signs. Therefor, Δ vanishes for $f_k = \delta_k |\mathbf{b}_k|$, implying that $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$.

For the converse, fix κ, and reduce f by requiring that $f_k = \delta_k |\mathbf{b}_k| = \mathbf{b}_k$ for $k \neq \kappa$. Let $\mathcal{I}$ denote the set of points $\theta \in \mathbf{T}^n$ such that $\hat{f}_k(\theta) = \delta_k$ for $k \neq \kappa$, which by Lemma 3.3 has cardinality V_κ. By Lemma 3.4, the set $\mathcal{I}$ is in a bijective correspondence with values of $\arg(f_\kappa)$ such that $\hat{f}_\kappa(\theta) = \delta_\kappa$. Therefor, we find that Δ vanishes at $f_\kappa = V_\kappa e^{i\varphi}$ for each $\varphi \in \mathcal{I}$. However, the discriminant Δ specializes, up to a constant, to the binomial

$$\Delta_\kappa(f_\kappa) = f_\kappa^{|\mathbf{b}_\kappa|} - \mathbf{b}_\kappa^{|\mathbf{b}_\kappa|} = f_\kappa^{V_\kappa} - \mathbf{b}_\kappa^{V_\kappa},$$

which has exactly V_κ-many solutions in $\mathbb{C}_*$ of distinct arguments. Hence, since $\Delta(f) = 0$ by assumption, and comparing the number of solutions, it holds that $\hat{f}_\kappa(\theta) = \delta_\kappa$ for one of the points $\theta \in \mathcal{I}$. □

4. The space of coamoebas

Let $U_k \subset \mathbb{C}_*^A$ denote the set of all f such that the number of connected components of the complement of $\overline{\mathcal{C}}_f$ is $\mathrm{Vol}(A) - k$. Describing the sets U_k is known as the problem of describing the *space of coamoebas* of $\mathbb{C}_*^A$. In this section, we will give explicit descriptions of the sets U_k in the case when A is a circuit. As a first observation we note that the image of the map $\mathrm{ord}_B(f)$ is at least of cardinality $\mathrm{Vol}(A) - 1$, implying that

$$\mathbb{C}_*^A = U_0 \cup U_1,$$

and in particular $U_k = \emptyset$ for $k \geq 2$. Hence, it suffices for us to give an explicit description of the set U_1. Our main result is the following two theorems, highlighting also the difference between vertex circuits and simplex circuits. Note that Δ is a real polynomial [9].

Theorem 4.1. *Assume that A is a nondegenerate simplex circuit, with $\mathbf{a}_{n+1}$ as an interior point. Choose B such that $\delta_{n+1} = -1$, and let Δ be as in* (8)*. Then, $f \in U_1$ if and only if* $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$ *and*

$$(-1)^{\mathrm{Vol}(A)}\,\Delta\big(\delta_0|f_0|, \ldots, \delta_{n+1}|f_{n+1}|\big) \leq 0. \tag{12}$$

Theorem 4.2. *Assume that A is a vertex circuit. Then, $f \in U_1$ if and only if* $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$.

The article [18] describes the *space of amoebas* in the case when A is a simplex circuit in dimension at least two. In this case, the number of connected components of the amoeba complement is either equal to the number of vertices of $\mathcal{N}_A$ or one greater. One implication of [18, thm. 4.4 and thm. 5.4] is that, if the amoeba complement has the minimal number of connected components, then

$$(-1)^{\mathrm{Vol}(A)}\,\Delta\big(\delta_0|f_0|, \ldots, \delta_{n+1}|f_{n+1}|\big) \geq 0.$$

Furthermore, this set intersect U_1 only in the discriminant locus $\Delta(f) = 0$. The space of amoebas in the case when A is a simplex circuit in dimension $n = 1$ has been studied in [19], and is a more delicate problem. On the other hand, if A is a vertex circuit, then each $f \in \mathbb{C}_*^A$ is maximally sparse and hence has a solid amoeba. That is, the components of the complement of the amoeba is in a bijective correspondance with the vertices of the Newton polytope $\mathcal{N}_A$. In particular, the number of connected components of the amoeba complement does not depend on f. From Theorems 4.1 and 4.2 we see that a similar discrepancy between simplex circuits and vertex circuits occurs for coamoebas.

Example 4.3. The reduced family

$$f(z) = 1 + z_1^3 + z_2^3 + \xi\, z_1 z_2.$$

was considered in [16, ex. 6, p. 59], where the study of the space of amoebas was initiated. We have drawn the space of amoebas and coamoebas jointly in the left picture in Figure 2. The blue region, whose boundary is a hypocycloid, marks values of ξ for which the amoeba complement has no bounded component. The set U_1 is seen in orange. The red dots is the discriminant locus $\Delta(\xi) = 0$, which is contained in the circle $|\xi| = 3$ corresponding to an equality in (12).

Example 4.4. The reduced family

$$f(z) = 1 + z_1 + z_2^3 + \xi\, z_1^3 z_2$$

is a vertex circuit. In this case, the topology of the amoeba does not depend on the coefficient ξ. The space of coamoebas is drawn in the right picture in Figure 2. The set U_1 comprises the three orange lines emerging from the origin. The red dots is the discriminant locus $\Delta(\xi) = 0$. It might seem like the set U_0 is disconnected, however this a consequence of that we consider f in reduced form. In $\mathbb{C}_*^A$ the set U_0 is connected, though not simply connected.

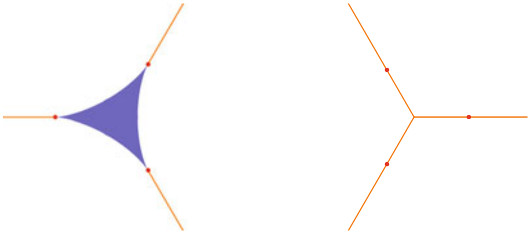


FIGURE 2. The amoeba and coamoeba spaces of Examples 4.3 and 4.4.

4.1. Proof of Theorem 4.1

Impose the assumptions of Theorem 4.1. Then,

$$\mathbf{b}_0 + \cdots + \mathbf{b}_n = -\mathbf{b}_{n+1} = \mathrm{Vol}(A),$$

and in particular $V_{n+1} = \mathrm{Vol}(A)$. By Lemma 3.3 there is a set $\mathcal{I}$ of cardinality $\mathrm{Vol}(A)$ consisting of all points θ such that $\hat{f}_0(\theta) = \cdots = \hat{f}_n(\theta) = \delta_k = 1$ In particular, f is colopsided at $\theta \in \mathcal{I}$ unless $\hat{f}_{n+1}(\theta) = -1$. It was shown in [7, sec. 5] that, if $f \notin \mathcal{C}_\Delta$, then $\mathcal{I}$ is an index set for the complement of $\overline{\mathcal{C}}_f$. In fact, $\mathcal{I}$ is an index set of the complement of $\overline{\mathcal{C}}_f$ for arbitrary f.

Proposition 4.5. *Let A be a simplex circuit. Assume that* $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$*, i.e., that there exists a $\theta \in \mathcal{I}$ with $\hat{f}_{n+1}(\theta) = \delta_{n+1}$. Then, the complement of $\overline{\mathcal{C}}_f$ has* $\mathrm{Vol}(A)$*-many connected components if and only if it contains θ.*

Proof. We can assume that $\theta = 0$. To prove the *if*-part, assume that $0 \in \Theta$ for some connected component Θ of the complement of $\overline{\mathcal{C}}_f$. We wish to show that f is never colopsided in Θ, for this implies that the complement of $\overline{\mathcal{C}}_f$ has $\mathrm{Vol}(A)$-many connected components. Assume to the contrary that there exist a point $\hat{\theta} \in \Theta$ such that f is colopsided at $\hat{\theta}$. Then, $\mathrm{ord}_B(f)(\hat{\theta}) = m\pi$ for some integer m, with $|m| < \mathrm{Vol}(A)$, see (7). Let $f^\varepsilon = (f_0, \ldots, f_n, f_{n+1}e^{i\varepsilon})$. Then f^ε is colopsided at 0 for $\varepsilon \notin 2\pi\mathbb{Z}$. By continuity of roots, for $\varepsilon > 0$ sufficiently small, the points 0 and $\hat{\theta}$ are contained in the same connected component of the complement of $\overline{\mathcal{C}}(f^\varepsilon)$. Hence, by [7, pro. 3.9], they are contained in the same connected component of the complement of $\overline{\mathcal{L}}(f^\varepsilon)$. However,

$$\mathrm{ord}_B(f^\varepsilon)(0) = \mathrm{Vol}(A)(\pi - \varepsilon) \neq m(\pi - \varepsilon) = \mathrm{ord}_B(f^\varepsilon)(\hat{\theta}),$$

contradicting that $\mathrm{ord}_B(f^\varepsilon)$ is constant on connected components of the complement of $\mathcal{L}(f^\varepsilon)$.

To prove the *only if*-part, assume that there exists a connected component Θ of the complement of $\overline{\mathcal{C}}_f$ in which f is never colopsided. We wish to prove that $0 \in \Theta$. As f^ε is colopsided at 0 for $\varepsilon > 0$ sufficiently small, we find that $0 \in \overline{\Theta}$. Indeed, if this was not the case, then the complement of $\overline{\mathcal{C}}(f^\varepsilon)$ has $(\mathrm{Vol}(A) + 1)$-many connected components, a contradiction. As $0 \notin \mathcal{H}_f$, and by [7, lem. 2.3], there exists a disc D_0 around 0 such that

$$D_0 \cap (\mathbf{T}^n \setminus \overline{\mathcal{C}}_f) = D_0 \cap \Theta.$$

Furthermore, $D_0 \cap \Theta \neq \emptyset$, since $0 \in \overline{\Theta}$. Let $\theta \in D_0 \cap \Theta$. As f is a real polynomial, conjugation yields that $-\theta \in D_0 \cap \Theta$. However, $\Theta \subset \mathbb{R}^n$ is convex, implying that $0 \in \Theta$. □

Proof of Theorem 4.1. If $\mathrm{Arg}(f) \notin \mathcal{C}_\Delta$ then the image of $\mathrm{ord}_B(f)$ is of cardinality $\mathrm{Vol}(A)$, and hence $f \in U_0$. Thus, we only need to consider $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$, where we can assume that $\hat{f}(0) = \delta_k$ for all k. In particular, f is a real polynomial. By Proposition 4.5, it holds that the complement of $\overline{\mathcal{C}}_f$ has $\mathrm{Vol}(A)$-many connected components if and only if it contains 0. Keeping $f_0, \ldots, f_n$ and $\arg(f_{n+1})$ fixed, let us consider f as a function of $|f_{n+1}|$. As f is a real polynomial, it restricts to a map $f\colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}$, whose image depends nontrivially on $|f_{n+1}|$. Notice that $0 \in \overline{\mathcal{C}}_f$ if and only if $f(\mathbb{R}^n_{\geq 0})$ contains the origin. Since $\hat{f}_k(0) = \delta_k = 1$ for $k \neq n+1$, and since $\mathbf{a}_{n+1}$ is an interior point of A, the map f takes the boundary of $\mathbb{R}^n_{\geq 0}$ to $[1, \infty)$. In particular, if $0 \in f(\mathbb{R}^n_{\geq 0})$, then $0 \in f(\mathbb{R}^n_+)$. The boundary of the set of all $|f_{n+1}|$ for which $0 \in f(\mathbb{R}^n_{\geq 0})$ is the set of all values of $|f_{n+1}|$ for which $f(\mathbb{R}^n_+) = [0, \infty)$. Furthermore, $f(\mathbb{R}^n_+) = [0, \infty)$ holds if and only if there exists an $r \in \mathbb{R}^n_+$ such that $f(r) = 0$, while $f(r) \geq 0$ in a neighborhood of r, implying that r is a critical point of f. That is,

$$\Delta\big(\delta_1 |f_1|, \ldots, \delta_{n+1} |f_{n+1}|\big) = 0.$$

Since Δ is a binomial, there is exactly one such value of $|f_{n+1}|$. Finally, we note that $0 \in \overline{\mathcal{C}}_f$ if $|f_{n+1}| \to \infty$, which finishes the proof. □

4.2. Proof of Theorem 4.2

If $\mathrm{Arg}(f) \notin \mathcal{C}_\Delta$, then the image of $\mathrm{ord}_B(f)$ is of cardinality $\mathrm{Vol}(A)$ and hence $f \in U_0$. Assume that $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$, and that $\hat{f}_k(0) = \delta_k$ for all k. It holds that $0 \in \mathcal{H}_f$ since there exists two adjacent vertices $\mathbf{a}_0$ and $\mathbf{a}_1$ of A such that $\delta_0 = 1$ and $\delta_1 = -1$. Let,

$$H = \{\theta \,|\, \langle \mathbf{a}_0 - \mathbf{a}_1, \theta \rangle = 0\}$$

be the hyperplane of $\mathcal{H}_f$ containing 0. Assume that exists connected component Θ of the complement of $\overline{\mathcal{C}}_f$ in which f is nowhere colopsided. As in the proof of Proposition 4.5, we conclude that $0 \in \overline{\Theta}$, for otherwise we could construct a coamoeba with $(\mathrm{Vol}(A)+1)$-many connected components of its complement. As $H \subset \overline{\mathcal{C}}_f$, we find that Θ is contained in one of the half-spaces

$$H_{\pm} = \{\theta \,|\, \pm\langle \mathbf{a}_0 - \mathbf{a}_1, \theta\rangle > 0\},$$

say that $\Theta \subset H_+$. Let $f^\varepsilon = (f_0 e^{i\varepsilon}, f_1, \dots, f_{n+1})$, and let H^ε denote the corresponding hyperplane

$$H^\varepsilon = \{\theta \,|\, \langle \mathbf{a}_0 - \mathbf{a}_1, \theta\rangle = -\varepsilon\}.$$

For $|\varepsilon|$ sufficiently small, continuity of roots implies that there is a connected component $\Theta^\varepsilon \subset H_+^\varepsilon$ in which f^ε is never colopsided. However, by choosing the sign of ε, we can force $0 \in H_-^\varepsilon$. This implies that the coamoeba $\overline{\mathcal{C}}_{f^\varepsilon}$ has $(\mathrm{Vol}(A)+1)$-many connected components of its complement, a contradiction.

5. The maximal area of planar circuit coamoebas

In this section, we will prove the following bound.

Theorem 5.1. *Let A be a planar circuit, and let $f \in \mathbb{C}_*^A$. Then* $\mathrm{Area}(\mathcal{C}_f) \leq 2\pi^2$.

Furthermore, we provide the following classification of for which circuits the bound of Theorem 5.1 is sharp.

Theorem 5.2. *Let A be a planar circuit. Then there exist a polynomial $f \in \mathbb{C}_*^A$ such that* $\mathrm{Area}(\mathcal{C}_f) = 2\pi^2$ *if and only if A admits an equimodular triangulation.*

Example 5.3. Let $f(z) = 1 + z_1 + z_2 - rz_1z_2$ for $r \in \mathbb{R}_+$. Notice that A admits a unimodular triangulation. The shell $\mathcal{H}_f$ consist of the families $\theta_1 = k_1\pi$ and $\theta_2 = k_2\pi$ for $k_1, k_2 \in \mathbb{Z}$. Hence, the shell $\mathcal{H}_f$ divides $\mathbf{T}^2$ into four regions of equal area. Exactly two of these regions are contained in the coamoeba, which implies that $\mathrm{Area}(\mathcal{C}_f) = 2\pi^2$. See the left picture of Figure 3.

Example 5.4. Let $f(z) = 1 + zw^2 + z^2w - rzw$ for $r \in \mathbb{R}_+$. Also in this case A admits a unimodular triangulation. Notice that $\mathrm{Arg}(f) \in \mathcal{C}_\Delta$. The coamoeba of the trinomial $g(z) = 1 + zw^2 + z^2w$ has three components of its complement, of which f is colopsided in two. We have that $\mathcal{H}_f = \mathcal{H}_g$. Thus, if the complement of $\overline{\mathcal{C}}_f$ has two connected components, i.e., if $r \geq 3$, then one of the three components of the complement of $\overline{\mathcal{C}}_g$ is contained in $\overline{\mathcal{C}}_f$, which in turn implies that $\mathrm{Area}(\mathcal{C}_f) = 2\pi^2$. See the right picture of Figure 3.

Let us compare our results to the corresponding results of planar circuit amoebas. It was shown in [16, Thm. 12, p. 30] that the sharp upper bound on the number of connected components of the amoeba complement is $\#A$. In [12], a bound on the area of the amoeba was given as $\pi^2\,\mathrm{Vol}(A)$, and it was shown that maximal area was obtained for Harnack curves. For coamoebas, to roles of the

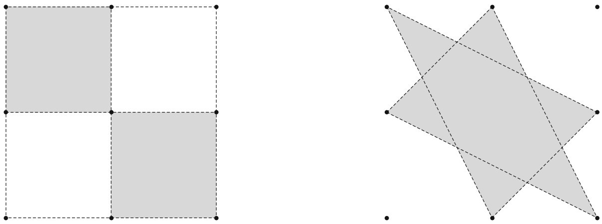


Figure 3. The coamoebas of Examples 5.3 and 5.4.

integers $\mathrm{Vol}(A)$ and $\#A$ are reversed. The upper bound on the number of connected components of the coamoeba complement is given by $\mathrm{Vol}(A)$. While, at least for codimension $m \leq 1$, the maximal area of the coameoba is $\pi^2(m+1) = \pi^2(\#A - n)$. Note also that the coamoebas of Examples 5.3 and 5.4 are both coamoebas of Harnack curves.

Consider a bivariate trinomial f, with one marked monomial. Let $\Sigma = \Sigma(f)$ denote the quadruple of polynomials obtained by flipping signs of the unmarked monomials. Furthermore, let

$$\mathcal{H}_\Sigma = \bigcup_{g \in \Sigma} \mathcal{H}_g,$$

which is a hyperplane arrangement in $\mathbb{R}^2$ (or $\mathbf{T}^2$). Let $\mathcal{P}_\Sigma$ denote the set of all intersection points of distinct hyperplanes in $\mathcal{H}_\Sigma$.

Proposition 5.5. *Let $f(z)$ be a bivariate trinomial. Then, the union*

$$\overline{\mathcal{C}}_\Sigma = \bigcup_{g \in \Sigma} \overline{\mathcal{C}}_g,$$

covers $\mathbb{R}^2$. To be specific, $\mathcal{P}_\Sigma$ is covered thrice, $\mathcal{H}_\Sigma \setminus \mathcal{P}_\Sigma$ is covered twice, and $\mathbb{R}^2 \setminus \mathcal{H}_\Sigma$ is covered once.

Proof. After applying an integer affine transformations, we reduce to the case when A consist of the vertices of the standard simplex. This case that follows from the description in [7] and [14], see also Figure 1. □

Corollary 5.6. *If $f(z)$ is a bivariate trinomial, then* $\mathrm{Area}(\overline{\mathcal{C}}_f) = \pi^2$.

Proof. The coamoebas appearing in the union $\overline{\mathcal{C}}_\Sigma$, when considered in $\mathbb{R}^2$, are merely translations of each other. Hence, they have equal area. As they cover the torus once a.e., and $\mathrm{Area}(\mathbf{T}^2) = 4\pi^2$, the result follows. □

Notice that a bivariate trinomial is not supported on a circuit, but on the vertex set of a simplex. Let $f_{\hat{k}}$ denote the image of f under the projection $\mathrm{pr}_k \colon \mathbb{C}_*^A \to \mathbb{C}_*^{A_k}$. As shown in [7] the family of trinomials $f_{\hat{k}}$, $k = 1, \ldots, 4$, contains all necessary information about the lopsided coamoeba $\mathcal{L}_f$.

Lemma 5.7. *Let A be a planar circuit, and let $f \in \mathbb{C}_*^A$. Assume that $\theta \in \mathbf{T}$ is generic in the sense that no two components of $\hat{f}(\theta)$ are antipodal, and assume* *furthermore that f is* not *colopsided at θ. Then, exactly two of the trinomials $f_{\hat{1}}, \dots, f_{\hat{4}}$ are colopsided at θ.*

Proof. Fix an arbitrary point $\mathbf{a}_1 \in A$, and let $\ell \subset \mathbb{C}$ denote the real subvector space containing $\hat{f}_1(\theta)$. As f is not colopsided at θ, both half-spaces relative ℓ contains at least one component of $\hat{f}(\theta)$. There is no restriction to assume that the upper half-space contains the two components $\hat{f}_2(\theta)$ and $\hat{f}_3(\theta)$, and that the latter is of greatest angular distance from $\hat{f}_1(\theta)$. Then, $f_{\hat{4}}$ is colopsided at θ. Furthermore, we find that $f_{\hat{2}}$ is not colopsided at θ, for if it where then so would f. As $\mathbf{a}_1 \in A_4$ and $\mathbf{a}_1 \in A_2$, there is at least one trinomial obtained from f containing $\mathbf{a}_1$ which is not colopsided at θ, and at least one which is colopsided at θ. As $\mathbf{a}_1$ was arbitrary, it follows that exactly two of the trinomials $f_{\hat{1}}, \dots, f_{\hat{4}}$ are colopsided at θ, and exactly two are not. $\square$

Proof of Theorem 5.1. By containment, it holds that $\mathrm{Area}(\mathcal{C}_f) \leq \mathrm{Area}(\mathcal{L}_f)$, and thus it suffices to calculate the area of $\mathcal{L}_f$. By [7, Prop. 3.4], we have that

$$\mathcal{L}_f = \bigcup_{k=1}^{4} \mathcal{C}_{f_{\hat{k}}}. \tag{13}$$

For a generic point $\theta \in \mathcal{L}_f$, Lemma 5.7 gives that θ (and, in fact, a small neighborhood of θ) is contained in the interior of exactly two out of the four coamoebas in the right-hand side of (13). Hence,

$$\mathrm{Area}(\mathcal{L}_f) = \frac{1}{2}\big(\mathrm{Area}(\mathcal{C}_{f_{\hat{1}}}) + \cdots + \mathrm{Area}(\mathcal{C}_{f_{\hat{4}}})\big) = 2\pi^2. \qquad \square$$

Proof of Theorem 5.2. To prove the *if* part, we will prove that A admits an equimodular triangulation only if, after applying an integer affine transformation, it is equal to the point configuration of either Example 5.3 or Example 5.4. Assume that $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$ are vertices of $\mathcal{N}_A$. After applying an integer affine transformation, we can assume that $\mathbf{a}_1 = k_1\mathbf{e}_1$, that $\mathbf{a}_2 = k_2\mathbf{e}_2$ with $k_1 \geq k_2$, and that $\mathbf{a}_3 = \mathbf{0}$. Notice that such a transformation rescales A, though it does not affect the area of the coamoeba $\mathcal{C}_f$ [7]. Let $\mathbf{a}_4 = m_1\mathbf{e}_1 + m_2\mathbf{e}_2$.

If A is a vertex circuit, then each triangulation of A consist of two simplices, which are of equal area by assumption. Comparing the areas of the subsimplices of A, we obtain the relations

$$|k_1k_2 - k_1m_2 - k_2m_1| = k_1k_2 \quad \text{and} \quad k_1m_2 = k_2m_1.$$

In m, this system has (k_1, k_2) as the only nontrivial solution, and we conclude that A is the unit square, up to integer affine transformations.

If A is a simplex circuit, then A has one triangulation with three simplices of equal area. Comparing areas, we obtain the relations

$$3k_1m_2 = 3k_2m_1 = k_1k_2.$$

Thus, $3m_1 = k_1$ and $3m_2 = k_2$, and we conclude that A is the simplex from Example 5.4, up to integer affine transformations.

To prove the *only if*-statement, consider $f \in \mathbb{C}_*^A$. Let $S = \{\mathbf{a}_1, \mathbf{a}_2\} \subset A$ be such that the line segment $[\mathbf{a}_1, \mathbf{a}_2]$ is interior to $\mathcal{N}_A$. Applying an integer affine transformation, we can assume that $[\mathbf{a}_1, \mathbf{a}_2] \subset \mathbb{R}\mathbf{e}_1$, and that $\mathbf{a}_3$ and $\mathbf{a}_4$ lies in the upper and lower half-space respectively. Then, the hyperplane arrangement $\mathcal{C}_{f_S} \subset \mathbf{T}$ consist of Length$[\mathbf{a}_1, \mathbf{a}_2]$-many lines, each parallel to the θ_2-axis. If $\mathbf{a}_3 = m_{31}\mathbf{e}_1 + m_{32}\mathbf{e}_2$ and $\mathbf{a}_4 = m_{41}\mathbf{e}_1 + m_{42}\mathbf{e}_2$, then $\hat{f}_3(\theta)$ and $\hat{f}_4(\theta)$ takes m_{32} respectively m_{42} turns around the origin when θ traverses once a line of $\mathcal{C}_{f_S}$. Notice that $\mathcal{C}_{f_S} \subset \overline{\mathcal{L}}_f$, as $\hat{f}_1(\theta)$ and $\hat{f}_2(\theta)$ are antipodal for $\theta \in \mathcal{C}_{f_S}$. That is, for such θ, $\hat{f}_S(\theta)$ is contained in a real subvector space $\ell_\theta \subset \mathbb{C}$.

Assume that f is colopsided for some $\theta \in \mathcal{C}_{f_S}$, so that in particular $\theta \notin \mathcal{C}_f$. If $\theta \in \mathcal{H}_f$, then at exactly one of the points $\hat{f}_3(\theta)$ and $\hat{f}_4(\theta)$ is contained in ℓ_θ, for otherwise f would not be colopsided at θ. By wiggling θ in $\mathcal{C}_{f_S}$ we can assume that $\theta \notin \mathcal{H}_f$. Under this assumption, we find that $\theta \notin \overline{\mathcal{C}}(f)$. Thus, there is a neighborhood N_θ which is separated from $\overline{\mathcal{C}}_f$. As $\theta \in \overline{\mathcal{L}}_f$, the intersection $N_\theta \cap \overline{\mathcal{L}}_f$ has positive area, implying that $\text{Area}(\overline{\mathcal{C}}_f) < \text{Area}(\overline{\mathcal{L}}_f)$.

Thus, if f is such that $\text{Area}(\overline{\mathcal{C}}_f) = 2\pi^2$, then f can never be colopsided in $\mathcal{C}_{f_S}$. In particular, for $\theta \in \mathcal{C}_{f_S}$ such that $\hat{f}_3(\theta) \in \ell$, it must hold $\hat{f}_4(\theta) \in \ell$, and vice versa. As there are $2m_{32}$ points of the first kind, and $2m_{42}$ points of the second kind, it holds that $m_{32} = m_{42}$. Hence, the simplices with vertices $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\}$ have equal area.

If A is a vertex circuit, this suffices in order to conclude that A admits an equimodular triangulation. If A is a simplex circuit, then we can assume that $\mathbf{a}_1$ is an interior point of $\mathcal{N}_A$. Repeating the argument for either $S = \{\mathbf{a}_1, \mathbf{a}_3\}$ or $S = \{\mathbf{a}_1, \mathbf{a}_4\}$ yields that A has a triangulation with three triangles of equal area. That is, it admits an equimodular triangulation. $\square$

6. Critical points

Let $C(f)$ denote the critical points of f, that is, the variety defined by (2). Let $\mathcal{I} = \text{Arg}(C(f))$ denote the coamoeba of $C(f)$. We will say that $\mathcal{I}$ is the set of *critical arguments* of f. In this section we will prove that, under certain assumptions on A, the set $\mathcal{I}$ is an index set of the coamoeba complement. That it is necessary to impose assumptions on A is related to the fact that an integer affine transformation acts nontrivially on the set of critical points $C(f)$.

Let A be a circuit, with the elements $\mathbf{a} \in A$ ordered so that it has a Gale dual $B = (B_1, B_2)^t$ such that $B_1 \in \mathbb{R}_+^{m_1+1}$ and that $B_2 \in \mathbb{R}_-^{m_2+1}$. That is, B_1 has only positive entries, while B_2 has only negative entries. We have that $m_1 + m_2 = n$. Let $A = (A_1, A_2)$ denote the corresponding decomposition of the matrix A. We will say that A is in *orthogonal form* if

$$A = \begin{pmatrix} 1 & 1 \\ \tilde{A}_1 & 0 \\ 0 & \tilde{A}_2 \end{pmatrix}, \tag{14}$$

where $\tilde{A}_1$ is an $m_1 \times (m_1 + 1)$-matrix and $\tilde{A}_2$ is an $m_2 \times (m_2 + 1)$-matrix. In particular, the Newton polytopes $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$ has $\mathbf{0}$ as a relatively interior point, and as their only intersection point.

With A in the form (14), we can act by integer affine transformations affecting $\tilde{A}_1$ and $\tilde{A}_2$ separately. Therefor, if A is in orthogonal form, then we can assume that

$$\tilde{A}_k = (-p_1\mathbf{e}_1, \dots, -p_{m_k}\mathbf{e}_{m_k}, \mathbf{a}_{m_k+1}), \tag{15}$$

where $p_1, \dots, p_{m_k}$ are positive integers, and hence $\mathbf{a}_{m_k+1}$ has only positive coordinates. We will say that A is in *special orthogonal form* if (15) holds. The main result of this section is the following lemma and theorem.

Lemma 6.1. *Each circuit A can be put in* (*special*) *orthogonal form by applying an integer affine transformation.*

Theorem 6.2. *Let A be a circuit in special orthogonal form. Then, for each $f \in \mathbb{C}_*^A$, the set of critical arguments is an index set of the complement of $\overline{\mathcal{C}}_f$.*

The conditions of Theorem 6.2 can be relaxed in small dimensions. When $n = 1$, it is enough to require that $\mathbf{0}$ is an interior point of $\mathcal{N}_A$. When $n = 2$, for generic f, it is enough to require that each quadrant Q fulfills that $\overline{Q} \setminus \{\mathbf{0}\}$ has nonempty intersection with A.

Proof of Lemma 6.1. Let $\mathbf{u}_1, \dots, \mathbf{u}_{m_2}$ be a basis for the left kernel $\ker(A_1)$, and let $\mathbf{v}_1, \dots, \mathbf{v}_{m_1}$ be a basis for the left kernel $\ker(A_2)$. Multiplying A from the left by

$$T = \big(\mathbf{e}_1, \mathbf{v}_1, \dots, \mathbf{v}_{m_1}, \mathbf{u}_1, \dots, \mathbf{u}_{m_2}\big)^t,$$

it takes the desired form. We need only to show that $\det(T) \neq 0$.

Notice that $\ker(A_1) \cap \ker(A_2) = 0$, since A is assumed to be of full dimension. Assume that there is a linear combination

$$\lambda_0\mathbf{e}_1 + \sum_{i=1}^{m_1} \lambda_i\mathbf{v}_i + \sum_{j=1}^{m_2} \lambda_j\mathbf{u}_j = 0.$$

Then, since B is a Gale dual of A,

$$0 = \left(\sum_{j=1}^{m_2} \lambda_j\mathbf{u}_j\right) AB = (0, \dots, 0, -\lambda_0, \dots, -\lambda_0)B = -\lambda_0 \sum_{\mathbf{a} \in A_2} \mathbf{b}_\mathbf{a} = \lambda_0 \operatorname{Vol}(A),$$

and hence $\lambda_0 = 0$. This implies that $\sum_{i=1}^{m_1} \lambda_i\mathbf{v}_i \in \ker(A_2)$, and hence $\sum_{i=1}^{m_1} \lambda_i\mathbf{v}_i = \mathbf{0}$. Thus, $\lambda_i = 0$ for all i by linear independence of the vectors $\mathbf{v}$. Then, linear independence of the vectors $\mathbf{u}_j$ imply that $\lambda_j = 0$. $\square$

Proof of Theorem 6.2. We find that

$$\begin{aligned} z_i\partial_i f(z) &= -p_i f_i z^{\mathbf{a}_i} + \langle \mathbf{a}_{m_k}, \mathbf{e}_i\rangle f_{m_1} z^{\mathbf{a}_{m_1}}, \quad i = 0, \dots, m_1 \\ z_j\partial_j f(z) &= -p_j f_i z^{\mathbf{a}_j} + \langle \mathbf{a}_{n+1}, \mathbf{e}_j\rangle f_{n+1} z^{\mathbf{a}_{n+1}}, \quad j = m_1 + 1, \dots, n. \end{aligned}$$

Hence, for each $\theta \in \mathcal{I}$, it holds that

$$\hat{f}_0(\theta) = \cdots = \hat{f}_{m_1}(\theta) \quad \text{and} \quad \hat{f}_{m_1+1}(\theta) = \cdots = \hat{f}_{n+1}(\theta). \tag{16}$$

In particular, f is colopsided at θ unless, after a rotation, $\hat{f}_k(\theta) = \delta_k$ for all k. In the latter case, we refer to Theorems 4.1 and 4.2.

To see that the points $\theta \in \mathcal{I}$ for which f is colopsided at θ are contained in distinct connected components of the complement of $\overline{\mathcal{L}}_f$, consider a line segment ℓ in $\mathbb{R}^n$ with endpoints in $\mathcal{I}$. Then, not all identities of (16) can hold identically along ℓ. Since the argument of each monomial is linear in θ, this implies that for a pair such that the identity in (16) does not hold identically along ℓ, there is an intermediate point $\theta \in \ell$ for which the corresponding monomials are antipodal, and hence $\theta \in \overline{\mathcal{L}}_f$. □

7. On systems supported on a circuit

In this section we will consider a system

$$F_1(z) = F_2(z) = 0 \tag{17}$$

of two bivariate polynomials. We will write $f(z) = 0$ for the system (17). The system is said to be generic if it has finitely many roots in $\mathbb{C}_*^2$, and it is said to be supported on a circuit A if the supports of F_1 and F_2 are contained in, but not necessarily equal to, A. That is, we allow coefficients in $\mathbb{C}$ rather than $\mathbb{C}_*$. By the Bernstein–Kushnirenko theorem, a generic system $f(z) = 0$ has at most $\mathrm{Vol}(A)$-many roots in $\mathbb{C}_*^2$. However, if f is real, then fewnomial theory states that a generic system $f(z)$ has at most three roots in $\mathbb{R}_+^2 = \mathrm{Arg}^{-1}(0)$. We will solve the complexified fewnomial problem, i.e., for $f(z)$ with complex coefficients we will bound the number of roots in each sector $\mathrm{Arg}^{-1}(\theta)$. Our intention is to offer a new approach to fewnomial theory. We will restrict to the case of simplex circuits, for the following two reasons. Firstly, it allows for a simpler exposition. Secondly, for vertex circuits our method recovers the known (sharp) bound, while for simplex circuits we obtain a sharpening of the fewnomial bound.

Theorem 7.1. *Let $f(z) = 0$ be a generic system of two bivariate polynomials supported on a planar simplex circuit $A \subset \mathbb{Z}^2$. Then, each sector $\mathrm{Arg}^{-1}(\theta)$ contains at most two solutions of $f(z) = 0$.*

7.1. Reducing $f(z)$ to a system of trinomials

A generic system $f(z)$ is, by taking appropriate linear combinations, equivalent to a system of two trinomials whose support intersect in a dupleton. That is, we can assume that $f(z)$ is in the form

$$\begin{cases} F_1(z) = & f_1 z^{\mathbf{a}_0} & + & z^{\mathbf{a}_2} & + & f_2 z^{\mathbf{a}_3} & = 0 \\ F_2(z) = & f_3 z^{\mathbf{a}_1} & + & z^{\mathbf{a}_2} & + & f_4 z^{\mathbf{a}_3} & = 0, \end{cases} \tag{18}$$

with coefficients in $\mathbb{C}_*$. We will use the notation

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \mathbf{a}_0 & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix} \quad \text{and} \quad \hat{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ A_1 & A_2 \end{pmatrix},$$

where A_k denotes the support of F_k (notice that this differs from the notation used in previous sections). Notice that we can identify a system $f(z)$ in the form (18) with its corresponding vector in $\mathbb{C}_*^{\hat{A}}$.

When reducing $f(z)$ to the form (18) by taking linear combinations, there is a choice of which monomials to eliminate in F_1 and F_2 respectively. In order for the arguments of the roots of $f(z) = 0$ to depend continuously on the coefficients, we need to be careful with which choice to make.

Lemma 7.2. *Let ℓ denote the line through $\mathbf{a}_2$ and $\mathbf{a}_3$, and let γ be a compact path in $\mathbb{C}_*^{\hat{A}}$. If ℓ intersect the interior of $\mathcal{N}_A$, then the arguments of the solutions to $f(z) = 0$ vary continuously along γ.*

Proof. It is enough to show that along a compact path γ, the set

$$\bigcup_{f\in\gamma} \mathcal{A}_f = \bigcup_{f\in\gamma} \operatorname{Log}(Z(f)) \tag{19}$$

is bounded, for it implies that for $f \in \gamma$, the roots of f are uniformly separated from the boundary of $\mathbb{C}_*^{\hat{A}}$.

We first claim that our assumptions imply that the normal fans of $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$ has no coinciding one-dimensional cones. Indeed, these fans has a coinciding one-dimensional cone if and only if the Newton polytopes $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$ has facets Γ_1 and Γ_2 with a common outward normal vector $\mathbf{n}$. As A is a circuit, it holds that $\Gamma_1 = \Gamma_2 = [\mathbf{a}_2, \mathbf{a}_3] \subset \ell$. Since the normal vector $\mathbf{n}$ is common for $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$, we find that Γ_1 (and Γ_2) is a facet of $\mathcal{N}_A$. But then ℓ contains a facet of $\mathcal{N}_A$, and hence it cannot intersect the interior of $\mathcal{N}_A$, a contradiction.

Consider a point $f \in \mathbb{C}_*^{A}$. Since the normal fans $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$ has no coinciding one-dimensional cones, the intersection of the amoebas $\mathcal{A}_{F_1}$ and $\mathcal{A}_{F_2}$ is bounded (this follows, e.g., from the fact the amoeba has finite Hausdorff distance from the Archimedean tropical variety, see [1]). Thus, the amoeba $\mathcal{A}_f$ is bounded, say that $\mathcal{A}_f \subset D(R_f)$ where $D(R_f)$ denotes the disk of radii R_f centered around the origin. By continuity of roots, $\mathcal{A}_{\tilde{f}} \subset D(R_f)$ for all $\tilde{f}$ in some neighborhood N_f of f. The compactness of γ implies our result. $\square$

In order for the assumptions of Lemma 7.2 to be fulfilled, for a simplex circuit A, we need that $\mathbf{a}_0$ and $\mathbf{a}_1$ are vertices of $\mathcal{N}(A)$, see Figure 4.

Proposition 7.3. *If f is nonreal at θ, then there is at most one zero of $f(z) = 0$ contained in the sector* $\operatorname{Arg}^{-1}(\theta)$.

Proof. If F_k is nonreal, then the fiber in $Z(F_k)$ over a point $\theta \in \mathcal{C}_{F_k}$ is a singleton. Hence, if the number of roots of $f(z) = 0$ in $\operatorname{Arg}^{-1}(\theta)$ is greater than one, then both F_1 and F_2 are real at θ. $\square$

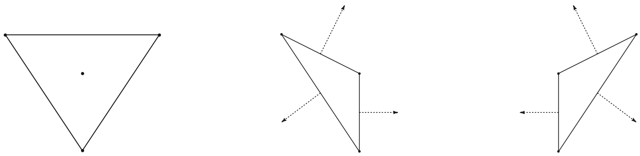


FIGURE 4. The Newton polytopes $\mathcal{N}_A$, $\mathcal{N}_{A_1}$, and $\mathcal{N}_{A_2}$.

The implication of Proposition 7.3 is that the complexified fewnomial problem reduces to the real fewnomial problem. However, our approach is dependent on allowing coefficients to be nonreal. In fact, we will consider a partially complexified problem, allowing $f_1, f_3 \in \mathbb{C}_*$ but requiring $f_2, f_4 \in \mathbb{R}_*$.

7.2. Colopsidedness

We define the colopsided coamoeba of the system $f(z)$ by

$$\mathcal{L}_f = \mathcal{L}_{F_1} \cap \mathcal{L}_{F_2} = \mathcal{C}_{F_1} \cap \mathcal{C}_{F_2},$$

where the last equality follows from [7, cor. 3.3]. That is, f is said to be colopsided at θ if either F_1 or F_2 is colopsided at θ. We will say that f is real at θ if both F_1 and F_2 are real at θ.

The lopsided coamoeba $\mathcal{L}_f$ consist of a number of polygons on $\mathbf{T}^2$, possibly degenerated to singletons. The following two lemmas will allow us to count the number of such polygons.

Lemma 7.4. *Assume that f is nonreal. Let g be a binomial constructed by choosing two monomials from* (18), *possibly alternating signs. If f_2 and f_4 are of opposite signs, then $\mathcal{C}_g \subset \mathbf{T}^2 \setminus \mathcal{L}_f$. If f_2 and f_4 are of equal signs, then $\mathcal{C}_g \subset \mathbf{T}^2 \setminus \mathcal{L}_f$* except *for $g(z) = \pm(f_1 z^{\mathbf{a}_0} - f_3 z^{\mathbf{a}_1})$.*

Proof. If, for $\theta \in \mathbf{T}^2$, two components of $\hat{F}_1(\theta)$ is contained in a real subvector space $\ell \subset \mathbb{C}$, then either F_1 is colopsided at θ or $\hat{F}_1(\theta) \subset \ell$. However, the latter implies that two components of $\hat{F}_2(\theta)$ are contained in ℓ. Repeating the argument yields that either f is real, or it is colopsided at θ.

Thus, the only binomials we need to consider is $g_\pm(z) = f_1 z^{\mathbf{a}_0} \pm f_3 z^{\mathbf{a}_1}$. For each $\theta \in \mathcal{C}_{g_+}$ the vectors $\hat{F}_1(\theta)$ and $\hat{F}_2(\theta)$ differ in sign in their first component, and hence at least one is colopsided at θ, unless f is real. For each $\theta \in \mathcal{C}_{g_-}$, the vectors $\hat{F}_1(\theta)$ and $\hat{F}_2(\theta)$ differ in signs in the the last component only if f_2 and f_4 differ in signs. If this is the case, then at least one is colopsided at θ unless f is real. $\square$

Lemma 7.5. *Let $\theta \in \mathcal{C}_{g_1} \cap \mathcal{C}_{g_2}$ for truncated binomials g_1 and g_2 of F_1 and F_2 respectively. If the Newton polytopes* (*i.e., line segments*) *of g_1 and g_2 are nonparallel, then $\theta \in \overline{\mathcal{L}}_f$.*

Proof. If F_1 and F_2 are both real at θ, then $\theta \in \mathcal{L}_f$. If F_1 is nonreal at θ, then for a sufficiently small neighborhood $N_\theta \subset \mathbb{R}^2$, it holds that

$$\mathcal{C}_{F_1} \cap N_\theta = \{\varphi \,|\, \langle \varphi, \mathbf{n}\rangle > \langle \theta, \mathbf{n}\rangle\} \cap N_\theta,$$

where $\mathbf{n}$ is a normal vector of $\mathcal{N}(g_1)$. Since connected components of the complement of $\mathcal{C}_{F_2}$ are convex, either $\mathcal{C}_{F_2}$ intersect $\mathcal{C}_{F_1}$ in N_θ, or the boundary of $\mathcal{C}_{F_2}$ is contained in the line $\ell = \{\varphi \,|\, \langle \varphi, \mathbf{n}\rangle = \langle \theta, \mathbf{n}\rangle\}$. As the boundary of $\mathcal{C}_{F_2}$ contains $\mathcal{C}_{g_2}$, it holds in the latter case that $\mathcal{C}_{g_2} \subset \ell$, which in turn implies that $\mathbf{n}$ is a normal vector of $\mathcal{N}(g_2)$, contradicting our assumptions. We conclude that $\mathcal{C}_{F_2} \cap \mathcal{C}_{F_1} \cap N_\theta \neq \emptyset$. Since this holds for any sufficiently small neighborhood N_θ, the result follows. $\square$

Example 7.6. Consider the system

$$f(z) = \begin{cases} f_1 z_1 z_2^2 + 1 + f_2 z_1 z_2 \\ f_3 z_1^2 z_2 + 1 + f_4 z_1 z_2. \end{cases}$$

We have that $\mathrm{Vol}(A) = 3$. Hence H divides $\mathbf{T}^2$ into three cells. The lopsided coamoeba $\mathcal{L}_f$, and the hyperplane arrangement H, can be seen in Figure 5. In the first two picture, the generic respectively real situation when f_2 and f_4 differs in signs. In last two pictures, the generic respectively real situation when f_2 and f_4 have equal signs. In the generic case, the lopsided coamoeba $\mathcal{L}_f$ consist of three polygons. When deforming from the generic to the real case, we observe the following behavior. Some polygons of $\mathcal{L}_f$ deform into single points – by necessity points contained in the lattice P. Some pairs of polytopes of $\mathcal{L}_f$ deforms to nonconvex polygons, typically with a single intersection point. Our proof of Theorem 7.1 is based on the observation that, when deforming from a generic to a real system, at most two polytopes of $\mathcal{L}_F$ deforms a nonconvex polygon intersecting H.

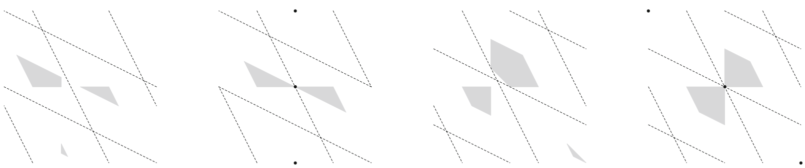


FIGURE 5. The lopsided coamoebas from Example 7.6.

7.3. Proof of Theorem 7.1

Let us consider the auxiliary binomials

$$\begin{aligned} g_1(z) &= f_1 z^{\mathbf{a}_0} - z^{\mathbf{a}_2}, & g_2(z) &= f_3 z^{\mathbf{a}_1} - z^{\mathbf{a}_2}, \\ h_1(z) &= f_1 z^{\mathbf{a}_0} + z^{\mathbf{a}_2}, \quad \text{and} & h_2(z) &= f_3 z^{\mathbf{a}_1} + z^{\mathbf{a}_2}. \end{aligned}$$

The vectors $\mathbf{a}_2 - \mathbf{a}_0$ and $\mathbf{a}_2 - \mathbf{a}_1$ span the simplex $\mathcal{N}_A$, hence the hyperplane arrangement $H = \mathcal{C}_{g_1} \cup \mathcal{C}_{g_2}$ divides $\mathbf{T}^2$ into $\mathrm{Vol}(A)$-many parallelograms with the points $P = \mathcal{C}_{h_1} \cap \mathcal{C}_{h_2}$ as their centers of mass.

If f is nonreal, then Lemma 7.4 shows that $H \subset \mathbf{T}^2 \setminus \mathcal{L}_f$, and Lemma 7.5 shows that $P \subset \overline{\mathcal{L}}_f$. By Lemma 7.2 we find that $\mathcal{L}_f$ has at most $\mathrm{Vol}(A)$-many connected components. Hence, $\mathcal{L}_f$ has at exactly one connected component in each of the cells of H, and the number of roots of $f(z) = 0$ projected by the argument map into each such component is exactly one.

Consider now the real case when f_2 and f_4 differs in signs. Then, at least one of F_1 and F_2 are colopsided at the intersection points $\mathcal{C}_{g_1} \cap \mathcal{C}_{g_2}$. Thus, if $\operatorname{Arg}^{-1}(\theta)$ contains a root of $f(z) = 0$, then a sufficiently small neighborhood N_θ intersect at most two of the cells of the hyperplane arrangement H. Hence, using Lemma 7.2 and wiggling the arguments of coefficients of f by ε, N_θ intersect at most two of the polygons of $\mathcal{L}_{f^\varepsilon}$. Hence, there can be at most two roots contained in $\operatorname{Arg}^{-1}(\theta)$.

Consider now the case when f real with f_2 and f_4 of equal signs. In this case, a point $\theta \in \mathcal{C}_{g_1} \cap \mathcal{C}_{g_2}$ can be contained in $\mathcal{L}_f$. See the left picture of Figure 6, where the hyperplane arrangement H is given in black, and the shells $\mathcal{H}_{F_1}$ and $\mathcal{H}_{F_2}$ are given in red and blue respectively, with indicated orientation. Wiggling the arguments of coefficient f_1 and/or f_3 by ε, we claim the we obtain a situation as in the right picture of Figure 6. That is, at most two polygons of $\mathcal{L}_{f^\varepsilon}$ will intersect a small neighborhood N_θ of θ. Let us prove this last claim.

Let f be generic, with f_2 and f_4 real and of equal signs. The hyperplanes $\mathcal{C}_{g_1}$ and $\mathcal{C}_{g_2}$ (locally) divides the plane into four regions. We can assume that $\mathbf{a}_2 = \mathbf{0}$. Then, $\mathcal{C}_{g_1}$ consist of all θ such that $\hat{f}_1(\theta) = 1$, and $\mathcal{C}_{g_1}$ consist of all θ such that $\hat{f}_3(\theta) = 1$. Thus, locally, the cells of H can be indexed by the signs of the imaginary parts of $\hat{f}_1(\theta)$ and $\hat{f}_3(\theta)$. Assume that $\tilde{\theta} \in \mathcal{L}_f \cap N_\theta$. Then neither F_1 nor F_2 is colopsided at $\tilde{\theta}$. Observe that $\hat{f}_2(\tilde{\theta}) = \hat{f}_4(\tilde{\theta})$, since f_2 and f_4 has equal sign. We find that

$$\operatorname{sgn}(\Im(\hat{f}_1(\tilde{\theta}))) = -\operatorname{sgn}(\Im(\hat{f}_2(\tilde{\theta}))) = -\operatorname{sgn}(\Im(\hat{f}_4(\tilde{\theta}))) = \operatorname{sgn}(\Im(\hat{f}_4(\tilde{\theta}))),$$

where the first and the last equality holds since neither F_1 nor F_2 is colopsided at $\tilde{\theta}$. This implies that polygons of $\mathcal{L}_f$ intersecting a small neighbourhood of θ are necessarily contained in the cells of H which corresponds to that the imaginary parts of $\hat{f}_1(\tilde{\theta})$ and $\hat{f}_4(\tilde{\theta})$ have equal signs. As there are two such cells, we find that there are at most two polygons of $\mathcal{L}_f$ intersecting a small neighbourhood of θ.

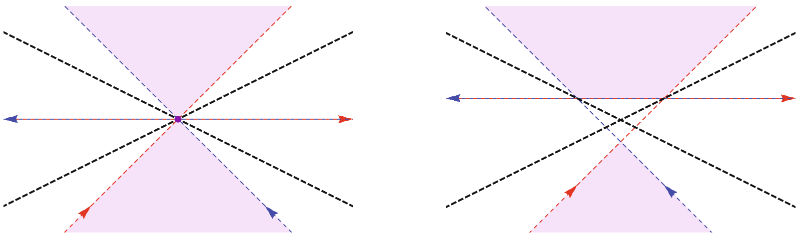


FIGURE 6. To the left: the coamoeba $\mathcal{L}_f$ close to a point of $\mathcal{C}_{g_1} \cap \mathcal{C}_{g_2}$ when f_2 and f_4 have equal in signs and f_1 and f_3 are real. To the right: the same picture after wiggling the argument of f_1 or f_3.

References

[1] Avendaño, M., Kogan, R., Nisse, M., and Rojas, J. M., Metric Estimates and Membership Complexity for Archimedean Amoebae and Tropical Hypersurfaces, *preprint*, 2013, arXiv:1307.3681.

[2] Berkesch, C., Forsgård, J., and Passare, M., Euler–Mellin integrals and A-hypergeometric functions, *Michigan Math. J.* **63** (2014), no. 1, pp. 101–123.

[3] Beukers, F., Monodromy of A-hypergeometric functions, *J. Reine Angew. Math.* **718** (2016), 183–206.

[4] Bihan, F., Polynomial systems supported on circuits and dessins d'enfants, *J. Lond. Math. Soc.* (2) **75** (2007), no. 1, 116–132.

[5] Bihan, F., Rojas, J. M., and Stella, C. E., Faster real feasibility via circuit discriminants, *ISSAC* 2009 – *Proceedings of the* 2009 *International Symposium on Symbolic and Algebraic Computation*, 39–46, *ACM, New York*, 2009.

[6] Bihan, F. and Sottile, F., Fewnomial bounds for completely mixed polynomial systems, *Adv. Geom.* **11** (2011), no. 3, 541–556.

[7] Forsgård, J. and Johansson, P., On the order map for hypersurface coamoebas, *Ark. Mat.* **53** (2015), no. 1, 79–104.

[8] Forsgård, J. and Johansson, P., Coamoebas and line arrangements in dimension two, *Math. Z.* **278** (2014), no. 1–2, 25–38.

[9] Gelfand, I. M., Kapranov, M. M., and Zelevinsky, A. V., *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser Boston Inc., Boston, MA, 1994.

[10] Jung, K. and van Straten, D., Arctic Computation of Monodromy, Report No. 40/2006, Workshop Komplexe Analysis, Oberwolfach (2006).

[11] Mikhalkin, G., Amoebas of half-dimensional varieties, *Analysis Meets Geometry: A Tribute to Mikael Passare*, 349–359, Trends Math., Springer Basel, 2017.

[12] Mikhalkin, G. and Rullgård, H., Amoebas of maximal area, *Internat. Math. Res. Notices* 2001, no. 9, 441–451.

[13] Nisse, M., Geometric and combinatorial structure of hypersurface coamoebas, *preprint*, 2009. arXiv:0906.2729.

[14] Nisse, M. and Sottile, F., The phase limit set of a variety, *Algebra Number Theory*, **7**-2 (2013), 339–352.

[15] Purbhoo, K., A Nullstellensatz for amoebas, *Duke Math. J.*, **141** (2008), 407–445.

[16] Rullgård, H., *Topics in geometry, analysis and inverse problems*, doctoral thesis, Stockholms universitet, Stockholm, 2003.

[17] Sturmfels, B., Gröbner bases and convex polytopes, University Lecture Series, 8, *American Mathematical Society, Providence*, RI, 1996.

[18] Theobald, T. and de Wolff, T., Amoebas of genus at most one, *Adv. Math.*, **239** (2013), 190–213.

[19] Theobald, T. and de Wolff, T., Norms of Roots of Trinomials, *Math. Ann.* **366** (2016), no. 1–2, 219–247.

[20] Ziegler, G. M., Lectures on polytopes, Graduate Texts in Mathematics, 152, *Springer-Verlag, New York*, 1995.

Jens Forsgård
Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
e-mail: jensf@math.tamu.edu

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 213–227

Limit of Green Functions and Ideals, the Case of Four Poles

Duong Quang Hai and Pascal J. Thomas

Abstract. We study the limits of pluricomplex Green functions with four poles tending to the origin in a hyperconvex domain, and the (related) limits of the ideals of holomorphic functions vanishing on those points. Taking subsequences, we always assume that the directions defined by pairs of points stabilize as they tend to 0. We prove that in a generic case, the limit of the Green functions is always the same, while the limits of ideals are distinct (in contrast to the three point case). We also study some exceptional cases, where only the limits of ideals are determined. In order to do this, we establish a useful result linking the length of the upper or lower limits of a family of ideals, and its convergence.

Mathematics Subject Classification (2010). 32U35, 32A27.

Keywords. Pluricomplex Green function, complex Monge–Ampère equation, ideals of holomorphic functions.

1. Introduction

The definition of multipole pluricomplex Green functions with logarithmic singularities [12], in the wake of Lempert's seminal work [6], was motivated by the nonlinearity of the complex Monge–Ampère equation, and generalizations of the Schwarz Lemma, see, e.g., Demailly [1], [12], Lelong [5].

Sometimes it is useful to study the limit case where poles tend to each other [10], an analogue of multiple zeroes for holomorphic functions, and this leads naturally to the more general notion of the Green function of an ideal of holomorphic functions:

Definition 1.1 ([8]). Let Ω be a hyperconvex bounded domain in $\mathbb{C}^n$, $\mathcal{O}(\Omega)$ the space of holomorphic functions on this domain.

This work, in a different form, is part of the first author's Ph.D. dissertation [3], defended at the Université Paul Sabatier, Toulouse, July 8th, 2013.

Let $\mathcal{I}$ be an ideal of $\mathcal{O}(\Omega)$, and ψ_j its generators. Then

$$G_{\mathcal{I}}^{\Omega}(z) := \sup \left\{ u(z) : u \in PSH_-(\Omega), u(z) \le \max_j \log |\psi_j| + O(1) \right\}.$$

Note that the condition is meaningful only near $a \in V(\mathcal{I}) := \{p \in \Omega : f(p) = 0, \forall f \in \mathcal{I}\}$. Since the domain is pseudoconvex, there are finitely many global generators $\psi_j \in \mathcal{O}(\Omega)$ such that for any $f \in \mathcal{I}$, there exists $h_j \in \mathcal{O}(\Omega)$ such that $f = \sum_j h_j \psi_j$, see, e.g., [4, Theorem 7.2.9, p. 190].

In the special case when S is a finite set in Ω and $\mathcal{I} = \mathcal{I}(S)$, the ideal of all functions vanishing on the set S (which we sometimes call point-based ideal), this reduces to a pluricomplex Green function with logarithmic singularities; we write $G_{\mathcal{I}(S)} = G_S$.

We want to study the limit of G_{S_ε} when S_ε is a set of points tending to the origin, and relate this to the limit of the ideals $\mathcal{I}(S_\varepsilon)$. It is a consequence of [9] that if convergence of those Green functions takes place in the (relatively weak) sense of $L^1_{\text{loc}}(\overline{\Omega})$, then that convergence is actually uniform on compacta of $\overline{\Omega} \setminus \{0\}$, so it will be understood that all convergence results are in this sense.

The case of 3 poles in dimension $n = 2$ was worked out in [7, Theorem 1.12, (i)]; a remaining subcase of that study was finally settled in [2].

In the present paper, we explore the case of 4 points tending to the origin in $\mathbb{C}^2$. Unlike in the three-point case, where the limit ideal was generically $\mathfrak{M}_0^2$ and the limits of the Green functions depended on the directions along which the points tended to 0, here we will see that, generically (in a sense to be made precise), $\lim G_{\mathcal{I}_\varepsilon} = G_{\lim \mathcal{I}_\varepsilon}$, and that this limit is the same, namely, $\lim G_{\mathcal{I}_\varepsilon} = 2 \max(\log |z_1|, \log |z_2|) + O(1)$ (Theorem 2.1), whereas the limit ideals very much depend on the directions of convergence to 0.

Some singular cases are studied in Theorems 2.2 and 2.3, although here we mostly compute limits of ideals, the Green functions of which cannot coincide with the limit of our Green functions because of Theorem 4.2 below. The results of [9] are used to yield some estimates of the Green functions in those cases, but the complete answer is not known.

In order to obtain those results, we establish Theorem 2.5, an auxiliary result about convergence of ideals which shortens the proofs, and should be of independent interest.

2. Statement of the results

2.1. Notations

As usual, $\mathfrak{M}_0 := \mathcal{I}(\{(0,0)\})$ stands for the maximal ideal at $(0,0)$, and $\mathfrak{M}_0^2, \mathfrak{M}_0^3 \dots$ for its successive powers. For an ideal $\mathcal{I} \subset \mathcal{O}(\Omega)$, its *length* (or co-length) is $\ell(\mathcal{I}) := \dim \mathcal{O}(\Omega) / \dim(\mathcal{I})$. For instance, $\ell(\mathfrak{M}_0^k) = \frac{1}{2}k(k+1)$.

We consider $S_\varepsilon := \{a_k^\varepsilon, 1 \le k \le 4\} \subset \Omega$, for $\varepsilon \in \mathbb{C}$, $\mathcal{I}_\varepsilon := \mathcal{I}(S_\varepsilon)$.

In general we should consider $A \subset \mathbb{C}$ such that $0 \in \bar{A} \backslash A$ and study limits along A; quite often we will use some compactness to ensure convergence and pass to a subsequence included in A. For simplicity, we will just write $\lim_{\varepsilon\to 0}$ or $\lim_\varepsilon$ instead of $\lim_{\varepsilon\to 0, \varepsilon\in A}$.

We will write several sufficient conditions about convergence of ideals and Green functions in terms of the asymptotic directions defined by pairs of poles:

$$v_{ij}^\varepsilon := [a_j^\varepsilon - a_i^\varepsilon] \in \mathbb{P}^1\mathbb{C},$$

where $[\cdot]$ denotes the class in $\mathbb{P}^1\mathbb{C}$ of an element of $\mathbb{C}^2 \setminus \{(0,0)\}$. Since $\mathbb{P}^1\mathbb{C}$ is compact, by restricting to an appropriate subsequence we assume $v_{ij} = \lim_{\varepsilon\to 0} v_{ij}^\varepsilon \in \mathbb{P}^1\mathbb{C}$, for $1 \leqslant i < j \leqslant 4$. When such convergence does not occur as $\varepsilon \to 0$ in an unrestricted fashion, one may consider the (possible) limits obtained from "convergent" subsequences, and conclude about global convergence by examining whether the partial limits coincide or not.

Let

$$\mathcal{D}^\varepsilon = \mathcal{D}(S_\varepsilon) := \{v_{ij}^\varepsilon \in \mathbb{P}^1\mathbb{C}, 1 \leqslant i < j \leqslant 4\}, \quad \mathcal{D} := \{v_{ij} \in \mathbb{P}^1\mathbb{C}, 1 \leqslant i < j \leqslant 4\}.$$

Given a subset $\tilde{S}_\varepsilon \subset S_\varepsilon$, we can define $\tilde{\mathcal{D}}_\varepsilon$ and $\tilde{\mathcal{D}}$ in a similar manner.

2.2. The generic 4-pole case

Theorem 2.1. *Let S_ε satisfy*

$$\forall \tilde{S}_\varepsilon \subset S_\varepsilon \text{ with } \#\tilde{S}_\varepsilon = 3, \text{ then } \#\tilde{\mathcal{D}} \geqslant 2. \tag{2.1}$$

and

$$\forall k \in \{1,2,3,4\}, \quad \#\big\{v_{km} \in \mathbb{P}^1\mathbb{C} : m \in \{1,2,3,4\}\backslash\{k\}\big\} \geqslant 2, \tag{2.2}$$

then there exists $\lim_\varepsilon \mathcal{I}_\varepsilon = \mathcal{I}$, *with* $\mathfrak{M}_0^3 \subset \mathcal{I} \subset \mathfrak{M}_0^2$ *and* $\ell(\mathcal{I}) = 4$*; and* $\lim_\varepsilon G_\varepsilon = G_J = 2\max(\log|z_1|, \log|z_2|) + O(1)$ *depends only on* Ω *and not on* $\mathcal{I}$.

2.3. Some singular cases

We will see how things change when we give up the second condition in Theorem 2.1.

Theorem 2.2. *Suppose that S_ε verifies condition* (2.1)*, and*

$$\exists i \in I := \{1,2,3,4\} \text{ s.t. } \#\big\{v_{ij} \in \mathbb{CP}^1 : j \in I\backslash\{i\}\big\} = 1, \tag{2.3}$$

then, after a linear change of variables, $\lim_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0 := \big\langle z_1z_2, z_2^2, z_1^3\big\rangle$, *and*

$$\liminf_\varepsilon G_\varepsilon \geq G_{\mathcal{I}_0}(z) = \max\big\{\log|z_1z_2|, 2\log|z_2|, 3\log|z_1|\big\} + O(1),$$

but there is no equality.

If the situation becomes even more singular, we can have more diverse limits for the ideals.

Theorem 2.3. *Suppose there exist a* 3 *point subset* $\tilde{S}_\varepsilon \subset S_\varepsilon$ *such that* $\#\tilde{\mathcal{D}} = 1$. *Again,* $\mathcal{I}_0 := \big\langle z_1z_2, z_2^2, z_1^3\big\rangle$. *Then*

1. *If* $\#\mathcal{D} \geqslant 3$, *then* $\lim_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0$, *after an appropriate linear change of variables.*

2. *If* $\#\mathcal{D} = 2$, *then, after passing to a subsequence and an appropriate linear change of variables,* $\lim_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0$ *or* $= \mathcal{J}_0 := \big\langle z_1z_2, z_1^2 + kz_2^2, z_1^3\big\rangle$, *for some* $k \in \mathbb{C}\setminus\{0\}$.

We suspect that the Green functions do admit a limit, but we haven't been able to determine it.

2.4. Upper and lower limits of ideals

We now formalize the notion of convergence of ideals using upper and lower limits.

Definition 2.4 ([7]).

(i) $\liminf_{A\ni\varepsilon\to 0} \mathcal{I}_\varepsilon$ is the ideal consisting of all $f \in \mathcal{O}(\Omega)$ such that $f_\varepsilon \to f$ locally uniformly on Ω, as $\varepsilon \to 0$, where $f_\varepsilon \in \mathcal{I}_\varepsilon$.

(ii) $\limsup_{A\ni\varepsilon\to 0} \mathcal{I}_\varepsilon$ is the ideal of $\mathcal{O}(\Omega)$ generated by all functions f such that $f_j \to f$ locally uniformly, as $j \to \infty$, for some sequence $\varepsilon_j \to 0$ in A and $f_j \in \mathcal{I}_{\varepsilon_j}$.

(iii) If the two limits are equal, we say that the family $\mathcal{I}_\varepsilon$ converges and write $\lim_{A\ni\varepsilon\to 0} \mathcal{I}_\varepsilon$ for the common value of the upper and lower limits.

This last notion of convergence is equivalent to convergence in the topology of the Douady space [7, Section 3]. Clearly, $\liminf_\varepsilon \mathcal{I}_\varepsilon \subset \limsup_\varepsilon \mathcal{I}_\varepsilon$ and so $\ell(\liminf_\varepsilon \mathcal{I}_\varepsilon) \geq \ell(\limsup_\varepsilon \mathcal{I}_\varepsilon)$. It also follows from [7, Lemmas 2.1 and 2.2] that $\ell(\limsup_\varepsilon \mathcal{I}_\varepsilon) \leq \limsup \ell(I_\varepsilon)$ and $\ell(\liminf_\varepsilon \mathcal{I}_\varepsilon) \geq \liminf \ell(I_\varepsilon)$.

Theorem 2.5. *Let* $\mathcal{I}_\varepsilon$ *be a family of ideals based on* N *distinct points, so that* $\ell(\mathcal{I}_\varepsilon) = N$, *for any* ε.

(i) *Let* $\mathcal{I} := \limsup_\varepsilon \mathcal{I}_\varepsilon$. *If* $\ell(\mathcal{I}) \geq N$ *(or equivalently* $= N$*), then* $\lim_\varepsilon \mathcal{I}_\varepsilon = \mathcal{I}$.

(ii) *Let* $\mathcal{I} := \liminf_\varepsilon \mathcal{I}_\varepsilon$. *If* $\ell(\mathcal{I}) \leq N$ *(or equivalently* $= N$*), then* $\lim_\varepsilon \mathcal{I}_\varepsilon = \mathcal{I}$.

3. Proof of Theorem 2.5

We will proceed by reducing everything to upper and lower limits of subspaces of a single finite-dimensional vector space.

We use multiindex notation, in particular if $\alpha, \beta \in \mathbb{N}^n$, $\alpha \leq \beta$ means $\alpha_j \leq \beta_j$, $1 \leq j \leq n$, and $\alpha < \beta$ means $\alpha_j < \beta_j$, $1 \leq j \leq n$.

Let π_j denote the projection to the jth coordinate axis. Passing to a subsequence if needed, $N_j := \#\pi_j(\{a_1^\varepsilon, \dots, a_N^\varepsilon\})$ is independent of ε. Let $\mathcal{N} := (N_1, \dots, N_n)$ and

$$P_\varepsilon := \pi_1(\{a_1^\varepsilon, \dots, a_N^\varepsilon\}) \times \cdots \times \pi_n(\{a_1^\varepsilon, \dots, a_N^\varepsilon\}) \text{ (Cartesian product)}$$

As in [7, Section 2], we now define a simpler sequence of ideals contained in each $\mathcal{I}_\varepsilon = \mathcal{I}(\{a_1^\varepsilon, \dots, a_N^\varepsilon\})$. Let $\mathcal{J}_\varepsilon := \mathcal{I}(P_\varepsilon)$. It is easy to see that $d := \ell(\mathcal{J}_\varepsilon) = \#P_\varepsilon = \prod_{j=1}^n N_j \leq N^n$, and [7, Lemma 2.3] gives

$$\lim_{\varepsilon\to 0} \mathcal{J}_\varepsilon = \mathcal{J} := \langle z_1^{N_1}, \dots, z_n^{N_n}\rangle = \left\{ f \in \mathcal{O}(\Omega) : \frac{\partial^\alpha f}{\partial z^\alpha} : \alpha < \mathcal{N} \right\}.$$

Claim. $\mathcal{O}/\mathcal{J}_\varepsilon \cong \mathbb{C}^d \cong \mathcal{O}/\mathcal{J}$.

Indeed, denote by $b_j^{i,\varepsilon}$ the elements of $\pi_j(\{a_1^\varepsilon, \dots, a_N^\varepsilon\})$. For $\alpha \le \mathcal{N}$, set $\Psi_\alpha(z) = z^\alpha$, and for $\zeta \in \mathbb{C}$, $1 \leqslant j \leqslant n$, $0 \leqslant k \leqslant N_j - 1$,

$$\varphi_{k,j}^\varepsilon(\zeta) := \prod_{i=1}^{k} (\zeta - b_j^{i,\varepsilon}).$$

Let

$$\Psi_\alpha^\varepsilon(z) := \prod_{j=1}^{n} \varphi_{\alpha_j,j}^\varepsilon(z_j).$$

Since all the $b_j^{i,\varepsilon}$ tend to 0, it is easy to see that for ε small enough (including $\varepsilon = 0$) the system $\{\Psi_\alpha^\varepsilon, \alpha < \mathcal{N}\}$ is linearly independent.

Let $[\cdot]_\varepsilon$ (resp. $[\cdot]$) denote the class of a function in $\mathcal{O}/\mathcal{J}_\varepsilon$ (resp. $\mathcal{O}/\mathcal{J}$). The natural projection from $\mathrm{Span}\{\Psi_\alpha^\varepsilon, \alpha < \mathcal{N}\}$ to $\mathcal{O}/\mathcal{J}_\varepsilon$ is injective, thus bijective, and $\{[\Psi_\alpha^\varepsilon]_\varepsilon, \alpha < \mathcal{N}\}$ is a basis of $\mathcal{O}/\mathcal{J}_\varepsilon$. Then the linear map defined by $\Phi_\varepsilon\Big(\big[\Psi_\alpha^\varepsilon\big]_\varepsilon\Big) = \big[\Psi_\alpha\big]$, for $\alpha < \mathcal{N}$, is the required isomorphism.

Lemma 3.1. *Suppose that* $\lim_\varepsilon f_\varepsilon = f$*, uniformly on compacta of* Ω*. Then, in the finite-dimensional vector space* $\mathcal{O}/\mathcal{J}$*,* $\big\{\Phi_\varepsilon\big([f_\varepsilon]_\varepsilon\big)\big\} \to [f]$ *as* $\varepsilon \to 0$.

Proof. There is a unique choice of coefficients $c_\alpha^\varepsilon(f)$ such that

$$f_\varepsilon = \sum_{\alpha < \mathcal{N}} c_\alpha^\varepsilon(f_\varepsilon)\Psi_\alpha^\varepsilon + h_\varepsilon,$$

with $h_\varepsilon \in \mathcal{J}_\varepsilon$. It will be enough to show that $c_\alpha^\varepsilon(f_\varepsilon) \to c_\alpha(f)$ as $\varepsilon \to 0$, for each α.

By rescaling, we might assume that $\overline{\mathbb{D}}^n \subset \Omega$. One can prove by induction on n (or deduce as an easy special case from the beginning of [11]) that if $|\varepsilon|$ is small enough, then

$$c_\alpha^\varepsilon(f_\varepsilon) = \frac{1}{(2i\pi)^n} \int_{(\partial\mathbb{D})^n} \frac{f_\varepsilon(z_1, \dots, z_n)}{\Psi_\alpha^\varepsilon(z)} \frac{dz_1}{z_1} \cdots \frac{dz_n}{z_n},$$

and one sees that those integrals converge towards the required limit. □

We define upper and lower limits for families of subspaces in a finite-dimensional vector space $\mathbb{C}^d$ by first choosing a norm on it. Since they are equivalent, we may as well choose a euclidean norm, and we do.

Then let L_ε be a family of subspaces of $\mathbb{C}^d$ such that $\dim L_\varepsilon = k$, for any ε. Let $K_\varepsilon := L_\varepsilon \cap \overline{B}(0;1)$. We can define the upper and lower limits of L_ε by $\liminf_\varepsilon L_\varepsilon := \mathrm{Span}\big(\liminf K_\varepsilon\big)$, and analogously $\limsup_\varepsilon L_\varepsilon := \mathrm{Span}\big(\limsup K_\varepsilon\big)$.

Here $\liminf K_\varepsilon$ and $\limsup K_\varepsilon$ are taken in the sense of the Hausdorff distance between compacta, namely if K_ε^δ stands for the δ-neighborhood of K_ε,

$$\liminf K_\varepsilon := \cap_{\delta>0} \cup_{r>0} \cap_{|\varepsilon|<r} K_\varepsilon^\delta \quad \text{and} \quad \limsup K_\varepsilon := \cap_{\delta>0} \cap_{r>0} \cup_{|\varepsilon|<r} K_\varepsilon^\delta.$$

Proposition 3.2.

1. $\limsup\limits_{\varepsilon\to 0} \Phi_\varepsilon\big(\mathcal{I}_\varepsilon/\mathcal{J}_\varepsilon\big) = (\limsup \mathcal{I}_\varepsilon)/\mathcal{J}$, 2. $\liminf\limits_{\varepsilon\to 0} \Phi_\varepsilon\big(\mathcal{I}_\varepsilon/\mathcal{J}_\varepsilon\big) = (\liminf \mathcal{I}_\varepsilon)/\mathcal{J}$.

Proof. To prove that $\limsup \mathcal{I}_\varepsilon/\mathcal{J} \subset \limsup(\Phi_\varepsilon(\mathcal{I}_\varepsilon/\mathcal{J}_\varepsilon))$, it is enough to consider elements $[f]$ where f is in a generating system of $\limsup \mathcal{I}_\varepsilon$. So there exist $(\varepsilon_j)_{j\in\mathbb{Z}_+}, \varepsilon_j \to 0$ as $j \to +\infty$ and $f_j \in \mathcal{I}_{\varepsilon_j}$ such that $f_j \to f$ uniformly on compacta of Ω. Proposition 3.1 implies that $\Phi_{\varepsilon_j}\big([f_j]_{\mathcal{J}_{\varepsilon_j}}\big) \to [f]$.

Conversely, take $g \in \mathcal{O}/\mathcal{J}$ such that there exists $(\varepsilon_j)_{j\in\mathbb{Z}_+}, \varepsilon_j \to 0$ as $j \to +\infty$ and $g_j \in \mathcal{I}_{\varepsilon_j}$ such that $\|\Phi_{\varepsilon_j}\big([g_j]_{\mathcal{J}_{\varepsilon_j}}\big) - [g]\| \to 0$ as $j \to +\infty$. We can write

$$\begin{aligned} g(z) &= \sum_{\alpha<\mathcal{N}} C_\alpha(g) z^\alpha + \sum_{j=1}^{n} z_j^{N_j} R_j(z) \text{ and} \\ \big[g_j(z)\big]_{\mathcal{J}_{\varepsilon_j}} &= \sum_{\alpha<\mathcal{N}} C_\alpha^{\varepsilon_j}(g_j)\big[\Psi_\alpha^{\varepsilon_j}(z)\big]_{\mathcal{J}_{\varepsilon_j}} \in \mathcal{I}_{\varepsilon_j}/\mathcal{J}_{\varepsilon_j}. \end{aligned}$$

The hypothesis says that $|C_\alpha^{\varepsilon_j}(g_j) - C_\alpha(g)| \to 0$ for any $\alpha < \mathcal{N}$. Set

$$f_j(z) := \sum_{\alpha<\mathcal{N}} C_\alpha^{\varepsilon_j}(g_j)\Psi_\alpha^{\varepsilon_j}(z) + \sum_{j=1}^{n}\prod_{i=1}^{N_j}\big(z_j - b_j^{i,\varepsilon_j}\big)R_j(z).$$

Then $f_j \in \mathcal{I}_{\varepsilon_j}$ and $f_j \to g$ uniformly on compacta of Ω.

Since the g's as above form a generating system for $\limsup(\Phi_\varepsilon(\mathcal{I}_\varepsilon/\mathcal{J}_\varepsilon))$, we are done.

The proof for lim inf is analogous and we omit it. □

The proof of our theorem then reduces to an elementary fact about families of finite-dimensional spaces.

Lemma 3.3. *Let (L_ε) be a family of vector subspaces of $\mathbb{C}^d$ such that $\dim L_\varepsilon = k \leq n$, for any ε.*

1. *If $\dim(\limsup\limits_{\varepsilon\to 0} L_\varepsilon) = k$, then $\liminf\limits_{\varepsilon\to 0} L_\varepsilon = \limsup\limits_{\varepsilon\to 0} L_\varepsilon$.*
2. *If $\dim(\liminf\limits_{\varepsilon\to 0} L_\varepsilon) = k$, then $\liminf\limits_{\varepsilon\to 0} L_\varepsilon = \limsup\limits_{\varepsilon\to 0} L_\varepsilon$.*

Proof. (1) Let L stand for $\limsup L_\varepsilon$. For any $\eta \in (0, \frac{1}{2})$, there exists $\varepsilon_\eta > 0$ such that $|\varepsilon| \leq \varepsilon_\eta$ implies that $L_\varepsilon \cap \overline{B}(0;1)$ is contained in an η-neighborhood of $L \cap \overline{B}(0;1)$. So the orthogonal projection of $L_\varepsilon \cap \overline{B}(0;1)$ to L must contain at least the ball $L \cap \overline{B}(0;(1-\eta^2)^{1/2})$, and any point of $L \cap \overline{B}(0;1)$ is a distance at most $\eta + 1 - (1-\eta^2)^{1/2}$ from $L_\varepsilon \cap \overline{B}(0;1)$, so $L \subset \liminf_\varepsilon L_\varepsilon$.

(2) Let $L := \liminf L_\varepsilon$. If we had $\limsup L_\varepsilon \not\subset L$, then $\limsup L_\varepsilon \supsetneq L$ and we can pick a unit vector $v \in \limsup L_\varepsilon \cap L^\perp$. We can find a sequence $\varepsilon_j \to 0$ and vectors $v_j \to v$, $v_j \in L_{\varepsilon_j}$. L_{ε_j} must also contain k vectors $e_1^{\varepsilon_j}, \dots, e_k^{\varepsilon_j}$ close to the vectors in an orthonormal basis $e_1, \dots, e_k$ of L. For j large enough, the system $e_1^{\varepsilon_j}, \dots, e_k^{\varepsilon_j}, v_j$ will have to be linearly independent, which contradicts $\dim L_{\varepsilon_j} = k$. □

4. Proofs of Theorems 2.1, 2.2 and 2.3

4.1. Previous results

Definition 4.1. A (point based) ideal is a *complete intersection ideal* if and only if it admits a set of n generators, where n is the dimension of the ambient space.

The main result of [7], Theorem 1.11, states:

Theorem 4.2. *Let* $\mathcal{I}_\varepsilon = \mathcal{I}(S_\varepsilon)$*, where* S_ε *is a set of* N *points all tending to* 0 *and assume that* $\lim_{\varepsilon\to 0}\mathcal{I}_\varepsilon = \mathcal{I}$*. Then* $(G_{\mathcal{I}_\varepsilon})$ *converges to* $G_{\mathcal{I}}$ *locally uniformly on* $\Omega\setminus\{0\}$ *if and only if* $\mathcal{I}$ *is a complete intersection ideal.*

The following was also defined in [7].

Definition 4.3. The family of ideals $(\mathcal{I}_\varepsilon)$ satisfies the *Uniform Complete Intersection Condition* if for any ε, there exists a map Ψ_0 and maps Ψ_ε from a neighborhood of $\overline{\Omega}$ to $\mathbb{C}^n$ such that Ψ_0 is proper from Ω to $\Psi_0(\Omega)$, and

1. $\{a_j^\varepsilon, 1\le j\le N\} = \Psi_\varepsilon^{-1}\{0\}$, for all ε;
2. For all $\varepsilon\neq 0$, $1\le j\le N$ and z in a neighborhood of a_j^ε,
$$\big|\log\|\Psi_\varepsilon(z)\| - \log\|z-a_j^\varepsilon\|\big| \le C(\varepsilon) < \infty;$$
3. $\lim_{\varepsilon\to 0}\Psi_\varepsilon = \Psi = (\Psi^1,\dots,\Psi^n)$, uniformly on $\overline{\Omega}$.

Notice that the first two conditions imply $\mathcal{I}_\varepsilon = \langle\Psi_\varepsilon^1,\dots,\Psi_\varepsilon^n\rangle$.
This is [7, Theorem 1.8]:

Theorem 4.4. *Let* $(\mathcal{I}_\varepsilon)$ *be a family of ideals satisfying the uniform complete intersection condition, set* $S_\varepsilon = V(\mathcal{I}_\varepsilon)$ *and* $\mathcal{I} = \langle\Psi^1,\dots,\Psi^n\rangle$*. Then*

1. $\lim\limits_{\varepsilon\to 0}\mathcal{I}_\varepsilon = \mathcal{I}$,
2. $\lim\limits_{\varepsilon\to 0}G_\varepsilon = G_{\mathcal{I}}$*, and the convergence is locally uniform on* $\Omega\setminus\{0\}$*.*

4.2. Proof of Theorem 2.1

Let $l_{ij}^\varepsilon, 1\leqslant i<j\leqslant 4$ be the (normalized) equations of the lines through $a_i^\varepsilon, a_j^\varepsilon$ and $l_{ij} := \lim\limits_{\varepsilon\to 0} l_{ij}^\varepsilon, 1\leqslant i<j\leqslant 4$. Set

$$\mathcal{L}^\varepsilon := \{f_1^\varepsilon := l_{12}^\varepsilon\cdot l_{34}^\varepsilon; f_2^\varepsilon := l_{13}^\varepsilon\cdot l_{24}^\varepsilon; f_3^\varepsilon := l_{14}^\varepsilon\cdot l_{23}^\varepsilon\}\subset\mathcal{I}(S_\varepsilon),$$

and $f_j := \lim\limits_{\varepsilon\to 0} f_j^\varepsilon$, $j=1,2,3$.

We will prove that under the hypotheses of the theorem, there exists $i\neq j\in\{1,2,3\}$ such that if $\Psi_0 := \big(f_i, f_j\big)$, then $\Psi_0^{-1}(0) = \{0\}$. (One can see that the hypotheses are necessary for this to happen [3, Remarque 4.1.2, p. 66]). Then we conclude using Theorem 4.4 with $\Psi_\varepsilon := f_i^\varepsilon f_j^\varepsilon$. Notice that since Ψ_0 is homogeneous of degree 2 and $\|\Psi_0\|$ is bounded and bounded away from 0 on the unit sphere, then $\log\|\Psi_0\| = \log\|z\|^2 + O(1)$, and the same estimate holds for $G_{\mathcal{I}}$. An application of the generalized maximum principle of Rashkovskii and Sigurdsson [8, Lemma 4.1] shows that the limit does not depend on the particular value of $\|\Psi_0\|$: there is only one maximal plurisubharmonic function with boundary values 0 on $\partial\Omega$ and a singularity equivalent to $\log \|z\|^2$.

We proceed with the proof that we can find an "independent" pair of f_i's.

Case 1: For any three point subset $\tilde{S}_\varepsilon \subset S_\varepsilon$, the set of limit directions satisfies $\#\tilde{\mathcal{D}} = 3$. So whenever $\{i,j\}$ and $\{i',j'\}$ have an element in common, l_{ij} is independent from $l_{i'j'}$ and so for any $1 \le k < k' \le 3$, f_k and $f_{k'}$ have no common factor. So $\Psi_0^{-1}(0) = \{0\}$.

Case 2: Suppose that there exists a three point subset $S'_\varepsilon \subset S_\varepsilon$ such that the set $\mathcal{D}'$ of limit directions satisfies $\#\mathcal{D}' = 2$. Without loss of generality, $S'_\varepsilon = \{a_1^\varepsilon, a_2^\varepsilon, a_3^\varepsilon\} \subset S_\varepsilon$.

Write v_{ij} for the direction in $\mathbb{P}^1$ defined by l_{ij}. With our hypothesis, we may assume $v_{23} = v_{12} \neq v_{13}$. It will be convenient to write

$$A_1 := \{v_{13}, v_{24}\} \cap \{v_{12}, v_{34}\},$$
$$A_2 := \{v_{13}, v_{24}\} \cap \{v_{14}, v_{23}\},$$
$$A_3 := \{v_{12}, v_{34}\} \cap \{v_{14}, v_{23}\}.$$

So here $A_3 \neq \emptyset$. We will show that there exists $p \in \{1,2\}$ such that $A_p = \emptyset$ (and thus the corresponding couple of function f_i will be without a common factor, and the proof concluded).

Suppose $A_1 \neq \emptyset$. Since $v_{23} = v_{12} \neq v_{13}$, by (2.2), $v_{12} \neq v_{24}$. Consequently, $v_{34} \in \{v_{13}, v_{24}\}$.

We study A_2. Since $v_{23} = v_{12} \neq v_{24}$, $v_{23} \notin \{v_{13}, v_{24}\}$. So we need to study v_{14}.

Case 2.1: $v_{34} = v_{13}$.

Then (2.1) implies that $v_{14} \neq v_{13} = v_{34}$. We will see that $v_{14} = v_{24}$ is impossible. For this, we need to take some coordinates.

Using translations, we may assume $a_1^\varepsilon = 0 \in \mathbb{D}^2$, for any ε. Choose vectors $\tilde{v}_{ij} \in \mathbb{C}^2$ such that $||\tilde{v}_{ij}|| = 1$ and $[\tilde{v}_{ij}] = v_{ij} \in \mathbb{P}^1\mathbb{C}, 1 \leqslant i < j \leqslant 4$. Since $v_{23} = v_{12} \neq v_{13}$, we can choose an invertible linear map Φ such that $[\Phi(\tilde{v}_{12})] = [1:0]$, $[\Phi(\tilde{v}_{13})] = [0:1]$. So we can study $\Phi(S_\varepsilon)$, where

$$\Phi(a_1^\varepsilon) = b_1^\varepsilon = (0,0),$$
$$\Phi(a_2^\varepsilon) = b_2^\varepsilon = (\rho_2(\varepsilon), \eta_2(\varepsilon)),$$
$$\Phi(a_3^\varepsilon) = b_3^\varepsilon = (\eta_3(\varepsilon), \rho_3(\varepsilon)),$$
$$\Phi(a_4^\varepsilon) = b_4^\varepsilon = (\alpha(\varepsilon), \beta(\varepsilon))$$

in which all coordinates tend to 0 and $\lim\limits_{\varepsilon \to 0} \eta_j(\varepsilon)/\rho_j(\varepsilon) = 0, j = 2,3$. We retain the notation $v_{ij} \in \mathbb{P}^1\mathbb{C}, 1 \leqslant i < j \leqslant 4$, and $v_{ij} := \lim_\varepsilon v_{ij}^\varepsilon$ where this last is the direction of the line through b_i^ε and b_j^ε. Let

$$\gamma(\varepsilon) := \frac{\rho_3(\varepsilon) - \eta_2(\varepsilon)}{\eta_3(\varepsilon) - \rho_2(\varepsilon)},$$

then $v_{23}^{\varepsilon} = [1 : \gamma(\varepsilon)]$. Since $v_{23} = v_{12} = [1 : 0]$, $\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0$. Thus

$$\lim_{\varepsilon \to 0} \frac{\rho_3(\varepsilon)}{\rho_2(\varepsilon)} = \lim_{\varepsilon \to 0} \frac{\gamma(\varepsilon) - \frac{\eta_2(\varepsilon)}{\rho_2(\varepsilon)}}{\gamma(\varepsilon) \cdot \frac{\eta_3(\varepsilon)}{\rho_3(\varepsilon)} - 1} = 0.$$

Assume now that $v_{14} = v_{24}$. Then $[1 : 0] = v_{12} \neq v_{14} = v_{24} \neq v_{34} = [0 : 1]$. Write $v_{14} = [1 : \ell]$, i.e., $\beta/\alpha \to \ell \neq 0, \infty$. Consider ρ_2/α. If $\|\rho_2/\alpha\| \leqslant C_2 < \infty$, as $\varepsilon \to 0$ (or even along a subsequence $\varepsilon_k \to 0$), then

$$\frac{\alpha - \eta_3}{\beta - \rho_3} = \frac{1 - \frac{\eta_3}{\rho_3} \cdot \frac{\rho_3}{\rho_2} \cdot \frac{\rho_2}{\alpha}}{\frac{\beta}{\alpha} - \frac{\rho_3}{\rho_2} \cdot \frac{\rho_2}{\alpha}} \to \ell^{-1} \neq 0, \quad \text{as} \quad \varepsilon \to 0.$$

This contradicts $\lim_{\varepsilon \to 0}[\alpha - \eta_3 : \beta - \rho_3] = v_{34} = v_{13} = [0 : 1]$. Therefore we have $\alpha/\rho_2 \to 0$, so

$$\frac{\beta - \eta_2}{\alpha - \rho_2} = \frac{\frac{\beta}{\alpha} \cdot \frac{\alpha}{\rho_2} - \frac{\eta_2}{\rho_2}}{\frac{\alpha}{\rho_2} - 1} \to 0, \text{ as } \varepsilon \to 0.$$

This contradicts $\lim_{\varepsilon \to 0}[\alpha - \rho_2 : \beta - \eta_2] = v_{24} = v_{14} \neq v_{12} = [1 : 0]$. This is the contradiction we sought.

Case 2.2: $v_{34} = v_{24}$.

In an analogous way, we will see that $A_2 = \emptyset$. We still have $v_{23} \notin \{v_{13}, v_{24}\}$. By condition (2.2), we have $v_{14} \neq v_{24} = v_{34}$. We still use the coordinates above.

Suppose that $v_{14} = v_{13} = [0 : 1]$. This implies $\alpha/\beta \to 0$. If $0 < \|\rho_2/\beta\| \leqslant C_4 < \infty$ as $\varepsilon \to 0$,

$$\frac{\alpha - \eta_3}{\beta - \rho_3} = \frac{\frac{\alpha}{\beta} - \frac{\eta_3}{\rho_3} \cdot \frac{\rho_3}{\rho_2} \cdot \frac{\rho_2}{\beta}}{1 - \frac{\rho_3}{\rho_2} \cdot \frac{\rho_2}{\beta}} \to 0, \text{ as } \varepsilon \to 0.$$

This contradicts $v_{34} = v_{24} \neq v_{23} = [0 : 1]$. Thus $\beta/\rho_2 \to 0$, therefore

$$\frac{\beta - \eta_2}{\alpha - \rho_2} = \frac{\frac{\beta}{\rho_2} - \frac{\eta_2}{\rho_2}}{\frac{\alpha}{\beta} \cdot \frac{\beta}{\rho_2} - 1} \to 0, \text{ as } \varepsilon \to 0.$$

This contradicts $v_{24} = v_{34} \neq v_{23} = [1 : 0]$. So $v_{14} \neq v_{13}$.

In a similar way, we can prove that if $A_2 \neq \emptyset$, then $A_1 = \emptyset$. $\square$

To finish the proof of Theorem 2.1, we need to prove the statements about the limit ideal. General properties of convergence show that $\ell(\mathcal{I}) = 4$ and the form of the generators show that $\mathcal{I} \subset \mathfrak{M}_0^2$. It remains to prove that $\mathcal{I} \supset \mathfrak{M}_0^3$, which is a consequence of a more general fact.

Proposition 4.5. *Suppose that all the directions in $\mathcal{D}(S_\varepsilon)$ admit a limit, and that $\#\mathcal{D} \geqslant 2$. Then $\mathfrak{M}_0^3 \subset \liminf_{\varepsilon \to 0} \mathcal{I}_\varepsilon$. Furthermore,*

$$\limsup_{\varepsilon} \mathcal{I}_\varepsilon \subset \mathfrak{M}_0^2.$$

Proof. $\mathfrak{M}_0^3$ is invariant under invertible linear maps. Since $\#\mathcal{D} \geqslant 2$, there exists $i \in \{1,2,3,4\}$ such that $v_{ik} \neq v_{ik'}$, with $k \neq k'$ and $k, k' \in \{1,2,3,4\} \setminus \{i\}$; otherwise it is easy to show that all directions are equal, in contradiction with the hypothesis.

Without loss of generality, assume $v_{12} \neq v_{13}$ and after a linear transformation, $v_{12} = [1:0], v_{13} = [0:1]$.

We reduce ourselves by translations to the case $a_1^\varepsilon = (0,0)$. Let

$$a_2^\varepsilon = (\rho_2(\varepsilon), \delta_2(\varepsilon)) \quad \text{and} \quad a_3^\varepsilon = (\delta_3(\varepsilon), \rho_3(\varepsilon)),$$

where $\delta_j(\varepsilon) = o(\rho_j(\varepsilon))$, $j = 2,3$. Let $a_4^\varepsilon = (x_4(\varepsilon), y_4(\varepsilon))$ tending to $(0,0)$. For any ε, set

$$\begin{aligned}
\psi_1^\varepsilon &:= \big[z_1 - \frac{\delta_3(\varepsilon)}{\rho_3(\varepsilon)} z_2\big]\big[z_1 - \rho_2(\varepsilon)\big]\big[z_1 - x_4(\varepsilon)\big],\\
\psi_2^\varepsilon &:= \big[z_1 - \frac{\delta_3(\varepsilon)}{\rho_3(\varepsilon)} z_2\big]\big[z_1 - \rho_2(\varepsilon)\big]\big[z_2 - y_4(\varepsilon)\big],\\
\psi_3^\varepsilon &:= \big[z_1 - \frac{\delta_3(\varepsilon)}{\rho_3(\varepsilon)} z_2\big]\big[z_2 - \frac{\delta_2(\varepsilon)}{\rho_2(\varepsilon)} z_1\big]\big[z_2 - y_4(\varepsilon)\big],\\
\psi_4^\varepsilon &:= \big[z_2 - \frac{\delta_2(\varepsilon)}{\rho_2(\varepsilon)} z_1\big]\big[z_2 - \rho_3(\varepsilon)\big]\big[z_2 - y_4(\varepsilon)\big].
\end{aligned}$$

Then $\psi_j^\varepsilon \in \mathcal{I}_\varepsilon, 1 \leqslant j \leqslant 4$, and, with uniform convergence on compacta of Ω,

$$\begin{aligned}
z_1^3 &= \lim_{\varepsilon\to 0} \psi_1^\varepsilon \in \liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon,\\
z_1^2 z_2 &= \lim_{\varepsilon\to 0} \psi_2^\varepsilon \in \liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon,\\
z_1 z_2^2 &= \lim_{\varepsilon\to 0} \psi_3^\varepsilon \in \liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon,\\
z_2^3 &= \lim_{\varepsilon\to 0} \psi_4^\varepsilon \in \liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon.
\end{aligned}$$

Thus $\mathfrak{M}_0^3 = \langle z_1^3, z_1^2 z_2, z_1 z_2^2, z_2^3 \rangle \subset \liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon$.

To get the other inclusion, we make the same normalizations (using the fact that $\mathfrak{M}_0^2$ is invariant under invertible linear transformations, too). Write $\tilde{S}_\varepsilon = \{a_1^\varepsilon, a_2^\varepsilon, a_3^\varepsilon\}$. By [7, Theorem 1.12, i], $\lim_{\varepsilon\to 0} \mathcal{I}(\tilde{S}_\varepsilon) = \mathfrak{M}_0^2$. Since $\mathcal{I}_\varepsilon \subset \mathcal{I}(\tilde{S}_\varepsilon)$,

$$\limsup_{\varepsilon\to 0} \mathcal{I}_\varepsilon \subset \limsup_{\varepsilon\to 0} \mathcal{I}(\tilde{S}_\varepsilon) = \mathfrak{M}_0^2. \qquad \square$$

4.3. Proof of Theorem 2.2

The fact that the limit inferior of the Green functions is greater than the Green function of the ideal, but not equal to it, follows from Theorem 4.2 since here $\mathcal{I}_0$ has 3 generators.

Remark. It would be desirable to have a better estimate of the limits of Green functions. Some explicit computations were carried out in [3, Section 4.3], using the methods from [9]. It concerned the family of poles given by $S_\varepsilon := \{(0;0), (\varepsilon;0), (0;\varepsilon), (\gamma\varepsilon;0)\}$, with $\gamma \neq 1$. Since the family is homogeneous in ε, in particular is given by a hyperplane section of a (singular) holomorphic curve, [9, Example 5.8] shows that the limit of the Green functions does exist.

The following estimates are obtained:

1. $\lim_{\varepsilon \to 0} G_{\mathcal{I}_\varepsilon}(z) \geqslant 2 \log \|z\| + O(1)$, for $z_2 \neq 0$;
2. $\lim_{\varepsilon \to 0} G_{\mathcal{I}_\varepsilon}(z) \geqslant \frac{5}{3} \log \|z\| + O(1)$, for $z_1 z_2^2 (z_1 + z_2)(z_1 + \gamma z_2) \neq 0$.

This is far from a complete answer, even in this case, but the computations involved are getting increasingly tedious.

We now proceed with the proof of convergence of the family of ideals.

As before, we may assume $a_1^\varepsilon = 0 \in \Omega$. Since $\#\tilde{\mathcal{D}} \geqslant 2$, for any three-point set $\tilde{S}_\varepsilon \subset S_\varepsilon$, $\#\mathcal{D} \geqslant 2$. Without loss of generality, assume $v_{12} \neq v_{13}$. By (2.3), we may assume that for $i = 2$, $v_{12} = v_{23} = v_{24}$.

Then we claim that $\#\mathcal{D} \geqslant 3$. Indeed, if we had $\#\mathcal{D} = 2$, then $\mathcal{D} = \{v_{12}, v_{13}\}$. Three cases may occur.

- If $v_{14} = v_{12}$, then $v_{12} = v_{14} = v_{24}$. This contradicts (2.1).
- If $v_{34} = v_{12}$, then $v_{23} = v_{34} = v_{24}$. This contradicts (2.1).
- If $v_{14} = v_{34} = v_{13}$, this contradicts (2.1) again.

This proves the claim.

We can chose an invertible linear map $\Phi : \mathbb{C}^2 \to \mathbb{C}^2$ such that

$$[\Phi(\tilde{v}_{12})] = [1 : 0] \text{ and } [\Phi(\tilde{v}_{13})] = [0 : 1],$$

where $\tilde{v}_{12}, \tilde{v}_{13} \in \mathbb{C}^2$ are chosen so that $\|\tilde{v}_{12}\| = \|\tilde{v}_{13}\| = 1$ and $[\tilde{v}_{12}] = v_{12}$, $[\tilde{v}_{13}] = v_{13}$. Then

$$\Phi\big(S_\varepsilon\big) = S'_\varepsilon = \{b_1^\varepsilon = (0,0), b_2^\varepsilon, b_3^\varepsilon, b_4^\varepsilon\}.$$

For this new system $v_{12} = [1 : 0] \neq v_{13} = [0 : 1]$. We can choose $l_{ij}^\varepsilon(z)$, normalized equations of the lines through the pairs of points b_i^ε and b_j^ε, $1 \leqslant i < j \leqslant 4$ such that $\lim_{\varepsilon \to 0} l_{12}^\varepsilon(z) = \lim_{\varepsilon \to 0} l_{23}^\varepsilon(z) = \lim_{\varepsilon \to 0} l_{24}^\varepsilon(z) = z_2$ and $\lim_{\varepsilon \to 0} l_{13}^\varepsilon(z) = z_1$. This implies

$$\begin{aligned} z_1 z_2 &= \lim_{\varepsilon \to 0} l_{13}^\varepsilon(z) l_{24}^\varepsilon(z) \in \liminf_{\varepsilon} \mathcal{I}_\varepsilon, \\ z_1^3 &= \lim_{\varepsilon \to 0} l_{13}^\varepsilon(z) \big[z_1 - z_1(b_2^\varepsilon)\big] \big[z_1 - z_1(b_4^\varepsilon)\big] \in \liminf_{\varepsilon} \mathcal{I}_\varepsilon. \end{aligned}$$

So $\langle z_1 z_2, z_1^3 \rangle \subset \liminf_\varepsilon \mathcal{I}_\varepsilon$.

Since $\#\mathcal{D} \geqslant 3$, there exists $(i,j) \in \{(1,3),(1,4),(3,4)\}$ such that $v_{ij}^\varepsilon \to [1 : t]$, with $t \neq 0, \infty$. So $\lim_{\varepsilon \to 0} l_{ij}^\varepsilon(z) = z_2 - t z_1 \in \liminf_\varepsilon \mathcal{I}_\varepsilon$. This implies

$$z_2^2 = \lim_{\varepsilon \to 0} \big(l_{ij}^\varepsilon(z) l_{km}^\varepsilon(z) + t l_{13}^\varepsilon(z) l_{24}^\varepsilon(z) \big) \in \liminf_{\varepsilon} \mathcal{I}_\varepsilon,$$

since $2 \in \{k, m\} := \{1,2,3,4\} \setminus \{i,j\}$. Thus $\mathcal{I}_0 := \langle z_1 z_2, z_2^2, z_1^3 \rangle \subset \liminf_\varepsilon \mathcal{I}_\varepsilon$, with $\ell\big(\mathcal{I}_0\big) = 4$. By Theorem 2.5, $\lim_{\varepsilon \to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0$. $\square$

4.4. Proof of Theorem 2.3

By the hypothesis $\#\mathcal{D} \geqslant 2$, we may assume $v_{12} \neq v_{13}$. Just as in the proof of Theorem 2.2, we perform a translation to reduce ourselves to $a_1^\varepsilon = (0,0)$, and we choose a linear map Φ so that we are reduced to $v_{12} = [1:0] \neq v_{13} = [0:1]$. We adopt the same notation $S'_\varepsilon = \{b_k^\varepsilon, 1 \leq k \leq 4\}$.

Since there is a 3 point subset $\tilde{S'}_\varepsilon \subset S'_\varepsilon$ such that $\#\tilde{\mathcal{D}}' = 1$, we may assume that $\tilde{S}'_\varepsilon = \{1,2,4\}$, so $v_{12} = v_{14} = v_{24} = [1:0]$. Again we may choose line equations so that $\lim_{\varepsilon\to 0} l_{12}^\varepsilon(z) = \lim_{\varepsilon\to 0} l_{14}^\varepsilon(z) = \lim_{\varepsilon\to 0} l_{24}^\varepsilon(z) = z_2$ and $\lim_{\varepsilon\to 0} l_{13}^\varepsilon(z) = z_1$.

The proof in case (1) can then be completed exactly as the proof of Theorem 2.2 above.

Case (2): $\#\mathcal{D} = 2$.

Either there exists $(i,j) \in \{(2,3),(3,4)\}$ such that $v_{ij}^\varepsilon \to v_{ij} = [1:0]$ or $v_{23} = v_{34} = [0:1]$.

Case (2.1): there exists $(i,j) \in \{(2,3),(3,4)\}$ such that $v_{ij}^\varepsilon \to v_{ij} = [1:0]$.

Then $\lim_{\varepsilon\to 0} l_{ij}^\varepsilon(z) = z_2$. Then, again,

$$z_2^2 = \lim_{\varepsilon\to 0} l_{ij}^\varepsilon(z) l_{km}^\varepsilon(z) \in \liminf_{\varepsilon} \mathcal{I}_\varepsilon,$$

since $2 \in \{k,m\} = \{1,2,3,4\} \setminus \{i,j\}$. Again, as in the proof of Theorem 2.2, we find that $\lim_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0$.

Case (2.2): $v_{23} = v_{34} = [0:1]$.

Thus $v_{13} = v_{23} = v_{34} = [0:1]$ and $v_{12} = v_{14} = v_{24} = [1:0]$.

As in [2], where systems of three points tending to the origin along a single direction are considered, we reparametrize $\{b_1^\varepsilon, b_2^\varepsilon, b_4^\varepsilon\}$ in such a way that $|\varepsilon| = \|b_2^\varepsilon - b_1^\varepsilon\|$ and choose a coordinate system depending on ε such that

$$b_1^\varepsilon = (0,0), b_2^\varepsilon = (\varepsilon, 0), b_4^\varepsilon = \big(\rho(\varepsilon), \delta(\varepsilon)\rho(\varepsilon)\big) \text{ where } 0 < |\rho(\varepsilon)| \leqslant \frac{1}{2}|\varepsilon|, \delta(\varepsilon) \to 0,$$

as $\varepsilon \to 0$. Denote $b_3^\varepsilon = \big(\alpha(\varepsilon), \beta(\varepsilon)\big)$. Since $v_{13}^\varepsilon = [\alpha(\varepsilon):\beta(\varepsilon)] \to [0:1]$, $\lim_{\varepsilon\to 0} \frac{\alpha(\varepsilon)}{\beta(\varepsilon)} = 0$. We will write $\rho = \rho(\varepsilon), \delta = \delta(\varepsilon), \alpha = \alpha(\varepsilon), \beta = \beta(\varepsilon)$. For δ small enough, set

$$\tilde{\delta} := \frac{\delta}{1 - \dfrac{\alpha}{\beta}\delta}, \quad \tilde{\rho} := \rho\big(1 - \frac{\alpha}{\beta}\delta\big).$$

Clearly $\tilde{\delta}, \tilde{\rho} \to 0$. Furthermore,

$$\frac{\tilde{\delta}}{\tilde{\rho} - \varepsilon} = \frac{\delta/(\rho - \varepsilon)}{\left(1 - \dfrac{\alpha}{\beta}\dfrac{\delta\rho}{\rho - \varepsilon}\right)\big(1 - \dfrac{\alpha}{\beta}\delta\big)},$$

so if $\lim\limits_{\varepsilon\to 0}\frac{\delta}{\rho-\varepsilon}=m$, then $\lim\limits_{\varepsilon\to 0}\frac{\tilde{\delta}}{\tilde{\rho}-\varepsilon}=m$. Consider the following biholomorphism (a small perturbation of the identity map):

$$\Phi_{1,\varepsilon}:\mathbb{C}^2\longrightarrow\mathbb{C}^2,\ z\mapsto\Phi_{1,\varepsilon}(z)=\left(z_1-\frac{\alpha}{\beta}z_2,z_2\right),$$

Then $\Phi_{1,\varepsilon}\big(S'_\varepsilon\big)=S_{1,\varepsilon}=\big\{(0,0),(\varepsilon,0),(\tilde{\rho},\tilde{\delta}\tilde{\rho}),(0,\beta)\big\}$. Since $v_{23}=[0:1]$ and $v_{13}=[0:1]$, $|\alpha-\varepsilon|\ll\frac{1}{2}|\beta|$ et $|\alpha|\ll\frac{1}{2}|\beta|$. So $|\varepsilon|\leqslant|\alpha-\varepsilon|+|\alpha|\ll|\beta|$.

The proof is concluded with the following result. Notice that this limit ideal in case (ii.2) is deduced from $\mathcal{I}_0$ by exchanging the coordinates z_1 and z_2, so is again equivalent to it by a linear invertible map.

Proposition 4.6. *Let* $S_\varepsilon=\{(0,0),(\varepsilon,0),(\rho,\delta\rho),(0,\beta)\}$ *tend to* $(0,0)$ *as* $\varepsilon\to 0$*, with* $\rho:=\rho(\varepsilon),\delta:=\delta(\varepsilon),\beta:=\beta(\varepsilon)$ *and* $0<|\rho|\leqslant\frac{1}{2}|\varepsilon|$, $|\varepsilon|\ll|\beta|$. *Then*

i) *If* $\lim\limits_{\varepsilon\to 0}\dfrac{\delta}{\rho-\varepsilon}=m\neq\infty$, $\lim\limits_{\varepsilon\to 0}\mathcal{I}\big(S_\varepsilon\big)=\mathcal{I}_0:=\big\langle z_1z_2,z_2^2,z_1^3\big\rangle$.

ii) *If* $\lim\limits_{\varepsilon\to 0}\dfrac{\delta}{\varepsilon}=\infty$, *we have two cases:*

1) *If* $\lim\limits_{\varepsilon\to 0}\dfrac{\rho-\varepsilon}{\delta\beta}=k\notin\{0,\infty\}$, *then* $\lim\limits_{\varepsilon\to 0}\mathcal{I}\big(S_\varepsilon\big)=\mathcal{J}_0:=\big\langle z_1z_2,z_1^2+kz_2^2,z_1^3\big\rangle$.

2) *If* $\lim\limits_{\varepsilon\to 0}\dfrac{\rho-\varepsilon}{\delta\beta}=k\in\{0,\infty\}$, *then* $\lim_\varepsilon\mathcal{I}\big(S_\varepsilon\big)=\mathcal{I}_0$, *if* $k=\infty$ *and* $\mathcal{I}_1:=\big\langle z_1z_2,z_1^2,z_2^3\big\rangle$, *if* $k=0$.

Proof. Since $|\varepsilon|\ll|\beta|$,

$$z_1z_2=\lim_{\varepsilon\to 0}\big(z_2-\rho z_1\big)\Big[z_1+\frac{\varepsilon}{\beta}z_2-\varepsilon\Big]\in\liminf_{\varepsilon}\mathcal{I}_\varepsilon,\ \text{and}$$

$$z_1^3=\lim_{\varepsilon\to 0}z_1(z_1-\rho)(z_1-\varepsilon)\in\liminf_{\varepsilon}\mathcal{I}_\varepsilon.$$

Thus $\big\langle z_1z_2,z_1^3\big\rangle\subset\liminf_\varepsilon\mathcal{I}_\varepsilon$.

Now we need to look at various cases separately.

i) Since $\lim\limits_{\varepsilon\to 0}\dfrac{\delta}{\rho-\varepsilon}=m\neq\infty$ and the polynomial

$$Q_\varepsilon(z):=\frac{\delta\varepsilon}{\rho-\varepsilon}(\delta\rho-\beta)z_1-\beta z_2-\frac{\delta}{\rho-\varepsilon}(\delta\rho-\beta)z_1^2+z_2^2\in\mathcal{I}\big(S_\varepsilon\big),$$

we obtain $z_2^2=\lim\limits_{\varepsilon\to 0}Q_\varepsilon(z)\in\liminf_\varepsilon\mathcal{I}_\varepsilon$. So $\mathcal{I}_0:=\big\langle z_1z_2,z_2^2,z_1^3\big\rangle\subset\liminf_\varepsilon\mathcal{I}_\varepsilon$. Since $\ell(\mathcal{I}_0)=4$, applying Theorem 2.5 $\lim\limits_{\varepsilon\to 0}\mathcal{I}\big(S_\varepsilon\big)=\mathcal{I}_0$.

ii) Since $0<|\rho|\leqslant\frac{1}{2}|\varepsilon|$, $\frac{|\varepsilon|}{2}\leqslant|\rho-\varepsilon|\leqslant|\varepsilon|$, so if $\lim\limits_{\varepsilon\to 0}\dfrac{\delta}{\varepsilon}=\infty$, $\lim\limits_{\varepsilon\to 0}\dfrac{\rho-\varepsilon}{\delta}=\lim\limits_{\varepsilon\to 0}\dfrac{\rho-\varepsilon}{\varepsilon}\dfrac{\varepsilon}{\delta}=0$.

We consider two subcases:

ii.1) Suppose $\lim\limits_{\varepsilon\to 0} \dfrac{\rho-\varepsilon}{\delta\beta} = k \notin \{0,\infty\}$. Consider the polynomial

$$P_\varepsilon(z) := -\varepsilon z_1 + \frac{\rho-\varepsilon}{\delta\beta}\frac{\beta}{\frac{\delta\rho}{\beta}-1} z_2 + z_1^2 - \frac{\rho-\varepsilon}{\delta\beta}\frac{1}{\frac{\delta\rho}{\beta}-1} z_2^2. \tag{4.1}$$

We can check that $P_\varepsilon(z) \in \mathcal{I}\big(S_\varepsilon\big)$. Since $|\delta\rho| \ll |\rho| \leqslant \frac{1}{2}|\varepsilon| \ll |\beta|$, we deduce $\frac{\delta\rho}{\beta} \to 0$. So

$$z_1^2 + kz_2^2 = \lim_{\varepsilon\to 0} P_\varepsilon(z) \in \liminf_\varepsilon \mathcal{I}_\varepsilon.$$

Thus $\mathcal{J}_0 := \big\langle z_1z_2, z_1^2 + kz_2^2, z_1^3\big\rangle \subset \liminf_\varepsilon \mathcal{I}_\varepsilon$. But the class $[z_1^2] = [z_1^2 + kz_2^2] - k[z_2^2] = -k[z_2^2] \in \mathcal{O}(\Omega)/\mathcal{J}_0$, thus $\mathcal{O}(\Omega)/\mathcal{J}_0 = Span\{[1],[z_1],[z_2],[z_2^2]\}$ and $\ell(\mathcal{J}_0) = 4$. Using Theorem 2.5, we conclude $\lim\limits_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{J}_0$.

ii.2) Suppose $\lim\limits_{\varepsilon\to 0} \dfrac{\rho-\varepsilon}{\delta\beta} = k \in \{0,\infty\}$. Analogously to (4.1), consider the polynomial

$$R_\varepsilon(z) := \frac{\delta\beta}{\varepsilon}\left(\frac{\delta\rho}{\beta} - 1\right)\frac{\varepsilon}{\rho-\varepsilon}\varepsilon z_1 - \beta z_2 - \frac{\delta\beta}{\varepsilon}\left(\frac{\delta\rho}{\beta} - 1\right)\frac{\varepsilon}{\rho-\varepsilon} z_1^2 + z_2^2.$$

We can check that $R_\varepsilon(z) \in \mathcal{I}\big(S_\varepsilon\big)$. If $k=\infty$, then $|\delta\beta| \ll |\rho-\varepsilon| \ll |\delta|$, and

$$\lim_{\varepsilon\to 0}\frac{\delta\beta}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{\delta\beta}{\rho-\varepsilon}\frac{\rho-\varepsilon}{\varepsilon} = 0.$$

Thus $z_2^2 = \lim\limits_{\varepsilon\to 0} P_\varepsilon(z) \in \liminf_\varepsilon \mathcal{I}_\varepsilon$. Then $\mathcal{I}_0 \subset \liminf_\varepsilon \mathcal{I}_\varepsilon$. Since $\ell(\mathcal{I}_0) = 4$, using Theorem 2.5, we conclude $\lim\limits_{\varepsilon\to 0} \mathcal{I}\big(S_\varepsilon\big) = \mathcal{I}_0$.

Finally, if $k=0$, $|\rho-\varepsilon| \ll |\delta\beta| \ll |\delta|$. From (4.1) we deduce $z_1^2 = \lim\limits_{\varepsilon\to 0} P_\varepsilon(z) \in \liminf_\varepsilon \mathcal{I}_\varepsilon$. In addition, $z_2^3 = \lim\limits_{\varepsilon\to 0} z_2(z_2-\delta\rho)(z_2-\beta) \in \liminf_\varepsilon \mathcal{I}_\varepsilon$.

Therefore $\mathcal{I}_1 := \big\langle z_1z_2, z_1^2, z_2^3\big\rangle \subset \liminf_\varepsilon \mathcal{I}_\varepsilon$. We conclude as before. □

References

[1] J.-P. Demailly, *Mesures de Monge–Ampère et mesures pluriharmoniques*, Math. Z. 194 (1987), 519–564.

[2] Duong Quang Hai, P. J. Thomas, *Limit Of Three-Point Green Functions: The Degenerate Case*, Serdica 40 (2014), 99–110.

[3] Duong Quang Hai, *Limites d'idéaux de fonctions holomorphes et de fonctions de Green pluricomplexes*, Ph. D. thesis, Université Toulouse III Paul Sabatier, July 8th, 2013, 94 pp.

[4] L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, Third Edition (revised), Mathematical Library, Vol. 7, North-Holland, Amsterdam-New York-Oxford-Tokyo, 1990.

[5] P. Lelong, *Fonction de Green pluricomplexe et lemmes de Schwarz dans les espaces de Banach*, J. Math. Pures Appl. 68 (1989), 319–347.

[6] L. Lempert, *La métrique de Kobayashi et la représentation des domaines sur la boule*, Bull. Soc. Math. France 109 (1981), no. 4, 427–474.

[7] J. I. Magnusson, A. Rashkovskii, R. Sigurdsson, P. J. Thomas, *Limits of multipole pluricomplex Green functions*, Int. J. Math. 23 (2012), no. 6, DOI: 10.1142/S0129167X12500656.

[8] A. Rashkovskii, R. Sigurdsson, *Green functions with singularities along complex spaces*, Int. J. Math. 16 (2005), no. 4, 333–355.

[9] A. Rashkovskii, P. J. Thomas, *Powers of ideals and convergence of Green functions with colliding poles*, Int. Math. Res. Not. IMRN (2014) 1253–1272.

[10] P. J. Thomas, *Green vs. Lempert functions: a minimal example*, Pacific J. Math., Vol. 257 (2012), no. 1, pp. 189–197.

[11] A. K. Tsikh, *Multidimensional Residues and Their Applications*, Translations of Mathematical Monographs, Vol. 103, American Mathematical Society, Providence, 1992.

[12] V. P. Zahariuta, *Spaces of analytic functions and maximal plurisubharmonic functions.* D. Sci. Dissertation, Rostov-on-Don, 1984.

Duong Quang Hai
Dai Hoc Su Pham Thai Nguyen
Thai Nguyen, Viet Nam
e-mail: duongquanghai@dhsptn.edu.vn

Pascal J. Thomas
Institut de Mathématiques de Toulouse, UMR 5219
Université de Toulouse & CNRS
Université Paul Sabatier
F-31062 Toulouse Cedex 9, France
e-mail: pascal.thomas@math.univ-toulouse.fr

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 229–238

Geodesics on Ellipsoids

Jens Hoppe

Dedicated to Mikael Passare and Joachim Reinhardt

Abstract. Various ways of describing geodesic motion on Ellipsoids are presented (intrinsic and constrained formulations) including Jacobi's solution, Weierstrass' solution, and level set Liouville integrability.

1. Introduction

175 years ago [1] Jacobi solved the problem of determining shortest paths on Ellipsoids.

Despite of many further contributions and articles related to the subject (see, e.g., [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]) most non-specialists would probably find it difficult to learn about this fascinating topic. The presentation[1] therefore aims at being rather elementary, and explicit.

Let us start with the trivial problem of determining geodesics in $\mathbb{R}^N$, considering the length L of paths from A to B as a functional of parametrized curves $\vec{x}(t)$ connecting $A = \vec{x}(\alpha)$ and $B = \vec{x}(\beta)$:

$$L = \int_\alpha^\beta \sqrt{\dot{\vec{x}}^{\,2}}\mathrm{d}t, \tag{1.1}$$

whose stationary points satisfy

$$\ddot{\vec{x}} - \frac{\dot{\vec{x}}}{\dot{\vec{x}}^{\,2}}\left(\dot{\vec{x}} \cdot \ddot{\vec{x}}\right) = \vec{0}. \tag{1.2}$$

Choosing the parameter t to be the arc length, i.e., $\dot{\vec{x}}^{\,2} = 1$, the reparametrization-invariant equation (1.2) reads

$$\ddot{\vec{x}} = 0, \tag{1.3}$$

[1] Based on lectures given at Boğaziçi University, ETH, Koç University, KTH, and Sogang University.

corresponding to the Lagrangian

$$\mathcal{L}_0 := \frac{1}{2}\dot{\vec{x}}^2 \tag{1.4}$$

whose integral, in contrast with (1.1), is *not* reparametrization-invariant.

Suppose now that the motion takes place on an M-dimensional hypersurface Σ, i.e., described parametrically by

$$\vec{x}\left(u^1(t),\ldots,u^M(t)\right). \tag{1.5}$$

As then $\dot{\vec{x}} = \dot{u}^a\partial_a\vec{x}$, hence $\dot{\vec{x}}^2 = \dot{u}^a\partial_a\vec{x}\cdot\partial_b\vec{x}\,\dot{u}^b =: \dot{u}^a g_{ab}\dot{u}_b$, the expression for the length becomes

$$L = \int_\alpha^\beta \sqrt{\dot{u}^a g_{ab}\dot{u}^b}\mathrm{d}t = L\left[u^a,\dot{u}^a\right], \tag{1.6}$$

where $g_{ab}\left(u^1,\ldots,u^M\right)$ could also be thought as intrinsically given, rather than being induced from $\mathbb{R}^N$ as $\partial_a\vec{x}\cdot\partial_b\vec{x}$.

Varying (1.6) gives

$$\ddot{u}^c + \gamma^c_{ab}\dot{u}^a\dot{u}^b = -\dot{u}^c\sqrt{\dot{u}^a g_{ab}\dot{u}^b}\partial_t\frac{1}{\sqrt{\dot{u}^a g_{ab}\dot{u}^b}} = -\frac{1}{2}\dot{u}^c\partial_t\ln\left(\dot{u}^a g_{ab}\dot{u}^b\right) \tag{1.7}$$

with

$$\gamma^c_{ab} := \frac{1}{2}g^{cd}\left(\partial_a g_{db} + \partial_b g_{ad} - \partial_d g_{ab}\right). \tag{1.8}$$

Again the (reparametrization-invariant) equations simplify significantly by choosing $\dot{\vec{x}}^2 = \dot{u}^a g_{ab}\dot{u}_b$ (cf. (1.6)) to be constant, i.e., the parameter t to be, up to constant rescaling, the arc length of the curve (making the r.h.s. of (1.7) vanish).

With this understanding, the coupled ODE:s

$$\ddot{u}^c + \gamma^c_{ab}\dot{u}^a\dot{u}^b = 0, \quad a,b,c = 1,\ldots,M, \tag{1.9}$$

are usually referred to as 'geodesic equations' for a Riemannian manifold $\mathcal{M}$ parametrized locally by parameters $u^a(a = 1,\ldots,M)$. In case

$$\mathcal{M} = \Sigma_M(\varphi) := \left\{\vec{x}\in\mathbb{R}^{M+1}\middle|\,\varphi\left(\vec{x}\right) = 0\right\}, \tag{1.10}$$

one could alternatively take

$$\mathcal{L} = \frac{1}{2}\dot{\vec{x}}^2 - \lambda\varphi\left(\vec{x}\right), \tag{1.11}$$

with Lagrangian equations of motion

$$\ddot{\vec{x}} = -\lambda\vec{\nabla}\varphi, \quad \varphi\left(\vec{x}(t)\right) = 0, \tag{1.12}$$

where λ can be obtained by noting that (differentiating $\varphi\left(\vec{x}(t)\right) = 0$ twice w.r.t. t)

$$\dot{\vec{x}}\cdot\vec{\nabla}\varphi\left(\vec{x}(t)\right) = 0, \quad \ddot{\vec{x}}\cdot\vec{\nabla}\varphi + \dot{x}^i\dot{x}^j\partial^2_{ij}\varphi = 0, \tag{1.13}$$

the first ensuring $\dot{\vec{x}}\cdot\ddot{\vec{x}} = 0$, the second implying

$$\lambda = -\frac{\ddot{\vec{x}}\cdot\vec{\nabla}\varphi}{\left(\nabla\varphi\right)^2} = +\frac{\dot{x}^i\dot{x}^j\partial^2_{ij}\varphi}{\left(\nabla\varphi\right)^2}, \tag{1.14}$$

so that

$$\ddot{\vec{x}} = -\frac{\dot{x}^i \dot{x}^j \partial_{ij}^2 \varphi}{(\nabla\varphi)^2} \vec{\nabla}\varphi \tag{1.15}$$

describes free motion on Σ_M (note that $\vec{\nabla}\varphi$ is normal to Σ_M so that there is no tangential acceleration, hence no tangential force).

2. Axially symmetric surfaces and Hamiltonian formulation

Before discussing how to solve (1.9), resp. (1.15), for the case of an ellipsoid, let us (Exercise I) note that for rotationally symmetric two-dimensional surfaces,

$$\vec{x}(u,v) = \begin{pmatrix} f(u)\cos v \\ f(u)\sin v \\ h(u) \end{pmatrix}, \tag{2.1}$$

(1.9) can easily be solved by quadrature, as $(1.9)_{a=2}$ (calculating g_{ab} and γ_{ab}^c from (2.1)),

$$\ddot{v} + 2\frac{f'}{f}\dot{u}\dot{v} = 0 \tag{2.2}$$

integrates to

$$\dot{v} = \frac{\text{const.}}{f^2\,(u(t))} =: \frac{l}{f^2}, \tag{2.3}$$

allowing one to eliminate v from $(1.9)_{a=1}$, resp. (simpler!)

$$\dot{u}^a g_{ab} \dot{u}^b = \left(f'^2 + h'^2\right)\dot{u}^2 + f^2\dot{v}^2 \overset{!}{=} \text{const.} =: 2E > 0. \tag{2.4}$$

Inserting (2.3) into (2.4) yields $u(t)$ by quadrature:

$$\pm \int \mathrm{d}u \sqrt{\frac{f'^2 + h'^2}{2E - \frac{l^2}{f^2}}} = t - t_0. \tag{2.5}$$

As Exercise II, note that (1.9) can be formulated in Hamiltonian form by considering

$$H = \frac{1}{2}\pi_a g^{ab}\pi_b = H\left[u^1, \ldots, u^M, \pi_1, \ldots, \pi_M\right] \tag{2.6}$$

with canonical Poisson structure, i.e.,

$$\begin{aligned} \dot{u}^a &= \frac{\delta H}{\delta \pi_a} = g^{ab}\pi_b \\ \dot{\pi}_c &= -\frac{\delta H}{\delta u^c} = -\frac{1}{2}\pi_a \partial_c g^{ab}\pi_b = \frac{1}{2}\pi_a g^{a'a}\partial_c g_{a'b'} g^{b'b}\pi_b = \frac{1}{2}\dot{u}^a\left(\partial_c g_{ab}\right)\dot{u}^b. \end{aligned} \tag{2.7}$$

3. Jacobi's solution

One way of stating Jacobi's seminal result is that for an ellipsoid, (2.6) separates in elliptic coordinates – which Jacobi originally [1838] defined (for $M = 2$) as angles φ and ψ in

$$\begin{aligned} x_1 &= \sqrt{\frac{\alpha_1}{\alpha_3-\alpha_1}}\sin\varphi\sqrt{\alpha_2\cos^2\psi+\alpha_3\sin^2\psi-\alpha_1} \\ x_2 &= \sqrt{\alpha_2}\cos\varphi\sin\psi \\ x_3 &= \sqrt{\frac{\alpha_3}{\alpha_3-\alpha_1}}\cos\psi\sqrt{\alpha_3-\alpha_1\cos^2\varphi-\alpha_2\sin^2\varphi} \end{aligned} \tag{3.1}$$

and then, for general M, as (apart from $u^0 = 0$) the zeros of

$$f(u) := \sum_{i=1}^{N}\frac{x_i^2}{\alpha_i-u} - 1 =: -\frac{\prod_{A=0}^{M}\left(u^A-u\right)}{\prod_{i=1}^{M+1}\left(\alpha_i-u\right)}; \tag{3.2}$$

that f fully factorizes into real factors, with

$$\alpha_1 < u^1 < \alpha_2 < \cdots < u^M < \alpha^{M+1=N} \tag{3.3}$$

is easily seen by noting that

$$f'(u) = +\sum_{i=1}^{N}\frac{x_i^2}{(\alpha_i-u)^2} > 0. \tag{3.4}$$

The (elliptic coordinates) u^a $(a = 1,\dots,M)$ coordinatize the M-dimensional ellipsoid

$$\mathbb{E}_M := \left\{\vec{x}\in\mathbb{R}^{M+1}\,\middle|\,\sum_{i=1}^{M+1=N}\frac{x_i^2}{\alpha_i}=1\right\}. \tag{3.5}$$

By a simple residue-argument

$$x_i^2 = \frac{\prod_A\left(\alpha_i-u^A\right)}{\prod_{j\neq i}\left(\alpha_i-\alpha_j\right)}, \tag{3.6}$$

hence

$$\begin{aligned} 4\mathrm{d}\vec{x}^{\,2} &= \sum_i x_i^2\left(\frac{2\mathrm{d}x_i}{x_i}\right)^2 = \sum_i x_i^2\left(-\sum_A\frac{\mathrm{d}u^A}{\alpha_i-u^A}\right)^2 \\ &= \sum_{i,A,B}\frac{\mathrm{d}u^A\mathrm{d}u^B}{\left(\alpha_i-u^A\right)\left(\alpha_i-u^B\right)}\frac{\prod_C\left(\alpha_i-u_C\right)}{\prod_{j\neq i}\left(\alpha_i-\alpha_j\right)} =: 4g_{AB}\mathrm{d}u^A\mathrm{d}u^B. \end{aligned} \tag{3.7}$$

Jacobi then used (four times!) that for any distinct numbers $z_1,\dots,z_{J>1}$

$$\sum_{j=1}^{J}\frac{z_j^s}{\prod_{k(\neq j)} z_j-z_k} = \begin{cases} 0 & \text{for } s=0,\dots,J-2, \\ 1 & \text{for } s=J-1, \\ \sum z_j & \text{for } s=J; \end{cases} \tag{3.8}$$

firstly (easy!) showing that the u^A are orthogonal coordinates, i.e., $g_{A\neq B} = 0$ (the factors $\alpha_i - u^A$ and $\left(\alpha_i - u^B\right)$ can then be cancelled in (3.7), leaving in the numerator a polynomial of degree $N-2$); secondly (writing, for $A = B$, each factor $\left(\alpha_i - u^{C\neq A}\right)$ as $\left(\alpha_i - u^A\right) + \left(u^A - u^C\right)$ and then having to always pick the second term, in order to avoid getting zero according to $(3.8)_{z_i=\alpha_i}$) to show that

$$g_{AA} = \frac{1}{4}\sum_i \frac{\prod_{C\neq A}\left(u^A - u^C\right)}{\left(\alpha_i - u^A\right)\prod_j'\left(\alpha_i - \alpha_j\right)}; \tag{3.9}$$

thirdly (with $J = N+1, z_i = \alpha_i, z_{N+1} = u^A$) to conclude that

$$4g_{AA} = -\frac{\prod_{C\neq A}\left(u^A - u^C\right)}{\prod_i\left(u^A - \alpha_i\right)} \overset{(A=a\neq 0)}{=} -u^a\frac{\prod'_{c(\neq a)}\left(u^a - u^c\right)}{\prod_i\left(u^a - \alpha_i\right)}. \tag{3.10}$$

Hence

$$H = -2\sum_{a=1}^{M}\pi_a\frac{q\left(u^a\right)}{\prod_{c\neq a}\left(u^a - u^c\right)}\pi_a$$

with

$$q(u) := \prod_{i=1}^{N}\frac{\left(u-\alpha_i\right)}{u} \tag{3.11}$$

describes geodesics on $\mathbb{E}_M$; the simplest non-trivial case being $N=3$, resp.

$$H = 2\frac{\pi_1^2 q\left(u^1\right)}{u^2 - u^1} - 2\frac{\pi_2^2 q\left(u^2\right)}{u^2 - u^1} \tag{3.12}$$

(note that $q\left(u^1\right) > 0$, while $q\left(u^2\right) < 0$).

The celebrated Hamilton–Jacobi method then solves the problem by first replacing the π_a by $\frac{\partial S}{\partial u^a}$ (transforming $H = E$ into a PDE) and making the separation Ansatz $S = \sum_{a=1}^{N-1} S_a\left(u^a\right)$, which indeed will produce solutions S depending on $N-1$ free constants $\beta_1, \ldots, \beta_{N-3}, \beta_{N-2} = \beta, \beta_{N-1} = E$, provided the S_a satisfy

$$\begin{aligned} 2S_a'\left(u^a\right)q\left(u^a\right) &= E\left(\beta + \beta_1 u^a + \cdots + \beta_{N-3}\left(u^a\right)^{N-3} + (-)^N\left(u^a\right)^{N-2}\right) \\ &=: T_{N-2}\left(u^a; \beta_1, \ldots, \beta_{N-3}, \beta_{N-2} = \beta, \beta_{N-1} = E\right); \end{aligned} \tag{3.13}$$

resp.

$$\pm\mathrm{d}S_a = \mathrm{d}u^a\sqrt{\frac{T_{N-2}\left(u^a\right)}{2q\left(u^a\right)}} \overset{(N=3)}{=} \sqrt{\frac{E}{2}}\sqrt{\frac{\left(\beta - u^a\right)u^a}{\left(u^a - \alpha_1\right)\left(u^a - \alpha_2\right)\left(u^a - \alpha_3\right)}}\mathrm{d}u^a, \tag{3.14}$$

hence

$$S = \sqrt{\frac{E}{2}}\sum_{a=1}^{N-1}\pm\int^{u^a}\sqrt{\frac{\frac{1}{E}T_{N-2}(u)}{q(u)}}\mathrm{d}u; \tag{3.15}$$

$\frac{\partial S}{\partial\beta} =$ const. (in accordance with action-angle coordinates) and $(N=3)$

$$u^1 = \alpha_1\cos^2\varphi + \alpha_2\sin^2\varphi, \quad u^2 = \alpha_3\sin^2\psi + \alpha_2\cos^2\psi \tag{3.16}$$

give Jacobi's celebrated solution [1] (note that his β is $\alpha_2 - \beta$ here).

4. Weierstrass' solution and conserved quantities

A simple and slightly more direct derivation (including relatively explicit formulae for the x_i as ratios of elliptic θ-functions) was presented by Weierstrass [3] (introducing conserved quantities that were discovered again 100 years later [5]). He noted that, as a consequence of the equations of motion (cf. (1.15))

$$\ddot{x}_i = -\frac{\sum_k \frac{\dot{x}_k^2}{\alpha_k}}{\sum_l \frac{x_l^2}{\alpha_l}} \frac{x_i}{\alpha_i} \tag{4.1}$$

$$\left(1 + \sum_i \frac{x_i^2}{u-\alpha_i}\right)\left(\sum_k \frac{\dot{x}_k^2}{u-\alpha_k}\right) - \left(\sum_l \frac{x_l \dot{x}_l}{u-\alpha_l}\right)^2 = \sum_i \frac{H_i}{u-\alpha_i} = \frac{W(u)}{Q(u)} \tag{4.2}$$

will be time-independent, hence defining $N-1$ constants of the motion via

$$\begin{aligned} W(u) &= cu \prod_{\alpha=1}^{N-2} (u-\delta_\alpha), \\ Q(u) &= \prod_{i=1}^{N} (u-\alpha_i). \end{aligned} \tag{4.3}$$

In accordance with (cf. (3.2))

$$P(u) := \left(1 + \sum_i \frac{x_i^2}{u-\alpha_i}\right) \prod_i (u-\alpha_i) =: u \prod_{a=1}^{N-1} (u-u^a), \tag{4.4}$$

$$\dot{P}\Big|_{u=u^a} = -u^a \dot{u}^a {\prod_c}' (u^a - u^c), \tag{4.5}$$

while (4.2), being of the form

$$\frac{P}{Q} \sum_k \frac{\dot{x}_k^2}{u-\alpha_k} - \frac{1}{4}\frac{\dot{P}^2}{Q^2} = \frac{W}{Q},$$

implying

$$\dot{P}(u^a) = \pm 2\sqrt{-QW}(u^a) =: \pm 2\sqrt{R}, \tag{4.6}$$

one deduces that

$$\mp \frac{u^a \mathrm{d}u^a}{2\sqrt{-QW}} = \frac{\mathrm{d}t}{\prod_c' (u^a - u^c)}, \tag{4.7}$$

hence (multiplying with $(u^a)^{s-1}$, and using (3.8))

$$\sum_{a=1}^{N-1} \mp \int^{u^a(t)} \frac{u^s}{\sqrt{R(u)}} \mathrm{d}u = \begin{cases} 0 & \text{for } s = 1, 2, \ldots, N-2, \\ 2(t-t_0) & \text{for } s = N-1, \end{cases} \tag{4.8}$$

with $R(u) = -cu \prod_{i=1}^{N} (u-\alpha_i) \prod_{\alpha=1}^{N-2} (u-\delta_\alpha)$ being a time-independent polynomial of degree $2N-1$.

Note that for $N = 3$ $(c > 0, u^1 - \delta_1 < 0)$ the integrability also follows from the (once observed [14] 'trivial') time-independence of

$$I = \sum_{i=1}^{N} \frac{x_i^2}{\alpha_i^2} \sum_{k=1}^{N} \frac{\dot{x}_k^2}{\alpha_k}. \tag{4.9}$$

5. Hamiltonian formulation with constraints

Among Hamiltonian treatments using the constrained embedding coordinates $x^i(t)$ rather than the intrinsic $u^a(t)$, let me first mention the one using Dirac's theory of constraints: consider

$$\begin{aligned} \varphi &:= \frac{1}{2}\left(\sum_i \frac{x_i^2}{\alpha_i} - 1\right) =: \varphi_1, \quad \pi := \sum_i \frac{x_i p_i}{\alpha_i} =: \varphi_2, \\ \{\varphi, \pi\} &= \sum_i \frac{x_i^2}{\alpha_i^2} =: J, \end{aligned} \tag{5.1}$$

leading to the Dirac bracket

$$\begin{aligned} \{f, g\}_D &:= \{f, g\} - \{f, \varphi_a\}\, \chi^{ab}\, \{\varphi_b, g\} \\ &= \{f, g\} + \{f, \varphi\}\, \frac{1}{J}\, \{\pi, g\} - \{f, \pi\}\, \frac{1}{J}\, \{\varphi, g\}, \end{aligned} \tag{5.2}$$

as the inverse of the constraint matrix

$$\big(\chi_{ab} := \{\varphi_a, \varphi_b\}\big) = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix} \text{ is } \frac{1}{J}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Exercise III (cf. [12]):

$$\begin{aligned} &\{x_i, x_j\}_D = 0, \quad \{x_i, p_j\}_D = \delta_{ij} - \frac{1}{J}\frac{x_i x_j}{\alpha_i \alpha_j}, \quad \{p_i, p_j\}_D = -\frac{L_{ij}}{\alpha_i \alpha_j J}, \\ &\qquad L_{ij} := x_i p_j - x_j p_i. \end{aligned} \tag{5.3}$$

Instead of using (5.3) to (tediously) show the Dirac–Poisson commutativity (i.e., on the constrained phase space) of the

$$F_i = p_i^2 + {\sum_j}' \frac{L_{ij}^2}{\alpha_i - \alpha_j} \tag{5.4}$$

it is much simpler to first show (Exercise IV)

$$\{F_i, F_j\} = 0 \tag{5.5}$$

and then note that due to $\{F_i, \varphi\} \approx 0$

$$\{F_i, F_j\}_D = 0 \tag{5.6}$$

trivially follows.

6. Level set Liouville integrability

Let me finish this excursion with a Hamiltonian description communicated to me by Martin Bordemann [15]: Let π be the projection operator onto the normal of $\mathbb{E}$, resp.

$$Q_{ij} = \delta_{ij} - \frac{\dfrac{x_i x_j}{\alpha_i \alpha_j}}{\displaystyle\sum_l \frac{x_l^2}{\alpha_l^2}} \tag{6.1}$$

the projection onto the tangent space of the ellipsoid. To verify that

$$H = \frac{1}{2}\langle \vec{p}, Q\vec{p}\rangle = \frac{1}{2}\langle \vec{p}, \vec{p}\rangle - \frac{1}{2}\frac{\langle \vec{p}, A\vec{x}\rangle^2}{\langle \vec{x}, A^2\vec{x}\rangle}, \quad A_{ij} := \delta_{ij}\frac{1}{\alpha_i}, \tag{6.2}$$

describes geodesic motion on $\mathbb{E}$ one can either prove that

$$\dot{\vec{x}} = Q\vec{p}, \quad \dot{\vec{p}} = -\frac{1}{2}\langle \vec{p}, \vec{\nabla} Q\vec{p}\rangle \tag{6.3}$$

implies $Q\ddot{\vec{x}} = \vec{0}$ (for this one can prove that for the general case of several constraints [15] $\varphi_1(\vec{x}) = 0, \ldots, \varphi_k(\vec{x}) = 0$ defining a submanifold, $h_{\alpha\beta} = \vec{\nabla}\varphi_\alpha \vec{\nabla}\varphi_\beta$ pos. def.,

$$\pi_{ij} := h^{\alpha\beta}\partial_i\varphi_\alpha\partial_j\varphi_\beta, \tag{6.4}$$

that $Q_{mi}(\partial_i Q_{kj})Q_{jn}$ is symmetric in $(m \leftrightarrow n)$); or (Exercise V) explicitly calculate $\ddot{\vec{x}}$ from

$$\begin{aligned} \dot{\vec{x}} &= \vec{p} - \frac{\langle \vec{p}, A\vec{x}\rangle}{\langle \vec{x}, A^2\vec{x}\rangle}A\vec{x} = \vec{p} - \gamma A\vec{x}, \\ \dot{\vec{p}} &= \frac{\langle \vec{p}, A\vec{x}\rangle}{\langle \vec{x}, A^2\vec{x}\rangle}A\vec{p} - \frac{\langle \vec{p}, A\vec{x}\rangle^2}{\langle \vec{x}, A^2\vec{x}\rangle^2}A^2\vec{x} = \gamma A\vec{p} - \gamma^2 A^2\vec{x}. \end{aligned} \tag{6.5}$$

With many terms canceling, one arrives at

$$\ddot{\vec{x}} = \left(-\frac{\langle \vec{p}, A\vec{p}\rangle}{\langle \vec{x}, A^2\vec{x}\rangle} + 2\frac{\langle \vec{p}, A\vec{x}\rangle\langle \vec{x}, A^2\vec{p}\rangle}{\langle \vec{x}, A^2\vec{x}\rangle^2} - \frac{\langle \vec{p}, A\vec{x}\rangle^2}{\langle \vec{x}, A^2\vec{x}\rangle}\langle \vec{x}, A^3\vec{x}\rangle \right) A\vec{x} = -\dot{\gamma}A\vec{x}. \tag{6.6}$$

Inserting $\dot{\vec{x}} = \vec{p} - \gamma A\vec{x}$ (cf. (6.5)) into $\langle \dot{\vec{x}}, A\dot{\vec{x}}\rangle$, (4.1) becomes (6.6).

To then show the integrability of (6.2), a canonical transformation

$$\tilde{\vec{x}} = \sqrt{A}\vec{x}, \quad p = \sqrt{A}\tilde{\vec{p}}$$

is made in [15], with

$$\begin{aligned} H(\vec{x}, \vec{p}) = \tilde{H}(\tilde{\vec{x}}, \tilde{\vec{p}}) &= \frac{1}{2}\frac{\langle \tilde{\vec{p}}, A\tilde{\vec{p}}\rangle\langle \tilde{\vec{x}}, A\tilde{\vec{x}}\rangle - \langle \tilde{\vec{p}}, A\tilde{\vec{x}}\rangle^2}{\langle \tilde{\vec{x}}, A\tilde{\vec{x}}\rangle} = \frac{1}{2}\frac{\check{H}(\tilde{\vec{x}}, \tilde{\vec{p}}) - 2E\langle \tilde{\vec{x}}, A\tilde{\vec{x}}\rangle}{\langle \tilde{\vec{x}}, A\tilde{\vec{x}}\rangle} + E \\ &= \hat{H}(\tilde{\vec{x}}, \tilde{\vec{p}}) + E, \end{aligned} \tag{6.7}$$

so that $H(\vec{x},\vec{p}) = E = \tilde{H}(\tilde{\vec{x}},\tilde{\vec{p}})$ corresponds to $\hat{H} = 0$, and then noted [15] that generally

$$H(\vec{x},\vec{p}) = \frac{G(\vec{x},\vec{p})}{Q(\vec{x},\vec{p})} \text{ on } G = 0 = H$$

for positive Q generates the same dynamics as G.

Finally,

$$G(\vec{x},\vec{p}) = \langle\vec{p}, A\vec{x}\rangle\langle\vec{x}, A\vec{x}\rangle - \langle\vec{p}, A\vec{x}\rangle^2 - 2E\langle\vec{x}, A\vec{x}\rangle = -\sum_i \frac{G_i}{\alpha_i}, \tag{6.8}$$

with Poisson commuting

$$G_i := 2Ex_i^2 + {\sum_j}' \frac{L_{ij}^2}{\alpha_i - \alpha_j} \tag{6.9}$$

is Liouville-integrable.

References

[1] C. G. J. Jacobi, *Note von der geodätischen Linie auf einem Ellipsoid und den verschiedenen Anwendungen einer merkwuerdigen analytischen Substitution*, JRAM 19, 309–313 (1839). See also his letter to the French Academy, written on December 28, 1838, as well as his "Vorlesungen über Dynamik", given at Königsberg University 1842/43.

[2] F. Joachimsthal, *Observationes de lineis brevissimis et curvis curvaturae in superficiebus secundi gradus*, JRAM 26, 155–171 (1843).

[3] K. Weierstrass, *Über die geodätischen Linien auf dem dreiachsigen Ellipsoid*, Monatsberichte der Königlichen Preuß. Akademie der Wissenschaften zu Berlin, S. 986–997 (1861).

[4] C. Neumann, *De problemate quodam mechanico, quod ad primam integralium ultraellipticorum classem revocatur*, Crelle Journal 56, 46 (1859).

[5] K. Uhlenbeck, *Equivariant harmonic maps into spheres*, Springer Lecture Notes in Mathematics 949, 146 (1982).

[6] J. Moser, *Various aspects of integrable Hamiltonian systems*, Progess in Mathematics 8, "Dynamical Systems", 233–287, Birkhäuser (1980); *Geometry of quadrics and spectral theory*, in *The Chern Symposium*, 147–188, Springer, New York-Berlin (1980).

[7] H. Knörrer, *Geodesics on the ellipsoid*, Inventiones math. 59, 119–143 (1980); *Geodesics on quadrics and a mechanical problem of C. Neumann*, J. Reine Angew. Math. 334, 69–78 (1982).

[8] T. Ratiu, *The C. Neumann problem as a completely integrable system on an adjoint orbit*, Trans. AMS. 263 no. 2 (1981).

[9] O. Babelon, O. and M. Talon, *Separation of variables for the classical and quantum Neumann model*, ArXiv: hep-th/9201035.

[10] A. P. Veselov, *Two remarks about the connection of Jacobi and Neumann integrable systems*, Math. Z. 216, 337–345 (1994).

[11] A. M. Perelomov, *A note on geodesics on ellipsoids*, ArXiv: math-ph/0203032.

[12] Chris. M. Davison, Holger R. Dullin, Alexey V. Bolsinov, *Geodesics on the Ellipsoid and Monodromy*, Journal of Geometry and Physics 57, 2437–2454 (2007).

[13] M. Dragovic and M. Radnovic, *On closed geodesics on ellipsoids*, ArXiv: math-ph/0512092.

[14] A. V. Boris and I. S. Mamaev, *Isomorphisms of geodesic flows on quadrics*, Regular and Chaotic Dynamics 14 no. 4–5 (2009).

[15] M. Bordemann, *Exercices pour le flot géodésique d'une sous-variété riemannienne de* $\mathbb{R}^N$, *Exercices pour l'intégrabilité du flot géodésique sur un ellipsoïd*, Systèmes Intégrables, (2012/13/14) UHA Mulhouse.

Jens Hoppe
Department of Mathematics
Royal Institute of Technology
SE-100 44 Stockholm, Sweden
e-mail: hoppe@math.kth.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 239–260

Welschinger Invariants Revisited

Ilia Itenberg, Viatcheslav Kharlamov and Eugenii Shustin

To the memory of Mikael Passare, remarkable mathematician and beautiful personality

Abstract. We establish the enumerativity of (original and modified) Welschinger invariants for every real divisor on any real algebraic del Pezzo surface and give an algebro-geometric proof of the invariance of that count both up to variation of the point constraints on a given surface and variation of the complex structure of the surface itself.

- Мы говорим с тобой на разных языках,
как всегда, - отозвался Воланд,
- но вещи, о которых мы говорим,
от этого не меняются.
*М. Булгаков. Мастер и Маргарита.**

Introduction

The discovery of Welschinger invariants [27, 28] has revolutionized real enumerative geometry. Since then much effort was devoted to the numerical study of Welschinger invariants, especially in the case of real del Pezzo surfaces, which allowed one to prove long time stated conjectures on existence of real solutions in corresponding enumerative problems and to observe a new, unexpected phenomena of abundance (see [2, 12, 14, 16, 17, 21]); it also led to introducing certain modified Welschinger invariants (see [16]). This development raised several natural questions: first, for which real del Pezzo surfaces the Welschinger invariants are strongly enumerative (*i.e.*, provided by a count, with weights ± 1, of real rational curves in a given divisor class, passing through a suitable number of real and complex conjugated points) and, second, to what extent such a count is invariant under deformations of the surface. The enumerative nature of the invariants in the

* "We speak different languages, as usual," responded Woland, "but this does not change the things we speak about."- *M. Bulgakov. The Master and Margarita.*

symplectic setting is the key point of [28], but it does not imply their enumerative nature in the algebro-geometric setting because of stronger genericity assumptions. The deformation invariance in the symplectic setting implies the deformation invariance in the algebro-geometric setting, but in [28] the symplectic deformation invariance is declared without proof. Therefore, our principal motivation has been to answer the question on algebro-geometric enumerativity of Welschinger invariants on real del Pezzo surfaces, and to prove the deformation invariance in the algebro-geometric setting. Our second motivation is an expectation that a good understanding of enumeration of real rational curves on real del Pezzo surfaces can help to extend the results to other types of surfaces and to curves of higher genus (such an expectation is confirmed now by [18, 24]). The algebro-geometric framework can be also helpful in the study of algorithmic and complexity aspects.

In most of the papers on the subject, the algebro-geometric enumerativity of Welschinger invariants on del Pezzo surfaces is considered as known. Indeed, it follows from enumerativity of Gromov–Witten invariants for such surfaces, and in the literature on Gromov–Witten invariants the latter enumerativity is considered as known. However, a careful analysis, see Lemma 9, has shown that there is one, and luckily only one, exception (apparently not mentioned in the literature): that is the case of del Pezzo surfaces of canonical degree 1 and $D = -K_\Sigma$; for any other pair of a real del Pezzo surface and a real divisor on it, the Welschinger invariants, original and modified, are strongly enumerative (in the above exceptional case, the number of solutions is still finite, but certain solutions may acquire some nontrivial multiplicity).

To prove the deformation invariance, we split the task into two parts. First, we fix the complex structure and vary the position of the points. Here, our strategy is close to that of the original proof of Welschinger in [28], but uses algebro-geometric tools instead of symplectic ones. In fact, already some time ago in [15] we have undertaken an attempt to give a purely algebro-geometric proof of such an invariance. However, that proof appears to be incomplete, since one type of local bifurcations in the set of counted curves was missing; it shows up for del Pezzo surfaces of canonical degree 1 and $D = -2K_\Sigma$ (see Lemma 11 (i) below, which states, in particular, that the closure of the one-dimensional family of rational curves in $|-2K_\Sigma|$ contains non-reduced curves). To the best of our knowledge, up to now this bifurcation has not been addressed in the literature, but it is unavoidable even in the symplectic setting (contrary to [28, Remark 2.12]). This step is summed up in Proposition 4, which states the invariance of the Welschinger count under the variation of points for any real divisor on each real del Pezzo surface.

The crucial point of the next step is the invariance under crossing the walls that correspond to, so-called, uninodal del Pezzo surfaces. Here, our proof is based on a real version of the Abramovich–Bertran–Vakil formula (note that adapting the formula to the symplectic setting one can prove the symplectic deformation invariance following the same lines). In addition, as in the study of the enumerativity, there appears a case not to miss and to investigate separately, here this is the case of del Pezzo surfaces of canonical degree 1 and $D = -K_\Sigma$.

The paper is organized as follows. In Section 1 we recall a few basic facts concerning del Pezzo surfaces and their deformations, introduce Welschinger invariants in their modified version and formulate the main results. Section 2 develops technical tools needed for the proof of the main results. There, we study moduli spaces of stable maps of pointed genus zero curves to del Pezzo surfaces and uninodal del Pezzo surfaces, describe generic elements of these moduli spaces and generic elements of the codimension one strata. We show also that Welschinger numbers extend by continuity from the case of immersions to the case of birational stable maps with arbitrary singularities. Section 3 is devoted to the proof of the main results.

1. Definitions and main statements

1.1. Surfaces under consideration

Over $\mathbb{C}$, a del Pezzo surface is either $(\mathbb{P}^1)^2$ or $\mathbb{P}^2$ blown up at $0 \le k \le 8$ points. Conversely, blowing up $0 \le k \le 8$ points of $\mathbb{P}^2$ yields a del Pezzo surface if and only if no 3 points lie on a straight line, no 6 lie on a conic, and no 8 points lie on a rational cubic having a singularity at one of these 8 points.

Del Pezzo surfaces of degree $d = K^2 = 9 - k \ge 5$ have no moduli. If $d = 9$ or $7 \ge d \ge 5$, then there is only one, up to isomorphism, del Pezzo surface of degree d and it can be seen as a blown up $\mathbb{P}^2$. If $d = 8$, then there are 2 isomorphism classes: $(\mathbb{P}^1)^2$ and $\mathbb{P}^2$ blown up at a point. The latter two surfaces are not deformation equivalent. For $4 \ge d \ge 1$ the moduli space of del Pezzo surfaces of degree $d = 9 - k$ is an irreducible $(2k - 8)$-dimensional variety.

All del Pezzo surfaces of given degree $d \ne 8$ are deformation equivalent to each other, and, for our purpose, it will be more convenient to use, instead of the moduli spaces, the deformation spaces, that is, to fix in each deformation class one of the del Pezzo surfaces (say, a blow up of $\mathbb{P}^2$ at a certain generic collection of points) and consider the Kodaira–Spencer–Kuranishi space, *i.e.*, the space of all complex structures on the underlying smooth 4-manifold factorized by the action of diffeomorphisms isotopic to identity. Naturally, we awake this space only when $d \le 4$. We denote it by $\mathcal{D}_d$. Del Pezzo surfaces of degree d form in $\mathcal{D}_d$ an open dense subset, which we denote by $\mathcal{D}_d^{DP}$.

The problem of deformations of complex structures on rational surfaces is not obstructed, since $H^2(X, \mathcal{T}_X) = 0$ for any smooth rational surface X (here and further on, we denote by $\mathcal{T}_X$ the tangent sheaf). In addition, for degree $d \le 4$ del Pezzo surfaces as well as for any generic smooth rational surface X with $K_X^2 \le 4$, we have $H^0(X, \mathcal{T}_X) = 0$, so that at such points the Kodaira–Spencer–Kuranishi space is smooth (but not necessarily Hausdorff).

In fact, the only properties of this space which we use further on are the following. We call a surface $\Sigma \in \mathcal{D}_d$ *uninodal del Pezzo* if it contains a smooth rational (-2)-curve E_Σ, and $-K_\Sigma C > 0$ for each irreducible curve $C \ne E_\Sigma$ (in particular, $C^2 \geq -1$). For $d \leq 4$, denote by $\mathcal{D}_d(A_1) \subset \mathcal{D}_d$ the subspace formed by uninodal del Pezzo surfaces.

Proposition 1. *All but finite number of surfaces in a generic one-parameter Kodaira–Spencer family of rational surfaces with $1 \leq K_\Sigma^2 \leq 4$ are unnodal (i.e., del Pezzo), while the exceptional members of the family are uninodal del Pezzo.*

Proof. Let us denote by $\mathcal{T}_{X\|D}$ the subsheaf of the sheaf $\mathcal{T}_X$ generated by vectors fields tangent to D, and by $\mathcal{N}'_{D/X}$ their quotient, so that we obtain the following short exact sequence of sheafs:

$$0 \to \mathcal{T}_{X\|D} \to \mathcal{T}_X \to \mathcal{N}'_{D/X} \to 0.$$

According to the well-known theory of deformations of pairs (see [22, Section 3.4.4]), and due to the long exact cohomology sequence associated to the above short sequence, it is sufficient to show that $h^1(\mathcal{N}'_{D/X}) \geq 2$ if D is either a rational irreducible curve with $D^2 \leq -3$ or $D = D_1 \cup D_2$ where $D_i^2 \leq -2$. In the first case, it follows from Serre–Riemann–Roch duality. In the second case, from the exactness of the fragment $H^0(\mathcal{N}_{D_2/X}) \to H^1(\mathcal{N}_{D_1/X}) \to H^1(\mathcal{N}'_{D/X}) \to H^1(\mathcal{N}_{D_2/X})$ of the long cohomology sequence associated with the exact sequence of sheaves $0 \to \mathcal{N}_{D_1/X} \to \mathcal{N}'_{D/X} \to \mathcal{N}_{D_2/X} \to 0$.[†] □

By a *real algebraic surface* we understand a pair (Y, c), where Y is a complex algebraic surface and $c : Y \to Y$ is an antiholomorphic involution. The classification of minimal real rational surfaces and the classification of real del Pezzo surfaces are well known: they are summarized in the two propositions below, respectively (see, e.g., [7, Theorems 6.11.11 and 17.3]).

Proposition 2. *Each minimal real rational surface Y is one of the following:*

(1) *$\mathbb{P}^2$ with its standard real structure ($d = 9$), the real part $\mathbb{R}Y$ of Y is homeomorphic to $\mathbb{RP}^2$.*
(2) *$\mathbb{P}^1 \times \mathbb{P}^1$ with one of its four nonequivalent real structures ($d = 8$): $\mathbb{R}Y = (S^1)^2$, $\mathbb{R}Y = S^2$, and two structures with $\mathbb{R}Y = \emptyset$;*
(3) *rational geometrically ruled surfaces $\mathcal{F}_a$, $a \geq 2$, with $\mathbb{R}Y = \#_2\mathbb{RP}^2$ and the standard real structure, if a is odd, and with $\mathbb{R}Y = (S^1)^2$ or $\emptyset$ and one of the two respective nonequivalent structures, if a is even ($d = 8$);*
(4) *real conic bundles over $\mathbb{P}^1$ with $2m \geq 4$ reducible fibers, which are all real and consist of pairs of complex conjugate exceptional curves ($d = 8 - 2m$), $\mathbb{R}Y = mS^2$;*
(5) *del Pezzo surfaces of degree $d = 1$ or 2: $\mathbb{R}Y = \mathbb{RP}^2 \sqcup 4S^2$, if $d = 1$, and $\mathbb{R}Y = 3S^2$ or $4S^2$, if $d = 2$.*

[†]In both cases, we use the equality $H^2(\mathcal{T}_{X\|D}) = 0$, which can be deduced, for example, from Serre duality, $H^2(\mathcal{T}_{X\|D}) = (H^0(\Omega_X^1(\log D) \otimes K))^*$, and Bogomolov–Sommese vanishing $H^0(\Omega_X^1(\log D) \otimes K) = 0$; the latter holds in our case since X is a rational surface with $K^2 \geq 1$, and thus its anticanonical Iitaka–Kodaira dimension is equal to 2.

Proposition 3. *With one exception, a real del Pezzo surface (Y, c) of degree $d \geq 1$ is determined up to deformation by the topology of $\mathbb{R}Y$. In the exceptional case $d = 8$ and $\mathbb{R}Y = \varnothing$, there are two deformation classes, distinguished by whether Y/c is Spin or not.*

The topological types of $\mathbb{R}Y$ are the following extremal types and their derivatives, which are obtained from the extremal ones by sequences of topological Morse simplifications of $\mathbb{R}Y$:

$$\begin{array}{ll}
d=9 & \mathbb{R}Y = \mathbb{RP}^2; \\
d=8 & \mathbb{R}Y = \#_2\mathbb{RP}^2 \ \textit{or}\ (S^1)^2; \\
d=7 & \mathbb{R}Y = \#_3\mathbb{RP}^2; \\
d=6 & \mathbb{R}Y = \#_4\mathbb{RP}^2 \ \textit{or}\ (S^1)^2; \\
d=5 & \mathbb{R}Y = \#_5\mathbb{RP}^2; \\
d=4 & \mathbb{R}Y = \#_6\mathbb{RP}^2,\ (S^1)^2,\ \textit{or}\ 2S^2; \\
d=3 & \mathbb{R}Y = \#_7\mathbb{RP}^2 \ \textit{or}\ \mathbb{RP}^2 \sqcup S^2; \\
d=2 & \mathbb{R}Y = \#_8\mathbb{RP}^2,\ 2\mathbb{RP}^2,\ \#_2\mathbb{RP}^2 \sqcup S^2,\ (S^1)^2,\ \textit{or}\ 4S^2; \\
d=1 & \mathbb{R}Y = \#_9\mathbb{RP}^2,\ \#_2\mathbb{RP}^2 \sqcup \mathbb{RP}^2,\ \#_3\mathbb{RP}^2 \sqcup S^2,\ \textit{or}\ \mathbb{RP}^2 \sqcup 4S^2.
\end{array}$$

1.2. Main results

Let us consider a real del Pezzo surface (Σ, c), and assume that its real point set $\mathbb{R}\Sigma = \mathrm{Fix}(c)$ is nonempty. Pick a real divisor class $D \in \mathrm{Pic}(\Sigma)$, satisfying $-DK_\Sigma > 0$ and $D^2 \geq -1$, and put $r = -DK_\Sigma - 1$. Fix an integer m such that $0 \leq 2m \leq r$ and introduce a real structure $c_{r,m}$ on Σ^r that maps $(w_1, \dots, w_r) \in \Sigma^r$ to $(w'_1, \dots, w'_r) \in \Sigma^r$ with $w'_i = c(w_i)$ if $i > 2m$, and $(w'_{2j-1}, w'_{2j}) = (c(w_{2j}), c(w_{2j-1}))$ if $j \leq m$. With respect to this real structure a point $\boldsymbol{w} = (w_1, \dots, w_r)$ is real, *i.e.*, $c_{r,m}$-invariant, if and only if w_i belongs to the real part $\mathbb{R}\Sigma$ of Σ for $i > 2m$ and w_{2j-1}, w_{2j} are conjugate to each other for $j \leq m$. In what follows we work with an open dense subset $\mathcal{P}_{r,m}(\Sigma)$ of $\mathbb{R}\Sigma^r = \mathrm{Fix}\, c_{r,m}$ consisting of $c_{r,m}$-invariant r-tuples $\boldsymbol{w} = (w_1, \dots, w_r)$ with pairwise distinct $w_i \in \Sigma$.

Observe that, if a real irreducible rational curve $C \in |D|$ can be traced through all the points w_i of $\boldsymbol{w}$ and $2m < r = -CK_\Sigma - 1$, the real points of $\boldsymbol{w}$ must lie on the unique one-dimensional connected component of the real part of C, hence must belong to the same connected component of $\mathbb{R}\Sigma$. In the case $2m = r$, each real rational curve $C \in |D|$ passing through a collection of m pairs of complex conjugate points of Σ has an odd intersection with the real divisor K_Σ, hence C has a homologically non-trivial real part in $\mathbb{R}\Sigma$.

Thus, we fix a connected component F of $\mathbb{R}\Sigma$ and put

$$\mathcal{P}_{r,m}(\Sigma, F) = \{\boldsymbol{w} = (w_1, \dots, w_r) \in \mathcal{P}_{r,m}(\Sigma) \ : \ w_i \in F \text{ for } i > 2m\}\ .$$

Denote by $\mathcal{M}_{0,r}(\Sigma, D)$ the set of isomorphism classes of pairs $(\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p})$, where $\nu : \mathbb{P}^1 \to \Sigma$ is a holomorphic map such that $\nu_*(\mathbb{P}^1) \in |D|$, and $\boldsymbol{p}$ is a sequence of r pairwise distinct points in $\mathbb{P}^1$. Put

$$\begin{aligned}\mathcal{R}(\Sigma, D, F, \boldsymbol{w}) \quad = \quad & \{[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}] \in \mathcal{M}_{0,r}(\Sigma, D) \ : \\ & \quad \nu \circ \mathrm{Conj} = c \circ \nu,\ \nu(\mathbb{RP}^1) \subset F,\ \nu(\boldsymbol{p}) = \boldsymbol{w}\} ,\end{aligned}$$

where $\mathrm{Conj} : \mathbb{P}^1 \to \mathbb{P}^1$ is the complex conjugation. If either the degree of Σ is greater than 1, or $D \ne -K_\Sigma$, then for any generic r-tuple $\boldsymbol{w} \in \mathcal{P}_{r,m}(\Sigma, F)$, the set $\mathcal{R}(\Sigma, D, F, \boldsymbol{w})$ is finite and presented by immersions (see Lemma 9). In such a case, pick a conjugation-invariant class $\varphi \in H_2(\Sigma \setminus F; \mathbb{Z}/2)$ and put

$$W_m(\Sigma, D, F, \varphi, \boldsymbol{w}) = \sum_{[\nu,\boldsymbol{p}]\in\mathcal{R}(\Sigma,D,F,\boldsymbol{w})} (-1)^{C_+\circ C_- + C_+\circ\varphi} , \tag{1}$$

where $C_\pm = \nu(\mathbb{P}^1_\pm)$ with $\mathbb{P}^1_+, \mathbb{P}^1_-$ being the two connected components of $\mathbb{P}^1 \setminus \mathbb{RP}^1$.

If the degree of Σ is equal to 1 and $D = -K_\Sigma$, then for any generic r-tuple $\boldsymbol{w} \in \mathcal{P}_{r,m}(\Sigma, F)$ and any conjugation-invariant class $\varphi \in H_2(\Sigma \setminus F; \mathbb{Z}/2)$ we define the number $W_m(\Sigma, D, F, \varphi, \boldsymbol{w})$ by the formula (1) retaining in it only the classes $[\nu, \boldsymbol{p}]$ presented by immersions.

If $\varphi = 0$, we get the original definition of Welschinger [27, 28].

Proposition 4. *The number $W_m(\Sigma, D, F, \varphi, \boldsymbol{w})$ does not depend on the choice of a generic element $\boldsymbol{w} \in \mathcal{P}_{r,m}(\Sigma, F)$.*

Proposition 4 is in fact a special case of more general deformation invariance statements. Consider a smooth real surface X_0 with $\mathbb{R}X_0 \ne \varnothing$, a real divisor class $D_0 \in \mathrm{Pic}(X_0)$, a connected component F_0 of $\mathbb{R}X_0$, a conjugation-invariant class $\varphi_0 \in H_2(X_0 \setminus F_0; \mathbb{Z}/2)$, and a conjugation invariant collection $\boldsymbol{w}_0$ of points in X_0. By an *elementary deformation* of the tuple $(X_0, D_0, F_0, \varphi_0)$ (respectively, $(X_0, D_0, F_0, \varphi_0, \boldsymbol{w}_0)$) we mean a one-parameter smooth family of smooth surfaces X_t, $t \in [-1, 1]$, extended to a continuous family of tuples $(X_t, D_t, F_t, \varphi_t)$ (respectively, $(X_t, D_t, F_t, \varphi_t, \boldsymbol{w}_t)$). Two tuples $T = (X, D, F, \varphi)$ and $T' = (X', D', F', \varphi')$ are called *deformation equivalent* if they can be connected by a chain $T = T^{(0)}$, $\dots$, $T^{(k)} = T'$ so that any two neighboring tuples in the chain are isomorphic to fibers of an elementary deformation.

Proposition 5. *Let $(\Sigma_t, D_t, F_t, \varphi_t, \boldsymbol{w}_t)$, $t \in [-1, 1]$, be an elementary deformation of tuples such that all surfaces Σ_t, $t \ne 0$, belong to $\mathcal{D}^{\mathrm{DP}}_d$ for some $1 \le d \le 9$, and the collections $\boldsymbol{w}_{\pm 1}$ belong to $\mathcal{P}_{r,m}(\Sigma_{\pm 1}, F_{\pm 1})$ and are generic. Then,*

$$W_m(\Sigma_{-1}, D_{-1}, F_{-1}, \varphi_{-1}, \boldsymbol{w}_{-1}) = W_m(\Sigma_1, D_1, F_1, \varphi_1, \boldsymbol{w}_1) . \tag{2}$$

We skip $\boldsymbol{w}$ in the notation of the numbers $W_m(\Sigma, D, F, \varphi, \boldsymbol{w})$ and call them *Welschinger invariants.*

Proposition 5 plays a central role in the proof of the following statement.

Theorem 6. *If tuples (Σ, D, F, φ) and $(\Sigma', D', F', \varphi')$ are deformation equivalent, then $W_m(\Sigma, D, F, \varphi) = W_m(\Sigma', D', F', \varphi')$.*

Proofs of Propositions 4 and 5, as well as the proof of Theorem 6, are found in Section 3.

2. Families of rational curves on rational surfaces

2.1. General setting

Let Σ be a smooth rational surface, and $D \in \mathrm{Pic}(\Sigma)$ a divisor class. Denote by $\overline{\mathcal{M}}_{0,n}(\Sigma, D)$ the space of the isomorphism classes of pairs $(\nu : \hat{C} \to \Sigma, \boldsymbol{p})$, where $\hat{C}$ is either $\mathbb{P}^1$ or a connected reducible nodal curve of arithmetic genus zero, $\nu_*\hat{C} \in |D|$, $\boldsymbol{p} = (p_1, \dots, p_n)$ is a sequence of distinct smooth points of $\hat{C}$, and each component of $\hat{C}$ contracted by ν contains at least three special points. This moduli space is a projective scheme (see [9]), and there are natural morphisms

$$\begin{aligned} \Phi_{\Sigma,D} : \overline{\mathcal{M}}_{0,n}(\Sigma, D) \to |D|, \quad & [\nu : \hat{C} \to \Sigma, \boldsymbol{p}] \mapsto \nu_*\hat{C} \ , \\ \mathrm{Ev} : \overline{\mathcal{M}}_{0,n}(\Sigma, D) \to \Sigma^n, \quad & [\nu : \hat{C} \to \Sigma, \boldsymbol{p}] \mapsto \nu(\boldsymbol{p}) \ . \end{aligned}$$

For any subscheme $\mathcal{V} \subset \overline{\mathcal{M}}_{0,n}(\Sigma, D)$, define the *intersection dimension* $\mathrm{idim}\mathcal{V}$ of $\mathcal{V}$ as follows:

$$\mathrm{idim}\mathcal{V} = \dim(\Phi_{\Sigma,D} \times \mathrm{Ev})(\mathcal{V}) \ ,$$

where the latter value is the maximum over the dimensions of all irreducible components.

Put

$$\begin{aligned} \mathcal{M}^{br}_{0,n}(\Sigma, D) &= \{[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}] \in \mathcal{M}_{0,n}(\Sigma, D) \ : \ \nu \text{ is birational onto } \nu(\mathbb{P}^1)\}, \\ \mathcal{M}^{im}_{0,n}(\Sigma, D) &= \{[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}] \in \mathcal{M}_{0,n}(\Sigma, D) \ : \ \nu \text{ is an immersion}\} \ . \end{aligned}$$

Denote by $\overline{\mathcal{M}^{br}_{0,n}}(\Sigma, D)$ the closure of $\mathcal{M}^{br}_{0,n}(\Sigma, D)$ in $\overline{\mathcal{M}}_{0,n}(\Sigma, D)$, and introduce also the space

$$\mathcal{M}'_{0,n}(\Sigma, D) = \{[\nu : \hat{C} \to \Sigma, \boldsymbol{p}] \in \overline{\mathcal{M}^{br}_{0,n}}(\Sigma, D) \ : \ \hat{C} \simeq \mathbb{P}^1\} \ .$$

The following statement will be used below.

Lemma 7. *For any element*

$$[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}] \in \mathcal{M}^{br}_{0,n}(\Sigma, D) \ \ \textit{such that} \ \ \nu(\boldsymbol{p}) \cap \mathrm{Sing}\,(\nu(\mathbb{P}^1)) = \emptyset,$$

the map $\Phi_{\Sigma,D} \times \mathrm{Ev}$ is injective in a neighborhood of that element, and, for the germ at $[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}]$ of any irreducible subscheme $\mathcal{V} \subset \mathcal{M}^{br}_{0,n}(\Sigma, D)$, we have

$$\dim \mathcal{V} = \mathrm{idim}\mathcal{V} \ .$$

Proof. The inequality $\mathrm{idim}\mathcal{V} \le \dim \mathcal{V}$ is immediate from the definition. The opposite inequality and the injecitvity of $\Phi_{\Sigma,D} \times \mathrm{Ev}$ follow from the observation that, for an irreducible rational curve $C \in |D|$ and a tuple $\boldsymbol{z} \subset C \setminus \mathrm{Sing}\,(C)$ of n distinct points, the normalization map $\nu : \mathbb{P}^1 \to C$ and the lift $\boldsymbol{p} = \nu^{-1}(\boldsymbol{z})$ represent the unique preimage of $(C, \boldsymbol{z}) \in |D| \times \Sigma^n$ in $\mathcal{M}^{br}_{0,n}(\Sigma, D)$. $\square$

2.2. Curves on del Pezzo and uninodal del Pezzo surfaces

We establish here certain properties of the spaces $\mathcal{M}^{im}_{0,0}(\Sigma, D)$, $\mathcal{M}^{br}_{0,0}(\Sigma, D)$, and $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$, notably, compute dimension and describe generic members of these spaces as well as of some divisors therein. These properties basically follow from [10, Theorem 4.1 and Lemma 4.10]. However, the cited paper considers the plane blown up at generic points, whereas we work with arbitrary del Pezzo or uninodal del Pezzo surfaces. For this reason, we supply all claims with complete proofs.

Through all this section we use the notation

$$r = -DK_\Sigma - 1.$$

Lemma 8. *If Σ is a smooth rational surface and $-DK_\Sigma > 0$, then the space $\mathcal{M}^{im}_{0,0}(\Sigma, D)$ is either empty, or a smooth variety of dimension r.*

Proof. Let $[\nu : \mathbb{P}^1 \to \Sigma] \in \mathcal{M}^{im}_{0,0}(\Sigma, D)$. The Zariski tangent space to $\mathcal{M}^{im}_{0,0}(\Sigma, D)$ at $[\nu : \mathbb{P}^1 \to \Sigma]$ can be identified with $H^0(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1})$, where $\mathcal{N}^\nu_{\mathbb{P}^1} = \nu^*\mathcal{T}\Sigma/\mathcal{T}\mathbb{P}^1$ is the normal bundle. Since

$$\deg \mathcal{N}^\nu_{\mathbb{P}^1} = -DK_\Sigma - 2 \ge -1 > (2g-2)\big|_{g=0} = -2 \ , \tag{3}$$

we have

$$h^1(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1}) = 0 \ , \tag{4}$$

and hence $\mathcal{M}^{im}_{0,0}(\Sigma, D)$ is smooth at $[\nu : \mathbb{P}^1 \to \Sigma]$ and is of dimension

$$h^0(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1}) = \deg \mathcal{N}^\nu_{\mathbb{P}^1} - g + 1 = -DK_\Sigma - 1 = r \ . \tag{5}$$

$\square$

Lemma 9.

(1) *Let $\Sigma \in \mathcal{D}^{\mathrm{DP}}_d$ and $-DK_\Sigma > 0$. Then, the following holds:*

   (i) *The space $\mathcal{M}^{br}_{0,0}(\Sigma, D)$ is either empty or a variety of dimension r, and* $\mathrm{idim}(\mathcal{M}_{0,0}(\Sigma, D) \setminus \mathcal{M}^{br}_{0,0}(\Sigma, D)) < r$.

   (ii) *If either $d > 1$ or $D \ne -K_\Sigma$, then $\mathcal{M}^{im}_{0,0}(\Sigma, D) \subset \mathcal{M}^{br}_{0,0}(\Sigma, D)$ is an open dense subset.*

   (iii) *There exists an open dense subset $U_1 \subset \mathcal{D}^{\mathrm{DP}}_1$ such that, if $\Sigma \in U_1$, then $\mathcal{M}_{0,0}(\Sigma, -K_\Sigma)$ consists of* 12 *elements, each corresponding to a rational nodal curve.*

(2) *Let $d \le 4$. There exists an open dense subset $U_d(A_1) \subset \mathcal{D}_d(A_1)$ such that if $\Sigma \in U_d(A_1)$ and $-DK_\Sigma > 0$, then*

   (i) $\mathrm{idim}\mathcal{M}_{0,0}(\Sigma, D) \le r$*;*

   (ii) *a generic element $[\nu : \mathbb{P}^1 \to \Sigma]$ of any irreducible component $\mathcal{V}$ of $\mathcal{M}_{0,0}(\Sigma, D)$ such that* $\mathrm{idim}\mathcal{V} = r$*, is an immersion, and the divisor $\nu^*(E_\Sigma)$ consists of DE_Σ distinct points.*

Proof. Let $\Sigma \in \mathcal{D}_d^{\rm DP} \cup \mathcal{D}_d(A_1)$. All the statements in the case of an effective $-K_\Sigma - D$ immediately follow from elementary properties of plane lines, conics, and cubics. Thus, we suppose that $-K_\Sigma - D$ is not effective.

Let $\mathcal{V}_1$ be an irreducible component of $\mathcal{M}_{0,0}^{br}(\Sigma, D)$ and $[\nu : \mathbb{P}^1 \to \Sigma]$ its generic element. Then by [19, Theorem II.1.2]

$$\dim \mathrm{Hom}(\mathbb{P}^1, \Sigma)_\nu \ge -DK_\Sigma + 2\chi(\mathcal{O}_{\mathbb{P}^1}) = -DK_\Sigma + 2 \ . \tag{6}$$

Reducing by the automorphisms of $\mathbb{P}^1$, we get

$$\dim \mathcal{V}_1 \ge -DK_\Sigma + 2 - 3 = r \ . \tag{7}$$

Hence, in view of Lemma 8, to prove that $\dim \mathcal{M}_{0,0}^{br}(\Sigma, D) = r$ and $\mathcal{M}_{0,0}^{im}(\Sigma, D)$ is dense in $\mathcal{M}_{0,0}^{br}(\Sigma, D)$, it is enough to show that $\dim(\mathcal{M}_{0,0}^{br}(\Sigma, D) \setminus \mathcal{M}_{0,0}^{im}(\Sigma, D)) < r$.

Notice, first, that, in the case $r = 0$, the curves $C \in \Phi_{\Sigma,D}(\mathcal{M}_{0,0}^{br}(\Sigma, D))$ are nonsingular due to the bound

$$-DK_\Sigma \ge (C \cdot C')(z) \ge s \ , \tag{8}$$

coming from the intersection of C with a curve $C' \in |-K_\Sigma|$ passing through a point $z \in C$, where C has multiplicity s. Thus, we suppose that $r > 0$. Let $\mathcal{V}_2$ be an irreducible component of $\mathcal{M}_{0,0}^{br}(\Sigma, D) \setminus \mathcal{M}_{0,0}^{im}(\Sigma, D)$, $[\nu : \mathbb{P}^1 \to \Sigma] \in \mathcal{V}_2$ a generic element, and let ν have $s \ge 1$ critical points of multiplicities $m_1 \ge \cdots \ge m_s \ge 2$. In particular, bound (8) gives

$$-DK_\Sigma \ge m_1 \ . \tag{9}$$

Then (*cf.* [5, First formula in the proof of Corollary 2.4]),

$$\dim \mathcal{V}_2 \le h^0(\mathbb{P}^1, \mathcal{N}_{\mathcal{P}^1}^\nu / \mathrm{Tors}(\mathcal{N}_{\mathbb{P}^1}^\nu)) \ ,$$

where the normal sheaf $\mathcal{N}_{\mathbb{P}^1}^\nu$ on $\mathbb{P}^1$ is defined as the cokernel of the map $d\nu : \mathcal{T}\mathbb{P}^1 \to \nu^*\mathcal{T}\Sigma$, and $\mathrm{Tors}(*)$ is the torsion sheaf. It follows from [5, Lemma 2.6] (*cf.* also the computation in [5, Page 363]) that $\deg \mathrm{Tors}(\mathcal{N}_{\mathbb{P}^1}^\nu) = \sum_i (m_i - 1)$, and hence

$$\deg \mathcal{N}_{\mathbb{P}^1}^\nu / \mathrm{Tors}(\mathcal{N}_{\mathbb{P}^1}^\nu) = -DK_\Sigma - 2 - \sum_{i=1}^s (m_i - 1) \tag{10}$$

which yields

$$\begin{aligned} \dim \mathcal{V}_2 &\le h^0(\mathbb{P}^1, \mathcal{N}_{\mathbb{P}^1}^\nu / \mathrm{Tors}(\mathcal{N}_{\mathbb{P}^1}^\nu)) \\ &= \max\{\deg \mathcal{N}_{\mathbb{P}^1}^\nu / \mathrm{Tors}(\mathcal{N}_{\mathbb{P}^1}^\nu) + 1, \ 0\} \overset{(9)}{\le} r - (m_1 - 1) < r, \end{aligned} \tag{11}$$

Let us show that $\mathrm{idim}\mathcal{V} < r$ for any irreducible component $\mathcal{V}$ of $\mathcal{M}_{0,0}(\Sigma, D) \setminus \mathcal{M}_{0,0}^{br}(\Sigma, D)$. Indeed, if a generic element $[\nu : \mathbb{P}^1 \to \Sigma] \in \mathcal{V}$ satisfies $\nu_*(\mathbb{P}^1) = sC$ for some $s \ge 2$, then

$$\mathrm{idim}\mathcal{V} \le -\frac{1}{s} DK_\Sigma - 1 < -DK_\Sigma - 1 = r \ .$$

To complete the proof of (2ii), let us assume that $\dim \mathcal{V} = r$ and the divisor $\nu^*(E_\Sigma)$ contains a multiple point sz, $s \ge 2$. In view of $DE_\Sigma \ge s$ and $(-K_\Sigma - E_\Sigma)D \ge 0$ (remind that D is irreducible and $-K - D$ is not effective), we have $-DK_\Sigma \ge s$. Furthermore, $T_{[\nu]}\mathcal{V}$ can be identified with a subspace of $H^0(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1}(-(s-1)z))$ (*cf.* [5, Remark in page 364]). Since

$$\deg \mathcal{N}^\nu_{\mathbb{P}^1}(-(s-1)z)) = -DK_\Sigma - 1 - s \ge -1 > -2 ,$$

we have

$$H^1(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1}(-(s-1)z)) = 0 ,$$

and hence

$$\dim \mathcal{V} \le h^0(\mathbb{P}^1, \mathcal{N}^\nu_{\mathbb{P}^1}(-(s-1)z)) = r - (s-1) < r$$

contrary to the assumption $\dim \mathcal{V} = r$. $\square$

Lemma 10. *There exists an open dense subset $U_2 \subset \mathcal{D}_1^{\mathrm{DP}}$ such that, for each $\Sigma \in U_2$, the set of effective divisor classes $D \in \mathrm{Pic}(\Sigma)$ satisfying $-DK_\Sigma = 1$ is finite, the set of rational curves in the corresponding linear systems $|D|$ is finite, and any two such rational curves C_1, C_2 either coincide, or are disjoint, or intersect in C_1C_2 distinct points.*

Proof. For any $\Sigma \in \mathcal{D}_1^{\mathrm{DP}}$, we have $\dim|-2K_\Sigma| = 3$. Hence, the condition $-DK_\Sigma = 1$ yields that $-2K_\Sigma - D$ is effective, which in turn implies the finiteness of the set of effective divisors such that $-DK_\Sigma = 1$. The finiteness of the set of rational curves in these linear systems $|D|$ follows from Lemma 9(i). At last, for a generic $\Sigma \in \mathcal{D}_1^{\mathrm{DP}}$, these curves are either singular elements in the elliptic pencil $|-K_\Sigma|$ or the (-1)-curves, and as it follows easily, for example, from considering Σ as a projective plane blown up at 8 generic points, any two of these curves intersect transversally and in distinct smooth points. $\square$

Lemma 11. *Let U_1, U_2 be the subsets of $\mathcal{D}_1^{\mathrm{DP}}$ introduced in Lemmas 9 and 10, respectively. For each $\Sigma \in U_1 \cap U_2$, each $D \in \mathrm{Pic}(\Sigma)$ with $-DK_\Sigma > 0$ and $D^2 \ge -1$, and for each irreducible component $\mathcal{V}$ of $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D) \setminus \mathcal{M}^{br}_{0,0}(\Sigma, D)$ with* idim$\mathcal{V}$ = $r-1$, *one has:*

(i) *A generic element $[\nu : \hat{C} \to \Sigma] \in \mathcal{V}$ is as follows*
- $\hat{C} = \hat{C}_1 \cup \hat{C}_2$ *with $\hat{C}_i \simeq \mathbb{P}^1$, $[\nu|_{\hat{C}_i} : \hat{C}_i \to \Sigma] \in \mathcal{M}^{im}_{0,0}(\Sigma, D_i)$, where $D_1D_2 > 0$ and $-D_iK_\Sigma > 0$, $D_i^2 \ge -1$ for each $i = 1, 2$;*
- $\nu(\hat{C}_1) \ne \nu(\hat{C}_2)$, *except for the only case when $D_1 = D_2 = -K_\Sigma$ and $\nu(\hat{C}_1) = \nu(\hat{C}_2)$ is one of the 12 uninodal curves in $|-K_\Sigma|$;*
- ν *is an immersion (i.e., a local isomorphism onto the image).*

Moreover, each element $[\nu : \hat{C} \to \Sigma] \in \overline{\mathcal{M}}_{0,0}(\Sigma, D)$ as above does belong to $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$.

(ii) *The germ of $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$ at a generic element of $\mathcal{V}$ is smooth.*

Proof. Show, first, that $\mathrm{idim}(\mathcal{M}'_{0,0}(\Sigma, D) \setminus \mathcal{M}^{br}_{0,0}(\Sigma, D)) \le r - 2$. Assume on the contrary that there exists a component $\mathcal{V}$ of $\mathcal{M}'_{0,0}(\Sigma, D) \setminus \mathcal{M}^{br}_{0,0}(\Sigma, D)$ with idim$\mathcal{V}$ =

$r-1$ (idim$\mathcal{V}$ cannot be bigger by Lemma 9(i)). Then its generic element $[\nu : \mathbb{P}^1 \to \Sigma]$ is such that $\nu_*(\mathbb{P}^1) = sC$ with C an irreducible rational curve, $s \ge 2$. Thus,

$$r-1 = -sCK_\Sigma - 2 \le -CK_\Sigma - 1 = \dim \mathcal{M}_{0,0}^{br}(\Sigma, C) \ ,$$

which yields $s = 2$ and $-CK_\Sigma = 1$. By adjunction formula, either $C^2 = -1$, or $C^2 \ge 1$. The former case is excluded by the assumption $D^2 \ge -1$. In the case $C^2 \ge 1$, since $K_\Sigma^2 = 1$ and $-CK_\Sigma = 1$, the only possibility is $C \in |-K_\Sigma|$. However, in such a case the map ν cannot be deformed into an element of $\mathcal{M}_{0,0}^{br}(\Sigma, -2K_\Sigma)$, since C has a node, and hence the deformed map would birationally send $\mathbb{P}^1$ onto a curve with δ-invariant ≥ 4, which is bigger than its arithmetic genus, $((-2K_\Sigma)^2 + (-2K_\Sigma)K_\Sigma)/2 + 1 = 2$.

Let $[\nu : \hat{C} \to \Sigma]$ be a generic element of an irreducible component $\mathcal{V}$ of $\overline{\mathcal{M}_{0,0}^{br}}(\Sigma, D) \setminus \mathcal{M}'_{0,0}(\Sigma, D)$ with idim$\mathcal{V} = r-1$. Then $\hat{C}$ has at least 2 components. On the other side, if $\hat{C}$ had ≥ 3 components, Lemma 9(1) would yield idim$\mathcal{V} \le -DK_\Sigma - 3 < r-1$. Hence $\hat{C} = \hat{C}_1 \cup \hat{C}_2$, $\hat{C}_1 \simeq \hat{C}_2 \simeq \mathbb{P}^1$, and, according to Lemma 8 and Lemma 9(1), for each $i = 1, 2$ we have: $\nu_i = \nu\big|_{\hat{C}_i}$ is an immersion, $\dim \mathcal{M}_{0,0}(\Sigma, D_i)_{[\nu_i]} = -D_iK_\Sigma - 1$, and $-D_iK_\Sigma > 0$, $D_i^2 \ge -1$.

If $-DK_\Sigma = 2$ and $\nu(\hat{C}_1) \ne \nu(\hat{C}_2)$, then the intersection points of these curves are nodes, which follows from the definition of the set U_2 (see Lemma 10), and hence ν is an immersion at the node $\hat{z}$ of $\hat{C}$.

If $-DK_\Sigma = 2$ and $\nu(\hat{C}_1) = \nu(\hat{C}_2)$, then $D_1 = D_2$ and $D_1^2 = D_2^2 \ge 1$ in view of the adjunction formula and the condition $D^2 \ge -1$. It is easy to see that this is only possible, when $D_1 = D_2 = -K_\Sigma$. In particular, by the definition of the set U_1 (see Lemma 9(iii)), the curve $C = \nu(\hat{C}_1) = \nu(\hat{C}_2) \in |-K_\Sigma|$ has one node z. We then see that, ν takes the germ $(\hat{C}, \hat{z})$ isomorphically onto the germ (C, z), since, otherwise we would get a deformed map ν with the image whose δ-invariant ≥ 4, which is bigger than its arithmetic genus, $((-2K_\Sigma)^2 + (-2K_\Sigma)K_\Sigma)/2 + 1 = 2$.

Suppose, now, that $-DK_\Sigma > 2$, thus, $-D_1K_\Sigma > 1$. Then

$$\dim \mathcal{M}_{0,0}(\Sigma, D_1)_{[\nu_1]} > 0,$$

and hence $C_1 \ne C_2$. To prove that ν is an immersion at the node $\hat{z} \in \hat{C}$, we will show that any two local branches of ν_1 and ν_2 either are disjoint, or intersect transversally. Indeed, assume on the contrary that there exist $z_i \in \hat{C}_i$, $i = 1, 2$, such that $\nu_1(z_1) = \nu_2(z_2) = z \in \Sigma$, and $\nu_1(\hat{C}_1, z_1)$ intersects $\nu_2(\hat{C}_2, z_2)$ at z with multiplicity ≥ 2. Then

$$\dim \mathcal{M}_{0,0}(\Sigma, D_1)_{[\nu_1]} \le h^0(\hat{C}_1, \mathcal{N}_{\hat{C}_1}^{\nu_1}(-z_1)) \ . \tag{12}$$

Since

$$\deg \mathcal{N}_{\hat{C}_1}^{\nu_1}(-z_1) = -D_1K_\Sigma - 2 - 1 = -D_1K_\Sigma - 3 > -2 \ ,$$

we get $h^1(\hat{C}_1, \mathcal{N}^{\nu_1}_{\hat{C}_1}(-z_1)) = 0$. Therefore,

$$\begin{aligned}\deg \mathcal{N}^{\nu_1}_{\hat{C}_1}(-z_1) &\le \deg \mathcal{N}^{\nu_1}_{\hat{C}_1}(-z_1) + 1 = -D_1 K_\Sigma - 2 \\ &< -D_1 K_\Sigma - 1 = \dim \mathcal{M}_{0,0}(\Sigma, D_1)_{[\nu_1]}\ ,\end{aligned}$$

which contradicts (12).

The smoothness of $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$ at $[\nu : \hat{C} \to \Sigma]$, where $\nu_* \hat{C}$ is a reduced nodal curve, follows from [25, Lemma 2.9], where the requirements are $D_i K_\Sigma < 0$, $i = 1, 2$. We will show that the same requirements suffice under assumption that ν is an immersion. Let us show that

$$T_{[\nu]} \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D) \simeq H^0(\hat{C}, \mathcal{N}^\nu_{\hat{C}})\ , \tag{13}$$

where the normal sheaf $\mathcal{N}^\nu_{\hat{C}}$ comes from the exact sequence

$$0 \to \mathcal{T}_{\hat{C}} \to \nu^* \mathcal{T}_\Sigma \to \mathcal{N}^\nu_{\hat{C}} \to 0\ , \tag{14}$$

$\mathcal{T}_\Sigma$ being the tangent bundle of Σ, and $\mathcal{T}_{\hat{C}}$ being the tangent sheaf of $\hat{C}$ viewed as the push-forward by the normalization $\pi : \hat{C}_1 \sqcup \hat{C}_2 \to \hat{C}$ of the subsheaf $\mathcal{T}'_{\hat{C}_1 \sqcup \hat{C}_2} \subset \mathcal{T}_{\hat{C}_1 \sqcup \hat{C}_2}$ generated by the sections vanishing at the preimages of the node $z \in \hat{C}$.

Indeed, the Zariski tangent space to $\mathrm{Hom}(\hat{C}, \Sigma)$ at ν is naturally isomorphic to $H^0(\hat{C}, \nu^* \mathcal{T}_\Sigma)$ (see [19, Theorem 1.7, Section II.1]). Next, we take the quotient by action of the germ of $\mathrm{Aut}(\hat{C})$ at the identity. This germ is smooth and acts freely on the germ of $\mathrm{Hom}(\hat{C}, \Sigma)$ at ν. The tangent space to $\mathrm{Aut}(\hat{C})$ at the identity is isomorphic to $H^0(\hat{C}, \mathcal{T}_{\hat{C}})$ (*cf.* [19, 2.16.4, Section I.2]). Since

$$H^1(\hat{C}, \mathcal{T}_{\hat{C}}) = H^1(\hat{C}_1 \sqcup \hat{C}_2, \mathcal{T}'_{\hat{C}_1 \sqcup \hat{C}_2}) = H^1(\hat{C}_1, \mathcal{O}_{\hat{C}_1}(1)) \oplus H^1(\hat{C}_2, \mathcal{O}_{\hat{C}_2}(1)) = 0\ , \tag{15}$$

the associated to (14) cohomology exact sequence yields

$$\begin{aligned}T_{[\nu]} \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D) &\simeq T_\nu \mathrm{Hom}(\hat{C}, \Sigma) / T_{\mathrm{Id}} \mathrm{Aut}(\hat{C}) \\ &\simeq H^0(\hat{C}, \nu^* \mathcal{T}_\Sigma) / H^0(\hat{C}, \mathcal{T}_{\hat{C}}) \simeq H^0(\hat{C}, \mathcal{N}^\nu_{\hat{C}})\ .\end{aligned}$$

We will verify that

$$h^0(\hat{C}, \mathcal{N}^\nu_{\hat{C}}) = r\ , \tag{16}$$

which in view of $\dim_{[\nu]} \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D) = r$ (see Lemma 9(i)) will imply the smoothness of $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$ at $[\nu]$. There exists a natural morphism of sheaves on $\hat{C}$:

$$\alpha : \pi_* \mathcal{N}^{\nu \circ \pi}_{\hat{C}_1 \sqcup \hat{C}_2} \longrightarrow \mathcal{N}^\nu_{\hat{C}}\ ,$$

where α is an isomorphism outside z and acts at z as follows: since ν embeds the germ of $\hat{C}$ at z into Σ, one can identify the stalk $\left(\pi_* \mathcal{N}^{\nu \circ \pi}_{\hat{C}_1 \sqcup \hat{C}_2}\right)_z$ with $\mathbb{C}\{x\} \oplus \mathbb{C}\{y\}$, the stalk $(\mathcal{N}^\nu_{\hat{C}})_z$ with $\mathbb{C}\{x, y\} / \langle xy \rangle$, and write

$$\alpha_z(f(x), g(y)) = x f(x) + y g(y) \in \left(\mathcal{N}^\nu_{\hat{C}}\right)_z \cong \mathbb{C}\{x, y\} / \langle xy \rangle\ .$$

Hence we obtain an exact sequence of sheaves

$$0 \to \pi_* \mathcal{N}^{\nu \circ \pi}_{\hat{C}_1 \sqcup \hat{C}_2} \xrightarrow{\alpha} \mathcal{N}^\nu_{\hat{C}} \to \mathcal{O}_z \to 0\ , \tag{17}$$

whose cohomology sequence vanishes at

$$h^1(z,\mathcal{O}_z)=0,\quad h^1(\hat{C}_1\sqcup\hat{C}_2,\mathcal{N}^{\nu\circ\pi}_{\hat{C}_1\sqcup\hat{C}_2})=0\ ,$$

(the latter one is equivalent to (4)); hence $h^1(\hat{C},\mathcal{N}^{\nu}_{\hat{C}})=0$ and, furthermore,

$$\begin{array}{rcl}h^0(\hat{C},\mathcal{N}^{\nu}_{\hat{C}}) & = & h^0(\hat{C}_1\sqcup\hat{C}_2,\mathcal{N}^{\nu\circ\pi}_{\hat{C}_1\sqcup\hat{C}_2}))+h^0(z,\mathcal{O}_z)\\ & = & h^0(\hat{C}_1,\mathcal{N}^{\nu_1}_{\hat{C}_1})+h^0(\hat{C}_2,\mathcal{N}^{\nu_2}_{\hat{C}_2})+h^0(z,\mathcal{O}_z)\\ & \overset{cf.\ (5)}{=} & (-D_1K_\Sigma-1)+(-D_2K_\Sigma-1)+1=r\end{array}$$

as predicted in (14).

Finally, let us show that any element $[\nu:\hat{C}\to\Sigma]\in\overline{\mathcal{M}_{0,0}}(\Sigma,D)$, satisfying conditions of 1(i)–1(iii), belongs to $\overline{\mathcal{M}^{br}_{0,0}}(\Sigma,D)$, or, equivalently, admits a deformation into a map $\mathbb{P}^1\to\Sigma$ birational onto its image. Indeed, it follows from [1, Theorem 15] under the condition $h^1(\hat{C},\nu^*\mathcal{T}_\Sigma)=0$, which one obtains from the cohomology exact sequence associated with (14) and from vanishing relations (15) and (16). □

Lemma 12. *Consider the subsets U_1, U_2 of $\mathcal{D}_1^{\mathrm{DP}}$ introduced in Lemmas* 9 *and* 10*, respectively, a surface $\Sigma\in U_1\cap U_2\subset\mathcal{D}_1^{\mathrm{DP}}$, and an effective divisor class $D\in\mathrm{Pic}(\Sigma)$ such that $-DK_\Sigma\ge 2$. Let $\boldsymbol{w}=(w_1,\dots,w_r)$ be a sequence of r distinct points in Σ, let σ_i be smooth curve germs in Σ centered at w_i, $r'<i\le r$, for some $r'<r$, $\boldsymbol{w}'=(w_i)_{1\le i\le r'}$, and let*

$$\begin{array}{l}\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D;\boldsymbol{w}',\{\sigma_i\}_{r'<i\le r})\\ \quad=\{[\nu:\hat{C}\to\Sigma,\boldsymbol{p}]\in\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D)\ :\\ \qquad\nu(p_i)=w_i\,\mathit{for}\ 1\le i\le r',\ \nu(p_i)\in\sigma_i,\ \mathit{for}\ r'<i\le r\}\ .\end{array}$$

(1) *Suppose that $[\nu:\mathbb{P}^1\to\Sigma,\boldsymbol{p}]\in\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D;\boldsymbol{w})\cap\mathcal{M}^{im}_{0,r}(\Sigma,D)$. Then* Ev *sends the germ of $\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D;\boldsymbol{w}',\{\sigma_i\}_{r'<i\le r})$ at $[\nu:\mathbb{P}^1\to\Sigma,\boldsymbol{p}]$ diffeomorphically onto $\prod_{r'<i\le r}\sigma_i$.*

(2) *Suppose that $[\nu:\hat{C}\to\Sigma,\boldsymbol{p}]\in\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D;\boldsymbol{w})$ is such that*

- *$[\nu:\hat{C}\to\Sigma]\in\overline{\mathcal{M}^{br}_{0,0}}(\Sigma,D)$ is as in Lemma* 11(i)*,*
- *$r'\ge -D_1K_\Sigma-1$, $\#(\boldsymbol{p}\cap\hat{C}_1)=-D_1K_\Sigma-1$, $\#(\boldsymbol{p}\cap\hat{C}_2)=-D_2K_\Sigma$, the point sequences $(w_i)_{1\le i<-D_1K_\Sigma}$, $(w_i)_{-D_1K_\Sigma\le i\le r}$ are generic on $C_1=\nu_*\hat{C}_1$, $C_2=\nu_*\hat{C}_2$, respectively, and the germs σ_i, $r'<i\le r$, cross C_2 transversally.*

Then Ev *sends the germ of $\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D;\boldsymbol{w}',\{\sigma_i\}_{r'<i\le r})$ at $[\nu:\hat{C}\to\Sigma,\boldsymbol{p}]$ diffeomorphically onto $\prod_{r'<i\le r}\sigma_i$.*

Proof. Both statements follow from the fact that Ev diffeomorphically sends the germ of $\overline{\mathcal{M}^{br}_{0,r}}(\Sigma,D)$ at $[\nu:\hat{C}\to\Sigma,\boldsymbol{p}]$ onto the germ of Σ^r at $\boldsymbol{w}=\nu(\boldsymbol{p})$.

In view of $\dim \overline{\mathcal{M}_{0,r}^{br}}(\Sigma, D) = 2r$ Lemma 9(i)), it is sufficient to show that the Zariski tangent space to $\mathrm{Ev}^{-1}(\boldsymbol{w})$ is zero-dimensional. In view of relation (13) this is equivalent to

$$h^0(\hat{C}, \mathcal{N}_{\hat{C}}^{\nu}(-\boldsymbol{p})) = 0 \ . \tag{18}$$

In the case of $[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}] \in \overline{\mathcal{M}_{0,r}^{br}}(\Sigma, D; \boldsymbol{w}', \{\sigma_i\}_{r'<i\le r}) \cap \mathcal{M}_{0,r}^{im}(\Sigma, D)$, we have

$$\deg \mathcal{N}_{\hat{C}}^{\nu}(-\boldsymbol{p}) = (-DK_\Sigma - 2) - (-DK_\Sigma - 1) = -1 > -2 \ ,$$

and hence (18) follows by Riemann–Roch.

In the second case, put $\widetilde{\boldsymbol{p}} = \boldsymbol{p} \setminus \{p_r\}$ and twist the exact sequence (17) to get

$$0 \to \pi_* \mathcal{N}_{\hat{C}_1 \sqcup \hat{C}_2}^{\nu\circ\pi}(-\widetilde{\boldsymbol{p}}) \to \mathcal{N}_{\hat{C}}^{\nu}(-\widetilde{\boldsymbol{p}}) \to \mathcal{O}_z \to 0.$$

Since

$$\deg \mathcal{N}_{\hat{C}_i}^{\nu_1}(-\widetilde{\boldsymbol{p}} \cap \hat{C}_i) = (-D_iK_\Sigma - 2) - (-D_iK_\Sigma - 1) = -1 > -2, \quad i = 1, 2 \ ,$$

we have $h^1(\pi_* \mathcal{N}_{\hat{C}_1 \sqcup \hat{C}_2}^{\nu\circ\pi}(-\widetilde{\boldsymbol{p}})) = 0$, and $h^0(\hat{C}, \pi_* \mathcal{N}_{\hat{C}_1 \sqcup \hat{C}_2}^{\nu\circ\pi}(-\widetilde{\boldsymbol{p}})) = 0$, which yields that $H^0(\hat{C}, \mathcal{N}_{\hat{C}}^{\nu}(-\widetilde{\boldsymbol{p}}))$ is isomorphically mapped onto $H^0(z, \mathcal{O}_z) \simeq \mathbb{C}$. It implies that a non-zero global section of the sheaf $\mathcal{N}_{\hat{C}}^{\nu}(-\widetilde{\boldsymbol{p}})$ does not vanish at z, and hence, it does not vanish at p_r chosen on $\hat{C}_2$ in a generic way. Thus, (18) follows. ☐

2.3. Deformation of isolated curve singularities

Let us recall a few facts on deformations of curve singularities (see, for example, [6]). Let Σ be a smooth algebraic surface, z an isolated singular point of a curve $C \subset \Sigma$, and $B_{C,z}$ the base of a semiuniversal deformation of the germ (C, z). This base can be viewed as a germ $(\mathbb{C}^N, 0)$ and can be identified with $\mathcal{O}_{C,z}/J_{C,z}$, where $J_{C,z} \subset \mathcal{O}_{C,z}$ is the Jacobian ideal.

The equigeneric locus $B_{C,z}^{eg} \subset B_{C,z}$ parametrizes local deformations of (C, z) with constant δ-invariant equal to $\delta(C, z)$. This locus is irreducible and has codimension $\delta(C, z)$ in $B_{C,z}$. The subset $B_{C,z}^{eg,im} \subset B_{C,z}^{eg}$ that parametrizes the immersed deformations is open and dense in $B_{C,z}^{eg}$, and consists only of smooth points of $B_{C,z}^{eg}$. The tangent cone $T_0 B_{C,z}^{eg}$ (defined as the limit of the tangent spaces at points of $B_{C,z}^{eg,im}$) can be identified with $J_{C,z}^{\mathrm{cond}}/J_{C,z}$, where $J_{C,z}^{\mathrm{cond}} \subset \mathcal{O}_{C,z}$ is the conductor ideal. The subset $B_{C,z}^{eg,nod} \subset B_{C,z}^{eg}$ that parameterizes the nodal deformations is also open and dense. Furthermore, $B_{C,z}^{eg} \setminus B_{C,z}^{eg,nod}$ is the closure of three codimension-one strata: $B_{C,z}^{eg}(A_2)$ that parameterizes deformations with one cusp A_2 and $\delta(C, z) - 1$ nodes, $B_{C,z}^{eg}(A_3)$ that parameterizes deformations with one tacnode A_3 and $\delta(C, z) - 2$ nodes, and $B_{C,z}^{eg}(D_4)$ that parameterizes deformations with one ordinary triple point D_4 and $\delta(C, z) - 3$ nodes.

If $C \subset \Sigma$ is a curve with isolated singularities, we consider the joint semiuniversal deformation for all singular points of C. The base of this deformation, the equigeneric locus, and the tangent cone to this locus at the point corresponding to C are as follows:

$$B_C = \prod_{z\in \mathrm{Sing}\,(C)} B_{C,z}, \quad B_C^{eg} = \prod_{z\in \mathrm{Sing}\,(C)} B_{C,z}^{eg}, \quad T_0 B_C^{eg} = \prod_{z\in \mathrm{Sing}\,(C)} T_0 B_{C,z}^{eg} \ .$$

Lemma 13. *Let $\nu : \mathbb{P}^1 \to \Sigma$ be birational onto its image $C = \nu(\mathbb{P}^1)$. Assume that $C \in |D|$, where D is a divisor class such that $r = -DK_\Sigma - 1 > 0$. Let $\boldsymbol{p}$ be an r-tuple of distinct points of $\mathbb{P}^1$ such that $\boldsymbol{w} = \nu(\boldsymbol{p})$ is an r-tuple of distinct nonsingular points of C. Let $|D|_{\boldsymbol{w}} \subset |D|$ be the linear subsystem of curves passing through $\boldsymbol{w}$, and $\Lambda(\boldsymbol{w}) \subset B_C$ be the natural image of $|D|_{\boldsymbol{w}}$.*

(1) *One has $\mathrm{codim}_{B_C}\Lambda(\boldsymbol{w}) = \dim B_C^{eg}$, and $\Lambda(\boldsymbol{w})$ intersects $T_0B_C^{eg}$ transversally.*

(2) *For any r-tuple $\widetilde{\boldsymbol{w}} \in \Sigma^r$ sufficiently close to $\boldsymbol{w}$ and such that $\Lambda(\widetilde{\boldsymbol{w}})$ intersects B_C^{eg} transversally and only at smooth points, the natural map from the germ $\mathcal{M}_{0,r}(\Sigma, D)_{[\nu,\boldsymbol{p}]}$ of $\mathcal{M}_{0,r}(\Sigma, D)$ at $[\nu : \mathbb{P}^1 \to \Sigma, \boldsymbol{p}]$ to B_C^{eg} gives rise to a bijection between the set of elements $[\widetilde{\nu} : \mathbb{P}^1 \to \Sigma, \widetilde{\boldsymbol{p}}] \in \mathcal{M}_{0,r}(\Sigma, D)_{[\nu,\boldsymbol{p}]}$ such that $\widetilde{\nu}(\widetilde{\boldsymbol{p}}) = \widetilde{\boldsymbol{w}}$ on one side and the set $\Lambda(\widetilde{\boldsymbol{w}}) \cap B_C^{eg}$ on the other side.*

Proof. (1) The dimension and the transversality statements reduce to the fact that the pull-back of $T_0B_C^{eg}$ to $|D|$ intersects $|D|_{\boldsymbol{w}}$ transversally and only at one point. In view of the identification of $T_0B_C^{eg}$ with $\prod_{z\in \mathrm{Sing}\,(C)} J_{C,z}^{\mathrm{cond}}/J_{C,z}$ [6, Theorem 4.15], both required claims read as

$$H^0(C, \mathcal{J}_C^{\mathrm{cond}}(-\boldsymbol{w}) \otimes \mathcal{O}_\Sigma(D)) = 0 \ , \tag{19}$$

where $\mathcal{J}_C^{\mathrm{cond}} = \mathrm{Ann}(\nu_*\mathcal{O}_{\mathbb{P}^1}/\mathcal{O}_C)$ is the conductor ideal sheaf, since $\mathcal{J}_C^{\mathrm{cond}}$ can be equivalently regarded as the ideal sheaf of the zero-dimensional subscheme of C defined at all singular points $z \in \mathrm{Sing}\,(C)$ by the conductor ideals $J_{C,z}^{\mathrm{cond}} = \mathrm{Ann}(\nu_* \bigoplus_{q\in\nu^{-1}(z)} \mathcal{O}_{\mathbb{P}^1,q})/\mathcal{O}_{C,z}$.

It is known that $\mathcal{J}_C^{\mathrm{cond}} = \nu_*\mathcal{O}_{\mathbb{P}^1}(-\Delta)$, where $\Delta \subset \mathbb{P}^1$ is the so-called double-point divisor, whose degree is $\deg \Delta = 2\sum_{z\in \mathrm{Sing}\,(C)} \delta(C, z)$ (see, for example, [5, Section 2.4] or [8, Section 4.2.4]). Hence, the relations (19) can be rewritten as

$$H^0(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(\boldsymbol{d} - \Delta - \boldsymbol{p})) = 0 \ , \tag{20}$$

where $\deg \boldsymbol{d} = D^2$. Since

$$\begin{aligned}\deg \mathcal{O}_{\mathbb{P}^1}(\boldsymbol{d} - \Delta - \boldsymbol{p})) &= D^2 - 2 \sum_{z\in \mathrm{Sing}\,(C)} \delta(C, z) - r \\ &= D^2 - 2\left(\frac{D^2 + DK_\Sigma}{2} + 1\right) - (-DK_\Sigma - 1) = -1 > -2 \ ,\end{aligned}$$

we obtain $H^1(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(\boldsymbol{d} - \Delta - \boldsymbol{p})) = 0$, and hence by Riemann–Roch

$$\dim H^0(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(\boldsymbol{d} - \Delta - \boldsymbol{p})) = \deg \mathcal{O}_{\mathbb{P}^1}(\boldsymbol{d} - \Delta - \boldsymbol{p})) + 1 = 0 \ .$$

(2) The second statement of Lemma 13 immediately follows from the first one due to the fact that the tangent spaces to the stratum B_C^{eg} at its smooth

points close to the origin converge to the same linear space of dimension $\dim B^{eg}_C$ [6, Theorem 4.15]. $\square$

Suppose now that Σ possesses a real structure, C is a real curve, and z is its real singular point. Let $b \in B^{eg,im}_{C,z}$ be a real point, and let C_b be the corresponding fiber of the semiuniversal deformation of the germ (C, z). Define the Welschinger sign W_b as follows. Let $\pi : \hat{C}_b \to C_b \hookrightarrow \Sigma$ be the normalization of C_b. Here $\hat{C}_b$ is the union of discs, some of them being real (*i.e.*, invariant with respect to the complex conjugation), the others forming complex conjugate pairs. Put $W_b = (-1)^{C_{b,+} \circ C_{b,-}}$, where $C_{b,\pm} = \pi(\hat{C}_{b,\pm})$ and $\hat{C}_b \setminus \mathbb{R}\hat{C}_b = \hat{C}_{b,+} \sqcup \hat{C}_{b,-}$ is a splitting into disjoint complex conjugate halves.

Lemma 14. *The Welschinger sign W_b is equal to $(-1)^s$, where s is the number of solitary nodes in a small real nodal perturbation of C_b.*

Proof. Straightforward from the definition. $\square$

Lemma 15. *Let L_t, $t \in (-\varepsilon, \varepsilon) \subset \mathbb{R}$, be a smooth one-parameter family of conjugation-invariant affine subspaces of $B_{C,z}$ of dimension $\delta(C, z)$ such that*

- *L_0 passes through the origin and is transversal to $T_0 B^{eg}_{C,z}$,*
- *$L_t \cap B^{eg}_{C,z} \subset B^{eg,im}_{C,z}$ for each $t \in (-\varepsilon, \varepsilon) \setminus \{0\}$.*

Then,

(i) *the intersection $L_t \cap B^{eg}_{C,z}$ is finite for each $t \in (-\varepsilon', \varepsilon') \setminus \{0\}$, where $\varepsilon' > 0$ is sufficiently small.*

(ii) *the function $W(t) = \sum_{b \in L_t \cap \mathbb{R}B^{eg}_{C,z}} W_b$ is constant in $(-\epsilon', \epsilon') \setminus \{0\}$, where $\varepsilon' > 0$ is sufficiently small.*

Proof. The finiteness of the intersection follows from the transversality of L_0 and $T_0 B^{eg}_{C,z}$ in $B_{C,z}$. To prove the second statement, assume, first, that the germ (C, z) represents an ordinary cusp A_2. Then $\mathbb{R}B_{C,z} = (\mathbb{R}^2, 0)$ and $\mathbb{R}B^{eg}_{C,z}$ is a semicubical parabola with vertex at the origin. For the points b belonging to one of the two connected components of $\mathbb{R}B^{eg}_{C,z} \setminus \{0\}$, the curve C_b has a non-solitary real node; for the points b from the other component, C_b has a solitary node. Since, in addition, the line L_0 crosses the tangent to the parabola at the origin transversally we have $W(t) = 0$ for each $t \in (-\varepsilon', \varepsilon') \setminus \{0\}$ for sufficiently small $\varepsilon' > 0$.

In the general case, if $\varepsilon' > 0$ is sufficiently small, then for any two points $t_1 < t_2$ in $(-\epsilon', \epsilon') \setminus \{0\}$ we can connect L_{t_1} with L_{t_2} by a family of $\delta(C, z)$-dimensional conjugation-invariant affine subspaces $L'_t \subset B_{C,z}$, $t \in [t_1, t_2]$, such that

- the subspaces L'_t, $t \in [t_1, t_2]$, are transversal to $B^{eg}_{C,z}$,
- the intersection number of L'_t and $B^{eg}_{C,z}$ is constant in $[t_1, t_2]$,
- for all but finitely many values of t the intersection $L'_t \cap B^{eg}_{C,z}$ is contained in $B^{eg,nod}_{C,z}$, and for the remaining values of t, the subspace L'_t intersects $\mathbb{R}B^{eg}_{C,z}$ within $B^{eg}_{C,z}(A_2) \cup B^{eg}_{C,z}(A_3) \cup B^{eg}_{C,z}(D_4)$.

The bifurcations through the immersed singularities A_3 and D_4 do not affect $W(t)$, as well as the cuspidal bifurcation, which we have treated above. $\square$

Remark 16. In fact, Lemma 15 allows one to extend the definition of Welschinger signs and attribute a *Welschinger weight* to any map $\nu : \mathbb{P}^1 \to \Sigma$ birational onto its image.

3. Proof of Theorem 6

3.1. Preliminary observations

We start with two remarks.

(1) If Y is an irreducible complex variety, equipped with a real structure, and $\mathbb{R}Y$ contains nonsingular points of Y, then $\mathbb{R}Y \cap U \neq \emptyset$ for any Zariski open subset $U \subset Y$. In particular, a generic element of $\mathcal{P}_{r,m}(\Sigma, F)$ is generic in Σ^r.

(2) By blowing up extra real points we can reduce the consideration to the case of del Pezzo surfaces of degree 1.

The following statement will be used in the sequel.

Lemma 17. *Let $t \in (\mathbb{R}, 0) \mapsto \Sigma_t$ be a germ of an elementary deformation $(\Sigma_t, D_t, F_t, \varphi_t, \boldsymbol{w}_t)$ of a tuple $(\Sigma_0, D_0, F_0, \varphi_0, \boldsymbol{w}_0)$, where Σ_0 is a del Pezzo surface of degree* 1*, $D_0 \in \mathrm{Pic}(\Sigma_0)$ is a real effective divisor such that $r = -D_0 K_{\Sigma_0} - 1 > 0$, and $\boldsymbol{w}_0$ belongs to $\mathcal{P}_{r,m}(\Sigma_0, F_0)$ and is generic. Then*

$$W_m(\Sigma_t, D_t, F_t, \varphi_t, \boldsymbol{w}_t) = W_m(\Sigma_0, D_0, F_0, \varphi_0, \boldsymbol{w}_0) \ .$$

Proof. Since $D_0 K_{\Sigma_0} > 1$ and $\boldsymbol{w}_0$ is generic, Lemma 9 implies that all the curves under count are immersed. Thus, each of these curves contributes 1 to the Gromov–Witten invariant, and the required equality follows from Lemma 14. $\square$

3.2. Proof of Proposition 4

The only situation to consider is the one where $\Sigma \in \mathcal{D}_1^{\mathrm{DP}}$ and $r = -DK_\Sigma - 1 > 0$. Due to Lemma 17, we can fix any dense subset in $\mathcal{D}_1^{\mathrm{DP}}$ and check the statement for the surfaces belonging to this subset. Throughout this section, we assume that $\Sigma \in U_1 \cap U_2$.

We prove the invariance of Welschinger numbers by studying wall-crossing events when moving either one real point of the given collection, or a pair of complex conjugate points.

3.2.1. Moving a real point of configuration. Suppose that $2m < r$. Let tuples $\boldsymbol{w}' \cup \{w^{(0)}\}, \boldsymbol{w}' \cup \{w^{(1)}\} \in \mathcal{P}_{r,m}(\Sigma, F)$, where $\boldsymbol{w}' \in \mathcal{P}_{r-1,m}(\Sigma, F)$, be such that the sets $\mathcal{R}(\Sigma, D, F, \boldsymbol{w}' \cup \{w^{(0)}\})$ and $\mathcal{R}(\Sigma, D, F, \boldsymbol{w}' \cup \{w^{(1)}\})$ are finite and presented by immersions (see Lemma 9). We prove that

$$W_m(\Sigma, D, F, \varphi, \boldsymbol{w}' \cup \{w^{(0)}\}) = W_m(\Sigma, D, F, \varphi, \boldsymbol{w}' \cup \{w^{(1)}\}) \ . \tag{21}$$

Due to Lemma 11, by a small deformation of $\boldsymbol{w}'$ we can reach the following: whenever an element $[\nu : \widehat{C} \to \Sigma] \in \overline{\mathcal{M}_{0,0}^{br}}(\Sigma, D) \setminus \mathcal{M}_{0,0}^{br}(\Sigma, D)$ is such that $\nu(\widehat{C}) \supset$

$\boldsymbol{w}'$, the element $[\nu : \widehat{C} \to \Sigma]$ satisfies the conditions of Lemma 11(i), $-D_1K_\Sigma - 1$ points of $\boldsymbol{w}'$ lie on $C_1 \setminus (\mathrm{Sing}\,(C_1) \cup C_2)$, and the remaining $-D_2K_\Sigma - 1$ points of $\boldsymbol{w}'$ lie on $C_2 \setminus (\mathrm{Sing}\,(C_2) \cup C_1)$.

There exists a smooth real-analytic path $\sigma : [0,1] \to F$ lying in the real part of some smooth real algebraic curve $\sigma(\mathbb{C}) \subset \Sigma$, such that σ is disjoint from all the points of $\boldsymbol{w}'$, $\sigma(0) = w^{(0)}$, $\sigma(1) = w^{(1)}$, and in the family

$$\overline{\mathcal{M}^{br}_{0,r}}(\Sigma, D; \boldsymbol{w}', \sigma) = \{[\nu : \hat{C} \to \Sigma, \boldsymbol{p}] \in \overline{\mathcal{M}^{br}_{0,r}}(\Sigma, D) \ : \ \nu(\boldsymbol{p}') = \boldsymbol{w}', \ \nu(p_r) \in \sigma\} \ ,$$

where $\boldsymbol{p}' = \boldsymbol{p} \setminus \{p_r\}$, all but finitely many elements belong to $\mathcal{M}^{im}_{0,r}(\Sigma, D)$, and the remaining elements $[\nu : \hat{C} \to \Sigma, \boldsymbol{p}]$ (corresponding to some values $t \in I_0 \subset [0,1]$, $|I_0| < \infty$) are such that:

($\mathrm{D1}_{re}$) either $[\nu : \hat{C} \to \Sigma] \in \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$ is as in Lemma 11(i), the point $w^{(t)} \in \sigma \cap C_2$ belongs to $C_2 \setminus (\mathrm{Sing}\,(C_2) \cup C_1 \cup \boldsymbol{w}')$, and the germ of $\sigma(\mathbb{C})$ at $w^{(t)} \in C$ intersects C_2 transversally;

($\mathrm{D2}_{re}$) or $[\nu : \hat{C} \to \Sigma] \in \mathcal{M}^{br}_{0,0}(\Sigma, D) \setminus \mathcal{M}^{im}_{0,0}(\Sigma, D)$, the point $w^{(t)} \in \sigma \cap C$, where $C = \nu(\hat{C})$, belongs to $C \setminus (\mathrm{Sing}\,(C) \cup \boldsymbol{w}')$, and the germ of $\sigma(\mathbb{C})$ at $w^{(t)}$ intersect C transversally.

Denote by $M_{[\nu,\boldsymbol{p}]}$ the germ of $\overline{\mathcal{M}^{br}_{0,r}}(\Sigma, D; \boldsymbol{w}', \sigma)$ at an element $[\nu : \hat{C} \to \Sigma, \boldsymbol{p}]$.

If $[\nu : \hat{C} \to \Sigma] \in \mathcal{M}^{im}_{0,0}(\Sigma, D)$, or $[\nu : \hat{C} \to \Sigma]$ satisfies condition ($\mathrm{D1}_{re}$), then, by Lemma 12, the germ $M_{[\nu,\boldsymbol{p}]}$ is diffeomorphically mapped by Ev onto the germ $(\sigma, w^{(t)})$. Moreover, the Welschinger sign $\mu(\nu, \varphi)$ does not change along $M_{[\nu,\sigma]}$. This is evident for $[\nu : \hat{C} \to \Sigma] \in \mathcal{M}^{im}_{0,0}(\Sigma, D)$, and, under condition ($\mathrm{D1}_{re}$), immediately follows from the fact that ν maps the germ of $\hat{C}$ at the node to a pair of real smooth branches that intersect transversally and undergo a standard smoothing in the considered bifurcation.

Under the hypotheses of condition ($\mathrm{D2}_{re}$), the required constancy of the Welschinger number $W_m(\Sigma, D, F, \varphi, \boldsymbol{w}' \cup \{w^{(t)}\})$ immediately follows from Lemmas 13, 14 and 15.

3.2.2. Moving a pair of imaginary conjugate points. Assume that $m \geq 1$. Let tuples $\boldsymbol{w}' \cup \{w^{(0)}, \mathrm{Conj}\, w^{(0)}\}, \boldsymbol{w}' \cup \{w^{(1)}, \mathrm{Conj}\, w^{(1)}\} \in \mathcal{P}_{r,m}(\Sigma, F)$, where $\boldsymbol{w}' \in \mathcal{P}_{r-2,m-1}(\Sigma, F)$, be such that the sets

$$\mathcal{R}(\Sigma, D, F, \boldsymbol{w}' \cup \{w^{(0)}, \mathrm{Conj}\, w^{(0)}\}) \ \text{ and } \ \mathcal{R}(\Sigma, D, F, \boldsymbol{w}' \cup \{w^{(1)}, \mathrm{Conj}\, w^{(1)}\})$$

are finite and presented by immersions (see Lemma 9). We prove that

$$W_m(\Sigma, D, F, \varphi, \boldsymbol{w}' \cup \{w^{(0)}, \mathrm{Conj}\, w^{(0)}\}) = W_m(\Sigma, D, F, \varphi, \boldsymbol{w}' \cup \{w^{(1)}, \mathrm{Conj}\, w^{(1)}\})\,. \tag{22}$$

Due to Lemma 11, by a small deformation of $\boldsymbol{w}'$ we can reach the following: for any point w of a certain Zariski open subset $\Sigma_{\boldsymbol{w}'} \subset \Sigma \setminus \boldsymbol{w}'$, whenever for an element $[\nu : \widehat{C} \to \Sigma] \in \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D) \setminus \mathcal{M}^{br}_{0,0}(\Sigma, D)$ we have $\nu(\widehat{C}) \supset \boldsymbol{w}' \cup \{w\}$, this element $[\nu : \widehat{C} \to \Sigma]$ satisfies the conditions of Lemma 11(i), $-D_1K_\Sigma - 1$ points of $\boldsymbol{w}'$ lie on $C_1 \setminus (\mathrm{Sing}\,(C_1) \cup C_2)$, and the remaining $-D_2K_\Sigma - 2$ points of $\boldsymbol{w}'$ and the point w lie on $C_2 \setminus (\mathrm{Sing}\,(C_2) \cup C_1)$. Further on, assuming this property of $\boldsymbol{w}'$, we can find a smooth real-analytic path $\sigma : [0,1] \to \mathrm{Sing} \setminus \mathbb{R}\Sigma$ lying in some smooth real algebraic curve $\sigma(\mathbb{C}) \subset \Sigma \setminus \mathbb{R}\Sigma$, such that σ starts at $w^{(0)}$ and ends up at $w^{(1)}$, avoids all the points of $\boldsymbol{w}'$, and satisfies the following condition (*cf.* section 3.2.1): for all but finitely many elements of the family

$$\overline{\mathcal{M}^{br}_{0,r}}(\Sigma, D; \boldsymbol{w}', \sigma, \mathrm{Conj}\,\sigma) = \{[\nu : \hat{C} \to \Sigma, \boldsymbol{p}] \in \overline{\mathcal{M}^{br}_{0,r}}(\Sigma, D) \ : \\ \nu(\boldsymbol{p}') = \boldsymbol{w}',\ \nu(p_{r-1}) \in \sigma,\ \nu(p_r) \in \mathrm{Conj}\,\sigma\}\ ,$$

where $\boldsymbol{p}' = \boldsymbol{p} \setminus \{p_{r-1}, p_r\}$, we have $[\nu : \hat{C} \to \Sigma] \in \mathcal{M}^{im}_{0,0}(\Sigma, D)$, while the remaining elements (which correspond to some values $t \in I_0 \subset [0,1]$, $|I_0| < \infty$) are such that:

($\mathrm{D1}_{im}$) either $[\nu : \hat{C} \to \Sigma] \in \overline{\mathcal{M}^{br}_{0,0}}(\Sigma, D)$ is as in Lemma 11(i), where $\nu_i : \hat{C}_i \to \Sigma$ commutes with the real structure, $-D_1K_\Sigma - 1$ points of $\boldsymbol{w}'$ lie on $C_1 \setminus (\mathrm{Sing}\,(C_1) \cup C_2)$, the remaining $-D_2K_\Sigma - 2$ points of $\boldsymbol{w}'$ lie in $C_2 \setminus (\mathrm{Sing}\,(C_2) \cup C_1)$, the point $w^{(t)} \in \sigma$ belongs to $C_2 \setminus (\mathrm{Sing}\,(C_2) \cup C_1 \cup \boldsymbol{w}')$, and the germ of $\sigma(\mathbb{C})$ at $w^{(t)} \in C_2$ intersects C_2 transversally;

($\mathrm{D2}_{im}$) or $[\nu : \hat{C} \to \Sigma] \in \mathcal{M}^{br}_{0,0}(\Sigma, D) \setminus \mathcal{M}^{im}_{0,0}(\Sigma, D)$, the point $w^{(t)} \in \sigma \cap C$, where $C = \nu(\hat{C})$, belongs to $C \setminus (\mathrm{Sing}\,(C) \cup \boldsymbol{w}')$, and the germ of $\sigma(\mathbb{C})$ at $w^{(t)}$ intersects C transversally.

Notice that, in ($\mathrm{D1}_{im}$), the case of $C_1 = C_2$ is not relevant due to $-DK_\Sigma > 2$, and the case of complex conjugate C_1 and C_2 does not occur either, since any real rational curve in $|D|$ must have a non-trivial one-dimensional real branch (see Section 1.2).

Then the proof of (22) literally follows the argument of the preceding section.

3.3. Proof of Proposition 5 and Theorem 6

In view of Proposition 4 and Lemmas 9(ii) and 17, Theorem 6 follows from Proposition 5, and, in its turn, to prove Proposition 5 it is sufficient to check the constancy of the Welschinger number in the following families:

- a germ of elementary deformation $\{\Sigma_t\}_{t \in (\mathbb{R},0)}$, where $\Sigma_0 \in U_1(A_1)$, $\Sigma_t \in U_1 \cap U_2$ for each $t \ne 0$, and $D_t \ne -K_{\Sigma_t}$;
- a germ of elementary deformation $\{\Sigma_t\}_{t \in (\mathbb{R},0)}$, where $\Sigma_0 \in \mathcal{D}^{\mathrm{DP}}_1 \setminus U_1$, $\Sigma_t \in U_1 \cap U_2$ for each $t \ne 0$, and $D_t = -K_{\Sigma_t}$.

Let $\Sigma_0 \in U_1(A_1)$, $\Sigma_t \in U_1 \cap U_2$ for $t \ne 0$, and $D_t \ne -K_{\Sigma_t}$. Extend the family $\{\Sigma_t\}_{t \in (\mathbb{R},0)}$ to a conjugation invariant family $\{\Sigma_t\}_{(\mathbb{C},0)}$. By Lemma 9(2), there exists $\boldsymbol{w}_0 \in \mathcal{P}_{r,m}(\Sigma_0, F_0)$ such that, for any $k \ge 0$, all elements $[\nu : \mathbb{P}^1 \to \Sigma_0, \boldsymbol{p}_0] \in \mathcal{M}_{0,r}(\Sigma_0, D - kE, \boldsymbol{w}_0)$ satisfy the properties indicated in Lemma 9(2ii). These elements appear only for a finite number of values of k and form a finite set. Let us associate with each of them a comb-like curve $[\nu : \hat{C} \to \Sigma_0, \boldsymbol{p}] \in \overline{\mathcal{M}_{0,r}(\Sigma_0, D, \boldsymbol{w}_0)}$ such that:

- either $\hat{C} \simeq \mathbb{P}^1$, or $\hat{C} = \hat{C}' \cup \hat{E}_1 \cup \cdots \cup \hat{E}_k$ for some $k > 0$, where $\hat{C}' \simeq \hat{E}_1 \simeq \cdots \simeq \hat{E}_k \simeq \mathbb{P}^1$, $\hat{E}_i \cap \hat{E}_j = \emptyset$ for all $i \neq j$, and $\#(\hat{C}' \cap \hat{E}_i) = 1$ for all $i = 1, \ldots, k$;
- $\boldsymbol{p} \subset \hat{C}'$ and $[\nu : \hat{C}' \to \Sigma_0, \boldsymbol{p}] \in \mathcal{M}^{im}_{0,r}(\Sigma_0, D - kE, \boldsymbol{w}_0)$, and each of $\hat{E}_1, \ldots, \hat{E}_k$ is isomorphically mapped onto E.

Then, complement $\boldsymbol{w}_0$ to a conjugation invariant family of r-tuples $\boldsymbol{w}_t \in (\Sigma_t)^r$, $t \in (\mathbb{C}, 0)$, so that $\boldsymbol{w}_t \in \mathcal{P}_{r,m}(\Sigma_t, F_t)$ for each real t. It follows from [26, Theorem 4.2] that each of the introduced elements $[\nu : \hat{C} \to \Sigma_0, \boldsymbol{p}] \in \overline{\mathcal{M}_{0,r}(\Sigma_0, D, \boldsymbol{w}_0)}$ extends to a smooth family $[\nu_t : \hat{C}_t \to \Sigma_t, \boldsymbol{p}_t] \in \overline{\mathcal{M}^{br}_{0,r}}(\Sigma_t, D, \boldsymbol{w}_t)$, $t \in (\mathbb{C}, 0)$, where $\hat{C}_t \simeq \mathbb{P}^1$ and ν_t is an immersion for all $t \neq 0$, and, furthermore, each element of $\mathcal{M}_{0,r}(\Sigma_t, D, \boldsymbol{w}_t)$, $t \in (\mathbb{C}, 0) \setminus \{0\}$ is included into some of the above families. Thus, the Welschinger number $W(\Sigma_t, D, F_t, \varphi_t, \boldsymbol{w}_t)$ remains constant in $t \in (\mathbb{R}, 0) \setminus \{0\}$, since the only change of the topology in the real loci of the curves under the count consists in smoothing of non-solitary nodes, while the difference between the homology classes of the halves $[C_\pm(t)]$ in $H_2(\Sigma_t, F_t; \mathbb{Z}/2) = H_2(\Sigma_0, F_0; \mathbb{Z}/2)$ with $t < 0$ and those with $t > 0$ belongs to $(1 + \mathrm{Conj}_*) H_2(\Sigma_0, F_0; \mathbb{Z}/2)$ and, hence $[C_\pm(t)] \circ \phi_t$ does not depend on t.

Assume that $\Sigma_0 \in \mathcal{D}^{\mathrm{DP}}_1 \setminus U_1$ $\Sigma_t \in U_1 \cap U_2$ for $t \neq 0$, and $D_t = -K_{\Sigma_t}$. In this case we deal with a family of real elliptic pencils $|-K_{\Sigma_t}|$, $t \in (\mathbb{R}, 0)$, such that the central one $|-K_{\Sigma_0}|$ has a real cuspidal curve $C_0 \in |-K_{\Sigma_0}|$ and, otherwise, the family is generic. As it can be seen from the local Weierstrass normal form, due to the above genericity the image of $|-K_{\Sigma_0}|$ in the base $(\mathbb{C}^2, 0)$ of the versal deformation of the cuspidal point intersects the tangent space to the discriminant locus, that is the cusp curve $27p^2 + 4q^3 = 0$ in terms of Weierstrass coordinates p, q, transversally at one point. Therefore, for $t \in (\mathbb{R}, 0)$ on one side of $t = 0$ the singular curves in $|-K_{\Sigma_t}|$ close to C_0 form a pair of complex conjugate curves, while for $t \in (\mathbb{R}, 0)$ on the opposite side of $t = 0$ they are real, one with a solitary node, and the other one with a cross point. Thus, the total Welschinger number is the same on the both sides.

Acknowledgement

An essential part of the work was done during the authors' stay at the Centre Interfacultaire Bernoulli at EPFL in the framework of the program "Tropical geometry in its complex and symplectic aspects". The authors are grateful to CIB for the support and excellent working conditions. The third author has also been supported from the German-Israeli Foundation grant no. 1174-197.6/2011 and from the Hermann-Minkowski-Minerva Center for Geometry at the Tel Aviv University. We are grateful to G.-M. Greuel and I. Tyomkin for helpful discussions of certain aspects of deformations of curves.

References

[1] Araujo, C., and Kollár, J.: Rational Curves on Varieties. *Higher Dimensional Varieties and Rational Points* (*Budapest,* 2001), Bolyai Society Mathematical Studies **12**, Springer, Berlin, 2003, pp. 13–68.

[2] Arroyo, A., Brugallé, E., and Lopéz de Medrano, L.: Recursive formulas for Welschinger invariants of the projective plane. *IMRN* **2011**, no. 5, 1107–1134.

[3] Bochnak, J., Coste, M., Roy, M.-F.: *Géométrie algébrique réelle.* Springer-Verlag, 1987.

[4] Brusotti, L.: Sulla "piccola variazione" di una curva piana algebrica reali. *Rend. Rom. Ac. Lincei* (5), **30** (1921), 375–379.

[5] Caporaso, L., and Harris, J.: Parameter spaces for curves on surfaces and enumeration of rational curves. *Compositio Math.* **113** (1998), no. 2, 155–208.

[6] Diaz, S., and Harris, J.: Ideals associated to deformations of singular plane curves. *Trans. Amer. Math. Soc.* **309** (1988), no. 2, 433–468.

[7] Degtyarev, A., Itenberg, I., and Kharlamov, V.: *Real Enriques surfaces.* Lect. Notes Math., vol. 1746, Springer-Verlag, Berlin Heidelberg, 2000.

[8] Dolgachev, I. V.: *Classical algebraic geometry: a modern view.* Cambridge Univ. Press, 2013.

[9] Fulton, W., and Pandharipande, R.: Notes on stable maps and quantum cohomology. *Algebraic geometry*, Santa Cruz 1995, Proc. Sympos. Pure Math, vol. 62, Amer. Math. Soc., Providence, RI (1997), 45–96.

[10] Göttsche, L., and Pandharipande, R.: The quantum cohomology of blow-ups of $\mathbb{P}^2$ and enumerative geometry. *J. Differential Geom.* **48** (1998), no. 1, 61–90.

[11] Greuel, G.-M., and Karras, U.: Families of varieties with prescribed singularities. *Compositio Math.* **69**, 83–110 (1989).

[12] Horev, A., and Solomon, J. P.: *The open Gromov–Witten–Welschinger theory of blowups of the projective plane.* Preprint at arXiv:1210.4034.

[13] Hartshorne, R.: *Algebraic Geometry.* Springer, NY, 1977. *Lect. Notes Math.* **1124** (1985), 98–131.

[14] Itenberg, I., Kharlamov, V., and Shustin, E.: Welschinger invariant and enumeration of real rational curves. *International Math. Research Notices* **49** (2003), 2639–2653.

[15] Itenberg, I., Kharlamov, V., and Shustin, E.: *Appendix to "Welschinger invariant and enumeration of real rational curves"*. Preprint at arXiv.math/0312142.

[16] Itenberg, I., Kharlamov, V., and Shustin, E.: Welschinger invariants of real del Pezzo surfaces of degree ≥ 3. *Math. Annalen* **355** (2013), no. 3, 849–878.

[17] Itenberg, I., Kharlamov, V., and Shustin, E.: Welschinger invariants of real del Pezzo surfaces of degree ≥ 2. *International J. Math.* **26** (2015), no. 6. DOI: 10.1142/S0129167X15500603.

[18] Itenberg, I., Kharlamov, V., and Shustin, E.: *Relative enumerative invariants of real nodal del Pezzo surfaces*, Preprint at arXiv:1611.02938.

[19] Kollár, J.: *Rational curves on algebraic varieties.* Springer, Berlin, 1996.

[20] Kontsevich, M., and Manin, Yu.: Gromov–Witten classes, quantum cohomology, and enumerative geometry. *Comm. Math. Phys.* **164** (1994), no. 3, 525–562.

[21] Mikhalkin, G.: Enumerative tropical algebraic geometry in $\mathbb{R}^2$. *J. Amer. Math. Soc.* **18** (2005), 313–377.

[22] Sernesi, E.: *Deformations of Algebraic Schemes.* Springer, 2006.

[23] Severi, F.: *Vorlesungen über algebraische Geometrie.* Teubner, 1921 (reprinted by Johnson, 1968).

[24] Shustin, E.: On higher genus Welschinger invariants of del Pezzo surfaces. *IMRN* **2015**, 6907–6940. DOI: 10.1093/imrn/rnu148.

[25] Tannenbaum, A.: Families of algebraic curves with nodes. *Compositio Math.* **41** (1980), 107–126.

[26] Vakil, R.: Counting curves on rational surfaces. *Manuscripta Math.* **102** (2000), 53–84.

[27] Welschinger, J.-Y.: Invariants of real rational symplectic 4-manifolds and lower bounds in real enumerative geometry. *C. R. Acad. Sci. Paris, Sér. I,* **336** (2003), 341–344.

[28] Welschinger, J.-Y.: Invariants of real symplectic 4-manifolds and lower bounds in real enumerative geometry. *Invent. Math.* **162** (2005), no. 1, 195–234.

Ilia Itenberg
Institut de Mathématiques de Jussieu – Paris Rive Gauche
Université Pierre et Marie Curie
4 place Jussieu
75252 Paris Cedex 5, France

and

Département de mathématiques et applications
Ecole Normale Supérieure
45 rue d'Ulm
75230 Paris Cedex 5, France
e-mail: ilia.itenberg@imj-prg.fr

Viatcheslav Kharlamov
Université de Strasbourg et IRMA
7, rue René Descartes
67084 Strasbourg Cedex, France
e-mail: kharlam@math.unistra.fr

Eugenii Shustin
School of Mathematical Sciences
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv
69978 Tel Aviv, Israel
e-mail: shustin@post.tau.ac.il

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 261–285

Some Results on Amoebas and Coamoebas of Affine Spaces

Petter Johansson

Abstract. We give some topological characteristics of the coamoeba of a generic k-dimensional affine space and two stronger versions, specific for the affine case, of a result by Nisse, Sottile and the author. We also give topological and partly algebraical characterizations of the amoeba and coamoeba in some special cases: $k = n - 1$, $k = 1$ and, when n is even, $k = n/2$, in the last case with a certain emphasis on the example $n = 4$.

1. Introduction

The complex n-torus $(\mathbb{C}^*)^n$ is split into $\mathbb{R}^n$ and $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$ under the mappings $\operatorname{Log} z = (\log|z_1|, \ldots, \log|z_n|)$ and $\operatorname{Arg} z = (\arg z_1, \ldots, \arg z_n)$. Given an algebraic variety V in $(\mathbb{C}^*)^n$, its *amoeba* $\mathcal{A}_V$ is the image of V under Log, while its *coamoeba* $\mathcal{A}'_V$ is the image of V under Arg. The amoeba was introduced by Gelfand, Kapranov and Zelevinsky in 1994, see [3].

Mikael Passare was one of the first mathematicians to study the amoeba. As a co-author of, e.g., [2] and [11], he played a key role in awakening the interest for this concept. Furthermore he coined the term coamoeba on which he held a seminar in early 2004, see [5].

During his last couple of years, Mikael Passare initiated a project to understand the amoebas and coamoebas of *affine spaces* in $(\mathbb{C}^*)^n$, that is, algebraic varieties that can be defined by linear equations. The plan was a common paper on the subject by Mounir Nisse, Passare and me. After Passare's premature death, Nisse and I chose to write two separate papers instead, since we approached the subject from two different angles. Thus, the papers remain complementary, although they both include ideas by Passare. For Nisse's paper, see [9].

As an introduction to the subject of this work, we will describe the amoebas and coamoebas of hyperplanes in $(\mathbb{C}^*)^n$, see Section 3. Hyperplanes are not only affine spaces but also hypersurfaces, for which the amoeba and coamoeba has been studied to a relatively large extent. Thus, the description of the amoeba we give in

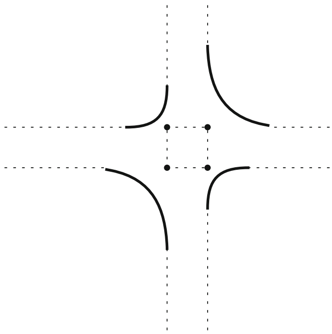


FIGURE 1. We sketch the coamoeba $\mathcal{A}'_f$ defined by the zero set of the polynomial $f = 1+i-z-w+zw$. The dotted line segments mark the parts of the boundary that are not contained in the coamoeba and the solid curves and points mark the parts of the boundary that are contained in the coamoeba and thus are part of the contour, see Section 2. The dotted line segments are contained in four lines on $\mathbb{T}^2$. These lines amount to the four proper initial coamoebas of $\mathcal{A}'_f$, cf. Theorem 1.2. The concave region in the figure corresponds to the interior of $\mathcal{A}'_f$.

this case is just a recollection of an already published result. A description of the closure of the coamoeba of a hyperplane first appeared in [1]. Our version is more or less the same, although we also specify which points on the boundary that are contained in the coamoeba.

The second simplest case is that of lines in $(\mathbb{C}^*)^n$, which we consider in Section 4. Passare suggested a statement about the appearance of the amoeba and coamoeba of a line, on which I have based Theorem 4.4. It should be mentioned that also Kuzvesov has results in this direction, see [6].

A third case of special interest, presented in Section 5, is when n is even and the codimension of the affine space is $n/2$. As we will see in Theorem 5.1, the volume of the coamoeba is then either π^n or zero. Furthermore, there is a local diffeomorphism that maps the interior of the coamoeba onto the interior of the amoeba. In a 4-dimensional case, Passare discovered how a certain inflated tetrahedron could be used to find the multiplicity of this diffeomorphism, and thus the volume of the amoeba. Also Nisse has studied the $n/2$-dimensional case, see [9].

In the previous examples, there are several recurrent properties of the coamoeba $\mathcal{A}'_L$ that do not hold for general varieties. On a topological level, the most striking property is perhaps the following, which we will prove in the final section, cf. Theorem 6.13.

Theorem 1.1. *Let L be a non-degenerate affine space of any codimension. Then there is no open set U intersecting $\partial\mathcal{A}'_L$ for which $U \cap \mathcal{A}'_L = U \cap \overline{\mathcal{A}}'_L$.*

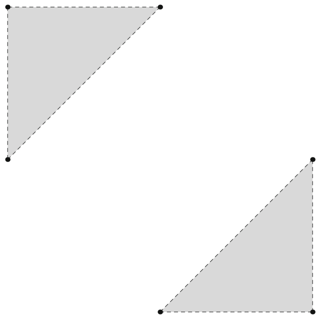


FIGURE 2. We sketch the coamoeba of the line in $(\mathbb{C}^*)^2$ defined by the polynomial $f = 1 + z_1 + z_2$ to exemplify Theorem 1.1 and 1.3. The marked points in the figure are actually only three, each point appearing two times on opposite sides of the square representing $\mathbb{T}^2$. The dashed lines marks the boundary of $\mathcal{A}'_f$, that coincide with the proper initial coamoebas that are defined by the polynomials $1+z_1$, $1+z_2$ and z_1+z_2 respectively. The contour, see Section 2, consists of the three intersection points of the proper initial coamoebas.

Note that for coamoebas of non-affine algebraic varieties, there typically *are* open sets U as in Theorem 1.1, see, e.g., Figure 1. The notion of degeneracy is explained in Section 5. Note also that our definition of $\partial\mathcal{A}'_L$ is not standard for the case of high codimension, see Section 2.

To give a context to Theorem 1.1, let us introduce the concept of *initial coamoebas*. Let f be a Laurent polynomial on $(\mathbb{C}^*)^n$ and $\omega \in \mathbb{R}^n$. Then the *initial form* f_ω is the sum of terms $a_\alpha z^\alpha$ of f such that $\alpha \cdot \omega = \alpha_1\omega_1 + \cdots + \alpha_n\omega_n$ is maximal. More generally, if I is an ideal of Laurent polynomials over $(\mathbb{C}^*)^n$ then the *initial ideal* I_ω of I at ω is the ideal $I_\omega := \langle f_\omega; f \in I\rangle$. The coamoeba of the variety defined by the ideal I_ω is called the *initial coamoeba* of I at ω and denoted by $\mathcal{A}'_\omega$ whenever it is clear which ideal we are considering. How the initial coamoebas affects the appearance of a coamoeba, is indicated by the next theorem.

Theorem 1.2 (Johansson, Nisse, Sottile). *The closure of the coamoeba of an ideal I equals the union of all initial coamoebas of I.*

Notice here that the initial coamoeba at the origin is the coamoeba itself and all other initial coamoebas will be called *proper* initial coamoebas. It is also noteworthy that it suffices to take the union in the theorem over a finite set, see, e.g., [12]. Theorem 1.2 was proved by Sottile and Nisse in [10] and independently for complete intersections by the author in [4].

We will prove two stronger versions of Theorem 1.2 specific for affine spaces: one for high codimensions and one for low codimensions.

Theorem 1.3. *The boundary of the coamoeba of a non-degenerate affine space of codimension at least $n/2$, equals the union of its proper initial coamoebas.*

For the precise meaning of boundary here, see Section 2. The assertion of Theorem 1.3 is very strong: for, e.g., hypersurfaces in dimension $n = 2$, there is only one known non-affine case satisfying the equality, see [4]. If we take away one inclusion in Theorem 1.3 and just demand the boundary of the coamoeba to be contained in the union of its proper initial coamoebas, then we get some additional examples of non-affine varieties fulfilling the assertion, e.g., when the variety is determined by certain discriminants, see [8]. Note that the bent segments of the boundary of the coamoeba in Figure 1, are *not* contained in the union of proper initial coamoebas.

Here comes the second stronger version of Theorem 1.2, specifically for generic affine spaces of low codimension.

Theorem 1.4. *The closure of the coamoeba of a non-degenerate affine space of codimension strictly less than $n/2$, equals the union of its proper initial coamoebas.*

Theorems 1.1, 1.3 and 1.4 has to my knowledge not been conjectured by anyone before, but they resonate well with the results presented in Sections 3–5.

As a step toward the proof of Theorem 1.1 and 1.3, we show in 6.3 some results on the *contour* of the amoeba and coamoeba, see Section 2, that might be of some interest in itself. On the final pages we also discuss an alternative algebraic criterion for degeneracy.

2. Preliminaries

If z is a regular point of V and the Jacobian of Log (or Arg) has full rank at z for some choice of local coordinates on V, then z is a *non-critical point* of Log (Arg) on V. Otherwise, z is a *critical point* of Log (Arg) on V. The contour $\mathcal{C}_V$ of the amoeba $\mathcal{A}_V$ is the set of critical values of Log, and the contour $\mathcal{C}'_V$ of the coamoeba $\mathcal{A}'_V$ is the set of critical values of Arg. Unless the (co)amoeba equals its contour, it is thus given by $\mathcal{C}' \cup U$, where U is a union of real submanifolds of dimension $\min(2 \dim V, n)$.

Let m denote the codimension of V. When $m > n/2$, the *boundary* $\partial\mathcal{A}_V$ of the amoeba is defined as the set $\mathcal{A}_V \backslash U$ and the boundary $\partial\mathcal{A}'_V$ of the coamoeba is defined as $\overline{\mathcal{A}}'_V \backslash U$. When $m \leq n/2$, the amoeba and coamoeba are in general full-dimensional and $\partial\mathcal{A}_V$ and $\partial\mathcal{A}'_V$ refers to the topological boundaries of the amoeba and coamoeba respectively. In Figure 1, the contour of a typical non-affine coamoeba is illustrated.

In several of the main results in this work, we need the distinction between *real* and *non-real* varieties. That V is real means that, possibly after some linear change of coordinates, it can be cut out by polynomials with real coefficients. In particular, a hyperplane is always real. Note that a linear change of coordinates corresponds to a translation of the amoeba and the coamoeba. One can also check that $\theta \in \mathcal{A}'_V$ if and only if $-\theta \in \mathcal{A}'_V$, whenever V is real. The notion of real varieties was used in [7] for the study of the amoeba.

If $\theta \in \mathcal{A}'_V$, then we refer to the set

$$G(\theta) := \mathrm{Log}\left(V \cap \mathrm{Arg}^{-1}(\theta)\right)$$

as the *fiber* in $\mathcal{A}_V$ over θ, and similarly we will consider the fiber $G^{-1}(x)$ in the coamoeba over a point x in the amoeba. The relation between points and subsets of the amoeba and coamoeba established by G, is not a primary object of study in this work. However, it is necessary to develop some theory around fibers to obtain the main results. A general fact is that $G(\mathcal{C}'_V) = \mathcal{C}_V$. This follows from the next proposition. Even though it was known to Passare, it has not been published before and therefore comes with a proof.

Proposition 2.1. *The critical points of* Log *and* Arg *on* V *coincide.*

Proof. Fix a regular point $z \in V$. Choose a local branch of the holomorphic function $\log = \mathrm{Log} + i\,\mathrm{Arg}$ in a neighborhood U of z and set $W = \log(U \cap V)$, $w = \log z$. It follows by linearity that

$$T_{\mathrm{Log}\,z}(\mathrm{Log}\,V) = \mathrm{Re}\,T_w(W), \quad T_{\mathrm{Arg}\,z}(\mathrm{Arg}\,V) = \mathrm{Im}\,T_w(W)$$

where $T_q(X)$ denotes the tangent space of X at q. Furthermore, the dimension of the real and imaginary parts of $T_w(W)$ are equal since $u + iv \in T_w(W)$ implies that $i(u + iv) = -v + iu \in T_w(W)$. The proposition follows. $\square$

Finally some notation: e_j is the jth standard basis vector for $\mathbb{R}^n$ for $1 \le j \le n$ and

$$e_0 = -\sum_{j=1}^{n} e_j.$$

We will sometimes consider the lifting of the amoeba of an affine space L to $(\mathbb{R}^+)^n$:

$$\begin{aligned}\mathrm{Exp}\,\mathcal{A} &= \{(e^{x_1}, \dots, e^{x_n}); (x_1, \dots, x_n) \in \mathcal{A}\} \\ &= \{(|z_1|, \dots, |z_n|); (z_1, \dots, z_n) \in L\}.\end{aligned}$$

The real projective line is denoted by $R\mathbb{P}$, and $(x : y) \in R\mathbb{P}$ is sometimes represented by $y/x \in]-\infty, \infty]$. On $\mathbb{T}^n$, we use the metric inherited from $\mathbb{R}^n$. A *line* in $\mathbb{T}^n$ is a geodesic, i.e., the natural projection on $\mathbb{T}^n$ of a line with rational slope in $\mathbb{R}^n$.

3. Hyperplanes

The simplest examples of amoebas and coamoebas are those of hyperplanes. For such an amoeba $\mathcal{A}_f := \mathcal{A}_{f^{-1}(0)}$, where f is a polynomial of degree 1, Forsberg, Passare and Tsikh showed in [2] the following result.

Theorem 3.1 (Forsberg, Passare, Tsikh). *Let $f = a_0 + \sum_{j=1}^{n} a_j z_j$. Then* $\operatorname{Exp} \mathcal{A}_f$ *is the polyhedron in* $(\mathbb{R}^+)^n$ *defined by the following generalized triangle inequalities:*

$$|a_0| \leq \sum_{j=1}^{n} |a_j| r_j,$$

$$|a_k| r_k \leq |a_0| + \sum_{j \neq k} |a_j| r_j \quad \forall k = 1, 2, \ldots, n.$$

There is an analogous result for coamoebas. Instead of considering the coamoeba on $\mathbb{T}^n$, we will here consider the coamoeba on a fundamental domain in $\mathbb{R}^n$ of the natural projection from $\mathbb{R}^n$ to $\mathbb{T}^n = (\mathbb{R}/2\pi\mathbb{Z})^n$.

Theorem 3.2. *Consider the hyperplane defined by $f = \sum_{j=0}^{n} a_j z_j$ for $a_j \neq 0$. On the domain $\bigcap_{j=1}^{n} \{\theta \in \mathbb{R}^n; -\pi - \arg a_j < \theta_j \leq \pi - \arg a_j\}$, we have that θ is contained in the complement of the closure of the coamoeba $\mathcal{A}' := \mathcal{A}'_f$ if and only if*

$$-\pi - \arg a_j + \arg a_k < \theta_j - \theta_k < \pi - \arg a_j + \arg a_k, \ \forall j, k. \tag{3.1}$$

Furthermore, the contour of $\mathcal{A}'$ is given by the $2^n - 1$ points $(p_1 - \arg a_1 + \arg a_0, \ldots, p_n - \arg a_n + \arg a_0)$ for which $p_j \in \{0, \pi\}$ and not all p_j are zero.

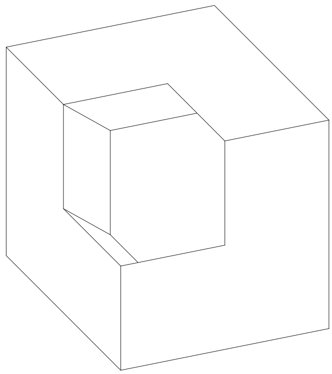


FIGURE 3. The coamoeba of a complex plane in $(\mathbb{C}^*)^3$. Its complement can be described as the convex hull of two cubes in opposite corners of the big cube representing $\mathbb{T}^3$.

As the contour is specified in Theorem 3.2, we actually get a description of the coamoeba itself and not only its closure.

For a shorter proof, we introduce the fixed number $z_0 = 1$. First we show the following lemma.

Lemma 3.3. *If $\mathcal{A}'$ is as in Theorem* 3.2, *then*

$$\overline{\mathcal{A}'} = \bigcup_{0 \leq j < k < l \leq n} \overline{\mathcal{A}'}_{e_j + e_k + e_l}.$$

Proof. The $\supseteq$-inclusion follows from Theorem 1.2. For the other inclusion, assume that $\theta \in \overline{\mathcal{A}'}$. Then the convex hull in $\mathbb{C}$ of the points a_0, $a_1 e^{i\theta_1}, \ldots, a_n e^{i\theta_n}$ must contain the origin. But then there are j, k, l such that the convex hull of $a_j e^{i\theta_j}$, $a_k e^{i\theta_k}$, $a_l e^{i\theta_l}$ contains the origin. By an arbitrarily small perturbation of either θ_j, θ_k or θ_l, the origin will be contained in the interior of the convex hull of $a_j e^{i\theta_j}$, $a_k e^{i\theta_k}$, $a_l e^{i\theta_l}$, and thus

$$0 = a_j e^{x_j + i\theta_j} + a_k e^{x_k + i\theta_k} + a_l e^{x_l + i\theta_l}$$

for some $x_j, x_k, x_l \in \mathbb{R}$. But this means exactly that $\theta \in \overline{\mathcal{A}'}_{e_j + e_k + e_l}$. □

Lemma 3.3 is a more specific version of Theorem 1.4 for the hyperplane case. There are similar specific versions of Theorem 1.4 for any specific low codimension, cf. Theorem 6.4 below.

The rest of the proof of Theorem 3.2 is mostly about cosmetic reformulations.

Proof of Theorem 3.2. We note that on the domain $\{-\pi < \theta_1, \theta_2 \leq \pi\}$, the complement of $\overline{\mathcal{A}'}_{a+bz_1+cz_2}$ in $\mathbb{R}^2$ is given by the inequalities

$$-\pi < \theta_1 - \theta_2 < \pi, \tag{3.2}$$

whenever a, b and c are strictly positive numbers, cf. Figure 2. The hyperplane in $(\mathbb{C}^*)^n$ given by $a_j z_j + a_k z_k + a_l z_l$ is also given by $a_j + a_k z_k / z_j + a_l z_l / z_j$. Hence on the domain

$$\{-\pi < \theta_k - \theta_j, \theta_l - \theta_j \leq \pi\},$$

the complement of the coamoeba of $a_j z_j + a_k z_k + a_l z_l$ in $\mathbb{R}^n$ is given by the inequalities

$$-\pi < \theta_k - \theta_l < \pi,$$

when $a_j, a_k, a_l > 0$. Since $a_j z_j + a_k z_k + a_l z_l = f_{e_j + e_k + e_l}$, Lemma 3.3 implies that the complement of $\overline{\mathcal{A}'_f}$ on the domain $\{-\pi < \theta_1, \ldots, \theta_n \leq \pi\}$ is given by the inequalities

$$-\pi < \theta_j - \theta_k < \pi, \ \forall j, k. \tag{3.3}$$

Furthermore it is easy to check the definition to see that the critical points are exactly the real points contained in $f^{-1}(0)$. Thus $\mathcal{C}' \cap \{-\pi < \theta_1, \ldots, \theta_n \leq \pi\}$ is given by the points $(p_1, \ldots, p_n)$ where $p_j \in \{0, \pi\}$ and not all p_j are zero. For the general case, just note that $f(z_1 e^{-i \arg a_1}, \ldots, z_n e^{-i \arg a_n}) = 0$ if and only if $\sum_{j=0}^{n} |a_j| z_j = 0$. □

4. Lines

The second case where the amoeba and coamoeba can be thoroughly described, is when L is a complex line. Throughout this section, we let L be given by the parametrization

$$z(t) = (t, a_2 + b_2 t, \ldots, a_n + b_n t), \quad t = x + iy,$$

where $a_j, b_j \in \mathbb{C}$ and $x, y \in R\mathbb{P}$ are chosen so that

$$t \notin \{0, \infty, -a_2/b_2, \dots, -a_n/b_n\}.$$

Furthermore, we assume that $b_j \neq 0$ for every $j \geq 2$. This is not really a restriction, since a line excluded in this way is the Cartesian product of a point in $(\mathbb{C}^*)^k$ and a line in $(\mathbb{C}^*)^{n-k}$ with $b_2, b_3, \dots, b_{n-k} \neq 0$, for some $k < n$.

Notice that L is real if and only if

$$\left(\frac{a_2}{b_2} : \frac{a_3}{b_3} : \cdots : \frac{a_n}{b_n}\right) \in R\mathbb{P}^{n-2}.$$

If $a_k \neq 0$, this inclusion is equivalent to the assertion that $(a_j b_k)/(a_k b_j) \in \mathbb{R}$ for every j, k. Hence L is real if and only if

$$\begin{aligned} 0 &= \mathrm{Im}\,(a_k b_j \bar{a}_j \bar{b}_k) = \mathrm{Re}\,(b_j \bar{a}_j)\mathrm{Im}\,(a_k \bar{b}_k) + \mathrm{Re}\,(a_k \bar{b}_k)\mathrm{Im}\,(b_j \bar{a}_j) \\ &= \mathrm{Re}\,(a_j \bar{b}_j)\mathrm{Im}\,(a_k \bar{b}_k) - \mathrm{Re}\,(a_k \bar{b}_k)\mathrm{Im}\,(a_j \bar{b}_j) \end{aligned} \tag{4.1}$$

for every j, k.

Let $\mathrm{Tan} : \mathbb{T}^n \to (R\mathbb{P})^n$ be the local diffeomorphism given by tan in each coordinate.

Proposition 4.1. *If L is not real, then $\mathcal{C}' = \emptyset$. If L is real, then* $\mathrm{Tan}\,\mathcal{C}'$ *consists of the single point whose jth coordinate equals* $\mathrm{Im}\,a_j/\mathrm{Re}\,a_j$ *if $a_j \neq 0$ and*

$$\mathrm{Im}\,(b_j a_k \bar{b}_k)/\mathrm{Re}\,(b_j a_k \bar{b}_k)$$

if $a_j = 0$, where k can be any number for which $a_k \neq 0$.

Thus, $\mathcal{C}'$ is either empty or consists of a finite number of points.

Proof of Proposition 4.1. To decide whether $z \in L$ is a critical point of Arg or not, it suffices to show that the rank of the $n \times 2$-matrix $A = \mathrm{Jac}(\mathrm{Tan} \circ \mathrm{Arg})_L$ equals 1 at z. Let

$$\tau_j = \tan \arg z_j = \mathrm{Im}\,z_j/\mathrm{Re}\,z_j.$$

By our parametrization, the first row of A is

$$\left(\frac{\partial \tau_1}{\partial x}, \frac{\partial \tau_1}{\partial y}\right) = \left(-\frac{y}{x^2}, \frac{1}{x}\right)$$

while the jth row, $2 \leq j \leq n$, equals

$$\left(\frac{\partial \tau_j}{\partial x}, \frac{\partial \tau_j}{\partial y}\right) = \left(\frac{\mathrm{Im}\,(\bar{a}_j b_j) - y|b_j|^2}{(\mathrm{Re}\,a_j + x\mathrm{Re}\,b_j - y\mathrm{Im}\,b_j)^2}, \frac{\mathrm{Re}\,(\bar{a}_j b_j) + x|b_j|^2}{(\mathrm{Re}\,a_j + x\mathrm{Re}\,b_j - y\mathrm{Im}\,b_j)^2}\right).$$

The two columns of A are linearly dependent exactly when every minor of A vanishes, that is $y\mathrm{Re}\,(\bar{a}_j b_j) - x\mathrm{Im}\,(\bar{a}_j b_j) = 0$ for every j. This means that

$$\tau_1 = y/x = -\frac{\mathrm{Im}\,(\bar{a}_k b_k)}{\mathrm{Re}\,(\bar{a}_k b_k)} = \frac{\mathrm{Im}\,(a_k \bar{b}_k)}{\mathrm{Re}\,(a_k \bar{b}_k)} \tag{4.2}$$

for every k for which $a_k \neq 0$, and by (4.1), this equality for every such k says exactly that L is real. From this we compute τ_j for j such that $a_j \neq 0$ by repeated use of (4.1):

$$\begin{aligned}
\tau_j &= \frac{\mathrm{Im}\, a_j + y\mathrm{Re}\, b_j + x\mathrm{Im}\, b_j}{\mathrm{Re}\, a_j + x\mathrm{Re}\, b_j - y\mathrm{Im}\, b_j} \\
&= \frac{\mathrm{Im}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) + x\mathrm{Re}\, b_j\mathrm{Im}\,(a_k\bar{b}_k) + x\mathrm{Im}\, b_j\mathrm{Re}\,(a_k\bar{b}_k)}{\mathrm{Re}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) + x\mathrm{Re}\, b_j\mathrm{Re}\,(a_k\bar{b}_k) - x\mathrm{Im}\, b_j\mathrm{Im}\,(a_k\bar{b}_k)} \\
&= \frac{\mathrm{Im}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) + x\mathrm{Im}\,(b_j a_k\bar{b}_k)}{\mathrm{Re}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) + x\mathrm{Re}\,(b_j a_k\bar{b}_k)} \cdot \frac{\mathrm{Re}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) - x\mathrm{Re}\,(b_j a_k\bar{b}_k)}{\mathrm{Re}\, a_j\mathrm{Re}\,(a_k\bar{b}_k) - x\mathrm{Re}\,(b_j a_k\bar{b}_k)} \\
&= \frac{\mathrm{Im}\, a_j\mathrm{Re}\, a_j\mathrm{Re}^{\,2}(a_k\bar{b}_k) - x^2\mathrm{Re}\,(b_j a_k\bar{b}_k)\mathrm{Im}\,(b_j a_k\bar{b}_k)}{\mathrm{Re}^{\,2} a_j\mathrm{Re}^{\,2}(a_k\bar{b}_k) - x^2\mathrm{Re}^{\,2}(b_j a_k\bar{b}_k)} \cdot \frac{\mathrm{Re}\, a_j}{\mathrm{Re}\, a_j} \\
&= \frac{\mathrm{Im}\, a_j\mathrm{Re}\, a_j\mathrm{Re}^{\,2}(a_k\bar{b}_k)\mathrm{Re}\, a_j - x^2\mathrm{Re}^{\,2}(b_j a_k\bar{b}_k)\mathrm{Im}\, a_j}{(\mathrm{Re}^{\,2} a_j\mathrm{Re}^{\,2}(a_k\bar{b}_k) - x^2\mathrm{Re}^{\,2}(b_j a_k\bar{b}_k))\mathrm{Re}\, a_j} \\
&= \frac{\mathrm{Im}\, a_j}{\mathrm{Re}\, a_j}.
\end{aligned} \tag{4.3}$$

For j such that $a_j = 0$, the result follows easily by use of (4.2). □

Proposition 4.1 implies that the topology of the amoeba and coamoeba differs between the real and non-real case. We will give a precise characterization of both cases, but first we need the following lemma.

Lemma 4.2. *For a line L, the fiber in $\mathcal{A}'$ over $x \in \mathcal{A}$ consists of two points if L is real and $x \notin \mathcal{C}$ and one point if L is not real or $x \in \mathcal{C}$.*

Proof. Assume that $s, t \in \mathbb{C}$, $s \neq t$, are such that $\mathrm{Log}\, z(s) = \mathrm{Log}\, z(t)$. Then in particular $|s| = |t|$. Furthermore we have for general complex numbers $a \neq b$ with $|a| = |b|$, that $|a_j + a| = |a_j + b|$ if and only if b is the reflection of a in the line through a_j and the origin, that is $\arg a_j - \arg b = -(\arg a_j - \arg a)$, so the only possibility for $\mathrm{Log}(z(s))$ to equal $\mathrm{Log}(z(t))$ is if

$$\arg(a_j) - \arg(b_j s) = \pm(\arg(a_j) - \arg(b_j t))$$

for every j. When the sign on the right-hand side is positive for some j, this implies that $s = t$. If it is negative for every coordinate j instead, then

$$2\arg(a_j/b_j) \equiv \arg s + \arg t \mod 2\pi \tag{4.4}$$

for every j, and thus L is real. Finally we check that (4.4) implies that $z(s) = z(t)$, that is $s = t$, if and only if $\arg t = \arg(a_j/b_j) \mod \pi$ for every j. But by Proposition 4.1 this is exactly when $\mathrm{Arg}\, z(t) \in \mathcal{C}'$, and hence by Proposition 2.1, $\mathrm{Log}\, z(t) \in \mathcal{C}$. □

Corollary 4.3. *If L is a real line, then $\partial\mathcal{A} = \mathcal{C}$.*

Proof. It follows by Lemma 4.2, that no non-critical points in L of Log can be mapped on the contour $\mathcal{C}$. Hence by definition, $\theta \in \partial\mathcal{A}$. □

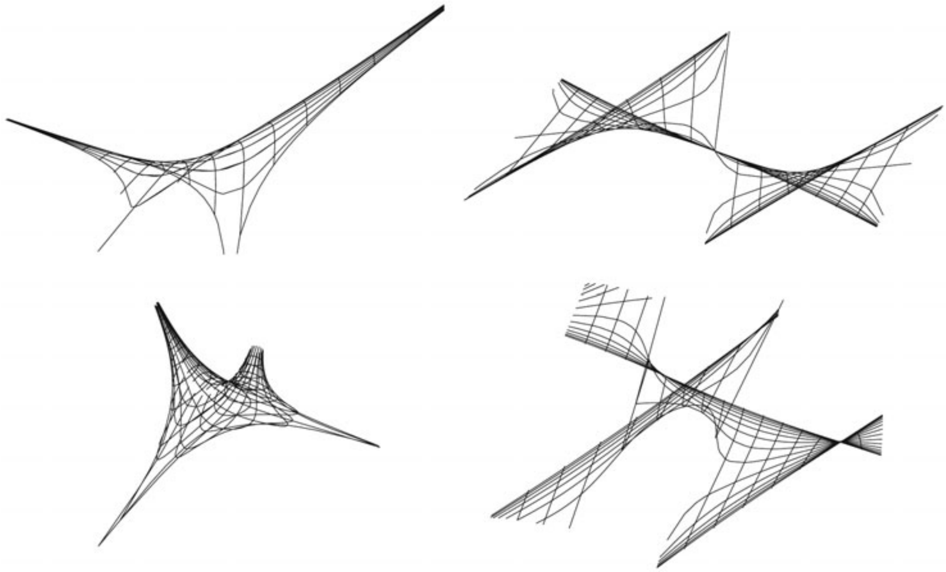


FIGURE 4. The amoeba (left) and coamoeba (right) of a real (up) and non-real (down) complex line in $(\mathbb{C}^*)^3$. The amoeba of a real line has boundary while the amoeba of a non-real line is a topological sphere minus generically four points. The boundary of a coamoeba of a line L, consists of four lines on $\mathbb{T}^3$. These lines intersect pairwise if and only if L is real. We will discuss this further in Example 6.9.

We are now ready to give a topological description of the amoeba and coamoeba of a line.

Theorem 4.4. *The amoeba and coamoeba of a non-real line L are diffeomorphic to the Riemann sphere minus k points where $4 \leq k \leq n+1$. If L is real, we have*

$$\mathcal{A}_L = \mathcal{A}_f \cap \{x; \operatorname{Exp} x \in Z\}, \tag{4.5}$$

$$\mathcal{A}'_L = \mathcal{A}'_f \cap \{\theta; \operatorname{Tan} \theta \in Z'\}, \tag{4.6}$$

where Z is a homogeneously quadratic surface in $\mathbb{R}^n$, Z' a cubic surface in $(R\mathbb{P})^n$ and f is any affine trinomial that vanishes on L.

Notice that an affine trinomial in $(\mathbb{C}^*)^n$ is independent of $n-2$ of $n+1$ projective coordinates when considering $(\mathbb{C}^*)^n$ as a torus in $\mathbb{P}^n$ with the zeroth coordinate chosen to be fixed. Hence the second part of Theorem 4.4 implies that the projection of the (co)amoeba of L on a subspace of $\mathbb{R}^n$ obtained by fixing some good choice of $n-2$ coordinates, is the (co)amoeba of a line in $(\mathbb{C}^*)^2$.

The number k in Theorem 4.4 is generically $n+1$, and otherwise, L is degenerate, see Section 5.

Proof of Theorem 4.4. Setting $r_j = |z_j|$ we have for $1 \le j \le n$ that

$$\begin{aligned} r_j^2 &= (\mathrm{Re}\, a_j + x\mathrm{Re}\, b_j - y\mathrm{Im}\, b_j)^2 + (\mathrm{Im}\, a_j + x\mathrm{Im}\, b_j + y\mathrm{Re}\, b_j)^2 \\ &= |a_j|^2 + 2\mathrm{Re}\,(a_j\bar{b}_j)x + 2\mathrm{Im}\,(a_j\bar{b}_j)y + |b_j|^2 r_1^2. \end{aligned}$$

If L is real we have by (4.1) that $\mathrm{Re}\,(a_j\bar{b}_j)/\mathrm{Re}\,(a_k\bar{b}_k) = \mathrm{Im}\,(a_j\bar{b}_j)/\mathrm{Im}\,(a_k\bar{b}_k)$ for every $j, k = 1, 2, \ldots, n$ such that $a_k \neq 0$. Hence there are for every j, k constants $\lambda_{jk} \in \mathbb{R}$ such that

$$r_j^2 + \lambda_{jk} r_k^2 - |a_j|^2 - \lambda_{jk}|a_k|^2 - (|b_j|^2 + \lambda_{jk}|b_k|^2) r_1^2 = 0.$$

Of these equations we choose $n-2$ that are algebraically independent and we see that $\mathrm{Exp}\,\mathcal{A}$ must lie on a quadratic surface Z of real dimension 2.

Let f be an affine trinomial that vanishes on L. By Theorem 3.1, $\mathrm{Exp}\,\mathcal{A}_f$ is a convex set. To show that $\mathrm{Exp}\,\mathcal{A}_L = Z \cap \mathrm{Exp}\,\mathcal{A}_f$, it hence suffices to show that $\partial\mathcal{A}_L \subset \partial\mathcal{A}_f$, recalling that Exp is a diffeomorphism. However, $f^{-1}(0)$ is the Cartesian product of a line in $(\mathbb{C}^*)^2$ and the subtorus of $(\mathbb{C}^*)^n$ obtained by excluding the coordinates occurring in f. Hence by Corollary 4.3, $\partial\mathcal{A}_f = \mathcal{C}_f$. Since furthermore $\partial\mathcal{A}_L = \mathcal{C}_L$ by the same corollary, it suffices to show that $\mathcal{C}_L \subset \mathcal{C}_f$.

By Proposition 4.1, $\theta \in \mathcal{C}'_L$ implies that

$$\theta_j \equiv \begin{cases} \arg a_j \quad \mod \pi & \text{if } a_j \neq 0 \\ \arg(b_j a_k/e_k) \quad \mod \pi \text{ for any } k \text{ such that } a_k \neq 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

Let $f(z) = az_j + bz_k + cz_l$, where $0 \le j < k < l \le n$ and $z_0 := 1$. We have that if f vanishes on L, then for every $t \in \mathbb{C}$,

$$\begin{aligned} 0 =& a(a_j + b_j t) + b(a_k + b_k t) + c(a_l + b_l t) \\ =& a a_j + b a_k + c a_l + (a b_j + b b_k + c b_l)t. \end{aligned}$$

We get a real system of two linear equations in a, b, c. By solving this, we see that $f(z)$ is a constant times

$$(b_l a_k - b_k a_l) z_j + (b_j a_l - b_l a_j) z_k + (b_j a_k - b_k a_j) z_l. \tag{4.8}$$

Since L is real, the arguments of the two terms in the factor of one of the coordinates of z, are congruent modulo π, whenever both are non-zero. Hence, if (4.7) holds, then $\mathrm{rank}\, C^{\theta}_{f^{-1}(0)} = 1$, that is $\theta \in \mathcal{C}'_f$. By Proposition 2.1, we have showed (4.5).

Next we look at the coamoeba of L in the real case. If we set $\tau_j := \tan\theta_j$, then we have $\tau_1 = y/x$ and so, for $j \ge 2$,

$$\tau_j = \frac{\mathrm{Im}\,(a_j + b_j t)}{\mathrm{Re}\,(a_j + b_j t)} = \frac{\mathrm{Im}\, a_j + \tau_1 x \mathrm{Re}\, b_j + x\mathrm{Im}\, b_j}{\mathrm{Re}\, a_j + x\mathrm{Re}\, b_j - \tau_1 x \mathrm{Im}\, b_j}, \quad \forall j. \tag{4.9}$$

If $a_j = 0$, then clearly (4.9) is a quadratic equation in τ_j and τ_1. If there is a k such that $a_k \neq 0$, then (4.9) implies that

$$x = \frac{\mathrm{Im}\,(a_k) - \tau_k \mathrm{Re}\, a_k}{\tau_1\tau_k \mathrm{Im}\, b_k + \tau_k \mathrm{Re}\, b_k - \tau_1 \mathrm{Re}\, b_k + \mathrm{Im}\, b_k}. \tag{4.10}$$

By exchanging x in (4.9) with the expression given in (4.10) whenever $a_j \neq 0$, we get $n-2$ algebraically independent equations in $\tau_1, \ldots, \tau_n$ of degree at most 3, cutting out a surface $Z' \subset (R\mathbb{P})^n$, quite analogous to the case of the amoeba.

The points on $\mathbb{P}$ which we exclude when we parametrize L are $-a_k/b_k$ for $0 \leq k \leq n$. If $j \neq k$ and $a_k \neq 0$, then we have from (4.1) that

$$\tau_j = \frac{\mathrm{Im}\,(a_j + b_j(-a_k/b_k))}{\mathrm{Re}\,(a_j + b_j(-a_k/b_k))} = \begin{cases} \mathrm{Im}\, a_j/\mathrm{Re}\, a_j & a_j \neq 0 \\ \mathrm{Im}\,(b_j a_k \bar{b}_k)/\mathrm{Re}\,(b_j a_k \bar{b}_k) & a_j = 0. \end{cases} \tag{4.11}$$

The complement of $\mathrm{Tan}\,\mathcal{A}'_L$ in Z' must hence be contained in the union of lines $l_0, l_1, \ldots, l_n$ where, if $j > 0$, all coordinates except the j:th in l_j are fixed as in Proposition 4.1 and the remaining coordinate can be any number, and where

$$l_0 = \left\{ (\lambda, \frac{\mathrm{Im}\, b_2 + \lambda \mathrm{Re}\, b_2}{\mathrm{Re}\, b_2 - \lambda \mathrm{Im}\, b_2}, \ldots, \frac{\mathrm{Im}\, b_n + \lambda \mathrm{Re}\, b_n}{\mathrm{Re}\, b_n - \lambda \mathrm{Im}\, b_n}); \lambda \in R\mathbb{P} \right\}.$$

Furthermore, $\dim \mathcal{C}' = 0$ by Proposition 4.1. Hence the boundary of $\mathcal{A}'_L$ is contained in $\mathrm{Tan}^{-1} \bigcup_{j=1}^n l_j$.

Now let f be a trinomial as in (4.8). The boundary of $\mathcal{A}'_f$ consists of the three hyperplanes given by the coamoebas of the sums of pairs of monomials of f, cf. Figure 2. Hence we are done if we show that each line l_i is contained in one of these hyperplanes. To this end, first assume that $i > 0$ and consider a binomial g obtained from the monomials as in (4.8) of f that do not depend on z_i. Then we verify that $l_i \subset \mathcal{A}'_g$. Next we verify that $l_0 \subseteq \mathrm{Tan}\,\mathcal{A}'_h$, where $h(z) = (b_j a_l - b_l a_j) z_k + (b_j a_k - b_k a_j) z_l$ (note that $1 \leq k, l$).

If L is not real, then by Proposition 4.1, Log and Arg are local diffeomorphisms. But one can also check that Arg is an injection (this follows, e.g., from a combination of Proposition 6.5 and 6.2 below), and hence a global diffeomorphism. The same is true for Log by Lemma 4.2. If there is a $c \in \mathbb{C}$ such that $-a_j/b_j$ either equals c or 0 for every j, then L is real. Since L is parametrized by $\mathbb{P} \backslash \{0, -a_2/b_2, \ldots, -a_n/b_n, \infty\}$, the theorem hence follows. □

5. Affine spaces of codimension $n/2$

An affine space $L \subset (\mathbb{C}^*)^n$ of codimension m is the restriction to $(\mathbb{C}^*)^n$ of an affine subspace P_L of $\mathbb{P}^n$ whenever such a restriction is non-empty. This means that P_L is defined by a system

$$Cz^t := C(z_0 : z_1 : \cdots : z_n)^t = 0 \tag{5.1}$$

of linearly independent equations, where $C = C_L$ is a complex $m \times (n+1)$-matrix for which every column is contained in the linear span of the other columns.

Let us rewrite (5.1) as a real system. Denote by C_j the j:th column of C, $0 \leq j \leq n$. For $\theta \in \mathbb{T}^n$, the projection of $e^{i\theta_j} C_j$ on $\mathbb{R}^{2m}$ sending the real and imaginary part of the kth coordinate to the $2k-1$th and $2k$th coordinate respectively, gives a vector C_j^θ. Letting the subscript j denote the column, these vectors together with C_0 define a real $2m \times (n+1)$-matrix C^θ. The system (5.1) is equivalent with $C^\theta r = 0$, if we set $z_j = r_j e^{i\theta}$ for $r \in R\mathbb{P}^n$ (we can assume that $\theta_0 = 0$).

The case $m = n/2$ is special in the sense that the system $C^\theta r = 0$ then is neither under- nor over-determined, unless the rows of C^θ are linearly dependent. Notice also that $C^\theta(1,r) = 0$ for $r \in (\mathbb{R}^*)^n$ if and only if $re^{i\theta} \in L$. If r is unknown, there are 2^n possible values of $\operatorname{Arg} re^{i\theta}$ of which exactly one is contained in $\mathcal{A}'$. Furthermore, the volume of $\mathbb{T}^n$ is $(2\pi)^n$. Hence the following result is plausible.

Theorem 5.1. *If L is an affine space of codimension $n/2$ for which $\mathcal{A}' \neq \mathcal{C}'$, then* $\operatorname{Vol}\mathcal{A}' = \pi^n$.

The proof for this will be given at the very end of Section 6. We remark that if $\mathcal{A}' = \mathcal{C}'$, then $\operatorname{Vol}\mathcal{A}' = 0$. To see this, first notice that the real dimension of L is n and that there is an additional real algebraic condition for θ to fulfil to be contained in $\mathcal{C}'$, see, e.g., Proposition 2.2 in [4].

The volume of the amoeba of a space as in Theorem 5.1, is not known in general. However, Nisse announces in [9] a proof for that the volume of the amoeba of a generic real affine space in $(\mathbb{C}^*)^n$ is π^n/n. In the case $n = 4$, we give an alternative proof for this. First we need the following definition: an affine space L is *degenerate* if any of the maximal minors of C vanishes. As we will see in the final section, L is degenerate if the condition in Theorem 5.1 is fulfilled, while the converse implication does not hold.

Theorem 5.2. *The volume of the amoeba of a non-degenerate, real plane in* $(\mathbb{C}^*)^4$, *equals* $\pi^4/4$.

Our proof leans on two lemmas. For the first one, notice that if $\theta \in \mathcal{A}'$ and the system $C^\theta r = 0$ does not have a unique solution, then there is a positively dimensional subspace of solutions and hence $\theta \in \mathcal{C}'$. This means that G, as defined in Section 2, can be considered as a mapping from $\mathcal{A}' \backslash \mathcal{C}'$ to $\mathcal{A}$, assigning to θ the value of $(\log r_1, \ldots, \log r_n)$, where $(r_1, \ldots, r_n)$ is the solution to $C^\theta(1, r_1, \ldots, r_n) = 0$.

Lemma 5.3. *The Jacobian of G equals* 1.

Proof. Denote the Jacobian matrix of Log, Arg and Arg^{-1} by J, J' and J'^{-1} respectively. The Jacobian matrix of G then equals JJ'^{-1}. The complex logarithm $\log = \operatorname{Log} + i \operatorname{Arg}$ is a holomorphic function, meaning that the Cauchy–Riemann equations

$$\frac{\partial \operatorname{Log} z_k}{\partial x_j} = \frac{\partial \operatorname{Arg} z_k}{\partial y_j}, \quad \frac{\partial \operatorname{Log} z_k}{\partial y_j} = -\frac{\partial \operatorname{Arg} z_k}{\partial x_j},$$

are satisfied. Thus J is obtained from J' by interchanging pairs of rows and changing sign on one row in every pair and then JJ'^{-1} is obtained from the unit matrix $E = J'J'^{-1}$ in the same way. By elementary linear algebra it follows that

$$|JJ'^{-1}| = |E| = 1. \qquad \square$$

We now give the second lemma.

Lemma 5.4. *Let L be a non-degenerate real plane in $(\mathbb{C}^*)^4$. Then the number of points in $G^{-1}(x)$ is 4 for every $x \in \operatorname{int}\mathcal{A}$ and 2 for every $x \in \partial\mathcal{A}$.*

Proof. To minimize the number of necessary unknown constants, we can as well consider a plane L for which C_L is given by

$$C_L = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & a & b \end{pmatrix}, \quad a, b \in \mathbb{R}. \tag{5.2}$$

Assume that $z \in L$. Setting $z_j = r_j e^{\theta_j i}$ we have that

$$r_1^2 = \operatorname{Re}^2(1 + z_3 + z_4) + \operatorname{Im}^2(1 + z_3 + z_4), \tag{5.3}$$

$$r_2^2 = \operatorname{Re}^2(1 + az_3 + bz_4) + \operatorname{Im}^2(1 + az_3 + bz_4), \tag{5.4}$$

which for fixed $r_1, \ldots, r_4$ gives us the system

$$\begin{aligned} r_1^2 &= 1 + r_3^2 + r_4^2 + 2r_3\cos\theta_3 + 2r_4\cos\theta_4 + 2r_3r_4\cos(\theta_3 - \theta_4), \\ r_2^2 &= 1 + a^2r_3^2 + b^2r_4^2 + 2ar_3\cos\theta_3 + 2br_4\cos\theta_4 + 2abr_3r_4\cos(\theta_3 - \theta_4), \end{aligned} \tag{5.5}$$

in $\theta_1, \ldots, \theta_4$. To solve it, set

$$\xi_1 = \cos\theta_3, \quad \xi_2 = \cos\theta_4, \quad \xi_3 = \cos(\theta_3 - \theta_4).$$

Then (5.5) determines a line L_r in $\{(\xi_1, \xi_2, \xi_3)\} = \mathbb{R}^3$. Furthermore, ξ_1, ξ_2, ξ_3 must satisfy

$$\xi_1^2 + \xi_2^2 + \xi_3^2 = 1 + 2\xi_1\xi_2\xi_3, \quad |\xi_1|, |\xi_2|, |\xi_3| \leq 1. \tag{5.6}$$

This describes the boundary of a convex region D – an *inflated tetrahedron*, see Figure 5. The intersection of ∂D and L_r determines the points $\theta \in \mathcal{A}'$ for which $G(\theta) = \operatorname{Log} r$.

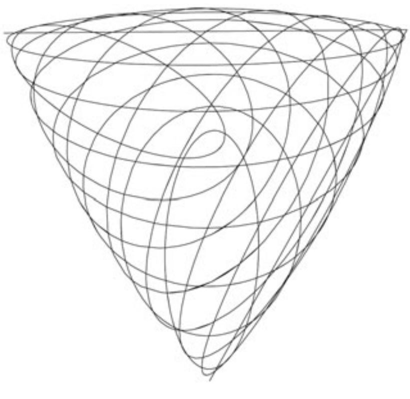


FIGURE 5. The inflated tetrahedron D.

The set of lines in $\mathbb{R}^3$ is a real 4-dimensional manifold $\mathcal{L}$. Since L is non-degenerate, a, b and 1 must be three distinct numbers. In view of this, it is straightforward to check that the mapping $r \mapsto L_r$ is a local diffeomorphism from $(\mathbb{R}^+)^4$ to $\mathcal{L}$. This means that $x \in \partial\mathcal{A}$ if and only if L_{e^x} tangents D. Furthermore, the set of lines L that intersects the interior of D can be parametrized by the two unique points contained in $L \cap \partial D$. Hence we have a mapping $\varphi : (\mathbb{R}^+)^4 \to \partial D$ with two locally diffeomorphic branches such that $G^{-1}(x)$ is given by the image of $x \in \operatorname{int}\mathcal{A}$ in following scheme:

$$\operatorname{int}\mathcal{A} \xrightarrow{\text{Exp}} \text{Exp}\mathcal{A} \xrightarrow{(\pi,\varphi)} (\mathbb{R}^+)^2 \times \partial D \xrightarrow{(i,\psi)} (\mathbb{R}^+)^2 \times \mathbb{T}^2 \xrightarrow{p} \mathcal{A}'. \tag{5.7}$$

Here π is given by $(r_1, r_2, r_3, r_4) \mapsto (r_3, r_4)$, i is the identity and ψ is the mapping that takes $\zeta \in \partial D$ to its corresponding points in $\mathbb{T}^2$. Also ψ has two branches, since

$$(\cos\theta_3, \cos\theta_4, \cos(\theta_3 - \theta_4)) = (\cos(-\theta_3), \cos(-\theta_4), \cos(-\theta_3 + \theta_4)).$$

The mapping p is given by

$$(r_3, r_4, \theta_3, \theta_4) \mapsto (\pi + \arg f(r_3 e^{i\theta_3}, r_4 e^{i\theta_4}), \pi + \arg g(r_3 e^{i\theta_3}, r_4 e^{i\theta_4}), \theta_3, \theta_4),$$

where $f(z) = 1 + z_1 + z_2$ and $g(z) = 1 + az_1 + bz_2$. It follows that $|G^{-1}(x)| = 4$ for every $x \in \operatorname{int}\mathcal{A}$.

When $x \in \partial\mathcal{A}$, L_{e^x} intersects ∂D at a unique point ζ. Except for this, $G^{-1}(x)$ is obtained just as in (5.7) and consequently, $|G^{-1}(x)| = 2$. □

Proof of Theorem 5.4. The volume of the contour of the amoeba or coamoeba of a plane in $(\mathbb{C}^*)^4$ is always 0, see the discussion above. By Proposition 5.1, the volume of $\mathcal{A}' \backslash \mathcal{C}'$ hence equals π^4. Since the Jacobian of G outside of $\mathcal{C}'$ equals 1, see Lemma 5.3, the result hence follows from Lemma 5.4. □

6. A general approach for the study of the affine coamoeba

In this section we will introduce a framework in which Theorem 1.1, 1.3 and 1.4 are better understood, and then we will prove them. Throughout the section, we consider an affine space L of codimension m defined by an $m \times (n+1)$-dimensional matrix C as in the previous section, with amoeba and coamoeba $\mathcal{A}$ and $\mathcal{A}'$, with contours $\mathcal{C}$ and $\mathcal{C}'$ respectively.

6.1. An indexation of the initial coamoebas

We start by looking at two different ways to construct complex spaces associated with L whose coamoebas are of interest for the understanding of $\mathcal{A}'$. Firstly, recall that we have the *initial spaces* L_ω of L, defined by the initial ideals I_ω as in the case of general algebraic varieties, see Section 2. It is well known that every I_ω is generated by polynomials of degree one, see, e.g., Proposition 1.6 in [12]. Thus it is straightforward to check that for any real n-vector ω, L_ω equals L_{ω_N} for some $N \subseteq \{0, 1, \ldots, n\}$, where

$$\omega_N = \sum_{j \notin N} e_j.$$

Secondly, for $N \subseteq \{0, 1, \ldots, n\}$ we may define an affine subspace L_N of $(\mathbb{C}^*)^n$ by the equation

$$C_N(1, z)^t = 0,$$

where C_N is the matrix obtained from C by setting all entries in the columns with indices in N to zero.

Lemma 6.1. *For any $N \subseteq \{0, 1, \ldots, n\}$, $L_{\omega_N} \subseteq L_N$. If L is not degenerate equality holds.*

Proof. Whenever the kth row of $C_N(1, z)^t$ is nonzero, it is given by

$$C_{kN}(1, z^t) := \sum_{j \notin N} c_{kj} z_j = \left(\sum_{j=0}^{n} c_{kj} z_j \right)_{\omega_N},$$

where $z_0 := 1$. Hence $I(L_{\omega_N}) \supseteq \langle C_{1N}(1, z^t), \ldots, C_{mN}(1, z^t) \rangle$, or equivalently $L_{\omega_N} \subseteq L_N$. When L is non-degenerate, $C_{kN}(1, z^t) = 0$ implies that $|N| > m$, which in turn implies that L_N, and hence also L_{ω_N}, is the empty set. $\square$

Notice that for L degenerate, the inclusion in Lemma 6.1 may be strict, since $C_{kN}(1, z)^t = 0$ means that

$$\left(\sum_{j=0}^{n} c_{kj} z_j \right)_{\omega_N} = \sum_{j \in N} c_{kj} z_j.$$

Thus the kth row of C corresponds to a non-zero polynomial in $I(L_{\omega_N})$, even if it does not correspond to a non-zero polynomial in $I(L_N)$.

By the discussion above, we have an indexation $\{\mathcal{A}'_N\}$ of the initial coamoebas of $\mathcal{A}'$ with $\mathcal{A}'_N := \mathcal{A}'_{L_N} = \mathcal{A}'_{\omega_N}$, whenever L is non-degenerate, and we will use this indexation throughout this section. However, $\mathcal{A}'_{L_N}$ has meaning even when L is degenerate, and hence some of the results below hold also in the degenerate case, although with a different meaning.

6.2. The compactified amoeba and the proof of Theorem 1.4

Given a polyhedron K, let $\operatorname{int} K$ denote its relative interior. Consider the nth unit simplex

$$\Delta_n = \left\{ s \in [0, 1]^{n+1}; \sum s_j = 1 \right\}$$

and set

$$L^\theta = \{ s \in \Delta_n; C^\theta s = 0 \}.$$

Then L^θ is the fiber over θ in the *compactified amoeba* of L, see [3]. Indeed, consider the diffeomorphism $\psi : \operatorname{int} \Delta_n \to \mathbb{R}^n$ given by

$$\psi(s) = (\log(s_1/s_0), \ldots, \log(s_n/s_0)).$$

It is easy to verify that for every $\theta \in \mathbb{T}^n$, the fiber in $\mathcal{A}$ over θ is given by $\psi(\operatorname{int} L^\theta)$, cf. the discussion preceding Theorem 5.1. Since furthermore $\Delta_n \cap K$ is a polygon of dimension $\dim K - 1$ whenever K is a subspace of $\mathbb{R}^{n+1}$ that intersects the interior of Δ_n, we have the following result.

Proposition 6.2. *The fiber in $\mathcal{A}$ over $\theta \in \mathcal{A}'$ is diffeomorphic to the interior of a polygon of dimension $n - \operatorname{rank} C^\theta$. In particular, the fiber is a single point if and only if $\operatorname{rank} C^\theta = n$.*

The faces of Δ_n can be indexed by the proper subsets of $\{0, 1, \ldots, n\}$ by setting $\Gamma_N = \{s \in \Delta_n; s_j = 0, j \in N\}$.

Lemma 6.3. *For $N \subseteq \{0, 1, \ldots, n\}$, $\theta \in \mathcal{A}'_N$ if and only if L^θ intersects $\operatorname{int} \Gamma_N$.*

Proof. The assertion says exactly that there is an $s \in \operatorname{int} \Delta_n$ with $C_N^\theta s = 0$ if and only if there is an $s' \in \operatorname{int} \Gamma_N$ with $C_N^\theta s' = C^\theta s' = 0$. If the former equation holds, then an $s' \in \operatorname{int} \Gamma_N$ satisfying the latter equation is obtained by setting every coordinate with index in N to zero and normalizing. If the latter equation holds, then an $s \in \operatorname{int} \Delta_n$ satisfying the former equation is obtained by the normalisation of $s' + w_N$. $\square$

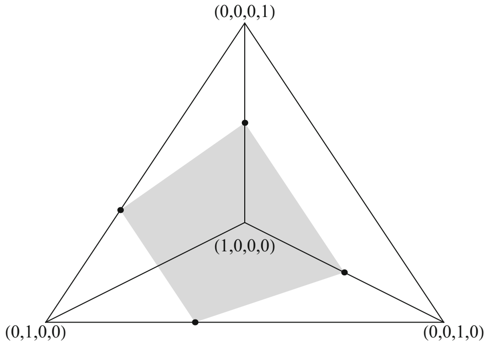


FIGURE 6. The unit simplex Δ_3 and L^θ with its four vertices marked with dots.

We will now state the result that lies behind Theorem 1.4.

Theorem 6.4. *If L is an affine space of codimension m, $0 < 2m < n$, then $\theta \in \mathcal{A}'$ if and only if $\theta \in \bigcap_{j=1}^k \mathcal{A}'_{N_j}$ for some index sets $N_1, \ldots, N_k$ with $\bigcap_{j=1}^k N_j = \emptyset$ and $|N_j| \geq n - 2m$.*

Proof. By Lemma 6.3 it suffices to show for every $\theta \in \mathbb{T}^n$ that L^θ intersects $\operatorname{int} \Delta_n$ if and only if L^θ intersects $\operatorname{int} \Gamma_{N_j}$ for $N_1, \ldots, N_k$ as in the formulation of the theorem. If the latter holds, with $s_j \in \operatorname{int} \Gamma_{N_j} \cap L^\theta$, then clearly the normalisation of $\sum_{j=1}^k s_j$ is contained in $\operatorname{int} \Delta_n \cap L^\theta$.

Now assume instead that there is an $s \in \operatorname{int} \Delta_n \cap L^\theta$. Since the maximal rank of C^θ is $2m < n$, L^θ must be a polytope of dimension at least $n - 2m > 0$. Its vertices are clearly contained in the interiors of faces Γ_{N_j} of Δ_n for $N_1, \dots, N_k$ as in the assertion, cf. Figure 6. $\square$

Remark 1. By studying the relation between the faces of L^θ and the faces of Δ_n, it is possible to choose much more specific classes $N_1, \dots, N_k$ in Theorem 6.4 and still obtain the whole coamoeba. $\square$

Proof of Theorem 1.4. By Theorem 6.4 and Lemma 6.1,

$$\mathcal{A}'_0 = \mathcal{A}' \subseteq \bigcup_{\{0,1,\dots,n\} \supseteq N \neq \emptyset} \mathcal{A}'_N = \bigcup_{\mathbb{R}^n \ni \omega \neq 0} \mathcal{A}'_\omega.$$

Hence the result follows from Theorem 1.2. $\square$

6.3. The contour of the coamoeba and partial proof of Theorem 1.3

In the affine case, the contour of the coamoeba basically contains one type of points, as indicated by the following result.

Proposition 6.5. *Let $\theta \in \mathcal{A}'$. When $2m \geq n$, θ belongs to the contour $\mathcal{C}'$ if and only if C^θ has rank strictly less than n. When $2m \leq n$, $\theta \in \mathcal{C}'$ if and only if the rank of C^θ is not maximal.*

Proof. First look at the case $2m \geq n$. If $\operatorname{rank} C^\theta < n$ for $\theta \in \mathcal{A}'$, then it follows from Proposition 6.2 that $\theta \in \mathcal{C}'$. If $\operatorname{rank} C^\theta > n$, then again by Proposition 6.2, $\theta \notin \mathcal{A}'$. Finally, assume that $\operatorname{rank} C^\theta = n$. Then there is an open neighborhood U in $\mathbb{T}^n$ of θ so that the column vectors $C_1^\varphi, \dots, C_n^\varphi$ are linearly independent whenever $\varphi \in U$. The set $\Omega = \operatorname{Arg}^{-1}(U) \cap L$ is non-empty since $\theta \in \mathcal{A}'$, and thus, since L lacks critical points, a differentiable manifold of real dimension $2(n-m)$. Furthermore, Arg maps Ω injectively on $\mathcal{A}'$ by Proposition 6.2. Hence θ is not contained in $\mathcal{C}'$.

Now assume that $2m \leq n$. If the rank of C^θ is not maximal, then it follows from Proposition 6.2 that $\theta \in \mathcal{C}'$. If it is maximal, then there is an open neighborhood U of θ such that, without loss of generality, $C_1^\varphi, \dots, C_{2m}^\varphi$ is a basis for $\mathbb{R}^{2m}$ for every $\varphi \in U$. Let

$$A_\varphi = \{(\log r_1, \dots, \log r_n); C^\varphi (1, r)^t = 0)\}.$$

Since C^φ depends continuously on φ, so does A_φ, and hence defines a vector bundle over U, that is a $2(n-m)$-dimensional open real manifold. Clearly L is parametrized locally at $\operatorname{Arg}^{-1}(\theta) \cap L$ by $(\varphi, x) \mapsto e^{x + i\varphi}$ for $x \in A_\varphi$, $\varphi \in U$. The proposition follows. $\square$

Remark 2. The second statement of Proposition 6.5 is a special case of Proposition 2.2 in [4]. $\square$

The following corollary points out an aspect of real affine spaces that we glimpsed in Theorem 4.4.

Corollary 6.6. *The amoeba and coamoeba of a real affine space, have contours.*

Proof. Denote the affine space in the lemma by L. We may assume that C_L has real coefficients. Thus there is an $r \in (\mathbb{R}^*)^n$ such that $C_L(1,r)^t = 0$. Letting θ_j equal 0 if $r_j > 0$ and π if $r < 0$, it follows that $\theta \in \mathcal{A}'_L$ and $\operatorname{rank} C^\theta = m < \min\{n, 2m\}$, that is, by Proposition 6.5, $\theta \in \mathcal{C}'_L$. It follows from Proposition 2.1 that also $\mathcal{C}_L$ is non-empty. $\square$

Just as we used the fact that $\dim L^\theta > 0$ for every $\theta \in \mathcal{A}'$ when $2m < n$ to get the formula of Theorem 6.4, we can by Proposition 6.5 get a similar formula for $\mathcal{C}'$, regardless of the codimension of L. However, for the proof of the remaining results, the following assertion suffices.

Lemma 6.7. *The contour of $\mathcal{A}'$ is contained in the union of sets $\mathcal{A}'_M \cap \mathcal{A}'_N$ for which $M \backslash N$ and $N \backslash M$ are non-empty and $|N|, |M| \geq \max\{1, n-2m+1\}$. If $2m \geq n$, we have furthermore that $\mathcal{C}' = \mathcal{A}' \cap \bigcup_{N \neq \emptyset} \mathcal{A}'_N$.*

Proof. By Proposition 6.5, $\theta \in \mathcal{C}'$ implies that $\operatorname{rank} C^\theta < \min\{2m, n\}$, that is

$$\dim L^\theta \geq n + 1 - \min\{2m, n\} \geq \max\{1, n+1-2m\}.$$

Hence L^θ intersects Γ_M and Γ_N for M, N as in the formulation of the lemma. The first part of the lemma now follows from Lemma 6.3.

For the second part, assume that $2m \geq n$ and let $N \neq \emptyset$. If $\theta \in \mathcal{A}' \cap \mathcal{A}'_N$, then by Lemma 6.3, L^θ intersects both $\operatorname{int}\Delta_n$ and $\operatorname{int}\Gamma_N$. Hence $\dim L^\theta > 0$ and it follows that $\theta \in \mathcal{C}'$. $\square$

Example 6.8. When $2m \geq n$, one can check that

$$\mathcal{C}' = \bigcup_{\substack{M,N \neq \emptyset \\ M \cap N = \emptyset}} \mathcal{A}'_L \cap \mathcal{A}'_L. \tag{6.1}$$

Example 6.9. When L is a non-degenerate line, the proper initial coamoebas are the real lines $\mathcal{A}_{\{j\}}$ for $0 \leq j \leq n$. These intersect pairwise if and only if L is real and $\mathcal{C}'$ is the set of isolated points given by these intersections, see Figure 4.

Example 6.10. Consider Figure 7. In (a) and (b), L^θ corresponds to an affine space L in complex dimension 2 that is cut out by some polynomials $a + bz_1 + cz_2$. In (a), $\theta \in \mathcal{A}'_{L_{\{0\}}} \cap \mathcal{A}'_{L_{\{2\}}}$, that is a typical situation when θ is contained in the contour of a line. In (b), $\theta \in \mathcal{A}'_{L_{\{01\}}} \cap \mathcal{A}'_{L_{\{2\}}}$. But $\theta \in \mathcal{A}'_{L_{\{01\}}}$ means that $c = 0$, and hence $\mathcal{A}'_{L_{\{01\}}} = \mathbb{T}^2$. In particular we have that $\mathcal{A}'_L = \mathcal{C}'_L$.

Example 6.11. Let us go back to the situation in Section 5 and consider a non-degenerate plane in $(\mathbb{C}^*)^4$. The contour $\mathcal{C}'$ of its coamoeba can by Example 6.8 be considered as more or less the union of the ten sets

$$\mathcal{C}'_{lm} := \mathcal{A}'_{L_{\{l\}}} \cap \mathcal{A}'_{L_{\{m\}}}.$$

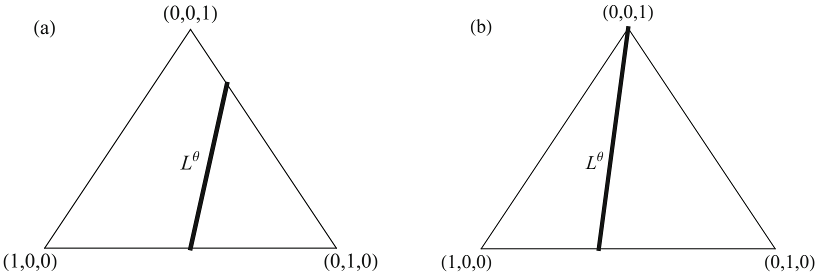


FIGURE 7. The characteristic positions possible for L^θ when $\theta \in \mathcal{C}'_L$ for a line L in $(\mathbb{C}^*)^2$.

Indeed, if the plane is non-real, then $\mathcal{C}'$ is precisely the union of these sets, which in this case are non-empty, connected and with disjoint closures. Most of this is rather straightforward to show. For the connectedness, let g, f be polynomials of degree one such that L is defined by the equations

$$f(z_3, z_4) - z_1 = 0, \quad g(z_3, z_4) - z_2 = 0.$$

The coamoebas $\mathcal{A}'_f$ and $\mathcal{A}'_g$ are different translates of the set described in Figure 2. One can show that $\theta \in \mathcal{C}'_{12}$ implies that (θ_3, θ_4) is contained in the open, connected set $\mathcal{A}'_f \cap \mathcal{A}'_g$ and that $\mathcal{C}'_{12}$ more or less is given by a parametrization over $\mathcal{A}'_f \cap \mathcal{A}'_g$. In particular, $\mathcal{C}'_{12}$ is connected.

If the plane is real and f, g are as above, then $\mathcal{A}'_f \cap \mathcal{A}'_g$ may be disconnected, in which case it consists of two points. These points correspond to certain points in $\mathcal{C}'$ that are contained in several connected components $\mathcal{C}'_{lm}$, and one of the points is in fact contained in all such components. As a consequence, $\mathcal{C}'$ is connected. One can use the mapping G, see Section 5, to show that also the contour of the amoeba of a real plane in $(\mathbb{C}^*)^4$ is connected.

By using the results in this section, we can prove one direction of Theorem 1.3.

Lemma 6.12. *If L is a non-degenerate affine space of codimension at least $n/2$, then*

$$\partial \mathcal{A}' \subseteq \bigcup_{\omega \neq 0} \mathcal{A}'_\omega \backslash \mathcal{A}'. \tag{6.2}$$

Proof. By Theorem 1.2,

$$\partial \mathcal{A}' \backslash \mathcal{A}' = \bigcup_{\omega \neq 0} \mathcal{A}'_\omega \backslash \mathcal{A}'. \tag{6.3}$$

Furthermore,

$$\partial \mathcal{A}' \cap \mathcal{A}' \subseteq \mathcal{C}' = \bigcup_{N \neq \emptyset} \mathcal{A}'_N \cap \mathcal{A}' = \bigcup_{\omega \neq 0} \mathcal{A}'_\omega \cap \mathcal{A}', \tag{6.4}$$

where the second equality follows from Lemma 6.7 and the third from Lemma 6.1. Together, (6.3) and (6.4) imply the desired result. □

6.4. Proof of Theorem 1.1 and the remaining proof of Theorem 1.3

We are now ready to finish the proving of the theorems from the introduction.

Theorem 6.13. *For a non-degenerate affine space L, the following assertions are true.*

1. *The interior* $\operatorname{int}\mathcal{A}'$ *of the coamoeba is a real, differentiable manifold of dimension* $\min\{2n-2m, n\}$*, and* $\overline{\operatorname{int}\mathcal{A}'} = \overline{\mathcal{A}'}$.
2. *If $U \subset \mathbb{T}^n$ is open and $U \cap \partial\mathcal{A}' \neq \emptyset$, then $U \cap (\partial\mathcal{A}' \backslash \mathcal{A}') \neq \emptyset$.*
3. *When the codimension of L is at least $n/2$, the following formula holds:*

$$\partial\mathcal{A}'_L = \bigcup_{N \neq \emptyset} \mathcal{A}'_{L_N}.$$

4. *When the codimension of L is at least $n/2$, then for every $k \in \{0, 1, \ldots, n\}$ there is a $\theta \in \overline{\mathcal{A}}'$ such that $e_k \notin T_\theta(\overline{\mathcal{A}}'_L)$.*

The second point is exactly Theorem 1.1. By Lemma 6.1, the third point implies Theorem 1.3, but in view of Lemma 6.7, it also implies that $\mathcal{C}' \subseteq \partial\mathcal{A}'$ whenever $2m \geq n$. The first point implies that the coamoeba does not equal its contour and hence tells us that our notion of degeneracy is not too strong. The last point is mostly included for technical reasons. Here θ does not have to be a smooth point, as $T_\theta(\overline{\mathcal{A}}'_L)$ refers to the space of tangent vectors at θ of any smooth curve $\gamma \subseteq \overline{\mathcal{A}'}$ passing through θ.

For clarity, we will from now on specify the affine space whose coamoeba we consider and, e.g., write $\mathcal{A}'_{L_N}$ instead of $\mathcal{A}'_N$. Let K_N be the affine subspace of $(\mathbb{C}^*)^{n-|N|}$ given by the quotient of L_N by the space generated by the vectors e_j, $j \in N$. Then $\mathcal{A}'_{K_N}$ is obtained from $\mathcal{A}'_{L_N}$ by taking the corresponding quotient in $\mathbb{T}^n$. Let $\pi_N : \mathbb{T}^n \to \mathbb{T}^{n-|N|}$ be the natural projection that takes $\mathcal{A}'_{L_N}$ to $\mathcal{A}'_{K_N}$.

Lemma 6.14. *For any M, N with $|M|, |N| \geq \max\{1, n-2m+1\}$ and any open set $U \subseteq \mathbb{T}^n$, the assertion*

$$U \cap \pi_M^{-1}\left(\mathcal{A}'_{K_M} \backslash \mathcal{C}'_{K_M}\right) = U \cap \pi_N^{-1}\left(\mathcal{A}'_{K_N} \backslash \mathcal{C}'_{K_N}\right) \neq \emptyset \tag{6.5}$$

implies that $\mathcal{A}'_{L_M} = \mathcal{A}'_{L_N}$.

Proof. Assume that there is an open set U for which (6.5) holds. Then we find, perhaps by cutting down U, open submanifolds V and W of K_M and K_N respectively such that $U \cap \pi_M^{-1}(\operatorname{Arg} V) = U \cap \pi_N^{-1}(\operatorname{Arg} W)$. Let V' be a large enough open submanifold of L_N that projects to V under the quotient of the vectors e_j for $j \in N$. Notice that $z \in \operatorname{Arg}^{-1}(U) \cap V'$ implies that

$$T_{\operatorname{Arg} z}\left(\pi_M^{-1}(\operatorname{Arg} V)\right) = T_{\operatorname{Arg} z}\left(\pi_N^{-1}(\operatorname{Arg} W)\right). \tag{6.6}$$

This amounts to a system of polynomial equations in the parameters of L that only depends on C_L. If the system is equivalent to $0 = 0$, it follows that $\mathcal{A}'_{L_M} = \mathcal{A}'_{L_N}$. Otherwise there is a point $z \in \operatorname{Arg}^{-1}(U) \cap V'$ such that (6.6) does not hold and we have a contradiction. The lemma follows. $\square$

Proof of Theorem 6.13. We start by replacing (2) with a stronger assertion:

2′. If $U \subset \mathbb{T}^n$ is open and $U \cap \overline{\mathcal{A}}'_{L_N} \neq \emptyset$, then $U \cap \overline{\mathcal{A}}'_{L_M} \neq U \cap \overline{\mathcal{A}}'_{L_N}$ for every $M \nsubseteq N$ with $|M| \geq \max\{1, n-2m+1\}$.

By Lemma 6.7 and the fact that $\partial\mathcal{A}'_L \cap \mathcal{A}'_L \subseteq \mathcal{C}'_L$, (2′) implies (2).

We show the theorem by using induction over the dimension of L. If $\dim L = 0$, then the result is immediate. Assume that the theorem is true for every non-degenerate affine space of dimension strictly less than $n-m$, and let $\dim L = n-m$. Then K_N is non-degenerate with $\dim K_N = n - |N| - m \leq n-m-1$ for every N with $|N| \geq \max\{1, n-2m+1\}$ and hence one can deduce, using the induction hypothesis and (1), that $\operatorname{int}\mathcal{A}'_{L_N}$ is a real, differentiable manifold with

$$\begin{aligned}\dim(\operatorname{int}\mathcal{A}'_{L_N}) &= \dim(\operatorname{int}\mathcal{A}'_{K_N}) + |N| \\ &= 2(n - |N| - m) + |N| \\ &\leq \min\{2n-2m-1, n-1\}.\end{aligned} \tag{6.7}$$

To show (2′), we can apply Lemma 6.7 and Theorem 1.2 to see that it suffices to show that

$$U \cap \pi_M^{-1}\left(\mathcal{A}'_{K_M} \backslash \mathcal{C}'_{K_M}\right) \neq U \cap \pi_N^{-1}\left(\mathcal{A}'_{K_N} \backslash \mathcal{C}'_{K_N}\right),$$

for any U as in the assertion. By Lemma 6.14 this is true unless $\mathcal{A}'_{L_M} = \mathcal{A}'_{L_N}$. But letting $k \in M \backslash N$ we then have that (4) does not hold for $\mathcal{A}'_{K_N}$, which by the induction hypothesis and the fact that $\operatorname{codim} K_N = m > n - |N| - m = \dim K_N$, means that K_N is degenerate, that is $\operatorname{rank} C_{M'} < m$ for some $M' \subseteq M$ with $|M'| = m$. But then also L is degenerate and we have a contradiction. We conclude that (2′) holds.

It follows from (2′) and Lemma 6.7 that $\mathcal{C}'_L \subseteq \overline{\partial\mathcal{A}'_L \backslash \mathcal{C}'_L}$, and thus also that $\mathcal{C}'_L \subseteq \overline{\mathcal{A}'_L \backslash \mathcal{C}'_L}$, and so (1) follows.

To show (3), let $m \geq n/2$ and assume for a contradiction that $\theta \in \operatorname{int}\mathcal{A}'_L \cap \mathcal{A}'_{L_N}$ for some $N \neq \emptyset$. By Theorem 1.2 and the fact that $\mathcal{A}'_{L_N}$ is an initial coamoeba of $\mathcal{A}_{L_M}$ whenever $M \subseteq N$, we can as well assume that $|N| = 1$. Let U be an open neighborhood of θ such that $U \subseteq \mathcal{A}'$. By the second part of Lemma 6.7, $U \cap \mathcal{A}'_{L_N} \subseteq \mathcal{C}'$. On the other hand we have by (2′) combined with the first part of Lemma 6.7, that $U \cap \mathcal{A}'_{L_N} \nsubseteq \mathcal{C}'$. We conclude that $\mathcal{A}'_{L_N} \subseteq \partial\mathcal{A}'$. The other direction in (3) is Lemma 6.12.

Given $k \in \{0, 1, \ldots, n\}$, choose N with $|N| = \min\{1, 2n-2m+1\}$, such that $k \notin N$. Since L is non-degenerate, so is K_N. Hence, by (4) and the induction hypothesis we can choose an open set U such that $e_k \notin T_\theta(\mathcal{A}'_{L_N})$ whenever $\theta \in U$. By (2′) and (3) we can furthermore assume that

$$U \cap \mathcal{A}'_{L_N} = U \cap \partial\mathcal{A}'_L \backslash \mathcal{C}' \neq \emptyset.$$

This means that $T_\theta(\overline{\mathcal{A}}'_L) = T_\theta(\mathcal{A}'_{L_N})$ for $\theta \in U \cap \mathcal{A}'_{L_N}$ and hence e_k is not contained in $T_\theta(\mathcal{A}'_L)$. Since k was arbitrary, this implies (4). The theorem follows by induction. $\square$

6.5. An alternative notion of degeneracy

Recall from Section 2 that the coamoeba of a variety of codimension m minus its contour, is either empty or equals a real manifold of dimension $\min\{n, 2(n-m)\}$. As we saw in Theorem 6.13, $\mathcal{A}' \neq \mathcal{C}'$ whenever L is a non-degenerate affine space (in fact the theorem implies that $\mathcal{C}'$ scarcely affects the appearance of the coamoeba). But could $\mathcal{A}' \neq \mathcal{C}'$ even if L is degenerate?

The answer to the question is yes. Consider, e.g., the line L in $(\mathbb{C}^*)^3$ defined by the equations $2 + z_1 + z_2 = 0$ and $1 + z_3 = 0$. We have that L is degenerate since $\det(C_1, C_2)$ vanishes. On the other hand, as

$$(\sqrt{2}e^{i3\pi/4}, \sqrt{2}e^{-i3\pi/4}, -1) \in L,$$

Proposition 6.5 implies that $(3\pi/4, -3\pi/4, \pi) \in \mathcal{A}' \backslash \mathcal{C}'$.

We say that L is *strongly degenerate* if there is an $N \neq \emptyset$ such that

$$\operatorname{rank} C_N \leq \min\{m, n/2\} - |N|/2.$$

It follows that L is degenerate if it is strongly degenerate. The motivation for the latter definition is partly the next result.

Proposition 6.15. *If L is a strongly degenerate affine space, then its amoeba and coamoeba equal their contours.*

Proof. Choose N so that $|N| + 2\operatorname{rank} C_N$ is minimal and fix $\theta \in \mathcal{A}'$. Since L is strongly degenerate, $\operatorname{rank} C_N^\theta \leq 2\operatorname{rank} C_N \leq n - |N|$. Thus the number of columns of C_N^θ is strictly larger than the rank of C_N^θ. Hence, for some $M \supseteq N$, every column of C_M^θ is contained in the linear span of the other columns, and $\operatorname{rank} C_M^\theta = \operatorname{rank} C_N^\theta - (|M| - |N|)$. Hence $G(\theta)$ has dimension at least

$$\begin{aligned} n + 1 - |M| - \operatorname{rank} C_M^\theta &= n + 1 - |N| - \operatorname{rank} C_N^\theta \\ &\geq n + 1 - |N| - 2\operatorname{rank} C_N. \end{aligned}$$

Since L is strongly degenerate, this means that $\dim G(\theta) > \max\{n - 2m, 0\}$ and it follows that $\theta \in \mathcal{C}'$. □

When $m = n/2$, the notion of strong degeneracy is particularly essential.

Proposition 6.16. *If the codimension of L is $n/2$, then $\mathcal{A}' = \mathcal{C}'$ if and only if L is strongly degenerate.*

Proof of Theorem 5.1 *and Proposition* 6.16. We want to show that the three assertions $\mathcal{A}' = \mathcal{C}'$, L is strongly degenerate and $\operatorname{Vol} \mathcal{A}' \neq \pi^n$ are all equivalent. If L is strongly degenerate, then the two other assertions follows by Proposition 6.15. Now assume that L is not strongly degenerate. It suffices to show that the two other assertions do not hold either.

For a fixed $\theta \in \mathbb{T}^n$, consider the system $C^\theta r = 0$ for $r \in R\mathbb{P}^n$. If C^θ has full rank, then there is a unique solution r to this system. Hence among all 2^n translations of θ by π in any set of coordinates, there is a unique translation $\tilde{\theta}$ such that the solution r of $C_{\tilde{\theta}} r = 0$ has the same sign in each coordinate, or equivalently, $\tilde{\theta} \in \mathcal{A}'_{L_N}$ for some $N \subset \{0, 1, \dots, n\}$. If on the other hand C^θ does not have full rank, then there is at least one translation $\tilde{\theta}$ as above, that is $\tilde{\theta} \in \mathcal{C}'_{L_N}$. Letting S be the subset of points $\theta \in \mathbb{T}^n$ for which $\operatorname{rank} C^\theta = n$, we conclude that

$$\operatorname{Vol}\Big(S \cap \bigcup_N \mathcal{A}'_{L_N} \Big) = \operatorname{Vol}(S)/2^n, \tag{6.8}$$

$$\operatorname{Vol}\Big(S^c \cap \bigcup_N \mathcal{C}'_{L_N} \Big) \geq \operatorname{Vol}(S^c)/2^n. \tag{6.9}$$

Hence we are done if we show that $\dim \mathcal{A}'_{L_N} < n$ for every $N \neq \emptyset$. Note that by Corollary 6.7, this also means that $\dim \mathcal{C}' < n$ which by (6.9) means that C^θ has full rank for almost every $\theta \in \mathbb{T}^n$, so that S in (6.8) can be exchanged with the whole torus and the right-hand side of (6.8) equals $(2\pi)^n/2^n = \pi^n$.

Recalling the space K_N from 6.4, we have for any set of indices N that

$$\begin{aligned}
\dim \mathcal{A}'_{L_N} &= \dim \mathcal{A}'_{K_N} + |N| \\
&= \min\{n - |N|, 2(n - |N| - \operatorname{rank} C_N)\} + |N| \\
&\leq 2n - |N| - 2 \operatorname{rank} C_N,
\end{aligned}$$

that is, $\operatorname{rank} C_N \leq (2n - |N| - \dim \mathcal{A}'_{L_N})/2$. But since L is not strongly degenerate and $N \neq \{0, 1, \dots, n\}$, we also have by definition that $\operatorname{rank} C_N \geq (n + 1 - |N|)/2$. Hence $\dim \mathcal{A}'_{L_N} \leq n - 1$ as desired. $\square$

In general, strong degeneracy is not necessary for the coamoeba to equal its contour. For example, the coamoeba of the not strongly degenerate line L with

$$C_L = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

is the line $\{(0, t, t); t \in \mathbb{R}\}$ and the amoeba $\{(0, x, x); x \in \mathbb{R}\}$ equals the fiber at any point of the coamoeba.

References

[1] I. A. Antipova. Inversion of multidimensional Mellin transforms. *Uspekhi Mat. Nauk*, 62(5(377)):147–148, 2007.

[2] Mikael Forsberg, Mikael Passare, and August Tsikh. Laurent determinants and arrangements of hyperplane amoebas. *Adv. Math.*, 151(1):45–70, 2000.

[3] I. M. Gel'fand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1994.

[4] Petter Johansson. The argument cycle and the coamoeba. *Complex Var. Elliptic Equ.*, 58(3):373–384, 2013.

[5] Christer O. Kiselman. Questions inspired by Mikael Passare's mathematics. *Afr. Mat.*, 25(2):271–288, 2014.

[6] K. V. Kuzvesov. *Amoebas of complex plane difference equations*. PhD thesis, Krasnoyarsk State University, Krasnoyarsk, 2007.

[7] Grigory Mikhalkin and Hans Rullgård. Amoebas of maximal area. *Internat. Math. Res. Notices*, 9:441–451, 2001.

[8] Lisa Nilsson and Mikael Passare. Discriminant coamoebas in dimension two. *J. Commut. Algebra*, 2(4):447–471, 2010.

[9] Mounir Nisse. Amoebas and coamoebas of linear spaces. Manuscript, 2012.

[10] Mounir Nisse and Frank Sottile. The phase limit set of a variety. *Algebra Number Theory*, 7(2):339–352, 2013.

[11] Mikael Passare and Hans Rullgård. Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope. *Duke Math. J.*, 121(3):481–507, 2004.

[12] Bernd Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.

Petter Johansson
Högskolan i Borås
S-501 90 Borås, Sweden
e-mail: petter.johansson@hb.se

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 287–309

Convexity of Marginal Functions in the Discrete Case

Christer O. Kiselman and Shiva Samieinia

Abstract. We define, using difference operators, classes of functions defined on the set of points with integer coordinates which are preserved under the formation of marginal functions. The duality between classes of functions with certain convexity properties and families of second-order difference operators plays an important role and is explained using notions from mathematical morphology.

Mathematics Subject Classification (2010). 39A12, 39A14, 52A41.

Keywords. Marginal function, discrete convexity, difference operators, A-lateral convexity, rhomboidal convexity, mathematical morphology, infimal convolution.

Prologue

This paper is dedicated to the memory of Mikael Passare: student, mentor, and friend; a great mathematician and a great human being.

After his brilliant achievements in the theory of several complex variables, in particular residue theory, Mikael turned his energy to amoebas and their spines, which are tropical hyperplanes. Tropical geometry was at the time a rather new research area, and he considered his change of focus as an important one, both mentally and scientifically. As far as we know, he did not work on digital geometry or discrete optimization, but he showed great respect for the problems encountered there, which was evident for instance during the preparation of his manuscript later published as Passare (2009). There are in fact strong analogies between tropicalization and discretization. The operation of taking the marginal function is a special case of infimal convolution, which in turn is a tropicalization of ordinary convolution – we have a link to Mikael's interest in tropical geometry. Euclidean geometry, digital geometry, and tropical geometry are three kinds of geometry with contrasting properties. They can support and enrich each other. Together with mathematical morphology and discrete optimization, they constitute research areas with many applications in technology and the sciences.

1. Introduction

1.1. The marginal function of a function of real variables

A simple everyday observation is that the shadow of a convex body is convex. Mathematically this means that the image under an affine mapping of a convex subset of a vector space is convex. It is convenient to express this in terms of marginal functions:

Definition 1.1. If F is a function defined on $\mathbf{R}^n \times \mathbf{R}^m$ and with values in the set of extended real numbers, which we denote by

$$\mathbf{R}_! = [-\infty, +\infty] = \mathbf{R} \cup \{-\infty, +\infty\},$$

then its *marginal function* $H\colon \mathbf{R}^n \to \mathbf{R}_!$ is defined by

$$H(x) = \inf_{y \in \mathbf{R}^m} F(x, y), \qquad x \in \mathbf{R}^n. \qquad \square$$

For completeness we also give the definition of a convex function:

Definition 1.2. A function $F\colon \mathbf{R}^n \to \mathbf{R}_!$ is said to be *convex* if it satisfies Jensen's inequality

$$\begin{cases} \text{For all real numbers } t \text{ with } 0 < t < 1 \text{ and all } x, y \in \mathbf{R}^n \\ \text{such that } F(x), F(y) < +\infty \text{ we have} \\ F((1-t)x + ty) \leqslant (1-t)F(x) + tF(y). \end{cases} \tag{1.1}$$

We shall denote the set of all convex functions by $CVX(\mathbf{R}^n, \mathbf{R}_!)$ and the subset of functions with finite values by $CVX(\mathbf{R}^n, \mathbf{R})$. $\square$

If F is convex, then so is its marginal function H. The proof of this result is completely elementary – and therefore usually mentioned only in passing in the textbooks. The result has nevertheless manifold uses in the applications of the theory for convex functions of real variables.

1.2. The marginal function of a function of integer variables

It would be of interest to establish a similar result for functions $f\colon \mathbf{Z}^n \times \mathbf{Z}^m \to \mathbf{R}_!$, i.e., functions defined at the points in $\mathbf{R}^n \times \mathbf{R}^m$ with integer coordinates. This is what we shall do here.

Definition 1.3. If $f\colon \mathbf{Z}^n \times \mathbf{Z}^m \to \mathbf{R}_!$, we define its *marginal function* h by

$$h(x) = \inf_{y \in \mathbf{Z}^m} f(x, y), \qquad x \in \mathbf{Z}^n. \qquad \square$$

The question now arises which kind of convexity we shall use. A first, seemingly most natural, definition is the following.

Definition 1.4. A function $f\colon \mathbf{Z}^n \to \mathbf{R}_!$ is said to be *convex extensible*[1] if it is the restriction to $\mathbf{Z}^n$ of a convex function defined in $\mathbf{R}^n$. The set of all convex-extensible functions will be denoted by $CVX(\mathbf{R}^n, \mathbf{R}_!)|_{\mathbf{Z}^n}$; the subset of functions which have a real-valued convex extension by $CVX(\mathbf{R}^n, \mathbf{R})|_{\mathbf{Z}^n}$. □

It should be noted that $CVX(\mathbf{R}^n, \mathbf{R})|_{\mathbf{Z}^n}$ is equal to the set of real-valued functions in $CVX(\mathbf{R}^n, \mathbf{R}_!)|_{\mathbf{Z}^n}$.

For $n = 1$, the convex-extensible functions are precisely those which satisfy the special case of Jensen's inequality (1.1) with $x \in \mathbf{Z}$, $y = x + 2$, $t = \frac{1}{2}$. While there are many different notions of discrete convexity in $\mathbf{Z}^n$, $n \geqslant 2$, there is only one reasonable notion of discrete convexity for $n = 1$: The one just described.

Let us now formulate the problem explicitly.

Problem 1.5. Define, for $n = 1, 2, \ldots$, classes $\mathscr{M}_n$ of functions defined in $\mathbf{Z}^n$ such that $\mathscr{M}_1 = CVX(\mathbf{R}, \mathbf{R})|_{\mathbf{Z}}$ and such that the successive marginal functions $h_{n-1}, h_{n-2}, \ldots, h_1$ of any function $f \in \mathscr{M}_n$ defined by $h_n = f$,

$$h_k(x) = \inf_{t \in \mathbf{Z}} h_{k+1}(x, t), \qquad x = (x_1, \ldots, x_k) \in \mathbf{Z}^k, \quad k = n-1, \ldots, 1,$$

belong to $\mathscr{M}_k$ whenever they do not take the value $-\infty$. □

We should also require that the classes are large enough so as to avoid trivial results, e.g., by taking $\mathscr{M}_n$ as the set of all functions $f\colon \mathbf{Z}^n \to \mathbf{R}$ such that $f(x_1, \ldots, x_n) = g(x_1)$ for some $g \in CVX(\mathbf{R}, \mathbf{R})|_{\mathbf{Z}}$.

We shall define in this paper classes of functions defined on the integer points which solve the problem completely; see Theorem 11.1. Since this theorem is very general, we have formulated a corollary (Corollary 11.3), which is perhaps easier to apply. The functions of interest are called *A-laterally convex*, where A is a subset of $\mathbf{Z}^n \times \mathbf{Z}^n$ (see Definition 6.2). This subset A determines a family of second-order difference operators; there is a duality between such families and classes of functions with certain convexity properties, which we explain in Section 7 using basic notions of mathematical morphology.

Moreover, we shall prove that the classes obtained are optimal in a natural sense (see Examples 9.2 and 9.3, and Section 12).

The most obvious attempt at defining a convex function of integer variables, i.e., taking $\mathscr{M}_n = CVX(\mathbf{R}^n, \mathbf{R})|_{\mathbf{Z}^n}$, fails in a very conspicuous way, even in low dimensions, as we shall see now.

Example 1.6. Define $f\colon \mathbf{Z} \times \mathbf{Z} \to \mathbf{Z}$ by

$$f(x, y) = |x - 2my|, \qquad (x, y) \in \mathbf{Z} \times \mathbf{Z},$$

where m is a positive integer. Then its marginal function

$$h(x) = \inf_{y \in \mathbf{Z}} f(x, y), \qquad x \in \mathbf{Z},$$

[1] This term has been used in a different, narrower sense by Murota (2003:93); for an example showing this, see Kiselman (2011: Example 3.5).

is a periodic function of period $2m$ which is equal to $|x|$ for $-m \leqslant x \leqslant m$. This means that it is a saw-tooth function with teeth as large as we like. We remark also that if we define f in $\mathbf{R} \times \mathbf{Z}$ by the same expression, then the same phenomenon appears. □

The function f in Example 1.6 is actually convex extensible; indeed, an extension is given by the same expression, while h is not convex extensible (or convex in any reasonable sense). Our conclusion is that the property of being convex extensible is too weak to be of use in this context. In view of this observation, one of us has studied a class of functions defined on $\mathbf{Z} \times \mathbf{Z}$ which is suitable for this and other important properties in convexity theory; see Kiselman (2008; 2010a).

The purpose of the present paper is to extend this study to higher dimensions, i.e., to functions on $\mathbf{Z}^n \times \mathbf{Z}^m$.

A kind of convexity called integral convexity was introduced by Favati and Tardella (1990) using locally convex functions. A function $f\colon \mathbf{Z}^n \to \mathbf{R}$ is called *integrally convex* if its convex extension over unit cubes is convex in all of $\mathbf{R}^n$. Integrally convex functions are all convex extensible, and their local minima are global. The class has the property of being invariant under simple coordinate transformations: If we put $g(x,y) = f(x,-y)$, $(x,y) \in \mathbf{Z}^2$, then f and g are integrally convex at the same time, and f and g have the same marginal function:

$$\inf_{y\in\mathbf{Z}} f(x,y) = \inf_{y\in\mathbf{Z}} g(x,y), \qquad x \in \mathbf{Z}.$$

Several of the other classes mentioned in Section 2 do not have this property, which implies that they are not suited for the study of marginal functions – and indeed provide poor analogues of convex functions of real variables, which are invariant under such simple coordinate transformations.

In the case of two integer variables, there are several equivalent ways to define integral convexity. In Kiselman (2008) integral convexity was introduced using difference operators. From this characterization it is obvious that the class is closed under addition.

The present paper is an elaborated version of our paper (2010), which was part of the second author's PhD thesis.

1.3. Relations between Minkowski addition, infimal convolution, and the operation of taking the marginal function

The *Minkowski sum* of two sets A and B is defined as

$$A + B = \{a + b;\ a \in A,\ b \in B\}, \qquad A, B \subset \mathbf{R}^n.$$

This very fundamental operation gives rise to *infimal convolution*, which is defined as the operation $(f,g) \mapsto f \sqcap g$, where

$$(f \sqcap g)(x) = \inf_{y\in\mathbf{Z}^n} \bigl(f(x-y) \dotplus g(y)\bigr), \qquad x \in \mathbf{R}^n,\ f,g\colon \mathbf{R}^n \to \mathbf{R}_!.$$

Here $x \dotplus y$, $x, y \in \mathbf{R}_!$, is the *upper sum* of x and y, which extends the sum of real numbers and takes the value $+\infty$ if one of x and y equals $+\infty$. As explained in, e.g., Kiselman (2015: §6), this is a tropicalization of the usual bilinear convolution product defined in (5.1) below.

If we choose $g(x)$ to be zero when $x_1 = x_2 = \cdots = x_m = 0$ and $+\infty$ elsewhere, then $f \sqcap g$ is the marginal function h of f defined as

$$h(x_1, \ldots, x_m) = \inf_{x_{m+1},\ldots,x_n} f(x), \qquad (x_1, \ldots, x_m) \in \mathbf{R}^m.$$

Thus marginal functions are a special case of infimal convolution.

In the other direction every infimal convolution is a marginal function, viz. the marginal function of the special function of $2n$ variables $(x, y) \mapsto f(x - y) \dot{+} g(y)$ when y varies.

So the two operations are actually equivalent, however at the expense of going up in dimension when viewing infimal convolution as a marginal function.

Infimal convolution in turn is a case of Minkowski addition. Indeed,

$$\mathbf{epi}_{\mathrm{s}}^{\mathrm{F}}(f \sqcap g) = \mathbf{epi}_{\mathrm{s}}^{\mathrm{F}}(f) + \mathbf{epi}_{\mathrm{s}}^{\mathrm{F}}(g),$$

where $\mathbf{epi}_{\mathrm{s}}^{\mathrm{F}}(f)$ is the *strict finite epigraph* of f, defined as

$$\mathbf{epi}_{\mathrm{s}}^{\mathrm{F}}(f) = \{(x, t) \in \mathbf{R}^n \times \mathbf{R};\ t > f(x)\}, \qquad f \colon \mathbf{R}^n \to \mathbf{R}_!.$$

2. Other notions of discrete convexity

Several kinds of discrete convexity have been studied. Miller (1971:168), introduced discretely convex functions for which local minima are global. These functions are not convex extensible – nor is the class closed under addition; see Murota & Shioura (2001:156, 161).

Two other concepts of convexity were introduced by Murota (1996; 1998). They are called M-convexity and L-convexity, respectively. For functions with either of these two properties, local minima are global. Two other classes of functions are obtained by a special restriction of M- and L-convex functions to a space of one dimension less. These functions are called M$^\natural$-convex and L$^\natural$-convex.[2] They were introduced by Murota & Shioura (1999:96) and Fujishige & Murota (2000:135), respectively. The class of M$^\natural$-convex (L$^\natural$-convex) functions properly contains the class of M-convex (L-convex) functions. These classes of functions have been studied with respect to some operations such as infimal convolution, addition, and addition by an affine function; see Murota & Shioura (2001). However, these classes are quite small (see Example 9.5).

[2]These expressions should be read, respectively, as "M-natural-convex" (Murota 2003:27, footnote 23), and "L-natural-convex" (Murota 2003:23, footnote 18). Here M stands for *matroid* and L for *lattice* (Murota 2003:xxi).

3. The convex hull and the convex envelope

Definition 3.1. The *convex hull* of a subset A of $\mathbf{R}^n$ is the smallest convex set containing A. It will be denoted by $\mathbf{cvxh}(A)$. □

Definition 3.2. The *convex envelope* of a function $f\colon A \to \mathbf{R}_!$, where A is any subset of $\mathbf{R}^n$, is the largest convex function $G\colon \mathbf{R}^n \to \mathbf{R}_!$ such that $G|_A \leqslant f$. We shall denote it by $\mathbf{cvxe}(f)$. □

The convex envelope is well defined because the supremum of all functions H which are convex and satisfy $H|_A \leqslant f$ has the same properties.

A function f is convex extensible if and only if $\mathbf{cvxe}(f)$ is an extension of f. Indeed, if f admits a convex extension, then also $\mathbf{cvxe}(f)$ is a convex extension. Equivalently, $\mathbf{cvxe}(f)|_A \geqslant f$.

4. The integer neighborhood and the canonical extension

Definition 4.1. We define the *integer neighborhood* of a real number a, denoted by $N(a)$, as the set $\{\lfloor a \rfloor, \lceil a \rceil\} \subset \mathbf{Z}$. We define the *integer neighborhood* of a point $a = (a_1, \dots, a_n) \in \mathbf{R}^n$ as the set

$$N(a) = N(a_1) \times \cdots \times N(a_n) \subset \mathbf{Z}^n. \qquad \square$$

The integer neighborhood has 2^k elements, where k is the number of indices j such that $a_j \in \mathbf{R} \smallsetminus \mathbf{Z}$. Equivalently,

$$N(a) = \big(a + B^{\infty}_{<}(0,1)\big) \cap \mathbf{Z}^n, \qquad a \in \mathbf{R}^n,$$

where $B^{\infty}_{<}(c,r)$ denotes the strict ball for the l^∞ norm with center at c and of radius r. The mapping

$$\mathbf{R}^n \supset A \mapsto \nu(A) = \bigcup_{a \in A} N(a) \subset \mathbf{Z}^n$$

is one of many digitizations of $\mathbf{R}^n$ and commutes with the formation of arbitrary unions, i.e.,

$$\nu\big(\textstyle\bigcup_{j\in J} A_j\big) = \displaystyle\bigcup_{j \in J} \nu(A_j), \qquad A_j \subset \mathbf{R}^n.$$

In mathematical morphology this is an important concept: a mapping with this property is said to be a *dilation*.

Definition 4.2. The *canonical extension* of a function $f\colon \mathbf{Z}^n \to \mathbf{R}_!$ is defined, for every $a \in \mathbf{R}^n$, as the value at a of the convex envelope of $f|_{N(a)}$, the restriction of f to the integer neighborhood of a. We shall denote it by $\mathbf{can}(f)\colon \mathbf{R}^n \to \mathbf{R}_!$. □

The canonical extension is actually an extension, since $N(a) = \{a\}$ for every $a \in \mathbf{Z}^n$.

Proposition 4.3. *For any function $f\colon \mathbf{Z}^n \to \mathbf{R}_!$, any point $a \in \mathbf{R}^n$, and any $p \in \mathbf{Z}^n$ such that the cube $p + [0,1]^n$ contains a, the value of the canonical extension at a is equal to the value at a of the convex envelope of $f|_{p+\{0,1\}^n}$.*

Proof. For brevity, let us denote by $C(p)$ the cube $p+[0,1]^n$ and by $V(p)$ its set of vertices, $p+\{0,1\}^n$, for any point p with integer coordinates.

If a point a belongs to only one cube $C(p)$, then $N(a)=V(p)$ and there is nothing to prove.

However, a point a may belong to two different cubes $C(p)$ and $C(q)$, $p,q\in\mathbf{Z}^n$, $p\neq q$. Then $N(a)$ is a subset of $V(p)\cap V(q)$. Since $N(a)$ is a subset of $V(p)$ if $a\in V(p)$, we have $\mathbf{cvxe}(f|_{N(a)})\geqslant\mathbf{cvxe}(f|_{V(p)})$. To prove the converse inequality, we define, given $a\in\mathbf{R}^n$ and $p\in\mathbf{Z}^n$ such that $a\in C(p)$, two sets of indices

$$J_k=\{j\in[1,n]_{\mathbf{Z}};\ a_j=p_j+k\},\qquad k=0,1,$$

and an affine function

$$G(x)=\sum_{j\in J_0}(x_j-p_j)+\sum_{j\in J_1}(p_j+1-x_j),\qquad x\in\mathbf{R}^n.$$

If both J_0 and J_1 are empty, then G is identically zero and $C(p)$ is the only cube to which a belongs. We now assume that this is not the case. Then the zero set of G is a hyperplane $Y(a,p)$ in $\mathbf{R}^n$. Obviously G is nonnegative in the cube $C(p)$, and $Y(a,p)$ is a supporting hyperplane of this cube. In general a supporting hyperplane intersects $V(p)$ in a set which contains other vertices than those in $N(a)$, but in view of our construction, the hyperplane has the important property that

$$Y(a,p)\cap V(p)=N(a).$$

This implies that any convex combination of points in $V(p)$ yielding a point in $Y(a,p)$ is already a convex combination of points in $N(a)$. This proves that the convex envelope of $f|_{V(p)}$ and the convex envelope of $f|_{N(a)}$ have the same value at a. We are done. □

Definition 4.4. We shall say with Favati and Tardella (1990:9), that a function $f\colon\mathbf{Z}^n\to\mathbf{R}_!$ is *integrally convex* if $\mathbf{can}(f)\colon\mathbf{R}^n\to\mathbf{R}_!$ is convex. □

We always have $\mathbf{cvxe}(f)\leqslant\mathbf{can}(f)$ with equality if and only if f is integrally convex. Every integrally convex function is convex extensible, since for such a function, $\mathbf{can}(f)$ is a convex extension.

5. Convolution and convex extensibility

The *convolution product* $f*g$ of two functions $f,g\colon\mathbf{Z}^n\to\mathbf{R}$ is defined by

$$(f*g)(x)=\sum_{y\in\mathbf{Z}^n}f(x-y)g(y),\qquad x\in\mathbf{Z}^n, \tag{5.1}$$

assuming some kind of convergence.

We define for $p=(p^{(1)},\dots,p^{(k)})\in(\mathbf{R}^n)^k$ and $\lambda=(\lambda_1,\dots,\lambda_k)\in\mathbf{R}^k$ satisfying $\lambda_j\geqslant 0$, $\sum_{j=1}^k\lambda_j=1$, and $\sum_{j=1}^k\lambda_jp^{(j)}=0$,

$$\mu_{p,\lambda}=\sum_{j=1}^k\lambda_j\delta_{p^{(j)}}.$$

Here δ_a denotes the *Kronecker delta* placed at a, defined by $\delta_a(a) = 1$ and $\delta_a(x) = 0$ when $x \neq a$. In particular, δ_0 is a neutral element: $f * \delta_0 = f$ for all functions f.

The convex envelope of a function defined on a subset A of $\mathbf{R}^n$ is given by

$$\mathbf{cvxe}(f)(x) = \inf_{p,\lambda}(\mu_{p,\lambda} * f)(x) = \inf_{p,\lambda} \sum_{y \in A} \mu_{p,\lambda}(x-y)f(y), \qquad x \in \mathbf{R}^n.$$

(In view of Carathéodory's theorem it suffices to take $k = n + 1$.)

This implies that convex extensibility of a function f defined on a subset A of $\mathbf{R}^n$ can be characterized by means of an infinite family of convolution operators, viz.

$$((\mu_{p,\lambda} - \delta_0) * f)(x) = \sum_{y \in A} \mu_{p,\lambda}(x-y)f(y) - f(x) \geqslant 0, \qquad x \in A,$$

for all p and λ of the kind mentioned.

When $n = 1$, the convex-extensible functions are those which satisfy the inequality $(\mu_{p,\lambda} - \delta_0) * f \geqslant 0$ for $p = (-1, 1)$ and $\lambda = (\frac{1}{2}, \frac{1}{2})$, thus defining the class using a single convolution operator.

6. Lateral convexity: Definition

The following definition extends that for two variables in Kiselman (2008: Definition 2.1); cf. Theorem 2.4 there. See also Kiselman (2011).

Definition 6.1. Given $a \in \mathbf{R}^n$, we define a difference operator $D_a \colon \mathbf{R}^{\mathbf{R}^n} \to \mathbf{R}^{\mathbf{R}^n}$ by

$$(D_a F)(x) = F(x+a) - F(x), \qquad x \in \mathbf{R}^n,\ F \in \mathbf{R}^{\mathbf{R}^n}. \tag{6.1}$$

□

If $a \in \mathbf{Z}^n$, D_a operates from $\mathbf{R}^{\mathbf{Z}^n}$ to $\mathbf{R}^{\mathbf{Z}^n}$ and from $\mathbf{Z}^{\mathbf{Z}^n}$ to $\mathbf{Z}^{\mathbf{Z}^n}$. In particular, $D_{e^{(j)}}$, where $e^{(j)}$ is the vector $(0, 0, \dots, 1, \dots, 0)$ with 1 at the j^{th} place, is the difference operator in the j^{th} coordinate.

The operator $f \mapsto D_a f$ is a convolution operator: $D_a f = \mu_a * f$ with $\mu_a = \delta_{-a} - \delta_0$. The composition of D_a and D_b is the convolution operator given by $D_b D_a f = (\mu_b * \mu_a) * f$ with $\mu_b * \mu_a = \delta_{-a-b} - \delta_{-a} - \delta_{-b} + \delta_0$.

The following definition generalizes several definitions used to define discrete convexity. As will be shown, it is highly relevant for problems concerning marginal functions.

Definition 6.2. Given a set $A \subset \mathbf{Z}^n \times \mathbf{Z}^n$, we shall say that a function $f \colon \mathbf{Z}^n \to \mathbf{R}$ is A-*laterally convex* if

$$(D_b D_a f)(x) \geqslant 0, \qquad x \in \mathbf{Z}^n, \quad (a, b) \in A. \tag{6.2}$$

We define $\Phi(A)$ as the set of all A-laterally convex functions.

In the other direction, given any subset F of $\mathbf{R}^{\mathbf{Z}^n}$, we define $\Psi(F)$ as the set of all pairs $(a, b) \in \mathbf{Z}^n \times \mathbf{Z}^n$ such that $D_b D_a f \geqslant 0$ for all $f \in F$. □

7. Lateral convexity: Morphological aspects

The notions of mathematical morphology are very helpful when it comes to understanding lateral convexity.

The mappings Φ and Ψ are decreasing and $\Psi \circ \Phi$ and $\Phi \circ \Psi$ are larger than the respective identity mappings. One expresses this fact by saying that the pair of mappings (Φ, Ψ) forms a *Galois connection.* This fact can also be expressed using the concept of the lower inverse of a mapping between ordered sets; see Kiselman (2010b: Subsection 4.1).

We define $\widetilde{A} = \Psi(\Phi(A))$ for any subset of $\mathbf{Z}^n \times \mathbf{Z}^n$. It is well known from Galois theory and easy to see that the operation $A \mapsto \widetilde{A}$ is increasing and idempotent, thus an *ethmomorphism* (a morphological filter). It is also larger than the identity, and so it is a *cleistomorphism* (a closure operator).

If a function is A-laterally convex, it is automatically $\widetilde{A}$-laterally convex; any set B satisfying $A \subset B \subset \widetilde{A}$ defines the same class of functions.

From the definition it is obvious that the class of A-laterally convex functions is closed under addition and multiplication by a nonnegative scalar. From the formulas

$$(D_{-a}f)(x) = -(D_a f)(x-a), \qquad (D_{-b}D_{-a}f)(x) = (D_b D_a f)(x-a-b)$$

it follows that

$$-A = \{(-a,-b); (a,b) \in A\}$$

is contained in $\widetilde{A}$. The same is true of

$$A^{\smile} = \{(b,a); (a,b) \in A\}.$$

We define

$$A^{\mathrm{sym}} = A \cup (-A) \cup A^{\smile} \cup (-A)^{\smile},$$

which may have up to four times as many elements as A but still defines the same class, i.e., $\Phi(A^{\mathrm{sym}}) = \Phi(A)$.

The formula

$$D_b D_{-a} f(x) = -D_b D_a f(x-a)$$

shows that f is $\{(-a,b)\}$-laterally convex if and only if $-f$ is $\{(a,b)\}$-laterally convex. So the concepts introduced will enable us to study also A-laterally concave functions and A-laterally affine functions.

The formula

$$(D_b f)(x) + (D_c f)(x+b) = (D_{b+c} f)(x)$$

applied to $D_a f$ yields

$$(D_b D_a f)(x) + (D_c D_a f)(x+b) = (D_{b+c} D_a f)(x), \tag{7.1}$$

which implies that if $D_b D_a f \geqslant 0$ and $D_c D_a f \geqslant 0$, then we also have $D_{b+c} D_a f \geqslant 0$. This means that the set of pairs $\{(a,b) \in \mathbf{Z}^n \times \mathbf{Z}^n\}$ such that the inequality holds is closed under *partial addition:*

$$(a,b) +_2 (a,c) = (a, b+c), \tag{7.2}$$

i.e., if the first elements agree, we may add the second elements. For sets we define

$$B +_2 C = \{(a, b+c); (a,b) \in B, (a,c) \in C\}.$$

Similarly we can define of course

$$(a,b) +_1 (c,b) = (a+c, b) \tag{7.3}$$

and

$$B +_1 C = \{(a+c, b); (a,b) \in B, (c,b) \in C\}$$

when the two second elements are the same.

By repeated use of these formulas we see that $\widetilde{A}$ contains the sets

$$A^{\mathrm{sym}} +_1 A^{\mathrm{sym}},\ A^{\mathrm{sym}} +_2 (A^{\mathrm{sym}} +_1 A^{\mathrm{sym}})$$

and so on. We sum up the discussion on $\widetilde{A}$ in the following lemma.

Lemma 7.1. *Let A be any subset of $\mathbf{Z}^n \times \mathbf{Z}^n$ and define $\widetilde{A} = \Psi(\Phi(A))$.*

1. *For any $a \in \mathbf{Z}^n$, $(a, \mathbf{0})$ and $(\mathbf{0}, a)$ belong to $\widetilde{A}$.*
2. *If $(a,b) \in \widetilde{A}$, then $(b,a), (-a,-b), (-b,-a)$ all belong to $\widetilde{A}$.*
3. *If $(a,b), (c,b) \in \widetilde{A}$, then $(a,b) +_1 (c,b) = (a+c, b)$ belongs to $\widetilde{A}$.*
4. *If $(a,b), (a,c) \in \widetilde{A}$, then $(a,b) +_2 (a,c) = (a, b+c)$ belongs to $\widetilde{A}$.*
5. *For any given set F of functions $\mathbf{Z}^n \to \mathbf{R}$, if $\Psi(F)$ contains a set A, it also contains $\widetilde{A}$.* □

When $n = 1$ and $A = \{(1,1)\}$, f is A-laterally convex if and only if it is convex extensible. As already mentioned, this is the only reasonable definition of convexity in one integer variable. We note that it is equivalent to B-lateral convexity for any B such that

$$(1,1) \in B \subset \widetilde{A} \text{ or } (-1,-1) \in B \subset \widetilde{A}.$$

In this case, $\widetilde{A}$ is easy to determine: It is equal to

$$\{(s,t) \in \mathbf{Z} \times \mathbf{Z}; st \geqslant 0\}.$$

More generally, for any n and any $j \in [1,n]_{\mathbf{Z}}$, if $A = \{(e^{(j)}, e^{(j)})\}$, then a function is A-laterally convex if and only if it is convex extensible in the variable x_j when the others are kept fixed. Since this is a convenient property, we shall normally require that

$$(e^{(j)}, e^{(j)}) \in A, \qquad j = 1, \dots, n. \tag{7.4}$$

If this is so, all A-laterally convex functions are $\{(1,1)\}$-laterally convex in each variable when the others are kept fixed.

8. Lateral convexity: Examples

Example 8.1. If f is the restriction to $\mathbf{Z}^n$ of a polynomial of degree at most two,

$$f(x) = \alpha + \sum_{j=1}^{n} \beta_j x_j + \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma_{jk} x_j x_k, \qquad x \in \mathbf{Z}^n,$$

with $\gamma_{jk} = \gamma_{kj}$, we see that

$$(D_b D_a f)(x) = 2 \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma_{jk} a_j b_k,$$

so that f is A-laterally convex if and only if the last expression is nonnegative for all $(a, b) \in A$.

In particular, the restriction to $\mathbf{Z}^n$ of an arbitrary affine function is A-laterally convex.

We also see that the special polynomial $f(x) = x_j^2$ is A-laterally convex if and only if $a_j b_j \geqslant 0$ for all $(a, b) \in A$. Conversely, if $a_j b_j \geqslant 0$ and g is any convex-extensible function of one variable, then the function $x \mapsto g(x_j)$ is $\{(a, b)\}$-laterally convex. $\square$

In view of this example we shall normally require that

$$(a, b) \in A \text{ implies } a_j b_j \geqslant 0, \qquad j = 1, \dots, n. \tag{8.1}$$

Example 8.2. A special kind of laterally convex functions are the L$^\natural$-convex functions, which are defined by Murota in (2003: 1.33) by the property

$$f(\lfloor \tfrac{1}{2}x + \tfrac{1}{2}y \rfloor) + f(\lceil \tfrac{1}{2}x + \tfrac{1}{2}y \rceil) \leqslant f(x) + f(y), \qquad x, y \in \mathbf{Z}^n. \tag{8.2}$$

A function $f \colon \mathbf{Z}^n \to \mathbf{R}$ is L$^\natural$-convex if and only if it is Λ-laterally convex with

$$\Lambda = \{(a, b) \in \mathbf{Z}^n \times \mathbf{Z}^n;\ b - a \in \{0, 1\}^n \cup \{-1, 0\}^n\}.$$

So, in all dimensions, L$^\natural$-convexity is a special case of lateral convexity.

We shall prove first that Λ-lateral convexity implies L$^\natural$-convexity. Let x and y be given and define

$$a = \lfloor \tfrac{1}{2}x + \tfrac{1}{2}y \rfloor - x \text{ and } b = \lceil \tfrac{1}{2}x + \tfrac{1}{2}y \rceil - x.$$

Note that $b - a \in \{0, 1\}^n \subset \Lambda$. Then $x + a + b = y$, so that, if f is Λ-laterally convex, we obtain $f(\lfloor \frac{1}{2}x + \frac{1}{2}y \rfloor) + f(\lceil \frac{1}{2}x + \frac{1}{2}y \rceil) = f(x+a) + f(x+b) \leqslant f(x) + f(x+a+b) = f(x) + f(y)$, proving (8.2).

Next we shall see that L$^\natural$ convexity implies Λ-lateral convexity. If x, a and b are given with $b - a \in \{0, 1\}^n \subset \Lambda$, we define $y = x + a + b$. Then $\lfloor \frac{1}{2}x + \frac{1}{2}y \rfloor = x + a$ and $\lceil \frac{1}{2}x + \frac{1}{2}y \rceil = x + b$ so that, if f is L$^\natural$-convex, we get $f(x + a) + f(x + b) = f(\lfloor \frac{1}{2}x + \frac{1}{2}y \rfloor) + f(\lceil \frac{1}{2}x + \frac{1}{2}y \rceil) \leqslant f(x) + f(y) = f(x) + f(x + a + b)$. If instead $b - a \in \{-1, 0\}^n$, we interchange a and b. This shows the implication. $\square$

9. Two variables: rhomboidal convexity

Let us see what Definition 6.2 means for functions of two variables.

Definition 9.1. We shall say that a function $f\colon \mathbf{Z}^2 \to \mathbf{R}$ is *rhomboidally convex* if it is P-laterally convex, where we define $P \subset \mathbf{Z}^2 \times \mathbf{Z}^2$ as

$$P = \{((1,0),(1,t)); t \in [-1,1]_{\mathbf{Z}}\} \cup \{((0,1),(s,1)); s \in [-1,1]_{\mathbf{Z}}\}. \quad \square \tag{9.1}$$

Given a function f, we consider the set $\Psi(\{f\})$ of all pairs $(a,b) \in \mathbf{Z}^2 \times \mathbf{Z}^2$ such that $D_b D_a f \geqslant 0$. Then we have to take into account several conditions, e.g., the two *one-variable conditions*

$$(e^{(1)}, e^{(1)}),\ (e^{(2)}, e^{(2)}) \in \Psi(\{f\}) \tag{9.2}$$

(which we usually require in order to avoid uninteresting cases – see (7.4)); the two *diagonal conditions*

$$((-1,1),(-1,1)),\ ((1,1),(1,1)) \in \Psi(\{f\}); \tag{9.3}$$

the *left* and *right horizontal lozenge conditions*[3]

$$((-1,0),(-1,1)),\ ((1,0),(1,1)) \in \Psi(\{f\}); \tag{9.4}$$

and finally the *left* and *right vertical lozenge conditions*,

$$((0,1),(-1,1)),\ ((0,1),(1,1)) \in \Psi(\{f\}). \tag{9.5}$$

We note that, by partial addition, $((1,0),(1,1)) +_1 ((0,1),(1,1)) = ((1,1),(1,1))$, which implies that the right horizontal lozenge condition and the right vertical lozenge condition yield the diagonal condition for $((1,1),(1,1))$. Thus we often do not need to consider the diagonal conditions.

To see which conditions are necessary for the marginal function to be convex extensible, it is instructive to look at the following examples.

Example 9.2. Let f be the function in Example 1.6 with $m = 1$. It does not satisfy $D_{(1,1)} D_{(1,0)} f(0,0) \geqslant 0$, which explains that $\frac{1}{2}h(0) + \frac{1}{2}h(2) = 0$ does not majorize $h(1) = 1$. It does satisfy all other conditions (9.2)–(9.5), i.e., it satisfies seven of the eight conditions, the only exception being the right horizontal lozenge condition $D_{(1,1)} D_{(1,0)} f \geqslant 0$. $\square$

Example 9.3. Let now f be the function defined as

$$f(x,y) = |3x - 2y|, \qquad (x,y) \in \mathbf{Z}^2.$$

It does not satisfy $D_{(1,1)} D_{(0,1)} f(0,0) \geqslant 0$, the right vertical lozenge condition. It does satisfy all other conditions (9.2)–(9.5), i.e., it satisfies seven of the eight conditions, the only exception being the right vertical lozenge condition $D_{(1,1)} D_{(0,1)} f \geqslant 0$. Its marginal function takes the value 0 at even integers and 1 at odd integers, and is thus not convex extensible. $\square$

[3] We are aware that *lozenge* and *rhombus* are considered to be synonyms, but we are brave enough to call a set like $\mathbf{cvxh}\{(0,0),(1,0),(1,1),(2,1)\}$ a lozenge, although its sides have Euclidean lengths 1 and $\sqrt{2}$. However, their l^∞ lengths are all equal, so it is actually a rhombus as well as a lozenge for the l^∞ metric.

By forming similar examples we can conclude that for the marginal function to be convex extensible, each of the four lozenge conditions (9.4) and (9.5) is necessary, even in the presence of the other three lozenge conditions, the two one-variable conditions, and the two diagonal conditions. So we conclude that all four lozenge conditions are needed, but that we can then omit the two diagonal conditions: We need six conditions for the marginal function to be convex extensible.

Example 9.4. When $n = 2$ and $A = \{(e^{(1)}, e^{(2)})\}$, a function is A-laterally convex if and only if it is submodular. Note that (7.4) is not satisfied in this case. (Cf. Murota (2003:26, 206–207).) $\square$

Example 9.5. Since, for $n = 2$, $\Lambda \supset P$ (see Example 8.2), we have $\Phi(\Lambda) \subset \Phi(P)$, i.e., every $\mathrm{L}^\natural$-convex function is rhomboidally convex. In fact, the $\mathrm{L}^\natural$-convex functions form a tiny fraction of the rhomboidally convex functions. To illustrate this fact, let us mention that a function $f(x_1, x_2) = g(x_1 + x_2)$, $(x_1, x_2) \in \mathbf{Z}^2$, is $\mathrm{L}^\natural$-convex if and only if $g\colon \mathbf{Z} \to \mathbf{R}$ is the restriction to $\mathbf{Z}$ of an affine function defined on $\mathbf{R}$, while it is rhomboidally convex if and only if g is convex extensible. (If $f(x_1, x_2) = h(x_1 - x_2)$, the situation is quite different.) $\square$

Proposition 9.6. *Consider the following conditions on a function* $f\colon \mathbf{Z}^2 \to \mathbf{R}$.

(A) f *is rhomboidally convex;*
(B) f *is integrally convex;*
(C) f *is convex extensible;*
(D) *The restriction of* f *to any digital line* $\{c + ta;\ t \in \mathbf{Z}\}$, $c, a \in \mathbf{Z}^2$, *is convex extensible.*

Then (A) $\Leftrightarrow$ (B) $\Rightarrow$ (C) $\Rightarrow$ (D), *and, in general,* (B) $\not\Leftarrow$ (C) $\not\Leftarrow$ (D).

Proof. (A) $\Leftrightarrow$ (B). See Kiselman (2008: Theorem 2.4).

(B) $\Rightarrow$ (C). See the comment after Definition 4.4.

(C) $\Rightarrow$ (D). If F is a convex extension of f, then $D_a D_a F \geqslant 0$ for all $a \in \mathbf{R}^2$. In particular $D_a D_a f \geqslant 0$ for all $a \in \mathbf{Z}^2$.

(B) $\not\Leftarrow$ (C). Example 1.6 with $m = 1$ shows this. Here $\mathbf{can}(f)(x, \frac{1}{2})$ takes the values 1, $\frac{1}{2}$, 1, $\frac{1}{2}$, 1 for $x = 0$, $\frac{1}{2}$, 1, $\frac{3}{2}$, 2, respectively, so $\mathbf{can}(f)$ is not convex.

(C) $\not\Leftarrow$ (D). Define

$$G(x, y) = (2y - x - 1)^+ \vee (2x - y - 1)^+ \vee (-x - y - 1)^+, \qquad (x, y) \in \mathbf{R}^2.$$

Here $s \vee t = \max(s, t)$ denotes the maximum of two numbers s and t, and $t^+ = t \vee 0$.

The function G is certainly convex, so its restriction $g = G|_{\mathbf{Z}^2}$ is convex extensible. Now define $f(x, y) = g(x, y)$ for $(x, y) \neq (0, 0)$ and $f(0, 0) = g(0, 0) + \frac{1}{2} = \frac{1}{2}$. For f to satisfy $D_a D_a f \geqslant 0$ it is enough to consider $D_a D_a f$ on a digital line

$$L = \{ta;\ t \in \mathbf{Z}\} = \{t(p, q);\ t \in \mathbf{Z}\}$$

which passes through the origin, since we have changed the value of g only at the origin. It is sufficient to prove that $\frac{1}{2} f(p, q) + \frac{1}{2} f(-p, -q) \geqslant f(0, 0) = \frac{1}{2}$ for two relatively prime integers p, q, since the points (p, q) and $(-p, -q)$ are the integer points closest to the origin on L. We see that

$$f(p, q) \vee f(-p, -q) \leqslant f(p, q) + f(-p, -q) < 1$$

only if

$$|2p - q| < 2, \ |2q - p| < 2, \ |p + q| < 2.$$

This happens only if $(p, q) = (0, 0)$. So the restriction of f to L is convex extensible, but f is not convex extensible. Indeed, the origin is the barycenter of the three points $(1, 1)$, $(-1, 0)$, $(0, -1)$:

$$(0, 0) = \tfrac{1}{3}(1, 1) + \tfrac{1}{3}(-1, 0) + \tfrac{1}{3}(0, -1),$$

but

$$f(0, 0) = \tfrac{1}{2} > \tfrac{1}{3}f(1, 1) + \tfrac{1}{3}f(-1, 0) + \tfrac{1}{3}f(0, -1) = 0,$$

so Jensen's inequality is not satisfied. □

10. The set where the infimum is attained

We shall first study the relation between A-lateral convexity and the interval (possibly empty) where the infimum defining the marginal function is attained.

Theorem 10.1. *Let us define, for any function* $f\colon \mathbf{Z}^n \to \mathbf{R}$,

$$\begin{aligned} M_f(x_1, \dots, x_{n-1}) &= M_f(x') \\ &= \Big\{ b \in \mathbf{Z};\ f(x_1, \dots, x_{n-1}, b) = \inf_{t \in \mathbf{Z}} f(x_1, \dots, x_{n-1}, t) \Big\}, \end{aligned}$$

where $x' = (x_1, \dots, x_{n-1}) \in \mathbf{Z}^{n-1}$. *We also define*

$$f_\beta(x) = f(x) - \beta x_n, \qquad x = (x_1, \dots, x_n) \in \mathbf{Z}^n, \quad \beta \in \mathbf{R}.$$

Now fix an element $a = (a', a_n)$ *of* $\mathbf{Z}^n$, *where* $a' = (a_1, \dots, a_{n-1})$ *and* $a_n \geqslant 0$, *and define*

$$A = \{(e^{(n)}, e^{(n)}), ((a', a_n), e^{(n)}), ((-a', a_n), e^{(n)})\},$$

a subset of $(\mathbf{R}^n)^2$ *with three elements. Then* f *is* A*-laterally convex if and only if* $t \mapsto f(x', t)$ *is convex extensible for every* x' *and a certain Lipschitz property holds:*

$$M_{f_\beta}(x' + a') \subset M_{f_\beta}(x') + [-a_n, a_n]_{\mathbf{Z}}, \qquad x' \in \mathbf{Z}^{n-1}, \quad \beta \in \mathbf{R}. \tag{10.1}$$

Proof. Assume first that f is A-laterally convex. Since A contains $(e^{(n)}, e^{(n)})$, $\mathbf{Z} \ni t \mapsto f(x', t)$ is convex extensible for every x'.

We note that for a function which is convex extensible in the last variable,

$$b \in M_f(x') \text{ if and only if } D_{e^{(n)}} f(x', b - 1) \leqslant 0 \leqslant D_{e^{(n)}} f(x', b). \tag{10.2}$$

Moreover

$$b, b + 1 \in M_f(x') \text{ if and only if } D_{e^{(n)}} f(x', b) = 0. \tag{10.3}$$

Let now f satisfy $D_a D_{e^{(n)}} f \geqslant 0$ and consider two points x' and $x' + a'$ in $\mathbf{Z}^{n-1}$. Then for any $b \in M_f(x')$ we have, since also $((-a', a_n), e^{(n)})$ is in A,

$$\begin{aligned} D_{e^{(n)}} f(x' + a', b - a_n - 1) &\leqslant D_{e^{(n)}} f(x', b - 1) \leqslant 0 \\ &\leqslant D_{e^{(n)}} f(x', b) \leqslant D_{e^{(n)}} f(x' + a', b + a_n), \end{aligned} \tag{10.4}$$

which implies that there is a point $c \in [b - a_n, b + a_n]_{\mathbf{Z}}$ with

$$D_{e^{(n)}} f(x' + a', c - 1) \leqslant 0 \leqslant D_{e^{(n)}} f(x' + a', c).$$

In view of (10.2), this means that $c \in M_f(x' + a')$. We have proved that $b \in c + [-a_n, a_n]_{\mathbf{Z}} \subset M_f(x' + a') + [-a_n, a_n]_{\mathbf{Z}}$, and, since b was any point in $M_f(x')$, that $M_f(x') \subset M_f(x' + a') + [-a_n, a_n]_{\mathbf{Z}}$. We are done, since the whole argument holds also for f_β.

Conversely, suppose that the function f satisfies $D_{e^{(n)}} D_{e^{(n)}} f \geqslant 0$ but is not A-laterally convex. Then it does not satisfy one of the two inequalities $D_a D_{e^{(n)}} f \geqslant 0$ and $D_{(-a', a_n)} D_{e^{(n)}} f \geqslant 0$. It suffices to consider one of these cases. We thus assume that there exist $(x', b) \in \mathbf{Z}^{n-1} \times \mathbf{Z}$ such that $D_{e^{(n)}} f(x' + a', b + a_n) < D_{e^{(n)}} f(x', b)$. We shall reach a contradiction to the Lipschitz property (10.1).

We take a real number β such that

$$D_{e^{(n)}} f(x' + a', b + a_n) < \beta < D_{e^{(n)}} f(x', b).$$

If we rewrite this for the function f_β, for which $D_{e^{(n)}} f_\beta = D_{e^{(n)}} f - \beta$, we obtain

$$D_{e^{(n)}} f_\beta(x' + a', b + a_n) < 0 < D_{e^{(n)}} f_\beta(x', b), \tag{10.5}$$

which implies that

$$M_{f_\beta}(x' + a') \subset [b + a_n + 1, +\infty[_{\mathbf{Z}} \text{ and that } M_{f_\beta}(x') \subset]-\infty, b]_{\mathbf{Z}}.$$

Hence

$$M_{f_\beta}(x' + a') + [-a_n, a_n]_{\mathbf{Z}} \subset [b + 1, +\infty[_{\mathbf{Z}}$$

and

$$M_{f_\beta}(x') + [-a_n, a_n]_{\mathbf{Z}} \subset]-\infty, b + a_n]_{\mathbf{Z}}.$$

Thus $M_{f_\beta}(x' + a')$ is not contained in $M_{f_\beta}(x') + [-a_n, a_n]_{\mathbf{Z}}$ unless it is empty, and $M_{f_\beta}(x')$ is not contained in $M_{f_\beta}(x' + a') + [-a_n, a_n]_{\mathbf{Z}}$ unless it is empty. As soon as one of them is nonempty, we get a contradiction to the Lipschitz property (10.1).

So the case when both sets are empty remains to be considered – so far, there is no contradiction in this case. Since $M_{f_\beta}(x' + a')$ is now empty by hypothesis, the function $t \mapsto D_{e^{(n)}} f_\beta(x' + a', t)$ can never change sign, and since $D_{e^{(n)}} f_\beta(x' + a', b + a_n)$ is negative, we must have $D_{e^{(n)}} f_\beta(x' + a', t) < 0$ for all $t \in \mathbf{Z}$. Now define $\gamma = D_{e^{(n)}} f(x', b) > \beta$. Then, by (10.3), $M_{f_\gamma}(x')$ is certainly nonempty; it contains b and $b + 1$. And since $\gamma > \beta$ we have

$$D_{e^{(n)}} f_\gamma(x' + a', t) = D_{e^{(n)}} f_\beta(x' + a', t) + \beta - \gamma < D_{e^{(n)}} f_\beta(x' + a', t) < 0$$

for all $t \in \mathbf{Z}$, so that (10.2) shows that $M_{f_\gamma}(x' + a')$ is empty. This contradicts the inclusion $M_{f_\gamma}(x') \subset M_{f_\gamma}(x' + a') + [-a_n, a_n]_{\mathbf{Z}}$. We are done. $\square$

By permuting the variables we easily obtain the following corollary.

Corollary 10.2. *Given a function* $f\colon \mathbf{Z}^n \to \mathbf{R}$*, we define, for* $1 \leqslant j \leqslant n$ *and* $x' = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \in \mathbf{Z}^{n-1}$*,*

$$\begin{aligned} M_{j,f}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_{n-1}) &= M_{j,f}(x') \\ &= \big\{b \in \mathbf{Z};\, f(x_1, \dots, x_{j-1}, b, x_{j+1}, \dots, x_{n-1}, x_n) = \inf_{x_j \in \mathbf{Z}} f(x)\big\}. \end{aligned}$$

We also define

$$f_{j,\beta}(x) = f(x) - \beta x_j,\ \ x = (x_1, \dots, x_n) \in \mathbf{Z}^n,\ \ j = 1, \dots, n,\ \ \beta \in \mathbf{R}.$$

Fix a set A *which contains* $(a, e^{(j)})$ *and* $(\bar{a}, e^{(j)})$*, where*

$$\bar{a} = 2a_j e^{(j)} - a = (-a_1, \dots, a_j, \dots, -a_n),$$

and satisfies (7.4) *and* (8.1)*. If* f *is* A*-laterally convex, then* f *is convex extensible in each variable separately and we have*

$$M_{j,f_{j,\beta}}(x' + a') \subset M_{j,f_{j,\beta}}(x') + [-a_j, a_j]_{\mathbf{Z}}, \qquad x' \in \mathbf{Z}^{n-1},$$

where now $a' = (a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n)$ *and similarly for* x'*.* □

11. Lateral convexity of marginal functions

11.1. Arbitrary dimensions

In Kiselman (2008: Theorem 3.1), it was shown that for integrally convex functions of two integer variables, the marginal function is convex extensible. We shall now study the marginal function of A-laterally convex functions in arbitrary dimension and for more general choices of A.

Theorem 11.1. *Let* $A \subset \mathbf{Z}^{n-1} \times \mathbf{Z}^{n-1}$ *and* $B \subset \mathbf{Z}^n \times \mathbf{Z}^n$ *be given. We assume that* (7.4) *and* (8.1) *hold for both* A *and* B*. Assume also that*

$$\text{If } (a, b) \in A \text{ and } s \in [-1, 1]_{\mathbf{Z}}, \text{ then } ((a, s), (b, 0)) \text{ belongs to } \widetilde{B}; \tag{11.1}$$

that

$$\text{If there exists } c \in \mathbf{Z}^{n-1} \text{ such that } (a, c) \in A,\ \text{ then } ((a, 1), e^{(n)}) \in \widetilde{B}; \tag{11.2}$$

and finally that

$$\text{If } ((a, 1), e^{(n)}) \in B,\ \text{ then } ((-a, 1), e^{(n)}) \in \widetilde{B}. \tag{11.3}$$

If $f\colon \mathbf{Z}^n \to \mathbf{R}$ *is* B*-laterally convex, then its marginal function*

$$h(x) = \inf_{t \in \mathbf{Z}} f(x, t), \qquad x \in \mathbf{Z}^{n-1},$$

is A*-laterally convex, provided that it does not take the value* $-\infty$*.*

Lemma 11.2. *Let A and B satisfy the hypotheses in Theorem* 11.1. *Then*

$$\text{If } (a,b) \in A, \text{ then } ((a,-1),(b,-1)),((a,1),(b,1)) \in \widetilde{B}. \tag{11.4}$$

Proof. From the conditions (11.1) and (11.2) we know that

$$\text{both } ((a,1),(b,0)) \text{ and } ((a,1),e^{(n)}) \text{ belong to } \widetilde{B}.$$

By partial addition $+_2$ we conclude that so does $((a,1),(b,1))$.

From the condition (11.2) we know that $((a,1),e^{(n)})$ and, consequently, in view of (11.3), also $((-a,1),e^{(n)})$ belongs to $\widetilde{B}$. So does the opposite pair $-((-a,1),e^{(n)}) = ((a,-1),-e^{(n)})$.

By condition (11.1) we find that $((a,-1),(b,0))$ is in $\widetilde{B}$, and we now only have to form the partial sum

$$((a,-1),-e^{(n)}) +_2 ((a,-1),(b,0)) = ((a,-1),(b,-1))$$

to conclude. □

By this lemma and (11.1) we know that if A and B satisfy the hypotheses of the theorem and if $(a,b) \in A$, then there are pairs of the form $((a,s),(b,t))$ in $\widetilde{B}$ with $-1 \leqslant s,t \leqslant 1$ and the sum $s+t$ taking any of the five values $-2,-1,0,1,2$.

Proof of Theorem 11.1. It is enough to prove the theorem for functions such that the infimum defining h is attained at a unique point. Indeed, if $t \mapsto f(x,t)$ is convex extensible, then for any positive $\varepsilon > 0$, the infimum defining the marginal function h_ε of $f_\varepsilon(x,t) = f(x,t) + \varepsilon t^2$ is attained at a unique integer $t = \varphi_\varepsilon(x)$, and h_ε tends to h as $\varepsilon \to 0$, preserving the A-lateral convexity of h_ε. We observe that f_ε is B-laterally convex with f provided that $(e^{(n)},e^{(n)}) \in B$, which we assume. We may therefore suppose that $h(x) = f(x,\varphi(x))$ for some function $\varphi\colon \mathbf{Z}^{n-1} \to \mathbf{Z}$. Moreover, we know that φ is Lipschitz in the sense that

$$|\varphi(x+a) - \varphi(x)| \leqslant 1, \qquad x \in \mathbf{Z}^{n-1}, \tag{11.5}$$

for certain values of $a \in \mathbf{Z}^{n-1}$, viz. when $((a,1),e^{(n)})$ and $((-a,1),e^{(n)})$ both belong to $\widetilde{B}$. For this to happen, it is enough that there exists a c such that $(a,c) \in A$.

Similarly, we know that

$$|\varphi(x+b) - \varphi(x)| \leqslant 1 \qquad x \in \mathbf{Z}^{n-1}, \tag{11.6}$$

for certain values of $b \in \mathbf{Z}^{n-1}$, viz. when $((b,1),e^{(n)})$ and $((-b,1),e^{(n)})$ both belong to $\widetilde{B}$. For this it is enough that there exists a d such that $(b,d) \in A$.

In particular, if (a,b) is in A, we can take $c = b$ and $d = a$ above to conclude that the two Lipschitz conditions (11.5) and (11.6) hold.

We have

$$\begin{aligned} D_bD_ah(x) &= f(x+a+b,\varphi(x+a+b)) \\ &\quad - f(x+a,\varphi(x+a)) - f(x+b,\varphi(x+b)) + f(x,\varphi(x)). \end{aligned} \tag{11.7}$$

The formula holds of course for all $x, a, b \in \mathbf{Z}^{n-1}$, but we shall need it only when $(a, b) \in A$. We shall compare (11.7) with

$$\begin{aligned} D_{(b,t)}D_{(a,s)}f(x, \varphi(x)) &= f(x + a + b, \varphi(x) + s + t) \\ &\quad - f(x + a, \varphi(x) + s) - f(x + b, \varphi(x) + t) + f(x, \varphi(x)) \end{aligned} \tag{11.8}$$

for suitable s and t. This expression is nonnegative if $((a, s), (b, t)) \in \widetilde{B}$.

By the definition of φ we have

$$-f(x + a, \varphi(x + a)) \geqslant -f(x + a, s) \text{ and } - f(x + b, \varphi(x + b)) \geqslant -f(x + b, t)$$

for any s and t, so we get $D_bD_ah(x) \geqslant D_{(b,t)}D_{(a,s)}f(x, \varphi(x))$ as soon as $s + t = \varphi(x + a + b) - \varphi(x)$.

In view of (11.5) and (11.6), which, as we have remarked, are applicable,

$$|\varphi(x + a + b) - \varphi(x)| \leqslant |\varphi(x + a + b) - \varphi(x + a)| + |\varphi(x + a) - \varphi(x)| \leqslant 2,$$

and we know from Lemma 11.2 that there are numbers s, t such that

$$s + t = \varphi(x + a + b) - \varphi(x) \text{ and } ((a, s), (b, t)) \in \widetilde{B}.$$

We are done. □

By iteration we easily obtain the following result.

Corollary 11.3. *Let us define* $B^{(0)} = \{(0, 0)\}$, $B^{(1)} = \{(1, 1)\}$, *and generally* $B^{(n)} \subset \mathbf{Z}^n \times \mathbf{Z}^n$ *such that* $B^{(n-1)}$ *and* $B^{(n)}$ *satisfy the conditions for* A *and* B *in Theorem* 11.1 *for* $n \geqslant 2$. *If* $f : \mathbf{Z}^n \to \mathbf{R}$ *is a given* $B^{(n)}$*-laterally convex function, then the successive marginal functions* $h_n = f$,

$$h_k(x) = \inf_{t \in \mathbf{Z}} h_{k+1}(x, t), \qquad x = (x_1, \ldots, x_k) \in \mathbf{Z}^k, \quad k = n - 1, \ldots, 1,$$

are $B^{(k)}$*-laterally convex, provided that* $h_1 > -\infty$. *In particular, the marginal function* h_1 *of one variable is* $\{(1, 1)\}$*-laterally convex, equivalently convex extensible.* □

In condition (11.1) it is often preferable to replace the pair

$$((a, -1), (b, 0)) \text{ by its opposite } ((-a, 1), (-b, 0)),$$

which determines the same condition. This is to be able to continue as in Corollary 11.3, where the last component should be nonnegative – this is needed in Theorem 10.1. We denote the set B so constructed by $\Theta^n(A)$. We can now define $B^{(n)} = \Theta^n(B^{(n-1)})$ and get Corollary 11.3 to work.

Thus taking $\mathscr{M}_n$ as the set of all $B^{(n)}$-laterally convex functions such that the marginal functions h_1 do not take the value $-\infty$ gives a satisfactory solution to Problem 1.5.

11.2. The case of two variables

Let us look in more detail at the construction of $\Theta^2(A)$. Then the corollary is about three functions: $h_2 = f$ defined on $\mathbf{Z}^2$, $h_1(x) = \inf_{y\in\mathbf{Z}} f(x,y)$ defined on $\mathbf{Z}^1$, and the constant $h_0 = \inf_{(x,y)\in\mathbf{Z}^2} f(x,y)$ defined as a function on $\mathbf{Z}^0 = \{0\}$. But here we do not say anything about the marginal function $k_1(y) = \inf_{x\in\mathbf{Z}} f(x,y)$. To do so, we should permute the variables. However, it turns out, perhaps surprisingly, that this is not necessary, for the conditions are symmetric in the two variables.

If we start with $A = \{(1,1)\} \subset \mathbf{Z}^1 \times \mathbf{Z}^1$ in one variable, the construction in Theorem 11.1 yields, in order,

$(e^{(1)}, e^{(1)})$, $(e^{(2)}, e^{(2)})$, applying (7.4);

$((1,-1),(1,0))$, $((1,1),(1,0))$, applying (11.1);

$((1,1),e^{(2)})$, applying (11.2); and

$((-1,1),e^{(2)})$, applying (11.3).

However, as already remarked, we should replace

$$((1,-1),(1,0)) \quad \text{by} \quad ((-1,1),(-1,0)).$$

We thus obtain

$$\begin{aligned} B = \{&(e^{(1)},e^{(1)}),\ (e^{(2)},e^{(2)}),\ ((-1,1),(-1,0)),\\ &((1,1),(1,0)),\ ((-1,1),(0,1)),\ ((1,1),(0,1))\}, \end{aligned}$$

This means that the two one-dimensional conditions and the four lozenge conditions are all satisfied, while the two diagonal conditions need not be listed since they follow from the others. We see now that if we permute the variables, the conditions remain the same.

We see that the set $B = \Theta^2(A) \subset \mathbf{Z}^2 \times \mathbf{Z}^2$, which defines rhomboidal convexity and corresponds to the six conditions (9.2), (9.4) and (9.5), consists of 6 pairs, and that $\Theta^3(B)$ consists of $6^2 = 36$ pairs.

11.3. Symmetric and asymmetric conditions

The condition on a function to have a convex-extensible marginal function is asymmetric. Indeed, the function $f(x,y) = (2x-y)^2$, $(x,y) \in \mathbf{Z}^2$, has a convex-extensible marginal function, whereas the marginal function of its reflection $g(x,y) = (2y-x)^2$ does not. Therefore a symmetric condition can never be necessary and sufficient.

For functions f such that $D_{(0,1)}D_{(0,1)}f \geqslant 0$, a necessary and sufficient condition for all functions $f_\beta(x,y) = f(x,y) - \beta y$, $\beta \in \mathbf{R}$, to have a convex-extensible marginal function is that all conditions

$$D_{(1,p)}D_{(1,p)}f \geqslant 0, \quad D_{(1,p)}D_{(1,p+1)}f \geqslant 0, \qquad p \in \mathbf{Z},$$

shall be satisfied. These are infinitely many conditions as opposed to the six conditions obtained in our construction: $\Theta^2(A)$ has six elements.

We conclude that there is a choice between a sufficient condition which is finite and symmetric but not necessary, and a sufficient and necessary condition which is infinite – and by necessity asymmetric.

12. Necessity of lateral convexity

As can be guessed from Examples 9.2 and 9.3, the convexity property we have defined is essentially best possible. Before showing this, two remarks are in order.

Let $\varphi\colon \mathbf{Z}^2 \to \mathbf{Z}$ be any function such that $\mathbf{Z} \ni y \mapsto \varphi(x,y) \in \mathbf{Z}$ is a surjection for every $x \in \mathbf{Z}$. Then the function $g(x,y) = f(x,\varphi(x,y))$, $(x,y) \in \mathbf{Z}^2$, has the same marginal function as f. In particular, the values on a vertical line can be arbitrarily scrambled. It follows that no reasonable conclusion concerning regularity of f can be drawn from knowledge of its marginal function. But if we consider the marginal functions h_β of the tilted functions $f_\beta(x,y) = f(x,y) - \beta y$, $\beta \in \mathbf{R}$, things are different.

For simplicity we now restrict attention to functions of two variables $(x,y) \in \mathbf{Z}^2$. We define the *partial Fenchel transform* of a function $f\colon \mathbf{Z} \times \mathbf{Z} \to \mathbf{R}_!$ by

$$f^*(x,\eta) = \sup_{y\in\mathbf{Z}}(\eta y - f(x,y)), \qquad (x,\eta) \in \mathbf{Z} \times \mathbf{R},$$

to be compared with the complete Fenchel transform,

$$\tilde{f}(\xi,\eta) = \sup_{(x,y)\in\mathbf{Z}^2}(\xi x + \eta y - f(x,y)), \qquad (\xi,\eta) \in \mathbf{R} \times \mathbf{R}.$$

Thus the marginal function of f is $h(x) = -f^*(x,0)$. Since the third transform f^{***} is equal to the first, the second transform f^{**} has the same marginal function as f. Therefore, again, it is not reasonable to expect that, from knowledge of a marginal function, one can conclude anything about f, only about its minorant f^{**}.

Proposition 12.1. *Let $f\colon \mathbf{Z} \times \mathbf{Z} \to \mathbf{R}$ be such that the marginal function h_β of $f_\beta(x,y) = f(x,y) - \beta y$ is convex extensible for all real numbers β. Then f^{**} satisfies the one-variable conditions* (9.2)*, the diagonal conditions* (9.3)*, and the horizontal lozenge conditions* (9.4)*.*

Proof. For brevity, let us write g instead of f^{**}.

By replacing $g(x,y)$ by $g(x,y)+\varepsilon y^2$, $\varepsilon > 0$, we may assume that the infimum of $y \mapsto g_\beta(x,y) = g(x,y) - \beta y$ is always attained at some point. Afterwards we let ε tend to zero; the properties are stable under this operation.

The vertical condition $g(x,y-1) + g(x,y+1) \geqslant 2g(x,y)$ is always satisfied by assumption.

Consider next the horizontal condition $g(x-1,y) + g(x+1,y) \geqslant 2g(x,y)$ for fixed x and y and define $\beta = g(x,y+1) - g(x,y)$. Then $h_\beta(x) = g_\beta(x,y) = g_\beta(x,y+1)$, and if h_β is convex extensible, we get

$$\begin{aligned} g_\beta(x-1,y) + g_\beta(x+1,z) &\geqslant h_\beta(x-1) + h_\beta(x+1) \\ &\geqslant 2h_\beta(x) = 2g_\beta(x,y) = 2g_\beta(x,y+1). \end{aligned}$$

Taking $z = y$ we see that the horizontal one-variable condition is satisfied; taking $z = y + 1$ we see that the right horizontal lozenge condition is satisfied; and taking $z = y + 2$ we see that one of the diagonal conditions is satisfied. For the left horizontal lozenge condition and the other diagonal condition we can argue similarly. □

Theorem 12.2. *Let $f\colon \mathbf{Z} \times \mathbf{Z} \to \mathbf{R}$ satisfy the one-variable conditions* (9.2)*. Define two marginal functions by*

$$h_\beta(x) = \inf_{y \in \mathbf{Z}} (f(x, y) - \beta y), \qquad x \in \mathbf{Z},\ \beta \in \mathbf{R},$$

and

$$k_\alpha(y) = \inf_{x \in \mathbf{Z}} (f(x, y) - \alpha x), \qquad y \in \mathbf{Z},\ \alpha \in \mathbf{R}.$$

Assume that h_β and k_α are convex extensible for all real numbers α and β. Then f is rhomboidally convex.

Proof. We apply Proposition 12.1 to f and to $(x, y) \mapsto f(y, x)$. □

13. Conclusion

We have studied a kind of convexity called *lateral convexity*, which is defined using second-order difference operators (a special kind of convolution operators). We have proved that this notion of convexity is perfectly adapted for proving that the marginal function of a real-valued function defined on the set of points with integer coordinates remains in the same class.

Notions of mathematical morphology proved to be helpful. We believe that the duality between classes of functions with a convexity property and classes of convolution operators studied here will have several applications in the future.

14. Hints for future work

14.1. Discrete convexity of infimal convolutions

Since, as remarked in Subsection 1.3, the operation of taking the marginal function is a special case of infimal convolution, it may be of interest to extend this study of discrete convexity to more general infimal convolutions.

14.2. Discrete convexity of p-marginal functions

Given a positive number p, we may define the *p-marginal function h_p* of a function $f\colon \mathbf{Z}^n \times \mathbf{Z}^m \to \mathbf{R}$ by

$$e^{-ph_p(x)} = \sum_{y \in \mathbf{Z}^m} e^{-pf(x,y)}, \qquad x \in \mathbf{Z}^n.$$

As p tends to $+\infty$ we get the usual marginal function. The question of finding suitable classes that are preserved under passage to the p-marginal function is not resolved. For such a class we would have a discrete analogue of Prékopa's theorem.

For more details on Prékopa's theorem for real variables and the problem for discrete variables, see Kiselman (2012, 2014).

14.3. Functions with integer values

It may be of interest also to consider functions $f\colon \mathbf{Z}^n \to \mathbf{Z}$ with integer values and their marginal functions. Then convex extensibility of the marginal function is too strong a condition. Instead it is relevant to require that the functions are $(\mathbf{Z}^n \times \mathbf{Z})$-convex, meaning that there exists a convex subset C of $\mathbf{R}^n \times \mathbf{R}$ such that

$$C \cap (\mathbf{Z}^n \times \mathbf{Z}) = \mathbf{epi}^{\mathrm{F}}(f) = \{(x,t) \in \mathbf{Z}^n \times \mathbf{Z};\ t \geqslant f(x)\}.$$

14.4. Duality defined by convolution inequalities

The duality studied in Sections 6 and 7 should be extended to a duality between sets M of functions μ and classes $\Phi(M)$ of functions f satisfying convolution inequalities $\mu * f \geqslant 0$ for all $\mu \in M$. Adama Koné has pursued this idea in his doctoral thesis (2016: Chapter 4).

References

Favati, Paola; Tardella, Fabio. 1990. Convexity in nonlinear integer programming. *Ricerca Operativa* **53**, 3–44.

Fujishige, Satoru; Murota, Kazuo. 2000. Notes on L-/M-convex functions and the separation theorems. *Math. Programming* **88**, 129–146.

Kiselman, Christer O. 2008. Minima locaux, fonctions marginales et hyperplans séparants dans l'optimisation discrète. *C. R. Acad. Sci. Paris, Ser. I* **346**, 49–52.

Kiselman, Christer O. 2010a. Local minima, marginal functions, and separating hyperplanes in discrete optimization. **In:** Bhatia, Rajendra (Ed.), *Abstracts: Short communications; Posters.* International Congress of Mathematicians, Hyderabad, August 19–27, 2010, pp. 572–573.

Kiselman, Christer O. 2010b. Inverses and quotients of mappings between ordered sets. *Image and Vision Computing* **28**, 1429–1442.

Kiselman, Christer O. 2011. Characterizing digital straightness and digital convexity by means of difference operators. *Mathematika* **57**, 355–380.

Kiselman, Christer O. 2012, 2014. Questions inspired by Mikael Passare's mathematics. Published online in *Afr. Mat.* on 2012 October 27. Print version published in *Afr. Mat.* **25**, issue 2 (2014), 271–288.

Kiselman, Christer O. 2015. Estimates for solutions to discrete convolution equations. *Mathematika* **61**, issue 02, 295–308.

Kiselman, Christer O.; Samieinia, Shiva. 2010. Convexity of marginal functions in the discrete case. **In:** Samieinia (2010).

Koné, Adama Arouna. 2016. *Géométrie digitale utilisée pour la discrétisation et le recouvrement optimal des objets euclidiens.* Thesis presented on 2016 January 14 at the Faculté des Sciences et Techniques de l'Université des Sciences, Techniques et des Technologies de Bamako. vi + 114 pp.

Miller, Bruce L. 1971. On minimizing nonseparable functions defined on the integers with an inventory application. *SIAM J. Appl. Math.* **21**, 166–185.

Murota, Kazuo. 1996. Convexity and Steinitz's exchange property. *Advances in Mathematics* **124**, 272–311.

Murota, Kazuo. 1998. Discrete convex analysis. *Math. Programming* **83**, 313–371.

Murota, Kazuo. 2003. *Discrete convex analysis.* SIAM monographs on discrete mathematics and applications.

Murota, Kazuo; Shioura, Akiyoshi. 1999. M-convex function on generalized polymatroid. *Math. Oper. Res.* **24**, 95–105.

Murota, Kazuo; Shioura, Akiyoshi. 2001. Relationship of M/L-convex functions with discrete convex functions by Miller and Favati–Tardella. *Discrete Appl. Math.* **115**, 151–176.

Passare, Mikael. 2009. Christer Kiselman's mathematics. **In:** Passare, Mikael, Ed., *Complex Analysis and Digital Geometry: Proceedings from the Kiselmanfest,* 2006, pp. 9–26. Uppsala, Uppsala University: Acta Universitatis Upsaliensis, C. Organisation och historia, No. 86, 364 pp. ISSN 0502-7454; ISBN 978-91-554-7672-4.

Samieinia, Shiva. 2010. *Digital Geometry, Combinatorics, and Discrete Optimization*, published on 2010 December 15. PhD Thesis defended at Stockholm University on 2011 January 21.

Christer O. Kiselman
Uppsala University
Department of Information Technology
P. O. Box 337
SE-751 05 Uppsala
e-mail: kiselman@it.uu.se
christer@kiselman.eu
URL: www.cb.uu.se/~kiselman/

Shiva Samieinia
Stockholm University
Department of Mathematics
SE-106 91 Stockholm
e-mail: shiva@math.su.se
samieinia@gmail.com
URL: sites.google.com/site/shivasamieinia/home

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 311–333

Modules of Square Integrable Holomorphic Germs

László Lempert

Abstract. This paper was inspired by Guan and Zhou's recent proof of the so-called strong openness conjecture for plurisubharmonic functions. We give a proof shorter than theirs and extend the result to possibly singular Hermitian metrics on vector bundles.

Mathematics Subject Classification (2010). 14F18, 32C35, 32L20, 32U05.

1. Introduction

This paper was written in March 2014, and submitted in October that year. What we referred to as "recent" in our introduction is not so recent anymore, but we have decided to keep the text as it was back then, and restricted ourselves to updating the bibliographical references when arxiv postings have appeared in the meantime.

Consider an open set $U \subset \mathbb{C}^m$ and a point $x \in U$. Given a holomorphic function f on some neighborhood V of x, we will denote by $\mathbf{f}_x$ its germ at x. A measurable function $u\colon U \to [-\infty, \infty]$ determines an ideal $I(u) = I(u, x)$ in the ring $\mathcal{O}_x = \mathcal{O}_{(\mathbb{C}^m,x)}$ of holomorphic germs at x,

$$I(u) = \{\mathbf{f}_x \colon f \in \mathcal{O}(V),\ \int_V |f|^2 e^{-u} < \infty,\ V \subset U \text{ open}, x \in V\}.$$

The integral is with respect to Lebesgue measure in $\mathbb{C}^m$. Clearly, if $v \le u + O(1)$ at x, then $I(u) \supset I(v)$. A conjecture, going back to Demailly and Kollár [DK,D2] had that if u is plurisubharmonic, then $I(u) = I(\eta u)$ with some $\eta \in (1, \infty)$. After partial results by Favre–Jonsson and Berndtsson, Guan and Zhou recently posted a proof of the conjecture, see [B3,FJ,GZ2-4]. A related posting is [Hi].

Our main purpose with this paper is to produce a proof, as we hope more transparent than Guan–Zhou's, by modifying their approach some, while keeping

Research partially supported by NSF grant DMS-1162070.

all their essential ideas. We will also discuss generalizations. The most immediate generalization replaces multiples of u by a sequence of plurisubharmonic functions:

Theorem 1.1. *If $u_1 \le u_2 \le \cdots$ are plurisubharmonic functions on U and $u = \lim_j u_j$ is locally bounded above on U, then $I(u) = I(u_j)$ for some j.*

Since in the original conjecture one can always assume that u is bounded above, and then in fact that $u \le 0$, the conjecture indeed follows from Theorem 1.1 if one puts $u_j = (1+1/j)u$. – That the ideas of Guan and Zhou also give a proof of Theorem 1.1 occurred to me while reading [GZ2] when it was first posted on arxiv in November 2013. Apparently at one point Guan and Zhou also noticed this, because after I communicated to them the generalization, Zhou sent me a preprint containing essentially the same result, dated earlier than my email to them, in fact even earlier than the submission date of [GZ2]. Subsequently this generalization was also mentioned in [GZ3]. A variant, in dimension 2, occurs already in [FJ, Proposition 2.6].

A further generalization involves, instead of ideals, modules of square integrable vector-valued holomorphic germs. The natural setting here is germs of holomorphic sections of a holomorphic Hilbert bundle $E \to U$, and the role of the weight e^{-u} is played by a possibly singular Hermitian metric h on E. The precise meaning of this and related notions will be explained in Section 4. For the time being, let $h\colon E \to [0,\infty]$ be any Borel measurable function on E. Indicating the space of holomorphic sections of a vector bundle by Γ, we are led to consider the sets

$$E(h,x) = \{\mathbf{f}_x \colon f \in \Gamma(V,E), \int_V h(f) < \infty,\ V \subset U \text{ open}, x \in V\}, \tag{1.1}$$

where $\mathbf{f}_x$ again stands for germ at x. This $E(h,x)$ is an $\mathcal{O}_x$–module for example if $\sqrt{h}$ is subadditive on the fibers of E and homogeneous in the sense that $\sqrt{h(\lambda e)} = |\lambda|\sqrt{h(e)}$ for $\lambda \in \mathbb{C}$ and $e \in E$. Assuming h has this property, as we let x vary, the modules $E(h,x)$ form the stalks of a sheaf of modules denoted $\mathcal{E}(h)$, a subsheaf of the sheaf of holomorphic sections of E.

Theorem 1.2. *Suppose $E \to U$ is a holomorphic Hilbert bundle and $h_1 \ge h_2 \ge \cdots$ are Hermitian metrics on E whose Nakano curvatures dominate 0. Suppose that $h = \lim_j h_j$ is bounded below by a continuous Hermitian metric. If rk $E < \infty$, or at least $\bigcup_j \mathcal{E}(h_j)$ is locally finitely generated, then $\bigcup_j \mathcal{E}(h_j) = \mathcal{E}(h)$.*

In Section 7 we will see that such a result fails if we simply drop the assumption of finite generation. Yet it seems to be an interesting problem to find weaker conditions on h_j to guarantee the conclusion of the theorem. Certain direct images of positively curved line bundles provide examples of Hilbert bundles and Hermitian metrics as in Theorem 1.2, see [B2], and while in these direct images $\mathcal{E}(h_j)$ and $\bigcup_j \mathcal{E}(h_j)$ are locally infinitely generated, the strong openness theorem of Guan–Zhou, or more generally Theorem 1.1 above, does provide a connection between $\bigcup_j \mathcal{E}(h_j)$ and $\mathcal{E}(h)$.

When E is of finite rank, $\mathcal{E}(h)$ is automatically locally finitely generated (= coherent, in this case), although we will not write out a proof. So the conclusion of Theorem 1.2 can be stated as $E(h, x) = E(h_j, x)$ for some j, in parallel with Theorem 1.1. Thus Theorem 1.1 is a special case of Theorem 1.2, and it would suffice to prove the latter. However, we will start by writing out the proof of the special case. The proof of Theorem 1.2 will follow the same line, but it will be burdened by auxiliary material that is not as readily available for vector bundles as for line bundles and that we will have to develop.

I am grateful to Bo Berndtsson for helpful discussions concerning [GZ2] and [C], and to Henri Skoda and Bernard Teissier for bibliographical information.

2. The proof of Theorem 1.1

In the setting of Theorem 1.1 we put $J = \bigcup_j I(u_j)$. Suppose $P \subset \mathbb{C}^m$ is a complex (affine) hyperplane and $W \subset P$ is relatively open. For a measurable $g\colon W \to \mathbb{C}$ define $||g|| \in [0, \infty]$ by

$$\|g\|^2 = \inf_j \int_W |g|^2 e^{-u_j},$$

the integral with respect to $2m-2$-dimensional Lebesgue measure. By the dominated convergence theorem

$$\|g\|^2 = \begin{cases} \infty & \text{or} \\ \lim_j \int_W |g|^2 e^{-u_j} = \int_W |g|^2 e^{-u}. \end{cases} \tag{2.1}$$

We denote by $\operatorname{dist}(x, P)$ the Euclidean distance between x and P, and write $P \| P_0$ to indicate that hyperplanes P, P_0 are parallel. The crux of the matter is the following characterization of J when $m \geq 1$:

Lemma 2.1. *Let $f \in \mathcal{O}(U)$. Its germ $\mathbf{f}_x$ is in J if and only if for any sufficiently small neighborhood $V \subset U$ of x and any hyperplane $P_0 \subset \mathbb{C}^m$*

$$\liminf \operatorname{dist}(x, P)\|f|V \cap P\| = 0, \quad \textit{as } P\|P_0 \textit{ and } \operatorname{dist}(x, P) \to 0. \tag{2.2}$$

The lemma is of interest even when each $u_j = u$ (in fact, once Theorem 1.1 is proved, this is the only interesting case). Indeed, assume V is a polydisc centered at $x = 0$ and $P_0 = \{z_1 = 0\}$. One direction of the lemma then says that, modulo shrinking, $\int_V |f|^2 e^{-u} < \infty$ provided

$$\liminf_{s \to 0} |s|^{-2} \int_{V \cap \{z_1 = s\}} |f|^2 e^{-u} = 0.$$

Since

$$\int_V |f|^2 e^{-u} = \int_{\mathbb{C}} \left(\int_{V \cap \{z_1 = s\}} |f|^2 e^{-u} \right) d\lambda_2(s),$$

for certain type of functions $\varphi(s)$ the lemma provides the convergence of the integral $\int_{\{|s|<r\}} \varphi(s) d\lambda_2(s)$ once $\liminf_{s \to 0} \varphi(s)/|s|^2 = 0$ is known, a rather perplexing connection.

For the proof we need the following simple result, a variant of [GZ2, Lemma 2.3]. Let $\Delta \subset \mathbb{C}$ denote the unit disc.

Proposition 2.2. *Let $k \in \mathbb{N}$ and let F be a holomorphic function in a neighborhood of $\overline{\Delta}$, which does not vanish on $\overline{\Delta}\backslash\{0\}$. Suppose $G \in \mathcal{O}(\Delta)$, $G = o(F)$ at 0, and with some $t \in \Delta\backslash\{0\}$*

$$F(\omega t) = G(\omega t) \text{ for all } k\text{th roots of unity } \omega.$$

Then

$$\sup_{\Delta} |G| \geq C_1 |t|^{-k}, \quad C_1 = \min_{|s|=1} |F(s)| > 0. \tag{2.3}$$

Proof. As in [GZ2] we start by writing $F(s) = s^p F_1(s)$, where $p \in \mathbb{N} \cup \{0\}$ and F_1 does not vanish on $\overline{\Delta}$. Upon dividing F and G by F_1 we reduce the proof to the case when $F(s) = s^p$. Consider the function

$$G_1(s) = \frac{1}{k} \sum_{\omega} \omega^{-p} G(\omega s),$$

the sum over kth roots of unity. Our G_1 has all the properties listed above for G, and in addition, its Taylor series contains only monomials s^q for which $q - p > 0$ is divisible by k. In particular, $q \geq p + k$, and so $G_1(s)/s^{p+k}$ is holomorphic on Δ. Hence

$$\sup_{s\in\Delta} |G_1(s)| = \sup_{s\in\Delta} |G_1(s)/s^{p+k}| \geq |G_1(t)/t^{p+k}| = |t|^{-k},$$

and (2.3) follows. $\square$

Proof of Lemma 2.1. We will assume $x = 0$. Suppose first that $\mathbf{f}_0 \in J$, and choose j and a neighborhood V of 0 so that $\int_V |f|^2 e^{-u_j} < \infty$. Given P_0, change coordinates to arrange that P_0 is parallel to the hyperplane $\{z \in \mathbb{C}^m : z_1 = 0\}$. By Fubini's theorem

$$\infty > \int_V |f|^2 e^{-u_j} = \int_{\mathbb{C}} \Big(\int_{V \cap \{z_1 = \sigma\}} |f|^2 e^{-u_j} \Big) d\lambda_2(\sigma). \tag{2.4}$$

Since $\int |\sigma|^{-2} d\lambda_2(\sigma)$ is a divergent integral over any neighborhood of $0 \in \mathbb{C}$, (2.4) implies

$$\liminf_{\sigma\to 0} |\sigma|^2 \int_{V\cap\{z_1=\sigma\}} |f|^2 e^{-u_j} = 0,$$

and (2.2) follows.

Conversely, we will show that if $\mathbf{f}_0 \notin J$ then, given any neighborhood V of 0, with some hyperplane P_0

$$\liminf \ \operatorname{dist}(x, P) \|f|V \cap P\| > 0, \quad \text{as } P \| P_0 \text{ and } \operatorname{dist}(x, P) \to 0. \tag{2.5}$$

Fix such V, that we can assume to be pseudoconvex and relatively compact in U. If $\alpha\colon \Delta \to U$ is holomorphic, $\alpha(0) = 0$, we write $J \circ \alpha$ for the pull back

$$\{\mathbf{g} \circ \alpha \colon \mathbf{g} \in J\} \subset \mathcal{O}_{(\mathbb{C},0)}.$$

Now $\mathbf{f}_0 \notin J$ implies there is a nonzero α such that $\mathbf{f}_0 \circ \alpha \notin \mathcal{O}_{(\mathbb{C},0)} J \circ \alpha$, see [LT, Théorème 2.1], by Lejeune–Jalabert and Teissier, or [GZ2, Remark 2.12]. (What matters here is that J is integrally closed, i.e., if $g^1,\dots,g^p,\psi$ are holomorphic functions in a neighborhood of $0\in U$, the germs $\mathbf{g}_0^i$ are in J, and $\psi=O(|g^1|+\cdots+|g^p|)$ at 0, then $\boldsymbol{\psi}_0\in J$. Teissier tells me that he and Lejeune–Jalabert most probably learned the result from Hironaka.) We choose a hyperplane P_0 through $0\in\mathbb{C}^m$ that does not contain $\alpha(\Delta)$. Upon adjusting the coordinates in $\mathbb{C}$ and in $\mathbb{C}^m$ we can assume that $P_0=\{z\in\mathbb{C}^m\colon z_1=0\}$, that $\alpha=(\alpha_1,\dots,\alpha_m)$ is holomorphic in a neighborhood of $\overline{\Delta}$, that $F=f\circ\alpha\neq 0$ on $\overline{\Delta}\backslash\{0\}$,

$$\alpha_1(s)=s^k,\qquad s\in\Delta, \tag{2.6}$$

and finally $\alpha(\overline{\Delta})\subset V$. This latter implies that there are constants C_2,C_2' such that for $g\in\mathcal{O}(V)$

$$\max_{\alpha(\overline{\Delta})}|g|^2\le C_2'\int_V|g|^2\le C_2^2\int_V|g|^2e^{-u}\le C_2^2\int_V|g|^2e^{-u_j} \tag{2.7}$$

for any j. Write P_σ for the hyperplane $\{z\in\mathbb{C}^m\colon z_1=\sigma\}$. We need to estimate $\|f|V\cap P_\sigma\|$ from below. Take an arbitrary $\sigma\in\Delta\backslash\{0\}$. Assume first that $\|f|V\cap P_\sigma\|<\infty$. By the Ohsawa–Takegoshi theorem there is a $g\in\mathcal{O}(V)$ that agrees with f on $V\cap P_\sigma$ and with some j

$$\int_V|g|^2e^{-u_j}\le C_3^2\|f|V\cap P_\sigma\|^2, \tag{2.8}$$

where C_3 is independent of j and σ. Indeed, with some j

$$\int_{V\cap P_\sigma}|f|^2e^{-u_j}\le 2\|f|V\cap P_\sigma\|^2,$$

cf. (2.1), and the Ohsawa–Takegoshi theorem, applied with this u_j, produces such a g. Set $G=g\circ\alpha$, whose germ at $0\in\mathbb{C}$ is in $J\circ\alpha$. As $\mathbf{F}_0=\mathbf{f}_0\notin\mathcal{O}_{(\mathbb{C},0)}J\circ\alpha$, it must be that $G=o(F)$ at $0\in\mathbb{C}$. Further, by (2.6) $F(\sqrt[k]{\sigma})=G(\sqrt[k]{\sigma})$ for any choice of kth root $\sqrt[k]{\sigma}$. Hence Proposition 2.2 gives

$$\max_{\alpha(\overline{\Delta})}|g|=\max_{\overline{\Delta}}|G|\ge C_1/|\sigma|. \tag{2.9}$$

Putting together (2.8), (2.7), and (2.9)

$$\|f|V\cap P_\sigma\|\ge\frac{C_1}{C_2C_3|\sigma|},\qquad \sigma\in\Delta\backslash\{0\}. \tag{2.10}$$

This we derived under the assumption that the left-hand side is finite, but of course (2.10) also holds when the left-hand side is infinite. (2.10) now implies (2.5), and the proof is complete. □

Proof of Theorem 1.1. Since $I(u)\supset J$ and $I(u)$ is finitely generated, all we need to prove is that $\mathbf{f}_x\in I(u)$ implies $\mathbf{f}_x\in J$. This we prove by induction on m, as in [GZ2]. The result is obvious when $m=0$; suppose it holds for $m-1$. Upon shrinking U we can assume $\mathbf{f}_x$ is the germ of some $f\in\mathcal{O}(U)$. First we apply the "only if" direction of Lemma 2.1, but with each u_j replaced by u. This provides a neighborhood $V_0 \subset U$ of x and for any hyperplane $P_0 \subset \mathbb{C}^m$ a sequence of hyperplanes $P_\nu \| P_0$ such that $\operatorname{dist}(x, P_\nu) \to 0$ and

$$\lim_{\nu\to\infty} \operatorname{dist}(x, P_\nu)^2 \int_{V_0\cap P_\nu} |f|^2 e^{-u} = 0. \tag{2.11}$$

Let now V be an arbitrary neighborhood of x, relatively compact in V_0. Since $\int_{V_0\cap P_\nu} |f|^2 e^{-u} < \infty$ for $\nu > \nu_0$, for these ν the induction hypothesis gives a $j = j_\nu$ such that

$$\int_{V\cap P_\nu} |f|^2 e^{-u_j} < \infty.$$

Hence $\|f|V \cap P_\nu\|^2 = \int_{V\cap P_\nu} |f|^2 e^{-u}$, cf. (2.1). Therefore (2.11) implies

$$\lim_{\nu} \operatorname{dist}(x, P_\nu)\|f|V_0 \cap P_\nu\| = 0,$$

and another application of Lemma 2.1, this time the "if" direction, proves $\mathbf{f}_x \in J$. □

3. Smooth Hermitian metrics and their curvature

In this section we review the basics of Hermitian metrics on holomorphic Hilbert bundles. Recall that a holomorphic Hilbert bundle is given by a holomorphic map $\pi : E \to X$ of complex manifolds modelled on complex Hilbert spaces, with each fiber $E_x = \pi^{-1}x$ endowed with the structure of a complex vector space ($x \in X$). It is required that for each $x \in X$ there be a neighborhood $U \subset X$, a Hilbert space H, and a biholomorphism (local trivialization) $E|U \to U \times H$ that maps the fibers E_y, $y \in U$, linearly on $\{y\} \times H$. While in what follows we will allow E to be infinite-dimensional, the base X will be kept finite-dimensional.

For the time being we restrict ourselves to trivial bundles $E = U \times H \to U$, with $U \subset \mathbb{C}^m$ open and $(H, \langle,\rangle)$ a complex Hilbert space. We write End H for the space of bounded linear operators on H, endowed with the operator norm. If $k = 0, 1, \dots, \infty$, a Hermitian metric on E of class C^k is a function $h\colon E \oplus E \to \mathbb{C}$ that can be represented as

$$h\left((z, \xi), (z, \eta)\right) = \langle P(z)\xi, \eta\rangle, \quad z \in U,\ \xi, \eta \in H, \tag{3.1}$$

with $P\colon U \to \operatorname{End} H$ a C^k map taking values in invertible, positive self adjoint operators. If $e \in E$, we write $h(e)$ for $h(e, e)$. Thus $\sqrt{h(e)}$ defines a norm on the fibers E_z of E. As usual, for two metrics $h \le k$ means $h(e) \le k(e)$ for all $e \in E$.

Just like in bundles of finite rank, a C^2 Hermitian metric h on E has a curvature R, a $(1, 1)$-form valued in End $E = \coprod_{z\in U}$ End E_z, see, e.g., [B2 or L3]. It is given by

$$R = \overline{\partial}(P^{-1}\partial P) = P^{-1}\overline{\partial}\partial P - P^{-1}\overline{\partial}PP^{-1} \wedge \partial P, \tag{3.2}$$

where P is from (3.1) and we identified $H = E_z$. Curvature determines a Hermitian form N on each space $T^{1,0}_z U \otimes E_z = T^{1,0}_z U \otimes H$ (tensor product over $\mathbb{C}$), given by

$$N(t \otimes \xi, u \otimes \eta) = h\left(R(t, \overline{u})\xi, \eta\right), \quad t, u \in T^{1,0}_z U, \quad \xi, \eta \in E_z. \tag{3.3}$$

Thus N is a Hermitian form on $T^{1,0}U \otimes E$, that we call the Nakano curvature of h. Instead of the usual terminology that h has Nakano semipositive curvature we can then say that the Nakano curvature of h is semipositive. – If $\tau \in T^{1,0}U \otimes E$ again we write $N(\tau)$ for $N(\tau, \tau)$.

In general, given Hermitian forms M, M' on $T^{1,0}U \otimes E$ we write $M \geq M'$ if $M - M'$ is positive semidefinite. Such forms can be written

$$M\Big(\sum_\nu \frac{\partial}{\partial z_\nu} \otimes \xi_\nu\Big) = \sum_{\mu,\nu} \langle M_{\mu\nu}\xi_\mu, \xi_\nu\rangle$$

with $M_{\mu\nu}\colon U \to \operatorname{End} H$ (or $M_{\mu\nu}$ sections of $\operatorname{End} E$). For example, if $M = N$ is the Nakano curvature of the metric h in (3.1), and $R = \sum R_{\mu\nu} dz_\mu \wedge d\overline{z}_\nu$, then $M_{\mu\nu} = PR_{\mu\nu}$. Thus $N \geq 0$, or N is semipositive if

$$\sum_{\mu,\nu} \langle PR_{\mu\nu}\xi_\mu, \xi_\nu\rangle \geq 0 \qquad \text{for arbitrary } \xi_1, \dots, \xi_m \in H.$$

4. Possibly singular Hermitian metrics

In this section we will introduce general Hermitian metrics on not necessarily trivial Hilbert bundles, but first we discuss degenerations of norms on an arbitrary complex Banach space B. Let $\| \ \|_1, \| \ \|_2, \dots$ be a sequence of norms on B, each generating the topology of B. For $x \in B$ set

$$\|x\| = \sup_j \|x\|_j \leq \infty, \qquad \text{and } A = \{x \in X\colon \|x\| < \infty\}.$$

Proposition 4.1. *$A \subset B$ is a linear subspace and $(A, \| \ \|)$ is a Banach space.*

Proof. That A is a subspace follows from the triangle inequality, and $\| \ \|$ is clearly a norm on A. To check completeness, consider a Cauchy sequence x_ν in $(A, \| \ \|)$. Then x_ν is Cauchy in $(B, \| \ \|_1)$ as well, hence $x = \lim x_\nu$ exists in the topology of B. For any j

$$\|x - x_\nu\|_j = \lim_{\mu\to\infty} \|x_\mu - x_\nu\|_j \leq \limsup_{\mu\to\infty} \|x_\mu - x_\nu\|, \text{ whence}$$

$$\|x - x_\nu\| \leq \limsup_{\mu\to\infty} \|x_\mu - x_\nu\| \to 0, \qquad \text{as } \nu \to \infty.$$

Thus $x \in A$ and $x_\nu \to x$ in $\| \ \|$. □

We will apply this construction to the fibers of a holomorphic Hilbert bundle $E \to X$ over a finite-dimensional complex manifold. Recall that a C^k Hermitian metric on E is a function $h\colon E \oplus E \to \mathbb{C}$ that in any local trivialization $E|U \simeq U \times H$ is a C^k Hermitian metric in the sense discussed in Section 3, cf. also [L3]. If $k \geq 2$, the Nakano curvature of h is a Hermitian form N on $T^{1,0}X \otimes E$ that can be computed in local trivializations by (3.2), (3.3). Given a continuous Hermitian metric h on E and a continuous real $(1,1)$ form ia represented in local coordinates as $i\sum a_{\mu\nu} dz_\mu \wedge d\overline{z}_\nu$, we define a Hermitian form $a \otimes h$ on $T^{1,0}X \otimes E$ by

$$(a \otimes h)\Big(\sum \frac{\partial}{\partial z_\nu} \otimes \xi_\nu\Big) = \sum a_{\mu\nu} h(\xi_\mu, \xi_\nu).$$

Definition 4.2. *For the purposes of this paper a function* $h\colon E \to [0,\infty]$ *is called a Hermitian metric if there is a sequence* $h_1 \le h_2 \le \cdots$ *of Hermitian metrics of class* C^2 *on* E *such that* $h(e) = \lim_j h_j(e)$ *for all* $e \in E$. *Given a continuous real* $(1,1)$ *form* ia *on* X, *we say that the Nakano curvature of* h *dominates* a *if the* h_j *can be chosen to have Nakano curvature* $N_j \ge a \otimes h_j$, *i.e., for* $\sum_\nu t_\nu \otimes \xi_\nu \in T^{1,0}X \otimes E$

$$N_j\Big(\sum_\nu t_\nu \otimes \xi_\nu\Big) \ge \sum_{\mu,\nu} a(t_\mu, \overline{t}_\nu) h_j(\xi_\mu, \xi_\nu).$$

The definition raises the obvious question whether for the Nakano curvature of a C^2 Hermitian metric semipositivity (as in Section 3) is the same as dominating 0. This boils down to asking whether among C^2 Hermitian metrics semipoisitive Nakano curvature is inherited under increasing limits. I do not know if the answer is yes or no, but either outcome would have interesting consequences. For example, a "no" answer would probably allow us to extend analytical results known when the Nakano curvature is semipositive (or bounded below in a certain way) to a class of smooth, and then also singular, metrics subject to a weaker curvature condition.

The notion of a possibly singular Hermitian metric on a vector bundle of finite rank has already been proposed by de Cataldo, Berndtsson–Paŭn, and Raufi [BP, dC, R]. Those notions are more general than ours. At the same time, of these authors only de Cataldo defines what Nakano positivity should mean for such a metric h or what it should mean that the Nakano curvature dominates a form a, and in this he is more restrictive than Definition 4.2. In his definition, like in ours, h should be the increasing limit of C^2 Hermitian metrics h_j, with lower estimates on their Nakano curvature, but he requires additionally that on an open subset of X of full measure the h_j should converge in the C^2 topology. (In other respects his definition is less restrictive than ours, namely in what sort of lower estimates are required on the Nakano curvature of h_j. We chose the condition in our definition because it is easy to formulate.)

By Proposition 4.1, a Hermitian metric h defines in each fiber E_x a Hilbert space $F_x = \{e \in E_x\colon h(e) < \infty\}$, endowed with the norm $\sqrt{h(e)}$. The F_x together form a field of Hilbert spaces $F = \coprod_{x\in X} F_x \to X$, see [G], a notion more general than a Hilbert bundle.

Let dV be a positive, continuous volume form on X and h, h_j as in Definition 4.2. For any measurable section f of E

$$\int_X h(f)dV = \lim_{j\to\infty} \int h_j(f)dV = \sup_j \int h_j(f)dV,$$

and we denote by $\|f\| = \|f\|_h \le \infty$ the square root of this quantity. We define $L^2(X,h) = L^2(X,h,dV)$ as the space of f for which $\|f\| < \infty$. The norm $\|\ \|$ on $L^2(X,h)$ is a Hilbertian norm, and $(L^2(X,h), \|\ \|)$ is complete, as Proposition 4.1 shows. Since $\|f\|^2 \ge \int_X h_1(f)dV$, for holomorphic sections convergence in $L^2(X,h)$ implies locally uniform convergence. Hence holomorphic sections form a closed subspace of $L^2(X,h)$.

Fix now a continuous Hermitian metric on X (i.e., on $T^{1,0}X$) with volume form dV. If φ is an E-valued (p,q) form, we define its norm $|\varphi| = |\varphi|_h \colon X \to [0,\infty]$ in the following way. Suppose $x \in X$, $z_1,\dots,z_m$ are local coordinates at x, and $\varphi = \sum \varphi_{IJ} dz^I \wedge d\overline{z}^J$. Then

$$|\varphi|^2(x) = |\varphi|_h^2(x) = \sum_{I,J} h(\varphi_{IJ}(x)) \le \infty, \tag{4.1}$$

provided $\partial/\partial z_\nu$ form an orthonormal basis of $T_x^{1,0}X$. If φ is measurable, we put

$$\|\varphi\| = \|\varphi\|_h = \left(\int_X |\varphi|^2 dV\right)^{1/2} \le \infty. \tag{4.2}$$

Clearly

$$\|\varphi\|_h = \lim_{j\to\infty} \|\varphi\|_{h_j} = \sup_j \|\varphi\|_{h_j}. \tag{4.3}$$

We define $L^2_{pq}(X,h)$ as the space of φ such that $\|\varphi\|_h < \infty$; thus $L^2_{00}(X,h) = L^2(X,h)$. Again $L^2_{pq}(X,h)$ with the norm $\|\ \|_h$ is a Hilbert space.

Consider a Hilbert space $(H, \langle,\rangle)$ and functions $f, f_j \colon X \to H$. We say that $f_j \to f$ uniformly weakly if $\langle f_j, v\rangle \to \langle f, v\rangle$ uniformly for every $v \in H$. For a trivialized Hilbert bundle $E = X \times H \to X$ we define uniform weak convergence of sections accordingly. Finally, sections f_j of a general Hilbert bundle $E \to X$ converge locally uniformly weakly to a section f if in some neighborhood of each $x \in X$ the bundle has a trivialization in which $f_j \to f$ uniformly weakly. In general, uniform weak convergence in one trivialization over some $U \subset X$ does not imply uniform weak convergence in some other trivialization, even over relatively compact $V \subset U$, but if the limit section is locally bounded, it does.

Lemma 4.3. *Let $E \to X$ be a holomorphic Hilbert bundle, $h_1 \le h_2 \le \cdots \to h$ Hermitian metrics on it, and suppose $\varphi_j \in L^2(X,h_j)$. If $\sup_j \|\varphi_j\|_{h_j} < \infty$, and $\varphi_j - \varphi_1$ is holomorphic for each j, then a subsequence of φ_j will converge locally uniformly weakly to a section φ such that $\|\varphi\|_h \le \sup_j \|\varphi_j\|_{h_j}$ and $\varphi - \varphi_1$ is holomorphic.*

Proof. We can assume $E = X \times H$ is trivialized. Thus (holomorphic) sections of E are in one-to-one correspondence with (holomorphic) functions $X \to H$, and one can talk about uniform boundedness and equicontinuity of a family of sections. For example, if we choose a C^2 Hermitian metric $k \le h_1$ on E, since $\sup_j \|\varphi_j\|_k < \infty$, the holomorphic sections $\varphi_j - \varphi_1$ are locally uniformly bounded and by Cauchy's formula, locally uniformly equicontinuous. Now closed balls in H are (even sequentially) compact in the weak topology. Hence for each $x \in$

X the sequence $\varphi_j(x) - \varphi_1(x)$ contains a weakly convergent subsequence. The Arzelà–Ascoli theorem therefore provides a locally uniformly weakly convergent subsequence $\varphi_{j_i} \to \varphi$. For any $v \in H$ the function $\langle \varphi - \varphi_1, v\rangle = \lim\langle \varphi_{j_i} - \varphi_1, v\rangle$ is holomorphic, whence $\varphi - \varphi_1$ is holomorphic (see, e.g., [M, Exercise 8E], whose solution rests on Cauchy's formula and the principle of uniform boundedness).

To estimate $\|\varphi\|$, let $s = \sup_j \|\varphi_j\|_{h_j}$. Fix j and choose a sequence $k_1 \le k_2 \le \cdots$ of C^2 Hermitian metrics that converge to h_j. For any $x \in X$ and $p = 1, 2, \ldots$, in the Hilbert space $(\{x\} \times H, k_p)$ the sequence $\varphi_{j_i}(x)$ weakly converges to $\varphi(x)$. Hence

$$k_p(\varphi(x)) \le \liminf_{i\to\infty} k_p(\varphi_{j_i}(x)) \quad \text{and}$$
$$\|\varphi\|_{k_p} \le \liminf_{i\to\infty} \|\varphi_{j_i}\|_{k_p} \le s$$

by Fatou's lemma. Therefore $\|\varphi\|_{h_j} \le s$ and $\|\varphi\| \le s$ by (4.3), and the proof is complete. □

5. The $\overline{\partial}$-equation and Hörmander–Skoda type estimates

Consider a holomorphic Hilbert bundle E over an m-dimensional complex manifold X. We denote by $\mathcal{D}^{pq}(X, E)$ the space of compactly supported smooth E-valued (p, q) forms. The Cauchy–Riemann operator $\overline{\partial}_E \colon \mathcal{D}^{pq}(X, E) \to \mathcal{D}^{p,q+1}(X, E)$ can be defined as in bundles of finite rank, for instance using local trivializations, see, e.g., [M or L1]. It can be extended to an operator defined on a larger subspace of $L^1_{pq,\mathrm{loc}}(X, E)$. By the latter space we mean the following. Take a continuous Hermitian metric h_0 on E and a smooth volume form dV on X. Then a measurable E-valued (p, q) form φ is in $L^1_{pq,\mathrm{loc}}(X, E)$ if

$$\int_C |\varphi|_{h_0} dV < \infty \text{ for any compact } C \subset X,$$

$|\varphi|_{h_0}$ defined in (4.1). Clearly $L^1_{pq,\mathrm{loc}}(X, E)$ is independent of the choice of h_0 and dV.

Let $(\ ,\ )$ denote the fiberwise pairing between E and its dual E^*. If $\sum \varphi_{IJ} dz^I \wedge d\overline{z}^J$ and $\sum \sigma_{KL} dz^K \wedge d\overline{z}^L$ are local expressions of E, respectively E^*-valued forms φ, σ, put

$$(\varphi, \sigma) = \sum_{I,J,K,L} (\varphi_{IJ}, \sigma_{KL}) dz^I \wedge d\overline{z}^J \wedge dz^K \wedge d\overline{z}^L.$$

Given $\varphi \in L^1_{pq,\mathrm{loc}}(X, E)$ and $\psi \in L^1_{p,q+1,\mathrm{loc}}(X, E)$, we write $\overline{\partial}\varphi = \psi$ if for any $\sigma \in \mathcal{D}^{m-p,m-q-1}(X, E^*)$

$$\int_X (\varphi, \overline{\partial}_{E^*}\sigma) = -\int_X (\psi, \sigma).$$

If such a ψ exists, it is uniquely determined a.e. by φ. Further, if $\varphi \in \mathcal{D}^{pq}(X, E)$ then $\overline{\partial}\varphi$ and $\overline{\partial}_E\varphi$ agree.

In this section we will reproduce the by now standard L^2 estimate, essentially due to Hörmander and Skoda, and streamlined and generalized by Demailly [D1,H,S], for solving $\overline{\partial}$ in the setting of Hilbert bundles. Our setting is slightly more general than the one in [D1, VIII. Theorem 6.1] because we allow bundles $E \to X$ of infinite rank and possibly singular metrics.

Fix a smooth Hermitian metric on X and a Hermitian metric on E. In addition to the Hilbert spaces $L^2_{pq}(X,h)$ and norms $\|\varphi\|_h$ introduced in Section 4, see (4.2), we will need one more piece of notation. Consider a continuous $(1,1)$ form a on X, $ia \geq 0$. If φ is an E-valued $(p,1)$ form, $0 \leq p \leq m$, its weighted norm $|\varphi|_{h,a}\colon X \to [0,\infty]$ is defined as follows. Given $x \in X$, choose local coordinates $z_1, \dots, z_m$ at x in which a diagonalizes: $a = \sum a_{\nu\nu} dz_\nu \wedge d\overline{z}_\nu$ at x. Then

$$|\varphi|^2_{h,a}(x) = \sum_\nu a^{-1}_{\nu\nu} |i_{\partial/\partial\overline{z}_\nu}\varphi|^2_h(x) \in [0,\infty], \tag{5.1}$$

where $i_{\partial/\partial\overline{z}_\nu}$ stands for contraction.

Recall that a complex manifold X is weakly pseudoconvex if it admits a smooth plurisubharmonic exhaustion function $X \to [0,\infty)$.

Theorem 5.1. *Let (X,ω) be an m-dimensional weakly pseudoconvex Kähler manifold, $ia \geq 0$ a continuous $(1,1)$ form on X, $E \to X$ a holomorphic Hilbert bundle, and h a Hermitian metric on E. Suppose that the Nakano curvature of h dominates a, cf. Definition 4.2. Given $\psi \in L^1_{m1,loc}(X,h)$,*

$$\overline{\partial}\psi = 0 \qquad \textit{and} \qquad \int_X |\psi|^2_{h,a}\omega^m < \infty,$$

there exists a $\varphi \in L^2_{m0}(X,h)$ such that

$$\overline{\partial}\varphi = \psi \qquad \textit{and} \qquad \int_X |\varphi|^2_h \omega^m \leq \int_X |\psi|^2_{h,a}\omega^m. \tag{5.2}$$

Proof. First we assume that h is of class C^2 and its Nakano curvature satisfies for $\sum t_\nu \otimes \xi_\nu \in T^{1,0}X \otimes E$

$$N\Big(\sum_\nu t_\nu \otimes \xi_\nu\Big) = \sum h\big(R(t_\mu,t_\nu)\xi_\mu,\xi_\nu\big) \geq \sum_{\mu,\nu} a(t_\mu,\overline{t}_\nu)h(\xi_\mu,\xi_\nu). \tag{5.3}$$

It is straightforward if perhaps tedious to check that the theory expounded in Demailly's book [D1], Chapters VII and VIII, is valid in Hilbert bundles of infinite rank. In particular, Theorem 6.1 in Chapter VIII holds for such bundles. The hypotheses of that theorem are obviously satisfied now, except possibly the one which involves an operator A on $\bigwedge^{\bullet,\bullet} T^*X \otimes E$, which we will have to check. Only the action of A on $\bigwedge^{m,1} T^*X \otimes E$ matters here. This was computed in [D1, VII.(7.1)], and in our notation it can be given as follows. Let $z_1, \dots, z_m$ be local coordinates at $x \in X$ such that $\partial/\partial z_\nu$ form a basis in $T^{1,0}_x X$, orthonormal for the inner product induced by ω. If the curvature of h is $R = \sum R_{\mu\nu} dz_\mu \wedge d\overline{z}_\nu$, then

$$h(A\psi, \psi) = \sum_{\mu,\nu} h(R_{\mu\nu}\psi_\mu, \psi_\nu), \quad \psi = \sum \psi_\nu dz \wedge d\overline{z}_\nu \in \bigwedge^{m,1} T_x^* X \otimes E_x,$$

where dz stands for $dz_1 \wedge \cdots \wedge dz_m$. If additionally we choose the coordinates so that $a = \sum a_{\nu\nu} dz_\nu \wedge d\overline{z}_\nu$ at x, then by (5.3)

$$h(A\psi, \psi) \geq \sum_\nu a_{\nu\nu} h(\psi_\nu) = \sum a_{\nu\nu} |i_{\partial/\partial\overline{z}_\nu}\psi|_h^2.$$

Hence if now ψ is an E-valued $(m, 1)$ form, at x

$$h(A^{-1}\psi, \psi) \leq \sum_\nu a_{\nu\nu}^{-1} |i_{\partial/\partial\overline{z}_\nu}\psi|_h(x) = |\psi|_{h,a}^2(x)$$

by (5.1).

The estimates in [D1, VIII. Theorems 6.1 and 4.5] are formulated in terms of $h(A^{-1}\psi, \psi)$, but clearly any greater function will also do. Replacing $\langle A^{-1}\psi, \psi\rangle = h(A^{-1}\psi, \psi)$ in Demailly's formulae by $|\psi|_{h,a}^2$ we obtain Theorem 5.1 when h is of class C^2.

A general h is the increasing limit of C^2 Hermitian metrics h_j with Nakano curvature $N_j \geq a \otimes h_j$. By what we have just seen, there are $\varphi_j \in L_{m0}^2(X, h_j)$ such that

$$\overline{\partial}\varphi_j = \psi \quad \text{and} \quad \|\varphi_j\|_{h_j} \leq \int_X |\psi|_{h_j,a}^2 \omega^m \leq \int_X |\psi|_{h,a}^2 \omega^m.$$

Denoting the canonical bundle of X by K, we can view φ_j as sections of $K \otimes E$ and apply Lemma 4.3 with E replaced by $K \otimes E$. Any subsequential weak limit φ will then satisfy (5.2). $\square$

6. An extension theorem of Ohsawa–Takegoshi type

The original publication [OT] of the Ohsawa–Takegoshi extension theorem sparked a lot of interest, various generalizations and alternative approaches have been proposed. Here we prove an extension for Hilbert bundles following an idea of Bo-Yong Chen [C], see also [Bl]. Guan and Zhou in [GZ1] already proved an extension theorem for finite rank vector bundles. Undoubtedly their proof could be generalized to Hilbert bundles as well, but Chen's approach is the simplest of all.

At the heart of all extension proofs are estimates for the solution of an equation

$$\overline{\partial}\varphi = \psi. \tag{6.1}$$

In his estimations Chen is inspired by an idea of Berndtsson that first appeared in [B1] and then in [BC]. Berndtsson's idea, in a context different from extensions, was as follows. Given (in [B1, BC] a scalar-valued) form ψ, suppose we find a solution φ whose L^2 norm with respect to a certain weight is minimal. This means that φ is orthogonal to Ker $\overline{\partial}$ in some weighted L^2 space. If u is a bounded function, then

$e^u\varphi$ will still be orthogonal to Ker $\overline{\partial}$, albeit with respect to a modified weight. So it will be the minimal solution of the "twisted" equation

$$\overline{\partial}(e^u\varphi) = e^u(\overline{\partial}u \wedge \varphi + u\psi). \tag{6.2}$$

If the weights involved are plurisubharmonic, then one can therefore use Hörmander's estimate (really, [D1, VIII. Theorem 6.1]) to bound the solution $e^u\varphi$ of (6.2) in terms of the right-hand side. True, this bound will involve φ itself, but if $\overline{\partial}u$ is sufficiently small, then the bound can be turned into one that involves ψ only, and provides much stronger estimates on φ then what follows directly from (6.1). To carry out this plan, Berndtsson assumed sup $|\overline{\partial}u|_{\partial\overline{\partial}u} < 1$. What Chen noticed was that useful estimates may follow even if sup $|\overline{\partial}u|_{\partial\overline{\partial}u} = 1$, and he produced a u that will do the trick for the $\overline{\partial}$ equation that arises in the extension problem.

Before introducing our version of the extension theorem we have to develop some notation. Suppose X is a smooth manifold, $Y \subset X$ a smooth submanifold, and $r\colon X \to [0,\infty)$ is a C^3 function, $Y = r^{-1}(0)$. Suppose further that, denoting by dist the distance induced by some Riemannian metric on X

$$r(x) \geq c_0 \operatorname{dist}^2(x, Y), \tag{6.3}$$

c_0 a positive constant. Given such an r, any continuous volume form $d\mu$ on X induces a volume form $d\mu_r$ on Y as follows. Suppose $\theta \in C(Y)$ is compactly supported. Extend it to a compactly supported $\tilde{\theta} \in C(X)$. Then $d\mu_r$ will satisfy

$$\int_Y \theta d\mu_r = \lim_{\varepsilon\to 0} \varepsilon^{-\operatorname{codim}_{\mathbb{R}} Y} \int_{\{x\in X\colon r(x)<\varepsilon^2\}} \tilde{\theta} d\mu. \tag{6.4}$$

Locally $d\mu_r$ can be computed if we introduce local coordinates $x_1, \ldots, x_l$ on X so that $r = \sum_1^k x_i^2$. If $d\mu = \alpha dx_1 \ldots dx_l$ and c_k is the volume of the unit ball in $\mathbb{R}^k$, then $d\mu_r = c_k \alpha dx_{k+1} \ldots dx_l$.

In the theorem below we will deal with a Hermitian holomorphic Hilbert bundle (E, h) over a Kähler manifold (X, ω). We write K for the canonical bundle of X and h^K for the metric on $K \otimes E$ induced by h and ω. That is, if $z_1, \ldots, z_m$ are local coordinates at $x \in X$ such that $\partial/\partial z_\nu$ form an orthonormal basis in $T_x^{1,0}X$, then

$$h^K(dz_1 \wedge \cdots \wedge dz_m \otimes \xi) = h(\xi), \qquad \xi \in E_x.$$

In other words, if we view a section g of $K \otimes E$ as an E-valued $(m, 0)$ form (and we will), then with notation (4.1)

$$h^K(g) = |g|_h^2. \tag{6.5}$$

Theorem 6.1. *Let (X, ω) be a weakly pseudoconvex m-dimensional Kähler manifold, $Y \subset X$ a complex submanifold of dimension $m - c$, $r\colon X \to [0, 1/(2e^2)]$ a C^3 function that vanishes on Y and satisfies* (6.3)*. Define a volume form $d\mu_r$ on Y by* (6.4) *using the volume form ω^m on X. Suppose that* $\log r$ *is plurisubharmonic. Let furthermore $E \to X$ be a holomorphic Hilbert bundle with a Hermitian metric* *h whose Nakano curvature dominates* 0*. If f is a holomorphic section of $K \otimes E$ over some neighborhood $U \subset X$ of Y and*

$$\int_Y h^K(f)d\mu_r < \infty,$$

then there is a holomorphic section g of $K \otimes E$ such that $f = g$ on Y and

$$\int_X \frac{h^K(g)}{r^c \log^2 r}\,\omega^m \le 4^{4+c}\int_Y h^K(f)d\mu_r. \tag{6.6}$$

Proof. We start by considering a holomorphic Hilbert bundle $F \to X$ with a Hermitian metric k, an upper semicontinuous $u\colon X \to [-\infty, \infty)$, and the Hermitian metric $k' = e^{-u}k$. When k, u are (finite and) of class C^2, one can compute that the curvatures R, R' of k and k' are related by

$$R' = R + I\partial\overline{\partial}u,$$

I denoting the identity endomorphism of F. Hence the Nakano curvatures N, N' satisfy

$$\begin{aligned} N'\Bigl(\sum_\nu t_\nu \otimes \xi_\nu\Bigr) &= \sum_{\mu,\nu} k'\left(R'(t_\mu, \overline{t}_\nu)\xi_\mu, \xi_\nu\right) \\ &= e^{-u}N\Bigl(\sum_\nu t_\nu \otimes \xi_\nu\Bigr) + \sum_{\mu,\nu}\partial\overline{\partial}u(t_\mu, \overline{t}_\nu)k'(\xi_\mu, \xi_\nu). \end{aligned} \tag{6.7}$$

It follows that if the Nakano curvature of k is semipositive and u is plurisubharmonic, then the Nakano curvature of k' is also semipositive. By approximation, this will also hold for possibly singular metrics k and plurisubharmonic functions u that are decreasing limits of C^2 plurisubharmonic functions. (6.7) also shows that if the Nakano curvature of a Hermitian metric k dominates 0 and $u \in C^2(X)$, then the Nakano curvature of k' will dominate $\partial\overline{\partial}u$.

Now let us put ourselves in the setting of Theorem 6.1. All integrals of functions on X and on open subsets of X will be with respect to the volume form ω^m; for brevity, from now on we will omit the volume form from the integrals. Since h can be approximated from below by C^2 Hermitian metrics, we can assume h is already such. We will also assume that X is a smoothly bounded, relatively compact open subset of a complex manifold X_0, that ω extends to a smooth Kähler form ω_0 on X_0, that $Y = Y_0 \cap X$, where $Y_0 \subset X_0$ is a complex submanifold intersecting ∂X transversely; that a C^3 extension of r to X_0 still satisfies $r(x) \ge c_0 \operatorname{dist}^2(x, Y_0)$; that (E, h) extends to a holomorphic Hermitian Hilbert bundle (also denoted (E, h)) over X_0, and that f extends to a holomorphic section of $E \otimes K_{X_0}$ over a neighborhood $U_0 \subset X_0$ of Y_0. Once this special case is handled, the general case will follow. We would take a smooth plurisubharmonic exhaustion function $\rho\colon X \to [0, \infty)$, solve the extension problem on generic sublevel sets of ρ and apply Lemma 4.3 combined with a diagonal selection procedure.

With all the extra assumptions above, we fix a smooth function $\chi\colon [0,\infty) \to [0,1]$,

$$\chi(t) = \begin{cases} 1, & \text{if } t < 1/4 \\ 0, & \text{if } t > 1 \end{cases}, \qquad |\chi'(t)| \le 2 \text{ for all } t.$$

Set, for $\varepsilon > 0$,

$$\Omega_\varepsilon = \{x \in X\colon r(x) < \varepsilon^2\}.$$

When ε is sufficiently small, $\Omega_\varepsilon \subset U_0$, so that

$$f' = f'_\varepsilon = \begin{cases} \chi(r/\varepsilon^2) f & \text{on } \Omega_\varepsilon \\ 0 & \text{on } X \setminus \Omega_\varepsilon \end{cases}$$

is a smooth section of $K \otimes E$. Let $\psi = \overline{\partial} f'$, a smooth closed $K \otimes E$-valued $(0,1)$ form. Presently we will check that the equation $\overline{\partial}\varphi = \psi$ has a solution $\varphi = \varphi_\varepsilon \in L^2(X, h^K/r^c)$. Accepting this for the moment, $g = f' - \varphi \in \Gamma(X, K \otimes E)$. Hence φ is smooth and

$$\varphi | Y = 0, \tag{6.8}$$

because $1/r^c$ is nowhere integrable at Y. Thus g agrees with f on Y, and what we need to do is estimate it. This amounts to estimating φ. Since we can freely add to φ any holomorphic section in $L^2(X, h^K/r^c)$, we can arrange that φ is orthogonal to the closed subspace of holomorphic sections in $L^2(X, h^K/r^c)$. It follows that with any $u \in C^2(\overline{X})$ the section $\theta = e^u\varphi$ is orthogonal to holomorphic sections in $L^2(X, e^{-u}h^K/r^c)$. – This latter space is the same as $L^2(X, h^K/r^c)$ but their inner products are different. – Therefore θ is the solution of

$$\overline{\partial}\theta = \overline{\partial}(e^u\varphi) = e^u(\varphi\overline{\partial}u + \psi) \tag{6.9}$$

that has minimal norm in $L^2(X, e^{-u}h^K/r^c)$.

Let us abbreviate $e^{-u}h/r^c = k$. As we observed above, the Nakano curvature of $h/r^c = e^{-c\log r}h$ dominates 0 (note that $\log(r+1/j) \searrow \log r$), and so the Nakano curvature of k dominates $\partial\overline{\partial}u$. We will choose a plurisubharmonic u so that

$$\int_X |\varphi\overline{\partial}u|^2_{k^K,\partial\overline{\partial}u} \quad , \quad \int_X |\psi|^2_{k^K,\partial\overline{\partial}u} < \infty. \tag{6.10}$$

Viewing $\varphi\overline{\partial}u$ and ψ as E-valued $(m,1)$ forms, the integrands above are the same as $|\varphi\overline{\partial}u|^2_{k,\partial\overline{\partial}u}$ and $|\psi|^2_{k,\partial\overline{\partial}u}$, hence by Theorem 5.1 and (6.9) we would conclude

$$\int_X k^K(\theta) \le \int_X |e^u(\varphi\overline{\partial}u + \psi)|^2_{k^K,\partial\overline{\partial}u},$$

or

$$\int_X \frac{e^u}{r^c} h^K(\varphi) \le \int_{X\setminus\Omega_\varepsilon} \frac{e^u}{r^c}|\varphi\overline{\partial}u|^2_{h^K,\partial\overline{\partial}u} + 2\int_{\Omega_\varepsilon} \frac{e^u}{r^c}|\varphi\overline{\partial}u|^2_{h^K,\partial\overline{\partial}u} + 2\int_{\Omega_\varepsilon} \frac{e^u}{r^c}|\psi|^2_{h^K,\partial\overline{\partial}u}.$$

Since $|\varphi\overline{\partial}u|^2_{h^K,\partial\overline{\partial}u} = h^K(\varphi)|\overline{\partial}u|^2_{\partial\overline{\partial}u}$, putting

$$\lambda(x) = \begin{cases} 1 - |\overline{\partial}u|^2_{\partial\overline{\partial}u}(x), & \text{if } x \in X\setminus\Omega_\varepsilon \\ 1 - 2|\overline{\partial}u|^2_{\partial\overline{\partial}u}(x), & \text{if } x \in \Omega_\varepsilon \end{cases}, \tag{6.11}$$

we obtain

$$\int_X \lambda\,\frac{e^u}{r^c}\,h^K(\varphi) \le 2\int_X \frac{e^u}{r^c}|\psi|^2_{h^K,\partial\overline{\partial}u}. \tag{6.12}$$

Remains to prove that $\overline{\partial}\varphi = \psi$ indeed has a solution in $L^2(X, h^K/r^c)$ and to choose u so that (6.10) holds and (6.12) implies the estimate (6.6) for $g = f' - \varphi$.

Following Chen's idea in a similar setting, we make sure that $\varepsilon < 1/(2e)$ and let

$$\rho = -\log(r+\varepsilon^2), \qquad \eta = \rho + \log\rho, \quad \text{and} \quad u = -\log\eta.$$

Thus $-\rho, -\eta$, and u are plurisubharmonic and C^3,

$$2 < \rho < \eta < 2\rho \qquad \text{and} \qquad u < 0.$$

We need to estimate $|\overline{\partial}u|_{\partial\overline{\partial}u}$. This will take a little bit of computation:

$$\overline{\partial}\rho = -\frac{\overline{\partial}r}{r+\varepsilon^2}, \quad \partial\overline{\partial}\rho = \frac{\partial r\wedge\overline{\partial}r}{(r+\varepsilon^2)^2} - \frac{\partial\overline{\partial}r}{r+\varepsilon^2}, \tag{6.13}$$

$$\eta\overline{\partial}u = -\overline{\partial}\eta = -\left(1+\frac{1}{\rho}\right)\overline{\partial}\rho = \left(1+\frac{1}{\rho}\right)\frac{\overline{\partial}r}{r+\varepsilon^2}, \tag{6.14}$$

$$\begin{aligned} i\eta^2\partial\overline{\partial}u &= i\left(\left(1+\frac{1}{\rho}\right)^2 + \frac{\eta}{\rho^2}\right)\partial\rho\wedge\overline{\partial}\rho - i\eta\left(1+\frac{1}{\rho}\right)\partial\overline{\partial}\rho \\ &\ge i\left(\left(1+\frac{1}{\rho}\right)^2 + \frac{1}{\rho}\right)\partial\rho\wedge\overline{\partial}\rho. \end{aligned} \tag{6.15}$$

From (6.13), (6.15), on X

$$i\partial r\wedge\overline{\partial}r \le \frac{i\eta^2(r+\varepsilon^2)^2}{(1+1/\rho)^2+1/\rho}\,\partial\overline{\partial}u, \quad |\overline{\partial}r|^2_{\partial\overline{\partial}u} \le \frac{\eta^2(r+\varepsilon^2)^2}{(1+1/\rho)^2+1/\rho}. \tag{6.16}$$

However, on Ω_ε the estimate can be improved. Indeed, (6.13) gives, when $r < \varepsilon^2$

$$\begin{aligned} (r+\varepsilon^2)^2(-i\partial\overline{\partial}\rho - i\partial\rho\wedge\partial\overline{\rho}) &= i(r+\varepsilon^2)\partial\overline{\partial}r - 2i\partial r\wedge\overline{\partial}r \\ &\ge 2i(r\partial\overline{\partial}r - \partial r\wedge\overline{\partial}r) \ge 0, \end{aligned}$$

the latter simply expressing that $i\partial\overline{\partial}\log r \ge 0$. Hence by (6.13), (6.15)

$$i\partial r\wedge\overline{\partial}r = i(r+\varepsilon^2)^2\partial\rho\wedge\overline{\partial}\rho \le -4i\varepsilon^4\partial\overline{\partial}\rho \le 4i\varepsilon^4\eta\partial\overline{\partial}u,$$

so that $|\overline{\partial}r|^2_{\partial\overline{\partial}u} \le 4\varepsilon^4\eta$. Therefore in view of (6.14), (6.16)

$$|\overline{\partial}u|^2_{\partial\overline{\partial}u} = \frac{(1+1/\rho)^2}{\eta^2(r+\varepsilon^2)^2}\,|\overline{\partial}r|^2_{\partial\overline{\partial}u} \le \begin{cases} \dfrac{16}{\eta} & \text{on } \Omega_\varepsilon \\ \dfrac{(\rho+1)^2}{\rho+1)^2+\rho} \le 1 - \dfrac{1}{4\rho} & \text{on } X, \end{cases} \tag{6.17}$$

as $\rho > 2$. On supp $\psi \subset \Omega_\varepsilon \backslash \Omega_{\varepsilon/2}$ we can estimate

$$|\psi|^2_{h^K,\partial\overline{\partial}u} = |f\chi'(r/\varepsilon^2)\overline{\partial}r/\varepsilon^2|^2_{h^K,\partial\overline{\partial}u} \le 4h^K(f)|\overline{\partial}r|^2_{\partial\overline{\partial}u}/\varepsilon^4 \le 16\eta h^K(f),$$

and so

$$\int_X \frac{e^u}{r^c}\,|\psi|^2_{h^K,\partial\overline{\partial}u} \le \frac{4^{2+c}}{\varepsilon^{2c}} \int_{\Omega_\varepsilon\setminus\Omega_{\varepsilon/2}} h^K(f) < \infty. \tag{6.18}$$

We apply Theorem 5.1 with the Hermitian metric $e^{-u}h/r^c$ to obtain a solution $\varphi \in L^2(X, e^{-u}h^K/r^c) = L^2(X, h^K/r^c)$ of the equation $\overline{\partial}\varphi = \psi$, which we choose to have the minimal norm in $L^2(X, h^K/r^c)$.

Now (6.18) and

$$\int_X r^{-c}|\varphi\overline{\partial}u|_{h^K,\partial\overline{\partial}u} = \int_X r^{-c}h^K(\varphi)|\overline{\partial}u|_{\partial\overline{\partial}u} \le \int_X r^{-c}\,h^K(\varphi) < \infty$$

prove (6.10), and it follows that φ satisfies (6.12). Looking up the definition of λ, (6.11), and comparing it with (6.17), we obtain

$$\lambda \ge \begin{cases} 1/2 & \text{on } \Omega_\varepsilon \\ 1/4\rho & \text{on } X\setminus\Omega_\varepsilon \end{cases} \ge \frac{1}{4\rho} \ge \frac{1}{4|\log r|}$$

when $\varepsilon > 0$ is sufficiently small. We also note that $e^u = 1/\eta \ge 1/(2\log r)$. Putting all this in (6.12) and taking (6.18) into account estimates $\varphi = \varphi_\varepsilon$:

$$\int_X \frac{h^K(\varphi_\varepsilon)}{r^c\log^2 r} \le \frac{4^{4+c}}{\varepsilon^{2c}} \int_{\Omega_\varepsilon} h^K(f). \tag{6.19}$$

We let $\varepsilon \to 0$. To estimate the limit on the right, set $\Omega^0_\varepsilon = \{x \in X_0 \colon r(x) < \varepsilon^2\}$ and choose a compactly supported continuous function $\sigma\colon X_0 \to [0,1]$ such that $\sigma|X \equiv 1$. Then

$$\limsup_{\varepsilon\to 0} \varepsilon^{-2c}\int_{\Omega_\varepsilon} h^K(f) \le \limsup_{\varepsilon\to 0}\varepsilon^{-2c}\int_{\Omega^0_\varepsilon} \sigma h^K(f) = \int_{Y_0} \sigma h^K(f)d\mu_r.$$

Approximating the characteristic function of $\overline{X}$ by such σ, this last integral gets as close to $\int_Y h^K(f)d\mu_r$ as we please, whence by (6.19)

$$\limsup_{\varepsilon\to 0} \int_X \frac{h^K(\varphi_\varepsilon)}{r^c\log^2 r} \le 4^{4+c}\int_Y h^K(f)d\mu_r.$$

At the same time, since $1/(r^c\log^2 r)$ is integrable on X,

$$\lim_{\varepsilon\to 0}\int_X \frac{h^K(f'_\varepsilon)}{r^c\log^2 r} = \lim_{\varepsilon\to 0}\int_{\Omega_\varepsilon} \frac{h^K\big(\chi(r/\varepsilon^2)f\big)}{r^c\log^2 r} = 0.$$

Therefore $g_\varepsilon = f'_\varepsilon - \varphi_\varepsilon$ is holomorphic, $g_\varepsilon|Y = f|Y$, and

$$\limsup_{\varepsilon\to 0} \int \frac{h^K(g_\varepsilon)}{r^c\log^2 r} \le 4^{4+c}\int_Y h^K(f)d\mu_r.$$

By Lemma 4.3 a subsequential weak limit g of g_ε will then satisfy the requirements. $\square$

7. The proof of Theorem 1.2

Let $U \subset \mathbb{C}^m$ be open and $E \to U$ a holomorphic Hilbert bundle as in Theorem 1.2. Let furthermore $h_1 \geq h_2 \geq \cdots$ be Hermitian metrics with Nakano curvatures dominating 0, and assume that $h = \lim_j h_j$ is bounded below by a continuous Hermitian metric. Much like in Section 2, if $P \subset \mathbb{C}^m$ is a complex affine hyperplane, $W \subset P \cap U$ is relatively open, and g is a measurable section of $E|W$, we define $\|g\| \in [0, \infty]$ by

$$\|g\|^2 = \inf_j \int_W h_j(g), \quad \text{so that}$$

$$\|g\|^2 = \begin{cases} \infty & \text{or} \\ \lim_j \int_W h_j(g) = \int_W h(g). \end{cases} \tag{7.1}$$

Let $x \in U$. First we prove the following characterization of $M = M_x = \bigcup_j E(h_j, x)$.

Lemma 7.1. *Suppose M is finitely generated as an $\mathcal{O}_x$-module. The germ $\mathbf{f}_x$ of an $f \in \Gamma(E)$ belongs to M if and only if for any sufficiently small neighborhood $V \subset U$ of x and any hyperplane $P_0 \subset \mathbb{C}^m$*

$$\liminf \operatorname{dist}(x, P)\|f|V \cap P\| = 0, \quad \text{as } P \| P_0 \text{ and } \operatorname{dist}(x, P) \to 0. \tag{7.2}$$

We need the following simple

Proposition 7.2. *Let W be a complex vector space, $(B, \| \ \|)$ a normed space, $L\colon W \to B$ and $l\colon W \to \mathbb{C}$ linear. If $|l(w)| \leq C\|L(w)\|$ for every $w \in W$ with some constant C, then there is a linear map $a\colon B \to \mathbb{C}$ of norm $\leq C$ such that $l = aL$.*

Proof. First we define a on $L(W) \subset B$. If $u = L(w) \in L(W)$, set $a(u) = l(w)$. This is independent of the choice of w, since $\operatorname{Ker} L \subset \operatorname{Ker} l$. Further,

$$|a(u)| = |l(w)| \leq C\|L(w)\| = C\|u\|.$$

By the Banach–Hahn theorem we extend a to a linear form on B satisfying the same estimate; this extension will clearly do. □

Proof of Lemma 7.1. We will assume $x = 0$. The "only if" direction follows from Fubini's theorem as in the proof of Lemma 2.1. Conversely, we will show that if $\mathbf{f}_0 \notin M$ then for any neighborhood $V \subset U$ of 0 and for some hyperplane P_0

$$\liminf \operatorname{dist}(x, P)\|f|V \cap P\| > 0, \quad \text{as } P \| P_0 \text{ and } \operatorname{dist}(x, P) \to 0. \tag{7.3}$$

Fix V. We can assume it is pseudoconvex, relatively compact in U, and there are $g^1, \ldots, g^p \in \Gamma(V, E)$ whose germs generate M. We can also assume $\int_V h_{j_0}(g^i) < \infty$ with some j_0 and $i = 1, \ldots, p$.

Write $\pi\colon E \to U$ for the bundle projection. If $\Delta \subset \mathbb{C}$ is the unit disc and $\alpha\colon \Delta \to E^*|V$ is holomorphic, mapping $0 \in \Delta$ to the zero vector in E_0^*, we can associate with sections $g \in \Gamma(V, E)$ functions $\alpha^\star g \in \mathcal{O}(\Delta)$ by evaluating $\alpha(s)$ on $g(\pi\alpha(s))$:

$$(\alpha^\star g)(s) = \alpha(s)g(\pi\alpha(s)).$$

Likewise we can pull back germs $\mathbf{g}_0$ of sections of E to germs $\alpha^\star \mathbf{g}_0 \in \mathcal{O}_{(\mathbb{C},0)}$. Set $\alpha^\star M = \{\alpha^\star \mathbf{g}_0 \colon \mathbf{g}_0 \in M\}$. We claim that there is an α such that

$$\alpha^\star \mathbf{f}_0 \notin \mathcal{O}_{(\mathbb{C},0)}\alpha^\star M. \tag{7.4}$$

Indeed, [L2, Lemma 5.1] implies that over some neighborhood of $0 \in U$, our $f, g^1, \ldots, g^p$ are in fact sections of a finite rank holomorphic subbundle of E. We will construct α as the pull back of a map into the dual of this subbundle, and so we can assume at this juncture that the rank of E itself is finite. Of course, we can also assume E is trivial, $E = U \times \mathbb{C}^n \to U$. With any $g \in \Gamma(V, E)$ given by $g(z) = (z, g_1(z), \ldots, g_n(z))$ we associate a function $\hat{g} \in \mathcal{O}(V \times \mathbb{C}^n)$,

$$\hat{g}(z, w) = \sum_\nu g_\nu(z) w_\nu.$$

We do likewise with germs of sections of E at $0 \in U$. Consider the integral closure $\hat{M} \subset \mathcal{O}_{(\mathbb{C}^{m+n},0)}$ of the ideal generated by $\hat{\mathbf{g}}_0^1, \ldots, \hat{\mathbf{g}}_0^p$. Thus, again by [LT, Théorème 2.1], $\hat{M}$ consists of germs $\boldsymbol{\varphi}_0 \in \mathcal{O}_{(\mathbb{C}^{m+n},0)}$ of functions φ that satisfy

$$|\varphi|^2 \le C^2 \left(|\hat{g}^1|^2 + \ldots + |\hat{g}^p|^2\right) \tag{7.5}$$

on some neighborhood of $0 \in \mathbb{C}^{m+n}$, with some constant C. Suppose $\varphi = \hat{g}$ satisfies (7.5) on some neighborhood; the neighborhood can be taken of form $\pi^{-1}(V_0)$, with $V_0 \subset V$ a neighborhood of $0 \in \mathbb{C}^m$. Let $z \in V_0$. We apply Proposition 7.2 with $W = \mathbb{C}^n$, B the Euclidean space $\mathbb{C}^p$, the components of $L : \mathbb{C}^n \to \mathbb{C}^p$ the functions $\hat{g}^i(z, \cdot)$, $i = 1, \ldots, p$, and $l = \hat{g}(z, \cdot)$. This produces $a_i \in \mathbb{C}$ such that $g(z) = \sum a_i g^i(z)$ and $\sum |a_i|^2 \le C^2$. By Schwarz' inequality

$$h_{j_0}(g(z)) \le \sum_i |a_i|^2 \sum_i h_{j_0}(g^i(z)) \le C^2 \sum_i h_{j_0}(g^i(z)).$$

We see that for a φ of form $\hat{g}$ (7.5) implies $\mathbf{g}_0 \in M$. Since our $\mathbf{f}_0 \notin M$, it follows that $\varphi = \hat{f}$ does not satisfy (7.5) on any neighborhood of 0, i.e., $\hat{\mathbf{f}}_0 \notin \hat{M}$. As in the proof of Lemma 2.1, according to [LT] this implies that there is a holomorphic $\alpha \colon \Delta \to V \times \mathbb{C}^n$, $\alpha(0) = 0$, such that $\hat{\mathbf{f}}_0 \circ \alpha \notin \mathcal{O}_{(\mathbb{C},0)}\hat{M} \circ \alpha$. Then α, viewed as a map $\Delta \to E^*|V = V \times \mathbb{C}^n$, satisfies (7.4).

There is quite some flexibility in the choice of α. To wit, $\hat{\mathbf{f}}_0 \circ \alpha \notin \mathcal{O}_{(\mathbb{C},0)}\hat{M} \circ \alpha$ means that at $0 \in \Delta$ the order of each $\hat{g}_i \circ \alpha$ is greater than the order of $\hat{f} \circ \alpha$. This will clearly persist if we perturb α by adding terms whose order is greater than the order of $\hat{f} \circ \alpha$. We use this flexibility to arrange that $\pi \circ \alpha \ne 0$.

Next choose a hyperplane P_0 through $0 \in \mathbb{C}^m$ that does not contain $\pi\alpha(\Delta)$. Again we can adjust coordinates in $\mathbb{C}$ and in $\mathbb{C}^n$ so that $P_0 = \{z \in \mathbb{C}^m \colon z_1 = 0\}$, that α is holomorphic in a neighborhood of $\overline{\Delta}$, that $F = \alpha^\star f \ne 0$ on $\overline{\Delta} \backslash \{0\}$, that the first component of $\pi\alpha$ satisfies with some $k \in \mathbb{N}$

$$\pi_1 \alpha(s) = s^k, \quad s \in \Delta,$$

and $\alpha(\overline{\Delta}) \subset V \times \mathbb{C}^n$. This latter implies for any $g \in \Gamma(V, E)$

$$\max_{\overline{\Delta}} |\alpha^\star g|^2 \le C_2^2 \int_V h(g) \le C_2^2 \int_V h_j(g), \qquad j = 1, 2 \ldots, \tag{7.6}$$

with some C_2 independent of g.

Let $\sigma \in \Delta\backslash\{0\}$ and let $P_\sigma = \{z \in \mathbb{C}^m \colon z_1 = \sigma\}$. Assume first $\|f|V \cap P_\sigma\| < \infty$. We apply Theorem 6.1 with $(X, \omega) = (V, \sum dz_\nu \wedge d\overline{z}_\nu)$, $Y = V \cap P_\sigma$ and $r(z) = c|z_1 - \sigma|^2$. If the constant $c > 0$ is sufficiently small, then $r \le 1/(2e^2)$ on V. The volume form $d\mu_r$ on $V \cap P_\sigma$ is a constant multiple of the Euclidean volume. Choose j such that

$$\int_{V \cap P_\sigma} h_j(f) \le 2\|f|V \cap P_\sigma\|^2.$$

Since the bundles $(K \otimes E, h_j^K)$ and (E, h_j) are isometrically isomorphic, Theorem 6.1 produces a $g \in \Gamma(V, E)$ with

$$f = g \text{ on } V \cap P_\sigma \qquad \text{and} \qquad \int_V h_j(g) \le C_3^2 \|f|V \cap P_\sigma\|^2, \tag{7.7}$$

where C_3 is independent of σ. Thus $\mathbf{g}_0 \in M$ and the germ of $G = \alpha^\star g$ is in $\alpha^\star M$. As the germ of $F = \alpha^\star g$ is not in $\mathcal{O}_{(\mathbb{C},0)}\alpha^\star M$, it follows that $G = o(F)$ at $0 \in \Delta$. Further, by (7.7) $F(\sqrt[k]{\sigma}) = G(\sqrt[k]{\sigma})$ for any choice of k'th root $\sqrt[k]{\sigma}$. Hence by Proposition 2.2

$$\max_{\overline{\Delta}} |\alpha^\star g| = \max_{\overline{\Delta}} |G| \ge C_1/|\sigma|. \tag{7.8}$$

(7.6), (7.7), and (7.8) together yield

$$\|f|V \cap P_\sigma\| \ge \frac{C_1}{C_2 C_3 |\sigma|}, \quad \sigma \in \Delta\backslash\{0\}.$$

As this also holds when $\|f|V \cap P_\sigma\| = \infty$, (7.3) has been proved and, with it, Lemma 7.1. $\square$

Proof of Theorem 1.2. We prove by induction on m. Again, the case $m = 0$ is obvious. Assume the statement for $m-1$, and consider the m-dimensional theorem. To apply the induction hypothesis we have to understand whether the restrictions $E|P$ to hyperplanes $P \subset \mathbb{C}^m$ satisfy the hypothesis of the theorem. On the one hand, if rk $E < \infty$ then of course rk $E|P < \infty$. On the other hand, if $\bigcup_j \mathcal{E}(h_j)$ is locally finitely generated, then the question becomes whether the sheaf $\mathcal{F} \to P$ whose stalk at $x \in P$ is

$$\left\{\boldsymbol{\varphi}_x \colon \varphi \in \Gamma(V \cap P, E), \int_{V \cap P} h_j(\varphi) < \infty \text{ for some } j \text{ and open } V \subset U,\ x \in V\right\} \tag{7.9}$$

is also locally finitely generated. This will be true for almost all P. For, at the price of shrinking U, we can assume there are $g^1, \ldots, g^p \in \Gamma(E)$ that generate each stalk of $\bigcup_j \mathcal{E}(h_j)$. We claim that whenever P is such that for some j

$$\int_{U \cap P} h_j(g^i) < \infty, \qquad i = 1, \ldots, p,$$

the sheaf $\mathcal{F}$ is generated by $g^i|P$. Indeed, let $\boldsymbol{\varphi}_x \in \mathcal{F}_x$ be the germ of a $\varphi \in \Gamma(V \cap P, E)$ as in (7.9). Assuming, as we may, that V is pseudoconvex, Theorem 6.1 can be applied as in the proof of Lemma 7.1 to extend φ to a $g \in \Gamma(V, E)$ such that $\int_V h_j(g) < \infty$. Hence $g_x \in \bigcup_j E(h_j, x) = M_x$. But since the latter module is generated by the g^i, it follows that $\boldsymbol{\varphi}_x = \mathbf{g}_x|P$ is in the module generated by $\mathbf{g}^i_x|P$.

We need to show $M_x = E(h, x)$ for arbitrary x, which we will take to be 0. As above, we assume that $g^1, \dots, g^p \in \Gamma(E)$ generate each stalk of $\mathcal{E}$. Fix a relatively compact neighborhood $V_0 \subset U$ of 0 and an $f \in \Gamma(V_0, E)$ such that $\int_{V_0} h(f) < \infty$. We will show that $\mathbf{f}_0 \in M_0$ using the characterization in Lemma 7.1. Take a neighborhood V of 0, relatively compact in V_0, and a hyperplane P_0. Again we assume $P_0 = \{t \in \mathbb{C}^m \colon z_1 = 0\}$ and for $s \in \mathbb{C}$ write $P_s = \{z \in \mathbb{C}^m \colon z_1 = s\}$. Fubini's theorem guarantees that there are a j_0 and a set $S \subset \mathbb{C}$ of full measure such that

$$\int_{V_0 \cap P_s} h(f) < \infty \quad \text{and} \quad \int_{V_0 \cap P_s} h_{j_0}(g^i) < \infty \text{ for } s \in S,\ i = 1, \dots, p,$$

and

$$\liminf_{S \ni s \to 0} |s|^2 \int_{V_0 \cap P_s} h(f) = 0. \tag{7.10}$$

The induction hypothesis implies that for each $s \in S$ there is a j such that $\int_{V \cap P_s} h_j(f) < \infty$, whence

$$\|f|V \cap P_s\|^2 = \int_{V \cap P_s} h(f), \qquad \text{cf. (7.1).}$$

Hence (7.10) implies (7.2) (with $x = 0$), and so by Lemma 7.1, $\mathbf{f}_0 \in M_0$ indeed. $\square$

Here is an example that shows that in Theorem 1.2 the condition of finite generation cannot be simply dropped. Let $U \subset \mathbb{C}$ be the unit disc and E the trivial bundle $U \times l^2 \to U$. We endow E with metrics that are determined by a sequence $\sigma = (\sigma_1, \sigma_2, \dots)$ of nonnegative numbers,

$$h(z, w) = h^\sigma(z, w) = \sum_\nu |w_\nu|^2 / |z|^{2\sigma_\nu}, \quad z \in U,\ w = (w_\nu) \in l^2.$$

We only consider σ_ν such that $\sup_\nu \sigma_\nu < 2$. Let $V \subset U$ be a neighborhood of 0. If $f \in \Gamma(V, E)$ given by $f(z) = (z, f_1(z), f_2(z), \dots)$ is in $L^2_{\text{loc}}(V, h)$, then $f_\nu(0) = 0$ whenever $\sigma_\nu \geq 1$. Assuming V is the disc $\{z \in \mathbb{C} \colon |z| < \rho\}$, with $\rho \in (0, 1]$, the $L^2(V, h)$ norm of such f is given in terms of the Taylor coefficients of $f_\nu(z) = \sum_k a_{\nu k} z^k$:

$$\int_V |f|_h^2 = \pi \sum_{\nu, k} |a_{\nu k}|^2 \, \frac{\rho^{2(k - \sigma_\nu + 1)}}{k - \sigma_\nu + 1}.$$

Let $\sigma_\nu = 1 - 1/\nu$. Then $f_\nu(z) \equiv 1/\nu^2$ defines a section f of E whose germ $\mathbf{f}_0 \in E(h^\sigma)$. However, for no $j \in \mathbb{N}$ is this germ in $E(h^{(1+1/j)\sigma})$, and so

$$\bigcup_{j=1}^{\infty} (E(h^{(1+1/j)\sigma}, 0) \subsetneq E(h^\sigma, 0).$$

At the same time, any holomorphic section in $L^2(V, h^\sigma)$ can be approximated by holomorphic sections in $\bigcup_j L^2(V, h^{(1+1/j)\sigma})$. To end this paper, we propose the following float. (According to the late Lee Rubel, a float, much like a conjecture, is a mathematical statement one would like to see proved; but while to make a conjecture the conjecturer should have substantial evidence in its favor, a float is allowed once it occurs to the floater that the statement might be true.)

Consider a holomorphic Hilbert bundle $E \to U$ over a pseudoconvex open $U \subset \mathbb{C}^m$ and let $h_1 \geq h_2 \geq \cdots$ be Hermitian metrics on E. Assume that the Nakano curvature of each h_j dominates 0 and that $h = \lim h_j$ dominates a continuous Hermitian metric. Let $f \in \Gamma(E) \cap L^2(U, h)$ and $V \subset U$ be a sublevel set of a plurisubharmonic exhaustion function. Then in the Hilbert space $L^2(V, h)$ the section $f|V$ can be approximated by holomorphic sections of $E|V$ that are in $\bigcup_j L^2(V, h_j)$.

References

[B1] B. Berndtsson, Weighted estimates for the $\overline{\partial}$-equation in complex analysis and geometry (Columbus, OH, 1999) 43–57, Ohio State Univ. Math. Res. Inst. Publ., 9, de Gruyter, Berlin, 2001.

[B2] B. Berndtsson, Curvature of vector bundles associated to holomorphic fibrations. Ann. of Math. (2) 169 (2009) 531–560.

[B3] B. Berndtsson, The openness conjecture for plurisubharmonic functions. arxiv: 1305.5781.

[BC] B. Berndtsson, Ph. Charpentier, A Sobolev mapping property of the Bergman kernel. Math. Z. 235 (2000) 1–10.

[BP] B. Berndtsson, M. Paŭn, Bergman kernels and the pseudoeffectivity of relative canonical bundles. Duke Math. J. 145 (2008) 341–377.

[Bl] Z. Błocki, On the Ohsawa–Takegoshi extension theorem. Univ. Iagellonicae Acta Math. L (2012) 53–61.

[C] B. Y. Chen, A simple proof of the Ohsawa–Takegoshi extension theorem. arxiv: 1105.2430.

[dC] M. A. de Cataldo, Singular Hermitian metrics on vector bundles. J. Reine Angew. Math. 502 (1998) 93–122.

[D1] J. P. Demailly, Complex analytic and differential geometry, http://www-fourier.ujf-grenoble.fr/~demailly/manuscripts/agbook.pdf

[D2] J. P. Demailly, Multiplier ideal sheaves and analytic methods in algebraic geometry. School on Vanishing Theorems and Effective Results in Algebraic Geometry. 1–145, ICTP lecture notes, 6, Trieste, 2001.

[DK] J. P. Demailly, J. Kollár, Semicontinuity of complex singularity exponents and Kähler–Einstein metrics on Fano orbifolds. Ann. Sci. École Norm. Sup. 34 (2001) 525–556.

[FJ] C. Favre, M. Jonsson, Valuations and multiplier ideals. J. Amer. Math. Soc. 18 (2005) 655–684.

[G] R. Godement, Sur la théorie des représentations unitaires. Ann. of Math. (2) 53 (1951) 68–124.

[GZ1] Q. Guan, X. Zhou, A solution of an L^2 extension problem with optimal estimate and applications. Ann. of Math. (2) 181 (2015) 35–59.

[GZ2] Q. Guan, X. Zhou, A proof of Demailly's strong openness conjecture. Ann. of Math. (2) 182 (2015) 605–616.

[GZ3] Q. Guan, X. Zhou, Strong openness conjecture and related problems for plurisubharmonic functions. arxiv:1401.7158.

[[GZ4] Q. Guan, X. Zhou, Effectiveness of Demailly's strong openness conjecture and related problems. Invent. Math. 202 (2015) 635–676.

[H] L. Hörmander, L^2 estimates and existence theorems for the $\overline{\partial}$ operator. Acta Math. 113 (1965) 89–152.

[Hi] P. H. Hiep, The weighted log canonical threshold. C.R. Math. Acad. Sci. Paris 351 (2014) 283–288.

[L1] L. Lempert, The Dolbeault complex in infinite dimensions. I. J. Amer. Math. Soc. 11 (1998) 485–520.

[L2] L. Lempert, Coherent sheaves and cohesive sheaves. Complex analysis, 227–244, Trends. Math., Birkhäuser, Basel 2010.

[L3] L. Lempert, A maximum principle for Hermitian (and other) metrics. Proc. Amer. Math. Soc. 143 (2015) 2193–2200.

[LT] M. Lejeune-Jalabert, B. Teissier, Clôture intégrale des idéaux et équisingularités. Séminaire à l'École Polytechnique 1974; Ann. Fac. Sci. Toulouse, XVII (2008) 781–859.

[M] J. Mujica, Complex analysis in Banach spaces. North-Holland, Amsterdam, 1986.

[OT] T. Ohsawa, K. Takegoshi, On the extension of L^2-holomorphic functions. Math. Z. 195 (1987) 197–204.

[R] H. Raufi, Singular Hermitian metrics on holomorphic vector bundles. Ark. Mat. 53 (2015) 359–382.

[S] H. Skoda, Morphismes surjectifs de fibrés vectoriels semi-positifs, Ann. Sci. École Norm. Sup. (4) 11 (1978) 577-611.

László Lempert
Department of Mathematics
Purdue University
West Lafayette, IN 47907-2067, USA
e-mail: lempert@purdue.edu

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 335–348

An Effective Uniform Artin–Rees Lemma

Johannes Lundqvist

Abstract. We prove a global uniform Artin–Rees lemma type theorem for sections of ample line bundles over smooth projective varieties. This result is used to prove an Artin–Rees lemma for the polynomial ring with uniform degree bounds. The proof is based on multidimensional residue calculus.

Mathematics Subject Classification (2010). 14Q20, 32A27.

Keywords. Artin–Rees, uniform, residue currents, polynomial.

1. Introduction

Assume that (X, x) is a germ of a reduced analytic variety. Let M be a finitely generated module over the local ring, $\mathcal{O}_{X,x}$, of germs of holomorphic functions at x. In [Szn12] it was proved by residue calculus that if N is a submodule of M, then there exists a constant μ such that the inclusion

$$I^{\mu+r}M \cap N \subset I^r N \tag{1}$$

holds for all ideals I of $\mathcal{O}_{X,x}$ and all non-negative integers r. This is the well-known uniform Artin–Rees lemma that was proved by Huneke in [Hun92] for much more general rings.

The uniform Artin–Rees lemma is related to the theorem of Briançon–Skoda, [BS74]. Since there are global versions of the latter, see [EL99] and [Hic01] for smooth X and [AW15] for singular X, it is reasonable to believe that there is a global version of the inclusion (1). In this paper we prove such a result when X is smooth.

Theorem 1.1. *Assume that X is a smooth projective variety of dimension n and that L is an ample line bundle over X. Assume moreover that $f^1, \dots, f^m$ are global holomorphic sections of L. Then there exist constants μ and s_0 such that for every set of global holomorphic sections $g^1, \dots, g^\ell$ of any ample line bundle M over X the following is true: If ϕ is a global section of*

$$M^{\otimes s} \otimes K_X \otimes L^{\otimes s_0}, \quad s \geq n + r, \quad r \geq 1,$$

such that $\phi \in \mathcal{J}(f)$ *and* $|\phi| \leq C|g|^{\mu+r-1}$ *for some* $C > 0$, *then*

$$\phi = \sum_{\substack{j=1,\dots,m \\ I_1+\cdots+I_\ell=r}} \alpha_{I,j}(g^1)^{I_1}\dots(g^\ell)^{I_\ell}f^j,$$

where $\alpha_{I,j}$ *are global sections of* $M^{\otimes(s-r)} \otimes K_X \otimes L^{\otimes(s_0-1)}$.

Here and throughout this paper $|g|$ is short for $|g^1| + \cdots + |g^\ell|$.

Remark 1.2. The constants μ and s_0 both depend on X and L. However, the point is that these constants are uniform in M and r. In general μ depends on Hironakas desingularization theorem but it can be related to exponents from Bernstein–Sato type formulas, cf. Remark 3.3.

Remark 1.3. We may replace the canonical bundle K_X in Theorem 1.1 by any bundle T such that $T \otimes K_X^{-1}$ is non-negative. This follows from the proof in Section 3.

By the theorem of Briançon–Skoda,

$$|\phi| \leq C|g|^{\mu+r+n-2}$$

implies that $\phi \in \mathcal{J}(g)^{\mu+r-1}$, and this certainly implies that $|\phi| \leq C'|g|^{\mu+r-1}$. Since μ is not specified in general we might as well use such an estimate instead of the membership condition. We choose to use the inequality in this paper for purely technical reasons. Also, we actually get a special case of the theorem of Briançon–Skoda from Theorem 1.1 with this setting.

If we assume that $M = L$ and $r = 1$ we get the following result.

Corollary 1.4. *Assume that* $f^1, \dots, f^m$ *and* L *are as in Theorem* 1.1. *Then there exist constants* μ *and* s_0 *such that for every set of global holomorphic sections* $g^1, \dots, g^\ell$ *of* L *the following holds: If* ϕ *is a global section of* $K_X \otimes L^{\otimes s_0}$, *that satisfies* $\phi \in \mathcal{J}(f)$ *and* $|\phi| \leq C|g|^\mu$, *then*

$$\phi = \sum_{ij} \alpha_{ij} g^i f^j, \tag{2}$$

where α_{ij} *are global sections of* $K_X \otimes L^{\otimes(s_0-2)}$.

Remark 1.5. If $\mathcal{J}(f) = \mathcal{J}(1)$, then it follows from the proof in Section 3 that we may take μ in Corollary 1.4 as $\min(n, \ell)$ and we get back a theorem of Briançon–Skoda type, cf. part (ii) of Corollary 2.2 in [EL99] and Theorem 7.1, and its proof, in [AW15]. That is, assume that X and L are as in Theorem 1.1 and $g^1, \dots, g^\ell$ are global holomorphic sections of L. Then if ϕ is a global section of

$$K_X \otimes L^{\otimes s}, \quad s \geq n+1,$$

such that $|\phi| \leq C|g|^{\min(n,\ell)}$, we may write

$$\phi = \sum_j \alpha_j g^j,$$

where α_j are global sections of $K_X \otimes L^{\otimes(s-1)}$.

Based on Theorem 1.1 and a geometric inequality in [EL99] we prove a theorem about polynomials, which can be regarded as an effective uniform Artin–Rees lemma for the polynomial ring.

Theorem 1.6. *Let $V \subset \mathbb{C}^N$ be an algebraic variety of dimension n and assume that X, the closure of V in $\mathbb{P}^N$, is smooth. Given polynomials $F_1, \dots, F_m$ on V there exists a constant μ such that the following holds: Assume that $G_1, \dots, G_\ell$ are polynomials on V of degree at most d, r is a positive integer, and Φ is a polynomial such that*

$$|\Phi| \leq C|G|^{\mu+r-1} \tag{3}$$

and

$$\Phi \in J(F_1, \dots, F_m).$$

Then there exist polynomials $P_{I,j}$ such that

$$\Phi = \sum_{\substack{j=1,\dots,m \\ I_1+\cdots+I_\ell=r}} P_{I,j} G_1^{I_1} \dots G_\ell^{I_\ell} F_j,$$

and

$$\begin{aligned} &\deg(P_{I,j} G_1^{I_1} \dots G_\ell^{I_\ell} F_j) \\ &\quad \leq \max\Big((\mu+r-1)d^{c_\infty^G} \deg X + \deg \Phi, (n+r)d + \kappa_1, \deg \Phi + \kappa_2\Big), \end{aligned} \tag{4}$$

where the constants κ_1 and κ_2 only depend on $J(F)$ and V.

Here $J(F)$ is the polynomial ideal generated by $F_1, \dots, F_m$. The constant c_∞^G is defined in Section 4; it is less than or equal to n. From this result we also derive a similar but weaker result in the case when X is singular, see Section 5.

If $X = \mathbb{P}^n$, $\ell = 1$, and $G_1 = 1$, then (4) becomes $\deg \Phi + \kappa$ for some κ. It is well known that in general κ is double exponential in the degree of the F_j:s, [May82], and it was proved already in [Her26] that one can choose κ as something like $2(2d')^{2^N-1}$, where $d' \geq \deg F_j$. This shows that the third entry in (4) is not only there for technical reasons. The same is true for the other entries as well. Assume for example that $r = 1$ and that the zero set of $J(G)$ does not intersect the hyperplane at infinity. In this case $c_\infty^G = -\infty$. However, if we let d tend to infinity it must be the case that the degree of $P_{i,j} G_i F_j$ tends to infinity linearly, so the second entry is necessary. Now, consider the case when $r = 1$, $J(F) = J(1)$, and assume that the zero set of $J(G)$ is empty. Then it was proved by Kollár, [Kol88], Sombra, [Som99], and Jelonek, [Jel05], that in general the degree of $P_{i,j} G_i F_j$ cannot be chosen less than $d^{\min(\ell,n)}$, so we need something like the first entry.

In special cases one can explicitly calculate the degree estimates and get back classical theorems of Macaulay and Max Noether. This is discussed in the end of Section 4.

2. Andersson–Wulcan currents and the diamond product

In this section we describe a residue current, introduced in [AW07], associated to a generically exact Hermitian complex of vector bundles and also an operation on such complexes introduced in [Szn12].

Assume that E_j are Hermitian vector bundles over an n-dimensional smooth variety X in $\mathbb{P}^N$ and that the complex

$$\cdots \xrightarrow{f_2} E_2 \xrightarrow{f_2} E_1 \xrightarrow{f_1} E_0 \tag{5}$$

is generically exact, i.e., pointwise exact outside some proper analytic subvariety, $\mathcal{Z}$, of X. Let $E = \bigoplus E_k$. Then there is a natural superstructure, i.e., a $\mathbb{Z}_2$-grading, on E, see [AW07]. From now on and throughout this paper we assume that E is equipped with that superstructure. Consider the sheaves, $\mathcal{E}^{p,q}(E)$, of smooth (p,q)-forms on X with values in E and the space, $\mathcal{D}'(E)$, of currents with values in E. The operator

$$\nabla_E = \sum f_j - \bar{\partial}$$

acts on $\mathcal{E}^{p,q}(E)$ and is naturally extended to $\mathcal{D}'(E)$ and the superstructure on E makes sure that $\nabla_E^2 = 0$, see [AW07].

If σ_k is the minimal inverse to f_k on $X \setminus \mathcal{Z}$, i.e.,

$$\sigma_k \xi = \begin{cases} \eta, \text{ where } f_k \eta = \xi \text{ and } \eta \text{ has minimal norm, if } \xi \in \operatorname{Im} f_k, \\ 0, \text{ if } \xi \in (\operatorname{Im} f_k)^{\perp}, \end{cases}$$

then the $\operatorname{Hom}(E_0, E)$-valued form

$$u := \sigma_1 + \sigma_2 \bar{\partial} \sigma_1 + \sigma_3 \bar{\partial} \sigma_2 \bar{\partial} \sigma_1 + \cdots$$

satisfies

$$\nabla_E u = 1_{E_0},$$

see [AW07]. Note that the component

$$u_k := \sigma_k \bar{\partial} \sigma_{k-1} \cdots \bar{\partial} \sigma_1$$

of u that takes values in $\operatorname{Hom}(E_0, E_k)$ has bidegree $(0, k-1)$. The form u can be extended across $\mathcal{Z}$ to a current U by letting

$$U := \lim_{\epsilon \to 0} \chi(|h|^2/\epsilon^2) u, \tag{6}$$

where $h_1, \dots, h_M$ are functions with $\mathcal{Z}$ as their common zero set. Here $\chi(t)$ is a smooth function on the reals that is 0 for $t < 1$ and 1 for $t > 2$. The existence of the limit (6) is nontrivial and requires the desingularization theorem of Hironaka.

We now define the residue current

$$R := 1_{E_0} - \nabla_E U. \tag{7}$$

It obviously has support on $\mathcal{Z}$. The current R is also a so-called pseudomeromorphic current as defined in [AW10]. We may restrict such currents to subvarieties in the following way. If T is a pseudomeromorphic current on X and V is a subvariety of X then the restriction of T to the complement of V has a natural extension to

X, denoted $1_{V^c}T$. The difference between the current T and that extension is a current with support on V denoted $1_V T$. That is,

$$T = 1_V T + 1_{V^c} T. \tag{8}$$

For details, see [AW10].

Remark 2.1. In this paper we have chosen to use the regularization $\chi(|h|^2/\epsilon^2)u$ when defining the current U. One might as well, which indeed often is done in the literature, use the analytic continuation of $|h|^{2\lambda}u$ to $\lambda = 0$ and get the same current, see, e.g., [BS10]. The reason that we use the cut-off function approach in this paper is simply because it is suitable in the proof of Theorem 1.1.

Remark 2.2. In the case when (5) is the Kozul complex the coefficients in the resulting current R in (7) is exactly the ones first introduced by Passare, Tsikh, and Yger generalizing the Coleff–Herrera current, see Theorem 1.1 in [PTY00].

The sheaf complex

$$\cdots \xrightarrow{f_2} \mathscr{O}(E_2) \xrightarrow{f_2} \mathscr{O}(E_1) \xrightarrow{f_1} \mathscr{O}(E_0), \tag{9}$$

associated to the complex (5), plays a key role in the following basic result, [AW07].

Theorem 2.3. *Assume that X is smooth and that E_0 in the complex* (5) *has rank one. Let $\mathcal{J}$ be the ideal sheaf* $\mathrm{Im}(f_1)$ *of the associated sheaf complex. If ϕ is a holomorphic section of E_0, then $\phi \in \mathcal{J}$ if $R\phi = 0$, and the converse is true if the associated sheaf complex is exact.*

Notice that even if the complex (5) is infinite the residue only takes values in $\mathrm{Hom}(E_0, E_0 \oplus \cdots \oplus E_{\dim(X)+1})$. This follows from the construction of u since the component u_k has bidegree $(0, k-1)$.

We would like to use Theorem 2.3 to draw the conclusion that a given section belongs to a certain product ideal. In order to do so we need an appropriate complex like (5) such that $\mathrm{Im}(f_1)$ lies in the product ideal in question. We use a construction due to [Szn12] and we give here the definition and basic properties.

Definition 2.4. Given r Hermitian complexes $E^1_\bullet, \ldots, E^r_\bullet$, with morphisms $f^k_j : E^k_j \to E^k_{j-1}$, the diamond product, denoted $E^1_\bullet \Diamond \cdots \Diamond E^r_\bullet$, is the complex $H_\bullet$, where

$$H_0 = E^1_0 \otimes \cdots \otimes E^r_0, \quad H_k = \bigoplus_{\substack{\alpha_1 + \cdots + \alpha_r \\ = k-1}} E^1_{1+\alpha_1} \otimes \cdots \otimes E^r_{1+\alpha_r},$$

and where the maps $h_j : H_j \to H_{j-1}$ are defined as

$$h_1 = f^r_1 f^{r-1}_1 \ldots f^1_1, \quad h_k = \sum_{1 \le s \le r, j \ge 2} f^s_j \big|_{H_k}.$$

Note that it follows directly from the definition that

$$E^1_\bullet \Diamond E^2_\bullet \Diamond E^3_\bullet = (E^1_\bullet \Diamond E^2_\bullet) \Diamond E^3_\bullet = E^1_\bullet \Diamond (E^2_\bullet \Diamond E^3_\bullet). \tag{10}$$

If r is odd, then $H_\bullet$ inherits its superstructure from the superstructures of the individual factors. However, if r is even, then one needs to do a trick by multiplying with the trivial complex

$$0 \to E \to E \to 0,$$

for any bundle E. For details, see [Szn12].

Let u^k be the $\mathrm{Hom}(E_0^k, E^k)$-valued form associated to the complex $E_\bullet^k$. It was shown in [Szn12] that the form

$$u^H := u^1 \otimes \cdots \otimes u^r \tag{11}$$

satisfies the equality

$$\nabla_H u = 1_{H_0}.$$

From u^H we define the currents U^H and R^H as in (6) and (7). One can describe the residue current R^H in terms of the individual building block complexes. Assume that $H_\bullet$ is the diamond complex of $M_\bullet$ and $L_\bullet$ and assume that U^L, R^L, U^M and R^M are the currents associated to $L_\bullet$ and $M_\bullet$. Assume also that $L_\bullet$ is exact outside an analytic set defined by a tuple, h_1, of analytic functions and let h_2 be a tuple that defines the corresponding set for $M_\bullet$. Then

$$R^H = R^M \wedge U^L - U^M \wedge R^L, \tag{12}$$

where

$$\begin{aligned} R^M \wedge U^L &= \lim_{\epsilon\to 0} \bar\partial\chi(|h_2|^2/\epsilon^2) \wedge u^M \wedge U^L \\ &= \lim_{\epsilon\to 0}\lim_{\delta\to 0} \bar\partial\chi(|h_2|^2/\epsilon^2) \wedge u^M \wedge \chi(|h_1|^2/\delta^2) u^L, \end{aligned}$$

and

$$\begin{aligned} U^M \wedge R^L &= \lim_{\epsilon\to 0} \chi(|h_2|^2/\epsilon^2) u^M \wedge R^L \\ &= \lim_{\epsilon\to 0}\lim_{\delta\to 0} \chi(|h_2|^2/\epsilon^2) u^M \wedge \bar\partial\chi(|h_1|^2/\delta^2) \wedge u^L, \end{aligned} \tag{13}$$

see Proposition 3.4 in [Szn12].

Products of more than two factors are defined in the same way. Once again, the existence of the limits is non-trivial. The order of the limits is important as we see in the one-variable principal value example

$$U = \frac{1}{z}, \quad R = \bar\partial\frac{1}{z}.$$

In this case we get

$$U \wedge R = 0, \quad R \wedge U = \bar\partial\frac{1}{z^2}.$$

Remark 2.5. There is a product of Koszul complexes in [And06] which is used to solve division problems for product ideals. That product is derived from so-called Eagon–Northcott complexes which is related to determinant ideals. However, products of Kuszul complexes can quite easily be computed directly. The diamond complex is a generalization of this product for general complexes of locally free sheaves. For products of Koszul complexes, see Section 3 in [And06].

3. The proof of Theorem 1.1

Our proof of Theorem 1.1 is based on the fact that ϕ annihilates a residue current R^H associated to the diamond product of appropriate choices of complexes.

Let X, L, M, f^j and g^j be as in Theorem 1.1. Since L is ample there exists an exact sequence like (9), with a direct sum of negative powers of L as E_k, such that $\operatorname{Im} f_1 = \mathcal{J}(f)$, see for example [Laz04]. Indeed, consider the sequence

$$\oplus\mathcal{O}(L^{-1}) \xrightarrow{f} \mathcal{O}_X \longrightarrow \mathcal{O}_X/\mathcal{J}(f) \longrightarrow 0,$$

where f is the mapping $(f^1, \ldots, f^m)$. Let F be the kernel of the surjection f. Then $F \otimes \mathcal{O}(L^{\otimes d_2})$ is generated by its global sections if d_2 is big enough by the Cartan–Serre–Grothendieck theorem. Fixing generating sections we get a surjective map $\mathcal{O}_X \to F \otimes \mathcal{O}(L^{\otimes d_2})$ and hence we have a surjection $\mathcal{O}(L^{-\otimes d_2}) \to F$. If we repeat this argument for the kernel of that map and so on we get a, possibly non-terminating, exact complex

$$\cdots \xrightarrow{f_3} \oplus\mathcal{O}(L^{-\otimes d_2}) \xrightarrow{f_2} \oplus\mathcal{O}(L^{-1}) \xrightarrow{f_1=f} \mathcal{O}_X \longrightarrow \mathcal{O}_X/J(f) \longrightarrow 0,$$

where $d_2, d_3, \ldots$ are positive integers. For a Hermitian vector bundle S_0 we get a Hermitian complex

$$\cdots \xrightarrow{f_3} S_0 \otimes (\oplus L^{-\otimes d_2}) \xrightarrow{f_2} S_0 \otimes (\oplus L^{-1}) \xrightarrow{f} S_0, \tag{14}$$

that is pointwise exact outside the zero set of $\mathcal{J}(f)$.

For $\mathcal{J}(g)$ we choose the Koszul complex, i.e., we let E^j be trivial line bundles over X with global frames e_j and set

$$E = M^{-1} \otimes E^1 \oplus \cdots \oplus M^{-1} \otimes E^l.$$

Then the Koszul complex is the Hermitian complex

$$0 \longrightarrow E_n \xrightarrow{\delta_n} \cdots \xrightarrow{\delta_2} E_1 \xrightarrow{\delta_1} E_0, \tag{15}$$

where

$$E_k = \Lambda^k E = M^{-k} \otimes \Lambda^k(E^1 \oplus \cdots \oplus E^\ell).$$

The maps $\delta_k : E_k \to E_{k-1}$ are interior multiplication with the section g of E^*, where $g = \sum g^j e_j^*$ and e_j^* is the dual frame. For details, see, e.g., Example 2.1 in [AW15].

Denote the complex (14) by $L_\bullet$ and by $M_\bullet$ the Koszul complex associated to $\mathcal{J}(g)$. For a Hermitian line bundle S let R^H be the residue current from Section 2 associated to the complex

$$H_\bullet := (S \otimes \underbrace{M_\bullet \Diamond M_\bullet \Diamond \cdots \Diamond M_\bullet}_{r \text{ times}}) \Diamond L_\bullet. \tag{16}$$

Then, according to (12) and (10), we can write

$$R^H = R^M \wedge U^L - U^M \wedge R^L,$$

where R^L, U^L, R^M and R^M are the currents associated to the complexes $L_\bullet$ and $S \otimes M_\bullet \Diamond \cdots \Diamond M_\bullet$.

The following proposition from [AW15] can be seen as a global version of the first part of Theorem 2.3.

Proposition 3.1. *Assume that* (5) *is a generically exact Hermitian complex over a smooth variety X and that ϕ is a holomorphic section of the bundle E_0. If R is the associated residue current, $R\phi = 0$, and*

$$H^{k-1}(X, \mathcal{O}(E_k)) = 0, \quad 1 \leq k \leq n+1,$$

then there is a global holomorphic section ψ of E_1 such that $f^1\psi = \phi$.

We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1. Assume that $\phi \in \mathcal{J}(f)$. Let $H_\bullet$ be the complex (16) and choose S_0 as $K_X \otimes L^{\otimes s_0}$ and S as $M^{\otimes s}$ in (14) and (16), respectively. If we can prove that $R^H\phi = 0$, then ϕ would be on the form (2) by Proposition 3.1 if all the relevant cohomology groups vanish.

We are interested in the cohomology groups of the bundles H_k in $H_\bullet$ for $1 \leq k \leq n+1$. Remember that H_k consists of a sum of tensor products of one bundle from the complex (14) and r bundles from (15) tensored by S. The possible bundles from (14) are

$$S_0 \otimes L^{-d_j}, \quad 1 \leq j \leq k,$$

and the possible bundles from (15) are

$$M^{-j} \otimes \Lambda^j(E^1 \oplus \cdots \oplus E^\ell), \quad 1 \leq j \leq k.$$

Note that the exponent of M in H_1 is $s - r$ and that the exponent decreases by at most 1 at every level in $H_\bullet$. In particular, since a tensor product of ample bundles is ample we can use Kodaira's vanishing theorem to see that the relevant cohomology groups vanish if

$$s_0 \geq \max_{1\leq j\leq n+1} d_j + 1 = d_{n+1} + 1$$

and

$$s \geq n + r.$$

Fix s_0 and s so that all the cohomology groups vanish. It then remains to show that there exists a constant μ such that ϕ annihilates the residue R^H, given that $|\phi| \leq C|g|^{\mu+r-1}$. Remember that R^H splits into the sum

$$R^M \wedge U^L - U^M \wedge R^L. \tag{17}$$

Since ϕ is assumed to belong to $\mathcal{J}(f)$ we get that $R^L\phi = 0$ by the second part of Theorem 2.3, and in view of (13) $U^M \wedge R^L\phi = 0$.

To see that the first term in (17) is annihilated we use that there exists a modification $\widetilde{X} \xrightarrow{\pi} X$ so that the pull back of U^L locally can be expressed as a finite sum of forms

$$\pi_*\left(\frac{\text{smooth}}{h}\pi^*\phi\right),$$

where h is a section to a line bundle $\widetilde{L}$ over $\widetilde{X}$ such that it locally is a monomial in some local coordinates, see [AW07]. In light of (11) we hence get that locally $R^M \wedge U^L\phi$ is the limit of the pushforward of a finite sum of terms on the form

$$\pi^*(\bar{\partial}\chi(|g|^2/\epsilon^2) \wedge (u^1 \otimes \cdots \otimes u^r)) \wedge \frac{\text{smooth}}{h}\pi^*\phi, \tag{18}$$

where every u^j is associated to $M_\bullet$. Since X is compact the divisor of h is a finite sum $\sum \tau_j D_j$ for positive integers τ_j and if $\tau = \sum \tau_j$ we get that h locally is a monomial of degree less than or equal to τ at every point in X. The arguments after expression (4.10) in the proof of Theorem 1.2 in [Szn12], which in turn is a variant of the proof of the main theorem in [ASS10], now show that $R^M \wedge U^L\phi = 0$, locally at a point x, if $\mu \geq \min(\ell, n) + \tau + 1$. Since n and τ do not depend on g or x the conclusion of the theorem follows if

$$|\phi| \leq C|g|^{\mu+r-1},$$

where $\mu \geq \min(\ell, n) + \tau + 1$. □

Remark 3.2. Note that if the f_j:s do not have any common zeros, i.e., $\mathcal{J}(f) = \mathcal{J}(1)$, then $\tau = 0$ and we may choose μ as $\min(\ell, n) + 1$. If one carefully reads the proof of Theorem 1.2 in [Szn12] one sees that $\mu = \min(\ell, n)$ does the trick in this case. We then get the result in Remark 1.5.

Remark 3.3. One can avoid the log resolution π in the proof above and follow the lines in [VY16] where the authors instead use normalized blow-ups and global Bernstein–Sato type formulas. In this way one can control the constant μ by the order of the differential operator in the Bernstein–Sato formulas, see Theorem 4.1 and its proof in [VY16] for details.

4. The proof of Theorem 1.6

Let X be a smooth projective variety of dimension n, L a nef line bundle over X and $\mathcal{J} \in \mathcal{O}_X$ an ideal sheaf. If Z_j are the distinguished subvarieties in the sense of Fulton–MacPherson of $\mathcal{J}$, see [EL99], and r_j are the coefficients associated to the Z_j:s, then

$$\sum r_j \deg_L Z_j \leq \deg_L X, \tag{19}$$

where

$$\deg Z_j = \int_{Z_j} c_1(L)^{\dim Z_j}$$

is the L-degree of Z_j. The geometric inequality (19) above is proved in [EL99, Proposition 3.1]. Note that if $L = \mathcal{O}(d)$, then

$$\deg_L X = d^n \deg X, \tag{20}$$

where $\deg X$ denotes $\deg_{\mathcal{O}(1)} X$.

If g_j is the d-homogenization of G_j, then for the ideal sheaf $\mathcal{J}(g)$ we associate a number c_∞^G defined to be the maximal codimension of the distinguished subvarieties Z_j contained in the hyperplane at infinity. If there is no distinguished subvariety at infinity we assign to c_∞^G the value $-\infty$. Using (20) and (19) we get that if $L = \mathcal{O}(d)$, then

$$r_j \leq d^{c_\infty^G} \deg X \tag{21}$$

for r_j associated to Z_j contained in the hyperplane at infinity.

Proof of Theorem 1.6. Let V, X, G_j, F_j and Φ be as in Theorem 1.6. Let d' be the maximum of the degrees of all the polynomials F_j and let f_j and g_j be the d' and d-homogenization of F_j and G_j, respectively. Let

$$\phi = z_0^{\rho - \deg \Phi} \Phi(z_0/z) z_0^{\deg \Phi} \tag{22}$$

be the ρ-homogenization of Φ. We consider f_j and g_j as sections of $\mathcal{O}(d')$ and $\mathcal{O}(d)$ restricted to X. The bundle $K_X^{-1} \otimes \mathcal{O}(k)$ is ample for k large enough, say $k \geq k_X$. By Remark 1.3 we may therefore use Theorem 1.1 on ϕ if ρ is big enough, ϕ belongs to $\mathcal{J}(f)$ even at the hyperplane at infinity, and the inequality

$$|\phi| \leq C|g|^{\mu+r-1} \tag{23}$$

is valid on the whole of X. Let us first show that ϕ belongs to $\mathcal{J}(f)$ provided that ρ is larger than some constant depending on $F_1, \ldots, F_m$ and V. If R^f is the residue associated to a locally free resolution of $\mathcal{J}(f)$, then by the second part of Theorem 2.3 we only need to prove that R^f is annihilated by ϕ. Remember that we may write

$$R^f = 1_V R^f + 1_{X \setminus V} R^f, \tag{24}$$

cf. Section 2. Since $\phi \in \mathcal{J}(f)$ on V it follows from Theorem 2.3 that ϕ annihilates $1_V R^f$. We know that $1_{X \setminus V} R^f$ has support on the hyperplane at infinity so z_0^ν annihilates $1_{X \setminus V} R^f$ if ν is large enough, say larger than ν_f. This means that if ρ in (22) is chosen so that

$$\rho \geq \deg \Phi + \nu_f, \tag{25}$$

then R^f is annihilated by ϕ and thus $\phi \in \mathcal{J}(f)$.

To make sure that (23) holds we consider the normalization

$$\widetilde{X} \xrightarrow{\pi} X,$$

of the blow-up of X along $\mathcal{J}(g)$. Let X_∞ be the part of X that intersect the hyperplane at infinity and write the exceptional divisor as $W = \sum r_j W_j$. Then, by definition, the distinguished subvarieties Z_j are the images of W_j, and hence

$$r_j \leq d^{c_\infty^G} \deg X$$

if $W_j \subseteq X_\infty$ by (21). The polynomial Φ satisfies (3) by hypothesis so we get that $\pi^*\phi$ vanishes to order $(\mu + r - 1) r_j$ on W_j if $\pi W_j \not\subseteq X_\infty$. If $\pi W_j \subseteq X_\infty$, then $\pi^*\phi$ vanishes to order $\rho - \deg \Phi$ on W_j. If we choose ρ such that

$$\rho \geq (\mu + r - 1) d^{c_\infty^G} \deg X + \deg \Phi, \tag{26}$$

we get that $\pi^*\phi$ vanishes to order $(\mu + r - 1) r_j$ on all W_j. This means that $|\pi^*\phi| \leq C|\pi^* g|^{\mu+r-1}$ on the whole of $\widetilde{X}$ and hence (23) holds.

If also

$$\rho \geq d(n+r) + (d' + k_X)s_0, \tag{27}$$

where s_0 is the same as the one in Theorem 1.1 we may apply that theorem on ϕ with

$$M = \mathcal{O}(d)\big|_X, \quad L = \mathcal{O}(d' + k_X)\big|_X.$$

To sum up, we may use Theorem 1.1 if ρ satisfies the inequalities (25), (26), and (27). The only thing left is that we need to make sure that the sections $\alpha_{I,j}$ that we get after applying Theorem 1.1 have extensions to global sections of $\mathcal{O}(\rho)$. However, that is true if ρ is larger than an absolute number η depending on X. The theorem follows with $\kappa_1 = (d' + k_X)s_0$ and $\kappa_2 = \nu_f + \eta$. □

If $V = \mathbb{C}^n$ and hence $X = \mathbb{P}^n$ so that $\deg X = 1$ and moreover $J(F) = J(1)$ and $r = 1$, then it follows from the proof of Theorem 1.1 and Theorem 1.6 that $\kappa_2 = \kappa_1 = 0$. However, one can actually take $\kappa_1 = -n$. To see this we just modify the proof of Theorem 1.1 slightly. Instead of taking $S = \mathcal{O}(sd)$ we could take $S = \mathcal{O}(s)$. In this case we get that s should be so large so that the cohomology groups $H^j(\mathbb{P}^n, \mathcal{O}(s - d(n+1)))$ vanishes. From Kodaira's vanishing theorem we see that $s \geq d(n+1) - n$ does the trick. Together with Remark 3.2 we get the following effective version of the Briançon–Skoda theorem.

Theorem 4.1. *For every set of polynomials $G_1, \dots, G_\ell$ on $\mathbb{C}^n$ with degree less than or equal to d the following holds: If Φ is a polynomial such that $|\Phi| \leq C|G|^{\min(\ell,n)}$, then there exist polynomials P_j such that*

$$\Phi = P_1 G_1 + \cdots + P_\ell G_\ell,$$

and the degree of $P_j G_j$ is at most

$$\max\left(\min(\ell, n) d^{c_\infty^G} + \deg \Phi, (n+1)d - n\right).$$

The theorem above was already proved in [AG11]. Note that if we also assume that the common zero set is empty we almost get back the optimal degree estimate, $d^{\min(\ell,n)}$, of Kollár and Jelonek, mentioned in Section 1. If we also assume that $G_1, \dots, G_\ell$ have no common zeros at infinity we do get back the classical theorem of Macaulay, [Mac16]. That is, we may write

$$1 = \sum P_j G_j,$$

where the degree of $P_j G_j$ is at most $(n+1)d - n$.

If we assume that $\deg G_j = 0$, the common zero set of $F_1, \dots, F_m$ is a discrete set, $m = n$, and that there are no zeros at the hyperplane at infinity, then we get back the theorem of Max Noether, i.e., we may write

$$\Phi = \sum P_j F_j,$$

where the degree of $P_j F_j$ is at most $\deg \Phi$, [Max73]. To see this we first note that $c_\infty^G = -\infty$ and that $\kappa_2 = 0$. This means that $\deg P_j F_j \leq \max(\deg \Phi, \kappa_1)$. From the proof of Theorem 1.6 we know that κ_1 is a multiple of s_0 from Theorem 1.1.

In this case this means that κ_1 is a number so that $H^{k-1}(\mathbb{P}^n, \mathcal{O}(\kappa_1 - d_k d')) = 0$, where d_k are the numbers in the proof of Theorem 1.1 and d' is the maximum degree of the F_j:s. Since $\mathcal{J}(f)$ is a complete intersection we may use the Koszul complex as the exact sequence that defines the residue associated with $\mathcal{J}(f)$. In particular, it has length n which means that we may choose κ_1 as 0.

5. The non-smooth case

One can deduce an Artin–Rees type theorem even for singular varieties from the smooth case.

Theorem 5.1. *Let $V \subset \mathbb{C}^N$ be a singular reduced algebraic variety of dimension n and let $F_1, \ldots, F_m$ be polynomials on V. Then there exist constants μ and ν such that the following holds: Assume that $G_1, \ldots, G_\ell$ are polynomials of degree at most d and that Φ is a polynomial such that*

$$|\Phi| \leq |G|^{\mu+\nu} \tag{28}$$

and

$$\Phi \in (F_1, \ldots, F_m)$$

on V. Then there exist polynomials $A_{i,j}$ such that

$$\Phi = \sum A_{i,j} G_i F_j$$

on V and

$$\deg(A_{j,\ell} G_j F_\ell) \leq \deg \Phi + (\nu + \mu) d^n \deg X + \mu d^N + O(d).$$

The degree estimate in this result is of type $O(d^N)$ and not as expected of type $O(d^n)$. It is probably true that there is an estimate of type $O(d^n)$ but we cannot prove any such result at this time.

Proof. Let $F_1 \ldots, F_m$ be polynomials on $V \subset \mathbb{C}^N$, let X be the closure of V in $\mathbb{P}^N$, and let $H_1, \ldots, H_t$ cut out V, i.e., $J_V = (H_1, \ldots, H_t)$.

First, Theorem 1.6 implies that there exists a constant μ such that for every set of polynomials $G_1, \ldots, G_\ell$ in $\mathbb{C}^N$ and every polynomial $\widehat{\Phi}$ in $\mathbb{C}^N$ we have that

$$\begin{aligned} &|\widehat{\Phi}| \leq |G|^\mu, \quad \widehat{\Phi} \in (F_1, \ldots, F_m, H_1, \ldots, H_t) \\ &\Longrightarrow \widehat{\Phi} = \sum A_{j,\ell} G_j F_\ell + \sum B_{j,\ell} G_j H_\ell, \end{aligned} \tag{29}$$

where $A_{j,\ell}, B_{j,\ell}$ are polynomials and

$$\deg A_{j,\ell} G_j F_\ell \leq \deg \widehat{\Phi} + \mu d^N + O(d).$$

Second, there is a Briançon–Skoda–Huneke constant ν on V, see [AW15, Theorem 6.4], such that if Φ and $G_1, \ldots, G_\ell$ are as in Theorem 5.1 and (28) holds on V, then

$$\Phi = \sum_{|I|=\mu} a_I G^I$$

on V with

$$\deg a_I G^I \leq \deg \Phi + (\nu + \mu)d^n \deg X + O(d).$$

Consider

$$\widehat{\Phi} = \sum_{|I|=\mu} a_I G^I$$

as a polynomial in $\mathbb{C}^N$. Then clearly $|\widehat{\Phi}| \leq |G|^\mu$ in $\mathbb{C}^N$ and moreover, $\widehat{\Phi} = \Phi$ on V which means that $\widehat{\Phi} \in (F_1, \ldots, F_m, H_1, \ldots, H_t)$. Therefore, by Theorem 1.6 as above we get that

$$\widehat{\Phi} = \sum A_{i,j} G_i F_j + \sum B_{i,j} G_i H_j,$$

with

$$\deg(A_{i,j} G_i F_j) \leq \deg \widehat{\Phi} + \mu d^N + O(d).$$

This means that

$$\Phi = \sum A_{ij} G_i F_j$$

on V with

$$\deg(A_{i,j} G_i F_j) \leq \deg \Phi + (\nu + \mu)d^n \deg X + \mu d^N + O(d).$$

Note that the linear term $O(d)$ is independent of Φ and the polynomials $G_1, \ldots, G_\ell$. $\square$

Acknowledgment

The author would like to thank Mats Andersson and Elizabeth Wulcan for valuable discussions and comments throughout the writing process of this paper. It should also be said that the referee had valuable comments that helped improve this paper.

References

[And06] Andersson M., *Explicit versions of the Briançon–Skoda theorem with variations*, Michigan Math. J. **54** (2006), 361–373.

[AG11] Andersson M., Götmark E., *Explicit representation of membership in polynomial ideals*, Math. Ann., **349** (2011), 345–365.

[ASS10] Andersson M., Samuelsson H., Sznajdman J., *On the Briançon–Skoda theorem on a singular variety*, Ann. Inst. Fourier **60** (2010), no. 2, 417–432.

[AW07] Andersson M., Wulcan E., *Residue currents with prescribed annihilator ideals*, Ann. Scient. Éc. Norm. Sup., (4), **40** (2007), no. 6, 985–1007.

[AW10] Andersson M., Wulcan E., *Decomposition of residue currents*, J. reine angew. math. **638** (2010), 103–118.

[AW15] Andersson M., Wulcan E., *Global effective versions of the Briançon–Skoda–Huneke theorem*, Invent. Math. **200** (2015), 607–651.

[BS10] Björk J.-E., Samuelsson H., *Regularizations of residue currents*, J. reine angew. Math. **649** (2010), 33–54.

[BS74] Briançon J., Skoda H., *Sur la clôture intégrale d'un idéal de germes de fonctions holomorphes en un point de* $\mathbb{C}^n$, C. R. Acad. Sci. Paris Sér. A. **278** (1974), 949–951.

[EL99] Ein L., Lazarsfeld R., *A geometric effective Nullstellensatz*, Invent. Math. **135** (1999), 427–448.

[Her26] Hermann G., *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann., **95** (1926), 736–788.

[Hic01] Hickel M., *Solution d'une conjecture de C. Berenstein–A. Yger et invariants de contact à l'infini*, Ann. Inst. Fourier **51** (2001), 707–744.

[Hun92] Huneke C., *Uniform bounds in Noetherian rings*, Invent. Math. **107** (1992), 203–223.

[Jel05] Jelonek Z., *On the Effective Nullstellensatz*, Invent. Math. **162** (2005), 1–17.

[Kol88] Kollár J., *Sharp effective Nullstellensatz*, J. American Math. Soc. **1** (1988), 963–975.

[Laz04] Lazarsfeld R., *Positivity in algebraic geometry I*, Springer-Verlag 2004.

[Mac16] Macaulay F. S., *The algebraic system of modular systems*, Cambridge Univ. Press 1916.

[May82] Mayer A., Mayr E., *The complexity of the word problem for commutative semigroups and polynomial ideals*, Adv. in math. **46** (1982), 305–329.

[Max73] Nöther M., *Über einen Satz aus der Theorie der algebraischen Funktionen*, Math. Ann. **6** (1873), 351–359.

[PTY00] Passare M., Tsikh A., Yger A., *Residue currents of Bochner–Martinelli type*, Publ. Mat. **44** (2000), 85–117.

[Som99] Sombra M., *A sparse effective Nullstellensatz*, Adv. in Appl. Math. **22** (1999), 271–295.

[Szn12] Sznajdman J., *A residue calculus approach to the uniform Artin–Rees lemma*, Israel J. Math. **196** (2013), 33–50.

[VY16] Vidras A., Yger A., *Briançon–Skoda theorem for a quotient ring*, Complex analysis and dynamical systems VI, Contemporary Math., Amer, Math. Soc. **667** (2016), 253–278.

Johannes Lundqvist
Department of Mathematics
Stockholm University
SE-106 91 Stockholm, Sweden
e-mail: lundqvist.johannes@gmail.com

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 349–359

Amoebas of Half-dimensional Varieties

Grigory Mikhalkin

To the memory of Mikael Passare

Abstract. An n-dimensional algebraic variety in $(\mathbb{C}^\times)^{2n}$ covers its amoeba as well as its coamoeba generically finite-to-one. We provide an upper bound for the volume of these amoebas as well as for the number of points in the inverse images under the amoeba and coamoeba maps.

1. Introduction

1.1. Definitions

Consider an n-dimensional algebraic variety $V \subset (\mathbb{C}^\times)^{2n}$.

Definition 1.1 (Gelfand–Kapranov–Zelevinsky [4]). The *amoeba* $\mathcal{A}$ of V is the image

$$\mathcal{A} = \operatorname{Log}(V) \subset \mathbb{R}^{2n}$$

of V under the coordinatewise logarithm map $\operatorname{Log} : (\mathbb{C}^\times)^{2n} \to \mathbb{R}^{2n}$,

$$\operatorname{Log}(z_1, \ldots, z_{2n}) = (\log|z_1|, \ldots, \log|z_{2n}|).$$

The restriction $\operatorname{Log}|_V$ is called the *amoeba map* for V.

Definition 1.2 (cf. Passare [15]). The *coamoeba* (or *alga*, cf. [3]) $\mathcal{B}$ of V is the image

$$\mathcal{B} = \operatorname{Arg}(V) \subset (S^1)^{2n}$$

of V under the coordinatewise argument map $\operatorname{Arg} : (\mathbb{C}^\times)^{2n} \to (\mathbb{R}/2\pi\mathbb{Z})^{2n} \approx (S^1)^{2n}$,

$$\operatorname{Arg}(z_1, \ldots, z_{2n}) = (\arg(z_1), \ldots, \arg(z_{2n})).$$

The restriction $\operatorname{Arg}|_V$ is called the *coamoeba map* for V.

Research is supported in part by the grant TROPGEO of the European Research Council, by the grants 140666 and 141329 of the Swiss National Science Foundation, and by the NCCR SwissMAP of the Swiss National Science Foundation.

For coamoebas it is often more convenient to use argument taken mod π instead of mod 2π, (cf. [12]). Namely, we denote $T_\pi = \mathbb{R}/\pi\mathbb{Z}$ and for $z \in \mathbb{C}^\times$ we define

$$\arg_\pi(z) = (\arg(z) \mathrm{mod}\ \pi) \in \mathbb{T}_\pi.$$

In other words, $\arg_\pi$ is the composition of arg and the double covering $\mathbb{R}/2\pi\mathbb{Z} \to T_\pi$. Then we define $\mathrm{Arg}_\pi : (\mathbb{C}^\times)^{2n} \to (\mathbb{R}/\pi\mathbb{Z})^{2n} = T_\pi^{2n}$, and

$$\mathrm{Arg}_\pi(z_1, \dots, z_{2n}) = (\arg_\pi(z_1), \dots, \arg_\pi(z_{2n})).$$

We call $\mathcal{B}_\pi = \mathrm{Arg}_\pi(V) \subset T_\pi^{2n}$ the *rolled coamoeba* of V.

We consider two antiholomorphic involutions

$$\mathrm{conj}, \mathrm{conj}' : (\mathbb{C}^\times)^{2n} \to (\mathbb{C}^\times)^{2n}$$

defined by

$$\mathrm{conj}(z_1, \dots, z_{2n}) = (\bar{z}_1, \dots, \bar{z}_{2n}),\ \mathrm{conj}'(z_1, \dots, z_{2n}) = \left(\frac{1}{\bar{z}_1}, \dots, \frac{1}{\bar{z}_{2n}}\right). \tag{1.1}$$

To each n-dimensional variety $V \subset (\mathbb{C}^\times)^{2n}$ we associate two integer numbers. Note that if $A, B \subset (\mathbb{C}^\times)^{2n}$ are two complex subvarieties of complimentary dimensions then for an open dense subset of $\epsilon \in (\mathbb{C}^\times)^{2n}$ all intersection points from $A \cap \epsilon B$ are transverse and their number does not depend on ϵ (as long as it is generic). Here ϵB stands for the coordinatewise multiplication of B by ϵ in $(\mathbb{C}^\times)^{2n}$ (in other words for the multiplicative translation). We define the toric intersection number $A.B \in \mathbb{Z}_{\geq 0}$ to be the number of points in $\#(A \cap \epsilon B)$ for a generic ϵ (times the corresponding multiplicities in the case when the corresponding components of A or B are not simple, i.e., if A or B are not reduced). Clearly, $A.B = B.A$.

Definition 1.3. We define the conj-*degree*

$$\alpha(V) = V. \mathrm{conj}(V) \in \mathbb{Z}_{\geq 0}$$

and the conj$'$-*degree*

$$\beta(V) = V. \mathrm{conj}'(V) \in \mathbb{Z}_{\geq 0}.$$

Definition 1.4. Let A and B be two smooth (differentiable) manifolds of the same dimension and $f : A \to B$ be a smooth map. We say that f covers its image at most m times if for any point $p \in B$ which is regular for f the inverse image $f^{-1}(p)$ consists of at most m points.

More generally, if A is a (not necessarily smooth) real or complex algebraic variety (such as $V \subset (\mathbb{C}^\times)^{2n}$ in the case when it is singular), it admits a stratification into smooth manifolds. Consider a map $f : A \to B$ whose restriction $f|_\Sigma$ to every stratum $\Sigma \subset A$ is smooth. Similarly, we say that f covers its image at most m times if for any point $p \in B$ the inverse image $f^{-1}(p)$ consists of at most m points unless p is a critical point for $f|_\Sigma$, where $\Sigma \subset A$ is a stratum of our stratification.

1.2. Statement of the results

The main results of this paper are contained in the following theorem.

Theorem 1. *Let* $V \subset (\mathbb{C}^\times)^{2n}$ *be an algebraic* n*-dimensional variety. Then the amoeba* $\mathcal{A}(V)$ *is covered by the map* $\operatorname{Log}|_V$ *at most* $\beta(V)$ *times, while the rolled coamoeba* $\mathcal{B}_\pi(V)$ *is covered by the map* $\operatorname{Arg}_\pi|_V$ *at most* $\alpha(V)$ *times. Furthermore,*

$$\operatorname{Vol}(\mathcal{A}) \le \frac{\pi^{2n}}{2}\alpha(V).$$

Note that the conventional (i.e., non-rolled) coamoeba $\mathcal{B}(V)$ cannot be covered more than the rolled coamoeba. Thus it is also covered by the coamoeba map at most $\alpha(V)$ times.

If V is a complete intersection then we can easily compute the conj-degree α as well as the conj$'$-degree β by means of the Bernstein–Kouchnirenko calculus as follows.

Definition 1.5. We say that

$$V = \bigcap_{j=1}^{n} V_j \subset (\mathbb{C}^\times)^{2n}$$

is a toric *complete intersection* of hypersurfaces $V_1, \dots, V_n \in (\mathbb{C}^\times)^{2n}$ if

$$V = \lim_{\epsilon_j \to 0} \bigcap_{j=1}^{n} \epsilon_j V_j.$$

Here the limit is taken in the sense of Hausdorff metric on the subsets of $(\mathbb{C}^\times)^{2n}$ (with a group invariant metric) and ϵ_j. In particular, we require this limit to exist.

Proposition 2. *Suppose that* $V = \bigcap_{j=1}^{n} V_j \subset (\mathbb{C}^\times)^{2n}$ *is a complete intersection of hypersurfaces* V_j *with Newton polyhedra* $\Delta_j \subset \mathbb{R}^{2n}$*,* $j = 1, \dots, n$*. Then we have*

$$\alpha(V) = \operatorname{Vol}(\Delta_1, \dots, \Delta_n, \Delta_1, \dots, \Delta_n)$$

and

$$\beta(V) = \operatorname{Vol}(-\Delta_1, \dots, -\Delta_n, \Delta_1, \dots, \Delta_n).$$

Here Vol *stands for the mixed volume of* $2n$ *polyhedra in* $\mathbb{R}^{2n}$*.*

Remark 1.6. Proposition 2 and Theorem 1 produce upper bounds for the volumes of amoebas in the case of toric complete intersections in terms of the mixed volumes of the corresponding Newton polyhedra. Such bounds were conjectured in the talk by Mounir Nisse on the memorial conference for Mikael Passare in Summer 2013. Finiteness of $\operatorname{Vol}(\mathcal{A})$ was observed in [9].

Proof. Note that $\operatorname{conj}(V)$ is a also a toric complete intersection defined by the polynomials with the same Newton polyhedra, but conjugate coefficients while $\operatorname{conj}'(V)$ is a toric complete intersection defined by the polynomials with $-\Delta_j$

as their Newton polyhedra as we need to make a substitution $z_j \mapsto \frac{1}{z_j}$ before conjugation. The proposition now follows from the Bezout theorem in the form of Bernstein–Kouchnirenko [1], [6]. $\square$

In particular, if V is a toric complete intersection with

$$\Delta_1 = \cdots = \Delta_n = \left\{ (x_1, \dots, x_{2n}) \in \mathbb{R}^{2n}_{\geq 0} \mid \sum_{j=1}^{2n} x_j \leq 1 \right\} \tag{1.2}$$

then $\alpha(V) = 1$ and $\beta(V) = \frac{(2n)!}{(n!)^2}$ so Theorem 1 has the following corollary.

Corollary 3. *If $V = \bar{V} \cap (\mathbb{C}^\times)^{2n}$, and $\bar{V}$ is a complete intersection of hypersurfaces of degrees $d_1, \dots, d_n$ in $\mathbb{CP}^{2n}$ then $\mathcal{A}(V)$ is covered by the amoeba map at most $\frac{(2n)!}{(n!)^2} \prod_{j=1}^{n} d_j^2$ while $\mathcal{B}_\pi(V)$ (as well as $\mathcal{B}(V)$ itself) is covered at most $\prod_{j=1}^{n} d_j^2$ times by the coamoeba map. Furthermore,*

$$\mathrm{Vol}(\mathcal{A}) \leq \frac{\pi^{2n}}{2} \prod_{j=1}^{n} d_j^2.$$

2. Proof of the theorem

2.1. Bounds for the number of inverse images for the amoeba and coamoeba maps

In this section we prove the first part of Theorem 1 establishing bounds for the number of inverse images of $\mathrm{Log}\,|_V$ and $\mathrm{Arg}_\pi\,|_V$.

Note that $(\mathrm{Arg}_\pi\,|_V)^{-1}(0) = V \cap (\mathbb{R}^\times)^{2n} \subset (\mathbb{C}^\times)^{2n}$. We may compare this with $(\mathrm{Arg}\,|_V)^{-1}(0) = V \cap (\mathbb{R}_{>0})^{2n}$ in the case of the conventional (not rolled) coamoeba map. Similarly, for $p \in T^{2n}_\pi = (\mathbb{R}/\pi\mathbb{Z})^{2n}$ we have

$$(\mathrm{Arg}_\pi\,|_V)^{-1}(p) = V \cap e^{ip}(\mathbb{R}^\times)^{2n} \subset (\mathbb{C}^\times)^{2n}$$

where $e^{ip} \in (\mathbb{C}^\times)^{2n}$ is obtained by coordinatewise exponentiating of ip. If p is a regular value of $\mathrm{Arg}_\pi\,|_V$ then V intersects $e^{ip}(\mathbb{R}^\times)^{2n}$ transversally. This means that every stratum in a stratification of V into smooth manifolds intersects $e^{ip}(\mathbb{R}^\times)^{2n}$ transversally. By the dimension considerations, the top-dimensional stratum intersects $e^{ip}(\mathbb{R}^\times)^{2n}$ in finitely many points while smaller-dimensional strata are disjoint from $e^{ip}(\mathbb{R}^\times)^{2n}$.

Furthermore, we have the inclusion

$$(\mathrm{Arg}_\pi\,|_V)^{-1}(p) = V \cap e^{ip}(\mathbb{R}^\times)^{2n} \subset V \cap e^{2ip}\,\mathrm{conj}(V) \tag{2.1}$$

as $e^{ip}(\mathbb{R}^\times)^{2n} \subset (\mathbb{C}^\times)^{2n}$ is the invariant locus for the complex conjugation

$$(z_1, \dots, z_{2n}) \mapsto e^{2ip}(z_1, \dots, z_{2n})$$

in $(\mathbb{C}^\times)^{2n}$. Thus the cardinality of $(\mathrm{Arg}_\pi\,|_V)^{-1}(p)$ for regular p is bounded by $\alpha(V)$ as stated in Theorem 1.

Similarly, for a regular value $q \in \mathbb{R}^{2n}$ of $\operatorname{Log}|_V$ we have a finite number of points in $(\operatorname{Log}|_V)^{-1}(q)$ as well as the inclusion

$$(\operatorname{Log}|_V)^{-1}(q) = V \cap e^q S \subset V \cap e^{2q} \operatorname{conj}'(V), \tag{2.2}$$

where $S \subset (\mathbb{C}^\times)^{2n}$ is the unit torus (the fixed point locus of conj$'$). Thus the cardinality of $(\operatorname{Log}|_V)^{-1}(q)$ for regular p is bounded by $\beta(V)$ as stated in Theorem 1.

2.2. Estimating the volume of amoeba

To finish the proof of Theorem 1 we consider the real-valued $2n$-form on $(\mathbb{C}^\times)^{2n}$

$$\omega = \prod_{j=1}^{2n} dx_j - \prod_{j=1}^{2n} dy_j. \tag{2.3}$$

Here product stands for the exterior product of differential forms $dx_j = \Re dz_j$, $dy_j = \operatorname{Im} dz_j$..

Lemma 4. *We have $\omega|_V \equiv 0$.*

Proof. We may write

$$\omega = \frac{1}{2^{2n}}\left(\prod_{j=1}^{2n} (dz_j + d\bar{z}_j) - (-1)^n \prod_{j=1}^{2n} (dz_j - d\bar{z}_j)\right). \tag{2.4}$$

The right-hand side of this expression is the sum of monomials of degree $2n$ in dz_j and $d\bar{z}_k$. Note that if n is odd then there are no monomials with odd number of $d\bar{z}_j$. Similarly, if n is even then there are no monomials with even number of $d\bar{z}_j$. Thus the right-hand side of (2.4) contains only monomials where either the number of dz_j is more than n or the number of $d\bar{z}_k$ is more than n. Thus, ω must vanish everywhere on a holomorphic n-variety V. □

We may consider the cardinality $\#(\operatorname{Log}|_V)^{-1}$ of the inverse image of the amoeba map as a measurable function on $\mathbb{R}^{2n}$ (since the critical locus of $\operatorname{Log}|_V$ is nowhere dense). Then

$$\operatorname{MultiVol}(\mathcal{A}) = \int_{\mathbb{R}^{2n}} \#((\operatorname{Log}|_V)^{-1}(x_1, \dots, x_{2n}))dx_1 \dots dx_{2n}$$

can be thought of as the volume of $\mathcal{A}$ taken with the multiplicities corresponding to the covering by the amoeba map. Similarly,

$$\operatorname{MultiVol}(\mathcal{B}_\pi) = \int_{T_\pi^{2n}} \#((Arg_\pi|_V)^{-1}(y_1, \dots, y_{2n}))dy_1 \dots dy_{2n}$$

can be thought of as the volume of $\mathcal{B}_\pi$ taken with the multiplicities corresponding to the covering by the coamoeba map.

Corollary 5. $\operatorname{MultiVol}(\mathcal{A}) = \operatorname{MultiVol}(\mathcal{B}_\pi)$.

Proof. Let $V_+ \subset V$ be the open subset of V where the real $2n$-form $dx_1 \wedge \cdots \wedge dx_{2n}$ is non-degenerate and defines the orientation that agrees with the complex orientation of V. Let $V_- \subset V$ be the open set where these orientations disagree. Note that by Lemma 4 the form $dy_1 \wedge \cdots \wedge dy_{2n}$ also agrees with the complex orientation on V_+ and disagrees on V_-. We have

$$\operatorname{MultiVol}(\mathcal{A}) = \int\limits_{V_+} dx_1 \wedge \cdots \wedge dx_{2n} - \int\limits_{V_-} dx_1 \wedge \cdots \wedge dx_{2n}$$

while

$$\operatorname{MultiVol}(\mathcal{B}_\pi) = \int\limits_{V_+} dy_1 \wedge \cdots \wedge dy_{2n} - \int\limits_{V_-} dy_1 \wedge \cdots \wedge dy_{2n}.$$

The two multivolumes are equal by Lemma 4. □

Lemma 6.

$$\operatorname{MultiVol}(\mathcal{A}) = \operatorname{MultiVol}(\mathcal{B}_\pi) \leq \alpha(V)\pi^{2n}.$$

Proof. By (2.1) the cardinality of $(\operatorname{Arg}_\pi|_V)^{-1}(p)$ is not greater than $\alpha(V)$ almost everywhere on T^{2n}_π while $\operatorname{Vol}(T^{2n}_\pi) = \pi^{2n}$. □

Note that $\operatorname{Log} : (\mathbb{C}^\times)^{2n} \to \mathbb{R}^{2n}$ is a proper map (inverse images of compact sets are compact) and thus $\operatorname{Log}|_V : V \to \mathbb{R}^{2n}$ is also proper. Since V and $\mathbb{R}^{2n}$ are oriented manifolds of the same dimension the map $\operatorname{Log}|_V$ has a well-defined degree. Recall that this degree is equal to the number of inverse images of a generic point $q \in \mathbb{R}^{2n}$ taken with the sign ± 1 depending whether $\operatorname{Log}|_V$ locally preserves the orientation.

Corollary 7. *The degree of the amoeba map is zero. We have*

$$\operatorname{MultiVol}(\mathcal{A}) = 2\int\limits_{V_+} dx_1 \wedge \cdots \wedge dx_{2n} = -2\int\limits_{V_-} dx_1 \wedge \cdots \wedge dx_{2n}.$$

Furthermore, $\operatorname{Vol}(\mathcal{A}) \leq \frac{1}{2}\operatorname{MultiVol}(\mathcal{A})$.

Proof. As the multivolume of V is bounded by Lemma 6 we have $\mathbb{R}^{2n} \smallsetminus \mathcal{A} \neq \emptyset$ and thus the degree of $\operatorname{Log}|_V$ must be zero. Each generic point $q \in \mathcal{A}$ is covered by V_+ and V_- the same number of times. □

Remark 2.1. Note that Corollary 7 immediately implies that $\beta(V)$ is always even as it coincides with the degree of the amoeba map $\operatorname{Log}|_V$ (this fact is also easy to deduce from symmetry reasons). However, as the maps $\operatorname{Arg}|_V$ and $\operatorname{Arg}_\pi|_V$ are not proper, we cannot apply the same reasoning. Note, in particular, that the parity of $(\operatorname{Arg}|_V)^{-1}(p)$ is different for different generic points of $(\mathbb{R}/2\pi\mathbb{Z})^{2n}$ already in the case when $V \subset \mathbb{CP}^2$ is a generic line (cf., e.g., [11]).

In the same time, (2.1) implies the parity of $(\operatorname{Arg}_\pi|_V)^{-1}(p)$ coincides with that of $\alpha(V)$ for generic points $p \in T^{2n}_\pi$ as non-real intersection points of $e^{-ip}V$

and $e^{ip}\,\mathrm{conj}(V)$ come in pairs. Thus the rolled coamoeba map has a well-defined degree mod 2 determined by $\alpha(V)$.

Corollary 7 implies that

$$\mathrm{Vol}(\mathcal{A}) = \mathrm{Vol}(\mathrm{Log}(V_+)) \le \frac{1}{2}\,\mathrm{MultiVol}(\mathcal{B}_\pi) \le \frac{\pi^{2n}}{2}\alpha(V).$$

This finishes the proof of Theorem 1.

3. Some remarks and open problems

3.1. Example: linear spaces in $\mathbb{CP}^{2n}$

Let $L \subset \mathbb{CP}^{2n}$ be an n-dimensional linear subspace that is generic with respect to the coordinate hyperplanes of $\mathbb{CP}^{2n}$. Then $V = L \cap (\mathbb{C}^\times)^{2n}$ can be presented as a complete intersection of hyperplanes with the Newton polyhedra given by (1.2). By Proposition 2 we have $\alpha(V) = 1$. By Remark 2.1 the set $(\mathrm{Arg}_\pi|_V)^{-1}(p)$ must consist of a single point for almost all values $p \in T_\pi^{2n}$, so the inequality of Lemma 6 turns into equality. We get the following proposition.

Proposition 8 (cf. [5], [14]). *If* $V = \bigcap_{j=1}^{n} H_j \subset (\mathbb{C}^\times)^{2n}$ *is a transverse intersection of n hyperplanes*

$$H_j = \left\{ (z_1, \dots, z_{2n}) \mid a_{j0} + \sum_{k=1}^{2n} a_{jk} z_k = 0 \right\}$$

with $\prod_{j=1}^{n} \prod_{k=0}^{2n} a_{jk} \neq 0$ *then*

$$\mathrm{MultiVol}(\mathcal{A}) = \mathrm{MultiVol}(\mathcal{B}_\pi) = \mathrm{Vol}(\mathcal{B}_\pi) = \pi^{2n}.$$

In the case of $n = 1$ we have $\beta(V) = 2$. By Corollary 7 we have

$$\mathrm{Vol}(\mathcal{A}) = \frac{1}{2}\,\mathrm{MultiVol}(\mathcal{A}) = \frac{\pi^2}{2}$$

in this case as in [17]. This equality was used by Passare [16] to give a new proof of Euler's formula $\zeta(2) = \frac{\pi^2}{6}$. In the case $n > 1$ we have $\beta(V) > 2$, so Proposition 8 only implies the inequalities

$$\frac{\pi^{2n}(n!)^2}{(2n!)^2} \le \mathrm{Vol}(\mathcal{A}) \le \frac{\pi^{2n}}{2}, \tag{3.1}$$

and $\mathrm{Vol}(\mathcal{A})$ might vary with V. Note that our linear subspace $V \subset (\mathbb{C}^\times)^{2n}$ varies in a (n^2-n)-dimensional family if we identify subspaces that can be obtained from each other by multiplication by $\epsilon \in (\mathbb{C}^\times)^{2n}$ (such multiplication corresponds to a translation of amoeba and thus does not change its shape or its volume).

Problem 3.1. What are the maximal and minimal possible values of $\mathrm{Vol}(\mathcal{A})$? It would be interesting to solve this problem already for $n = 2$.

3.2. MultiHarnack varieties in $\mathbb{RP}^{2n}$

Definition 3.2. We say that an n-dimensional variety $V \subset (\mathbb{C}^\times)^{2n}$ is *multiHarnack* if

$$\mathrm{MultiVol}(\mathcal{A}) = \pi^{2n}\alpha(V).$$

Let us recall the notion of simple Harnack curves in $(\mathbb{C}^\times)^2$ (introduced in [10]). According to the maximal volume characterization given in [13] a curve $V \subset (\mathbb{C}^\times)^2$ can be presented as $V = \epsilon C$ for a simple Harnack curve $C \subset (\mathbb{C}^\times)^2$ and a multiplicative vector $\epsilon \in (\mathbb{C}^\times)^2$ if and only if we have $\mathrm{Vol}(\mathcal{A}) = \frac{\pi^2}{2}\alpha(V)$.

This class of curves was generalized to a larger class of curves in $(\mathbb{C}^\times)^2$, (also called multiHarnack curves) by Lionel Lang ([8], [7]). Definition 3.2 gives the multiHarnack curves in the case $n = 1$.

According to Proposition 8 all generic linear spaces in $\mathbb{CP}^{2n}$ are multiHarnack.

Problem 3.3. Do there exist multiHarnack varieties of higher degree?

Note that once $n > 1$ being multiHarnack no longer implies being real even after multiplication by $\epsilon \in (\mathbb{C}^\times)^{2n}$ already for linear spaces.

Remark 3.4. It might be instructive to compare Definition 3.2 against another attempt to generalize the definition of simple Harnack curves from [10] to higher dimensions. The survey [11] gave a definition of torically maximal hypersurfaces of dimension n generalizing the Definition from [10]. However, it was recently shown (see [2]) that all torically maximal hypersurfaces in $\mathbb{RP}^{n+1}$ for $n > 1$ have degree 1.

3.3. Foliation of $\partial\mathcal{A}$

In this subsection we suppose for simplicity that $V \subset (\mathbb{C}^\times)^{2n}$ is smooth (otherwise we may restrict ourselves to the smooth part of V). Let us look at the critical locus $C \subset V$ of the map $\mathrm{Log}|_V$ and its image $D = \mathrm{Log}(C) \subset \mathcal{A} \subset \mathbb{R}^{2n}$ (also called the *discriminant locus*). We have the following generalization of Lemma 3 from [10].

Proposition 9. *The set C consists of the points $z \in V$ where V and $z(\mathbb{R}^\times)^{2n}$ are tangent.*

As usual, $z(\mathbb{R}^\times)^{2n}$ stands for the coordinatewise multiplication of $(\mathbb{R}^\times)^{2n}$ by $z \in (\mathbb{C}^\times)^{2n}$.

Proof. We have $z \in C$ iff there are vectors in T_zV tangent to the argument torus $\mathrm{Log}^{-1}(q)$, where $q = \mathrm{Log}(z)$. Any such vector multiplied by i gives a vector tangent both to V and $z(\mathbb{R}^\times)^{2n}$, and vice versa. $\square$

Definition 3.5. Let $z \in C$. Denote

$$F(z) = T_z(V) \cap T_z(z(\mathbb{R}^\times)^{2n}) \subset T_z((\mathbb{C}^\times)^{2n}).$$

It is a real vector subspace of the tangent space $T_z((\mathbb{C}^\times)^{2n})$. The *rank* of $z \in C$ is $\dim_{\mathbb{R}} F(p)$.

We denote with $C_r \subset C$ the locus of critical points of rank at least r. The following proposition follows immediately from the injectivity of $d\,\mathrm{Log}$ on the tangent space to $z(\mathbb{R}^\times)^{2n}$.

Proposition 10. *The subspace*

$$(d\,\mathrm{Log})(F(z)) \subset T_q(\mathbb{R}^{2n})$$

has dimension r for $z \in C_r$, $q = \mathrm{Log}(z)$.

Thus we get a preferred r-dimensional subspace in the tangent space of $\mathrm{Log}(z)$ for each $z \in C_r$.

Let us choose a stratification of the discriminant locus D to k-dimensional (non-closed) subvarieties Σ_k,

$$D = \bigcup_{k=0}^{2n-1} \Sigma_k.$$

By a k-dimensional multidistribution on an open set U of a manifold we mean specifying a finite set of k-dimensional subspaces of T_qU for every $q \in U$ so that they depend on q smoothly.

Lemma 11. *For a generic point q of Σ_{2n-k} the set $\mathrm{Log}^{-1}(q) \cap C_k$ is finite and disjoint from C_{k+1}. Furthermore, for each point $z \in \mathrm{Log}^{-1}(q) \cap C_k$ the k-dimensional space $(d\,\mathrm{Log})(F(z))$ is tangent to Σ_{2n-k}. We have a k-dimensional multidistribution (perhaps empty) of an open dense subset of Σ_{2n-k}.*

Proof. Since $\dim \Sigma_k = k$ and the rank of $d(\mathrm{Log}\,|_{C_{k+1}})$ is at most $2n-k-1$, the image $\mathrm{Log}(C_{k+1})$ is nowhere dense in Σ_k. Also the critical values of $\mathrm{Log}\,|_{C_k \cap \mathrm{Log}^{-1}(\Sigma_k)}$ (treated as a map from any of its smooth stratum to the open manifold Σ_{2n-k}) are nowhere dense in Σ_k. If $(d\,\mathrm{Log})(F(z))$ is not tangent to $\Sigma_k \ni \mathrm{Log}(F(z))$ at a regular point of $\mathrm{Log}\,|_{C_k}$ then the rank of $d(\mathrm{Log}\,|_V)$ is at least $2n-k+1$ which contradicts to the definition of C_k. □

Suppose that $\mathcal{A} \subset \mathbb{R}^{2n}$ is non-degenerate, i.e., the interior of $\mathcal{A}$ is nonempty. (This condition is equivalent to the condition $C \neq V$, i.e., to the condition that $\mathrm{Log}\,|_V$ has a regular point.) Then the amoeba boundary $\partial\mathcal{A}$ is a $(2n-1)$-dimensional subset of D. Let us denote with $\partial_1\mathcal{A}$ the subset of $\partial\mathcal{A}$ formed by points q such that $\mathrm{Log}^{-1}(q) \cap C$ consists of a single point.

Corollary 12. *We have a 1-dimensional non-empty multifoliation on an open dense set in $\partial\mathcal{A}$. This is a genuine 1-dimensional foliation on an open dense set in $\partial_1\mathcal{A}$.*

Remark 3.6. Since V is n-dimensional (over $\mathbb{C}$) and $z(\mathbb{R}^\times)^{2n}$ is $2n$-dimensional (over $\mathbb{R}$) and totally real, the maximal dimension of $F(z)$ is n.

Suppose that $V = \operatorname{conj} V$, i.e., V is defined over $\mathbb{R}$. Then we have $\mathbb{R}V = V \cap (\mathbb{R}^\times)^{2n} \subset C_n$.

3.4. Dimensions greater than a half

Suppose that V is a k-dimensional algebraic variety in $(\mathbb{C}^\times)^n$ with $k > \frac{n}{2}$. Then the generic fibers of $\operatorname{Log}|_V$ and $\operatorname{Arg}_\pi|_V$ are $(2k-n)$-dimensional varieties as V is $2k$-dimensional (over $\mathbb{R}$) and $\mathbb{R}^n$ is n-dimensional. We may still present generic fibers of $\operatorname{Arg}_\pi|_V$ and $\operatorname{Log}|_V$ as real algebraic varieties in a way similar to the half-dimensional case (where those fibers were points). For $p \in T^n_\pi$ and $q \in \mathbb{R}^n$ we consider the antiholomorphic involutions $\operatorname{conj}_p, \operatorname{conj}'_q : (\mathbb{C}^\times)^n \to (\mathbb{C}^\times)^n$ defined by

$$\operatorname{conj}_p(z) = e^{ip}(\operatorname{conj}(e^{-ip}z)),\ \operatorname{conj}'_q(z) = e^{q}(\operatorname{conj}'(e^{-q}z)), \tag{3.2}$$

where conj and conj′ are defined as in (1.1):

$$\operatorname{conj}(z_1,\dots,z_n) = (\bar z_1,\dots,\bar z_n),\ \operatorname{conj}'(z_1,\dots,z_n) = \left(\frac{1}{\bar z_1},\dots,\frac{1}{\bar z_n}\right).$$

Note that the fixed point set of conj_p is $\operatorname{Arg}_\pi^{-1}(p) = e^{ip}(\mathbb{R}^\times)^n$ while the fixed point set of conj'_q is $\operatorname{Log}^{-1}(q)$. The following proposition is straightforward.

Proposition 13. *The antiholomorphic involutions* conj_p, conj'_q *act on algebraic varieties* $V \cap \operatorname{conj}_p(V)$ *and* $V \cap \operatorname{conj}'_q(V)$ *so that the fixed point sets are* $(\operatorname{Arg}_\pi|_V)^{-1}(p)$ *and* $(\operatorname{Log}|_V)^{-1}(q)$.

Thus we may think of the fibers of $\operatorname{Arg}_\pi|_V$ *and* $\operatorname{Log}|_V$ *as real algebraic varieties whose complexification is* $V \cap \operatorname{conj}_p(V)$ *and* $V \cap \operatorname{conj}'_q(V)$. *For regular fibers these varieties are non-singular* $(2k-n)$*-dimensional varieties near their real points.*

Example 3.7. Consider the plane

$$V = \{(x,y,z) \in (\mathbb{C}^\times)^3 \mid 1 + x + y + z = 0\} \tag{3.3}$$

in $(\mathbb{C}^\times)^3$. Both V and $\operatorname{conj}_p(V)$ are planes, so fibers of $\operatorname{Arg}_\pi|_V$ are intersections of two real planes in $(\mathbb{R}^\times)^3$ after a multiplicative translation by e^{ip}. For generic fibers these two planes are transversal, so their intersection is a line. For special fibers these planes might be parallel planes, or two copies of the same plane. These special cases correspond to empty or two-dimensional fibers of $\operatorname{Arg}\pi|_V$.

The surface $\operatorname{conj}'_q(V)$ is the image of a plane under the Cremona transformation $x \mapsto \frac{1}{x}$, $y \mapsto \frac{1}{y}$, $z \mapsto \frac{1}{z}$. For a generic q the intersection of V and $\operatorname{conj}'_q(V)$ is a smooth elliptic curve. Its real locus may be empty, or consist of one or two circles. All three cases are realized as generic fibers of $\operatorname{Log}|_V$ for V given by (3.3).

References

[1] D. N. Bernstein. The number of roots of a system of equations. *Funkcional. Anal. i Priložen.*, 9(3):1–4, 1975.

[2] Erwan Brugallé, Grigory Mikhalkin, Jean-Jacques Risler, and Kristin Shaw. Non-existence of torically maximal hypersurfaces. arXiv:1506.02813.

[3] Bo Feng, Yang-Hui He, Kristian D. Kennaway, and Cumrun Vafa. Dimer models from mirror symmetry and quivering amoebae. *Adv. Theor. Math. Phys.*, 12(3):489–545, 2008.

[4] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants.* Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2008. Reprint of the 1994 edition.

[5] Peter Johansson. Coamoebas, 2010. Licentiate Thesis, Stockholm University.

[6] A. G. Kouchnirenko. Newton polyhedra and Bezout's theorem. *Funkcional. Anal. i Priložen.*, 10(3, 82–83.), 1976.

[7] Lionel Lang. A generalization of simple Harnack curves. arXiv:1504.07256.

[8] Lionel Lang. Geometry of tropical curves and application to real algebraic geometry. PhD Thesis, Geneva 2014.

[9] Farid Madani and Mounir Nisse. On the volume of complex amoebas. *Proceedings of the AMS*, 141:1113–1123, 2013.

[10] G. Mikhalkin. Real algebraic curves, the moment map and amoebas. *Ann. of Math. (2)*, 151(1):309–326, 2000.

[11] Grigory Mikhalkin. Amoebas of algebraic varieties and tropical geometry. In *Different faces of geometry*, volume 3 of *Int. Math. Ser. (N. Y.)*, pages 257–300. Kluwer/Plenum, New York, 2004.

[12] Grigory Mikhalkin and Andrei Okounkov. Geometry of planar log-fronts. *Mosc. Math. J.*, 7(3):507–531, 575, 2007.

[13] Grigory Mikhalkin and Hans Rullgård. Amoebas of maximal area. *Internat. Math. Res. Notices*, (9):441–451, 2001.

[14] Mounir Nisse and Mikael Passare. Amoebas and coamoebas of linear spaces. arXiv:1205.2808.

[15] Mikael Passare. Amoebas, convexity and the volume of integer polytopes. In *Complex analysis in several variables – Memorial Conference of Kiyoshi Oka's Centennial Birthday*, volume 42 of *Adv. Stud. Pure Math.*, pages 263–268. Math. Soc. Japan, Tokyo, 2004.

[16] Mikael Passare. How to compute $\sum 1/n^2$ by solving triangles. *Amer. Math. Monthly*, 115(8):745–752, 2008.

[17] Mikael Passare and Hans Rullgård. Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope. *Duke Math. J.*, 121(3):481–507, 2004.

Grigory Mikhalkin
Université de Genève
Mathématiques
Villa Battelle
CH-1227 Carouge, Suisse
e-mail: grigory.mikhalkin@unige.ch

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 361–368

A log Canonical Threshold Test

Alexander Rashkovskii

To the memory of Mikael Passare

Abstract. In terms of log canonical threshold, we characterize plurisubharmonic functions with logarithmic asymptotical behaviour.

1. Introduction and statement of results

Let u be a plurisubharmonic function on a neighborhood of the origin of $\mathbb{C}^n$. Its *log canonical threshold* at 0,

$$c_u = \sup\{c > 0 : e^{-c\,u} \in L^2_{\rm loc}(0)\},$$

is an important characteristic of asymptotical behavior of u at 0. The log canonical threshold $c(\mathcal{I})$ of a local ideal in $\mathcal{I} \subset \mathcal{O}_0$ can be defined as c_u for the function $u = \log|F|$, where $F = (F_1, \dots, F_p)$ with $\{F_j\}$ generators of $\mathcal{I}$. (Surprisingly, the latter notion was introduced later than its plurisubharmonic counterpart.) For general results on log canonical thresholds, including their computation and applications, we refer to [9], [17], [18].

A classical result due to Skoda [23] states that

$$c_u \ge \nu_u^{-1}, \tag{1}$$

where ν_u is the Lelong number of u at 0. A more recent result is due to Demailly [8]: if 0 is an isolated point of $u^{-1}(-\infty)$, then

$$c_u \ge F_n(u) := n\, e_n(u)^{-1/n}. \tag{2}$$

Here $e_k(u) = (dd^c u)^k \wedge (dd^c \log|z|)^{n-k}(0)$ are the Lelong numbers of the currents $(dd^c u)^k$ at 0 for $k = 1, \dots, n$, and $d = \partial + \bar\partial$, $d^c = (\partial - \bar\partial)/2\pi i$; note that $e_1(u) = \nu_u$. This was extended by Zeriahi [24] to all plurisubharmonic functions with a well-defined Monge–Ampère operator near 0.

In [20], inequality (2) was used to obtain the 'intermediate' bounds

$$c_u \ge F_k(u) := k\, e_k(u)^{-1/k}, \quad 1 \le k \le l, \tag{3}$$

l being the codimension of an analytic set A containing the unbounded locus $L(u)$ of u. None of the bounds for different values of k can be deduced from the others.

It is worth mentioning that relation (2) was proved in [8] on the base of a corresponding result for ideals[1] obtained in [6]:

$$c(\mathcal{I}) \geq n\, e(\mathcal{I})^{-1/n}, \tag{4}$$

where $e(\mathcal{I})$ is the Hilbert–Samuel multiplicity of the (zero-dimensional) ideal $\mathcal{I}$. Furthermore, it was shown in [6] that an equality in (4) holds if and only if the integral closure of $\mathcal{I}$ is a power of the maximal ideal $\mathfrak{m}_0$. Accordingly, the question of equality in (2) has been raised in [8] where it was conjectured that, similarly to the case of ideals, the extremal functions would be those with logarithmic singularity at 0.

The conjecture was proved in [21] where it was shown that

$$c_u = F_n(u) \tag{5}$$

if and only if the *greenification* g_u of u has the asymptotics $g_u(z) = e_1(u)\log|z| + O(1)$ as $z \to 0$. Here the function g_u is the upper semicontinuous regularization of the upper envelope of all negative plurisubharmonic functions v on a bounded neighborhood D of 0, such that $v \leq u + O(1)$ near 0, see [19]. Note that if $u = \log|F|$, then $g_u = u + O(1)$ [22, Prop. 5.1].

The equality situation in (1) (i.e., in (3) with $k = 1$) was first treated in [5] and [11] for the dimension $n = 2$: the functions satisfying $c_u = \nu_u^{-1}$ were proved in that case to be of the form $u = c\log|f| + v$, where f is an analytic function, regular at 0, and v is a plurisubharmonic function with zero Lelong number at 0. In [16], the result was extended to any n. This was achieved by a careful slicing technique reducing the general case to the aforementioned two-dimensional result. In addition, it used a regularization result for plurisubharmonic functions with keeping the log canonical threshold (see Lemma 1 below).

Concerning inequalities (3), it was shown in [20] that the only multi-circled plurisubharmonic functions $u(z) = u(|z_1|, \ldots, |z_n|)$ satisfying $c_u = F_l(u)$ are essentially of the form $c\max_{j\in J}\log|z_j|$ for an l-tuple $J \subset \{1, \ldots, n\}$. Here we address the question on equalities in the bounds (3) in the general case.

We present an approach that is different from that of [16] and which actually works also for the 'intermediate' equality situations. It is based on a recent result of Demailly and Pham Hoang Hiep [10]: if the complex Monge–Ampère operator $(dd^c u)^n$ is well defined near 0 and $e_1(u) > 0$, then

$$c_u \geq E_n(u) := \sum_{1\leq j\leq n} \frac{e_{j-1}(u)}{e_j(u)},$$

where $e_0(u) = 1$. In particular, this implies (2) and sharpens, for the case of functions with a well-defined Monge–Ampère operator, inequality (1). Moreover,

[1] A direct proof was given later in [2].

it is this bound that was used in [21] to prove the conjecture from [8] on functions satisfying (5).

Given $1 < l \le n$, let $\mathcal{E}_l$ be the collection of all plurisubharmonic functions u whose unbounded loci $L(u)$ have zero $2(n-l+1)$-dimensional Hausdorff measure. For such a function u, the currents $(dd^c u)^k$ are well defined for all $k \le l$ [12]. In particular, $u \in \mathcal{E}_l$ if $L(u)$ lies in an analytic variety of codimension at least l. Furthermore, we set $\mathcal{E}_1$ to be just the collection of all plurisubharmonic functions near 0.

Let $c_u(z)$ denote the log canonical threshold of u at z and, similarly, let $e_k(u,z)$ denote the Lelong number of $(dd^c u)^k$ at z; in our notation, $c_u(0) = c_u$ and $e_k(u,0) = e_k(u)$. As is known, the sets $\{z : \; c_u(z) \le c\}$ are analytic for all $c > 0$. Our first result describes, in particular, regularity of such a set for $c = c_u$, provided $c_u = F_l(u)$.

For $u \in \mathcal{E}_l$ we set

$$E_k(u) = \sum_{1\le j\le k} \frac{e_{j-1}(u)}{e_j(u)}, \quad k \le l.$$

Note that, by the arithmetic-geometric mean inequality, we have

$$E_k(u) \ge F_k(u), \quad 1 \le k \le l. \tag{6}$$

Theorem 1. *Let $u \in \mathcal{E}_l$ for some $l \ge 1$, and let $e_1(u) > 0$. Then*

(i) *$c_u \ge E_k(u)$ for all $k \le l$;*
(ii) *$c_u \ge F_k(u)$ for all $k \le l$;*
(iii) *if u satisfies $c_u = F_k(u)$ for some $k \le l$, then $k = l$ and there is a neighborhood V of the origin such that the set $A = \{z : \; c_u(z) \le c_u\}$ is an l-codimensional manifold in V. Furthermore, $A = \{z : \; e_l(u,z) \ge e_l(u)\}$.*

For $l = 1$, assertion (iii) re-proves the aforementioned result from [16]. Let $A = \{z_1 = 0\}$, then the function $u - c_u \log|z_1|$ is locally bounded from above near A and thus extends to a plurisubharmonic function v; evidently, $\nu_v = 0$. On the other hand, all the functions $u = c_u \log|z_1| + v$ with $\nu_v = 0$ satisfy $c_u = \nu_u$.

When $l > 1$, there are functions u such that $\{z : \; c_u(z) \le c_u\}$ is an l-codimensional manifold, but $c_u > F_l(u)$. Indeed, let us take $u(z_1, z_2, z_3) = \max\{\log|z_1|, 2\log|z_2|\} \in \mathcal{E}_2$. Then $A = \{z \in \mathbb{C}^n : \; c_u(z) \le c_u\} = \{z_1 = z_2 = 0\}$, while $F_2(u) = \sqrt{2} < 3/2 = c_u$. (Note that $c_u = E_2(u)$ in this case.)

Furthermore, the same example shows that the equality $(dd^c u)^2 = \delta^2 \, [z_1 = z_2 = 0]$ does not imply $u = \delta \log|(z_1, z_2)| + v$ with plurisubharmonic v and $\nu_v = 0$.

Therefore, in the higher-dimensional situation we need to deduce a more precise information on asymptotical behavior of u near A. By analogy with the case $l = n$, it is tempting to make the following conjecture.

Let $u \in \mathcal{E}_l$, then

$$c_u = F_l(u) \tag{7}$$

if and only if, for a choice of coordinates $z=(z',z'')\in\mathbb{C}^l\times\mathbb{C}^{n-l}$, *the greenification* g_u *of* u *near* 0 *satisfies*

$$g_u = e_1(u)\log|z'| + O(1) \text{ as } z\to 0.$$

The 'if' direction is obvious in view of $c_u = c_{g_u}$ [21] and the trivial fact $c_{\log|z'|}=l$, however the reverse statement might be difficult to prove even in the case $l=1$ because that would imply non-existence of a plurisubharmonic function ϕ with $e_1(\phi)=0$ and $e_n(\phi)>0$, which is a known open problem. Namely, let such a function ϕ exist, and set $u=\phi+\log|z_1|$. Then $1=\nu_u\le c_u\le c_{\log|z_1|}=1$. On the other hand, for $D=\mathbb{D}^n$, $g_u=g_\phi+\log|z_1|$ and the relation $e_n(\phi)>0$ implies $g_\phi\ne 0$ and thus $\liminf(g_u-\log|z_1|)=-\infty$ when $z\to 0$.

What we can prove is the following, slightly weaker statement.

Theorem 2. *If* $u\in\mathcal{E}_l$ *satisfies* (7)*, then* $e_k(u)=e_1(u)^k$ *for all* $k\le l$ *and, for a choice of coordinates* $z=(z',z'')\in\mathbb{C}^l\times\mathbb{C}^{n-l}$*, the function* u *satisfies* $u\le e_1(u)\log|z'|+O(1)$ *near* 0*, while the greenification* g_{u_N} *of* $u_N=\max\{u,N\log|z|\}$ *with any* $N\ge e_1(u)$ *satisfies*

$$g_{u_N}=\max\{e_1(u)\log|z'|,N\log|z''|\}+O(1),\quad z\to 0. \tag{8}$$

Let us fix a neighborhood $D\subset V$ of 0 to be the product of unit balls in $\mathbb{C}^l$ and $\mathbb{C}^{n-l}$ and consider the greenifications with respect to D. Then the functions g_{u_N} are equal to $\max\{e_1(u)\log|z'|,N\log|z''|\}$ and they converge, as $N\to\infty$, to $e_1(u)\log|z'|\ge g_u$.

Denote, for any bounded neighborhood D of 0 and any u plurisubharmonic in D,

$$\tilde g_u=\lim_{N\to\infty} g_{u_N}.$$

where $u_N=\max\{u,N\log|z|\}$. Evidently, $\tilde g_u\ge g_u$.

Theorem 3. *Let* $u\in\mathcal{E}_l$ *be such that* $\tilde g_u=g_u$*. Then it satisfies* (7) *if and only if, for a choice of coordinates* $z=(z',z'')\in\mathbb{C}^l\times\mathbb{C}^{n-l}$, $g_u=e_1(u)\log|z'|+O(1)$ *as* $z\to 0$.

In particular, this is true for $u=\alpha\log|F|+O(1)$*, where* F *is a holomorphic mapping,* $F(0)=0$*. Moreover, in this case we also have* $u=e_1(u)\log|z'|+O(1)$*.*

The statement on $\alpha\log|F|$ can be reformulated in algebraic terms as follows. Let $\mathcal{I}$ be an ideal of the local ring $\mathcal{O}_0$, and let $V(\mathcal{I})$ be its variety: $V(\mathcal{I})=\{z: f(z)=0\ \forall f\in\mathcal{I}\}$. If $\mathrm{codim}_0 V(\mathcal{I})\ge k$, then the mixed Rees multiplicity $e_k(\mathcal{I},\mathfrak{m}_0)$ of k copies of $\mathcal{I}$ and $n-k$ copies of the maximal ideal $\mathfrak{m}_0$ is well defined [4]. If $k=n$, then, as shown in [8], the Hilbert–Samuel multiplicity $e(\mathcal{I})$ of $\mathcal{I}$ equals $e_n(u)$, where, as before, $u=\log|F|$ for generators $\{F_p\}$ of $\mathcal{I}$. By the polarization formula, $e_k(\mathcal{I},\mathfrak{m}_0)=e_k(u)$ for all k; by a limit transition, this holds true for all $k\le l$ if $\mathrm{codim}_0 V(\mathcal{I})=l$.

Bounds (3) specify for this case as

$$c(\mathcal{I})\ge k\, e_k(\mathcal{I},\mathfrak{m}_0)^{-1/k},\quad 1\le k\le l.$$

From Theorems 1 and 3 we thus derive

Corollary 1. *If* $\operatorname{codim}_0 V(\mathcal{I}) = l$ *and* $c(\mathcal{I}) = k\, e_k(\mathcal{I}, \mathfrak{m}_0)^{-1/k}$ *for some* $k \le l$, *then* $k = l$, $V(\mathcal{I})$ *is an* l*-codimensional hypersurface, regular at* 0, *and there exists an ideal* $\mathfrak{n}_0$ *generated by coordinate* (*smooth transversal*) *germs* $f_1, \dots, f_l \in \mathcal{O}_0$ *such that* $\overline{\mathcal{I}} = \mathfrak{n}_0^s$ *for some* $s \in \mathbb{Z}_+$.

2. Proofs

In what follows, we will use the mentioned regularization result by Qi'an Guan and Xiangyu Zhou. Note that its proof rests on the strong openness conjecture from [9], proved in [13] and [14], see also [3] and [15].

Lemma 1 ([16, Prop. 2.1]). *Let* u *be a plurisubharmonic function near the origin,* $c_u = 1$. *Then there exists a plurisubharmonic function* $\tilde u \ge u$ *on a neighborhood of* 0 *such that* $e^{-2u} - e^{-2\tilde u}$ *is integrable on* V *and* $\tilde u$ *is locally bounded on* $V \setminus \{z : c_u(z) \le 1\}$.

We will also refer to the following uniqueness theorem.

Lemma 2 ([19, Lem. 6.3][2] and [21, Lem. 1.1]). *If* u *and* v *are two plurisubharmonic functions with isolated singularity at* 0, *such that* $u \le v + O(1)$ *near* 0 *and* $e_n(u) = e_n(v)$, *then their greenifications coincide.*

Proof of Theorem 1. Since all the functionals $u \mapsto c_u$, $E_k(u)$, $F_k(u)$ are positive homogeneous of degree -1, we can always assume $c_u = 1$.

Let $\tilde u$ be the function from Lemma 1. Its unbounded locus $L(\tilde u)$ is contained in the analytic variety $A = \{z : c_u(z) \le 1\}$. Since $A \subset L(u)$ and $u \in \mathcal{E}_l$, $\operatorname{codim} A \ge l$.

For $\tilde u$, statement (i) is proved in [21, Thm. 1.4]. Note that the relation $u \le \tilde u$ implies $e_k(u) \ge e_k(\tilde u)$ for all $k \le l$ and thus $E_l(u) \le E_l(\tilde u)$ [10]. Since $c_u = c_{\tilde u}$, this gives us (i).

Assertion (ii) follows from (i) by (6).

To prove (iii), we first note that (i) implies $c_u \ge E_l(u) > E_k(u) \ge F_k(u)$ for any $k < l$, so we cannot have $c_u = F_k(u)$ unless $k = l$.

Next, if the analytic variety A has codimension $m > l$, then $\tilde u \in \mathcal{E}_m$, so $c_u = c_{\tilde u} \ge E_m(\tilde u) > E_l(\tilde u) \ge E_l(u) \ge F_l(u)$, which contradicts the assumption, so $\operatorname{codim} A = l$.

Now we prove that 0 is a regular point of the variety A. By Siu's representation formula,

$$(dd^c u)^l = \sum p_j [A_j] + R$$

on a neighborhood V of 0, where $p_j > 0$, $[A_j]$ are integration currents along l-codimensional analytic varieties containing 0, and R is a closed positive current such that for any $a > 0$ the analytic variety $\{z \in V : \nu(R, z) \ge a\}$ has codimension

[2]For the general case of non-isolated singularities, see [1, Thm. 3.7]

strictly greater than l. If $\nu(R,0)>0$, then for almost all points $z\in A$ we have $e_l(u,z)<e_l(u)$. This implies, by (ii), $c_u(z)>c_u$ for all such points z, which is impossible. The same argument shows that the collection $\{A_j\}$ consists of at most one variety and 0 is its regular point. □

Proof of Theorem 2. By (6) and Theorem 1(i), the condition $c_u=F_l(u)$ implies $E_l(u)=F_l(u)$. Therefore, by the arithmetic-geometric mean theorem, we get the relations

$$\frac{e_{k-1}(u)}{e_k(u)}=\frac{e_{j-1}(u)}{e_j(u)}$$

for any $k,j\le l$, which gives us $e_k(u)=[e_1(u)]^k$ for all $k\le l$.

Since relation (8) for $e_1(u)=0$ is obvious (in this case $g_{u_N}\equiv 0$), we can assume $e_1(u)=1$.

Note that for any z, we have $e_k(u,z)\ge[e_1(u,z)]^k$. As follows from the proof of (iii), the relation $c_u=F_l(u)$ implies then, on a neighborhood V of 0,

$$A\cap V=\{z\in V:\ c_u(z)\le 1\}=\{z\in V:\ F_l(u,z)\le 1\}=\{z\in V:\ e_k(u,z)\ge 1\}$$

for all $k\le l$. Moreover, we have $e_k(u,z)=e_1(u,z)^k=1$ for almost all $z\in A\cap V$.

Let us choose, according to Theorem 1, a coordinate system such that $A\cap V=\{z\in V:\ z_k=0,\ 1\le k\le l\}$. Denote $v(z)=\log|z'|$, $z=(z',z'')\in\mathbb{C}^l\times\mathbb{C}^{n-l}$, then $A\cap V=\{z:\ e_k(u,z)\ge e_k(v,z)\}$, with equalities almost everywhere.

In particular, we have $u(z)\le\log|z-(0,\zeta'')|+C(\zeta'')$ as $z\to(0,\zeta'')$ for all $z\in\mathbb{C}^n$ and $\zeta''\in\mathbb{C}^{n-l}$ that are close enough to 0. Assuming $u(z)\le 0$ for all z with $\max|z_k|<2$, we get $u(z)\le\log|z-(0,\zeta'')|$ for all $z\in V$ and $\zeta''\in\mathbb{C}^{n-l}$ with $(0,\zeta'')\in V$. By choosing $\zeta''=z''$ this gives us $u(z)\le v(z)$ on V.

Let $u_N=\max\{u,N\log|z|\}$ and $v_N=\max\{v,N\log|z|\}$. Then $u_N\le v_N$, while for $N\ge 1$ we get, by Demailly's comparison theorem for the Lelong numbers [7],

$$e_n(u_N)\le(dd^cu)^l\wedge(dd^cN\log|z|)^{n-l}(0)=N^{n-l}e_l(u)=N^{n-l}=e_n(v_N).$$

By Lemma 2, $g_{u_N}=g_{v_N}$. □

Proof of Theorem 3. The only part to prove is the one concerning $u=\alpha\log|F|+O(1)$; we assume $\alpha=1$. As follows from Theorem 2, one can choose coordinates such that the zero set Z_F of F is $\{z:\ z'=0\}\cap V\subset\{0\}\times\mathbb{C}^{n-l}$. Observe that for such a function u we have $e_k(u,z)=e_1(u)^k$ for all $z\in Z_F$ near 0.

Let $\mathcal{I}$ be the ideal generated by the components of the mapping F. Then, as mentioned in Section 1, $e_l(u)$ equals $e_l(\mathcal{I},\mathfrak{m}_0)$, the mixed multiplicity of l copies of the ideal $\mathcal{I}$ and $n-l$ copies of the maximal ideal $\mathfrak{m}_0$. By [4, Prop. 2.9], $e_l(\mathcal{I},\mathfrak{m}_0)$ can be computed as the multiplicity $e(\mathcal{J})$ of the ideal $\mathcal{J}$ generated by generic functions $\Psi_1,\dots,\Psi_l\in\mathcal{I}$ and $\xi_1,\dots,\xi_{n-l}\in\mathfrak{m}_0$. Since $e(\mathcal{J})=e_l(w)$, where $w=\log|\Psi|$, we have $e_l(u)=e_l(w)$.

Let now

$$v=e_1(u)\log|z'|,\quad w_N=\max\{w,N\log|z''|\},\quad\text{and}\quad v_N=\max\{v,N\log|z''|\}.$$

Since $w \le \log|F| + O(1)$, we have from Theorem 2 the inequality $w \le v + O(1)$ and thus $w_N \le v_N + O(1)$. Note that the mapping Ψ satisfies the Łojasiewicz inequality $|\Psi_0(z)| \ge |z'|^M$ near 0 for some $M > 0$. Therefore, for sufficiently big N we have $w_N = w'_N = \max\{w, N\log|z|\}$. Then, as in the proof of Theorem 2, we compute

$$\begin{aligned} e_n(w_N) = e_n(w'_N) &\le (dd^c w)^l \wedge (dd^c N \log|z|)^{n-l}(0) \\ &= N^{n-l} e_l(w) = N^{n-l} e_l(u) = e_n(v_N), \end{aligned}$$

which, by Lemma 2, implies $g_{w_N} = g_{v_N}$ for the greenifications on a bounded neighborhood D of 0.

We can assume $D = \{|z'| < 1\} \times \{|z''| < 1\}$, then $g_{v_N} = v_N$, while $g_{w_N} \le w_N$ because the latter function is maximal on D and nonnegative on ∂D. Letting $N \to \infty$ we get $w \ge v$.

Since $w \le u + O(1)$, we have, in particular, $u \ge v + O(1)$, which, in view of Theorem 2, completes the proof. □

References

[1] P. Åhag, U. Cegrell, R. Czyz, Phạm Hoàng Hiệp, *Monge–Ampère measures on pluripolar sets*, J. Math. Pures Appl. (9) **92** (2009), no. 6, 613–627.

[2] P. Åhag, U. Cegrell, S. Kolodziej, H. H. Pham, A. Zeriahi, *Partial pluricomplex energy and integrability exponents of plurisubharmonic functions*, Adv. Math. 222 (2009), no. 6, 2036–2058.

[3] B. Berndtsson, *The openness conjecture for plurisubharmonic functions*, arXiv: 1305.5781

[4] C. Bivià-Ausina, *Joint reductions of monomial ideals and multiplicity of complex analytic maps*, Math. Res. Lett. 15 (2008), no. 2, 389–407.

[5] M. Blel and S. K. Mimouni, *Singularités et intégrabilité des fonctions plurisousharmoniques*, Ann. Inst. Fourier (Grenoble) **55** (2005), no. 2, 319–351.

[6] T. de Fernex, L. Ein, M. Mustaţă, *Multiplicities and log canonical threshold*, J. Algebraic Geom. **13** (2004), no. 3, 603–615.

[7] J.-P. Demailly, *Monge–Ampère operators, Lelong numbers and intersection theory,* Complex Analysis and Geometry (Univ. Series in Math.), ed. by V. Ancona and A. Silva, Plenum Press, New York 1993, 115–193.

[8] J.-P. Demailly, *Estimates on Monge–Ampère operators derived from a local algebra inequality*, Complex Analysis and Digital Geometry. Proceedings from the Kiselmanfest, 2006, ed. by M. Passare. Uppsala University, 2009, 131–143.

[9] J.-P. Demailly and J. Kollár, *Semi-continuity of complex singularity exponents and Kähler–Einstein metrics on Fano orbifolds*, Ann. Sci. Ecole Norm. Sup. (4) **34** (2001), no. 4, 525–556.

[10] J.-P. Demailly and Pham Hoang Hiep, *A sharp lower bound for the log canonical threshold*, Acta Math. **212** (2014), no. 1, 1–9.

[11] C. Favre and M. Jonsson, *Valuations and multiplier ideals*, J. Amer. Math. Soc. **18** (2005), no. 3, 655–684.

[12] J. E. Fornaess and N. Sibony, *Oka's inequality for currents and applications*, Math. Ann. **301** (1995), no. 3, 399-419.

[13] Qi'an Guan, Xiangyu Zhou, *Strong openness conjecture for plurisubharmonic functions*, arXiv:1311.3781

[14] Qi'an Guan, Xiangyu Zhou, *Strong openness conjecture and related problems for plurisubharmonic functions*, arXiv:1403.7247

[15] Qi'an Guan, Xiangyu Zhou, *A proof of Demailly's strong openness conjecture*, Ann. Math. (2) **182** (2015), No. 2, 605–616.

[16] Qi'an Guan, Xiangyu Zhou, *Classification of multiplier ideal sheaves with Lelong number one*, Adv. Math. **285** (2015), 1688–1705.

[17] R. Lazarsfeld, *A short course on multiplier ideals.* Analytic and algebraic geometry, 451–494, IAS/Park City Math. Ser., 17, Amer. Math. Soc., Providence, RI, 2010.

[18] M. Mustaţă, *IMPANGA lecture notes on log canonical thresholds*. EMS Ser. Congr. Rep., Contributions to algebraic geometry, 407–442, Eur. Math. Soc., Zürich, 2012.

[19] A. Rashkovskii, *Relative types and extremal problems for plurisubharmonic functions*, Int. Math. Res. Not., 2006, Art. ID 76283, 26 pp.

[20] A. Rashkovskii, *Multi-circled singularities, Lelong numbers, and integrability index*, J. Geom. Anal. **23** (2013), no. 4, 1976–1992.

[21] A. Rashkovskii, *Extremal cases for the log canonical threshold*, C. R. Acad. Sci. Paris, Ser. I, **353** (2015), 21–24.

[22] A. Rashkovskii and R. Sigurdsson, *Green functions with singularities along complex spaces*, Internat. J. Math. **16** (2005), no. 4, 333–355.

[23] H. Skoda, *Sous-ensembles analytiques d'ordre fini ou infini dans* $\mathbf{C}^n$, Bull. Soc. Math. France **100** (1972), 353–408.

[24] A. Zeriahi, *Appendix: A stronger version of Demailly's estimate on Monge–Ampère operators*, Complex Analysis and Digital Geometry. Proceedings from the Kiselmanfest, 2006, ed. by M. Passare. Uppsala University, 2009, 144–146.

Alexander Rashkovskii
Tek/Nat
University of Stavanger
N-4036 Stavanger, Norway
e-mail: alexander.rashkovskii@uis.no

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 369–438

Root-counting Measures of Jacobi Polynomials and Topological Types and Critical Geodesics of Related Quadratic Differentials

Boris Shapiro and Alexander Solynin

To Mikael Passare, in memoriam

Abstract. Two main topics of this paper are asymptotic distributions of zeros of Jacobi polynomials and topology of critical trajectories of related quadratic differentials. First, we will discuss recent developments and some new results concerning the limit of the root-counting measures of these polynomials. In particular, we will show that the support of the limit measure sits on the critical trajectories of a quadratic differential of the form $Q(z)\,dz^2 = \frac{az^2+bz+c}{(z^2-1)^2}\,dz^2$. Then we will give a complete classification, in terms of complex parameters a, b, and c, of possible topological types of critical geodesics for the quadratic differential of this type.

Mathematics Subject Classification (2010). 30C15, 31A35, 34E05.

Keywords. Jacobi polynomials, asymptotic root-counting measure, quadratic differentials, critical trajectories.

1. Introduction: From Jacobi polynomials to quadratic differentials

Two main themes of this work are asymptotic behavior of zeros of certain polynomials and topological properties of related quadratic differentials. The study of asymptotic root distributions of hypergeometric, Jacobi, and Laguerre polynomials with variable real parameters, which grow linearly with degree, became a rather hot topic in recent publications, which attracted attention of many authors [15], [16], [17], [18], [19], [23], [25], [26], [28]. In this paper, we survey some known results in this area and present some new results keeping focus on Jacobi polynomials.

Recall that the Jacobi polynomial $P_n^{(\alpha,\beta)}(z)$ of degree n with complex parameters α,β is defined by

$$P_n^{(\alpha,\beta)}(z)=2^{-n}\sum_{k=0}^{n}\binom{n+\alpha}{n-k}\binom{n+\beta}{k}(z-1)^k(z+1)^{n-k},$$

where $\binom{\gamma}{k}=\frac{\gamma(\gamma-1)\dots(\gamma-k+1)}{k!}$ with a non-negative integer k and an arbitrary complex number γ. Equivalently, $P_n^{(\alpha,\beta)}(z)$ can be defined by the well-known Rodrigues formula:

$$P_n^{(\alpha,\beta)}(z)=\frac{1}{2^n n!}(z-1)^{-\alpha}(z+1)^{-\beta}\left(\frac{d}{dz}\right)^n[(z-1)^{n+\alpha}(z+1)^{n+\beta}].$$

The following statement, which can be found, for instance, in [25, Proposition 2], gives an important characterization of Jacobi polynomials as solutions of second-order differential equation.

Proposition 1. *For arbitrary fixed complex numbers α and β, the differential equation*

$$(1-z^2)y''+(\beta-\alpha-(\alpha+\beta+2)z)y'+\lambda y=0$$

with a spectral parameter λ has a non-trivial polynomial solution of degree n if and only if $\lambda=n(n+\alpha+\beta+1)$. This polynomial solution is unique (up to a constant factor) and coincides with $P_n^{(\alpha,\beta)}(z)$.

Working with root distributions of polynomials, it is convenient to use root-counting measures and their Cauchy transforms, which are defined as follows.

Definition 1. For a polynomial $p(z)$ of degree n with (not necessarily distinct) roots $\xi_1,\dots,\xi_n$, its *root-counting measure* μ_p is defined as

$$\mu_p=\frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i},$$

where δ_ξ is the Dirac measure supported at ξ.

Definition 2. Given a finite complex-valued Borel measure μ compactly supported in $\mathbb{C}$, its *Cauchy transform* $\mathcal{C}_\mu$ is defined as

$$\mathcal{C}_\mu(z)=\int_{\mathbb{C}}\frac{d\mu(\xi)}{z-\xi}. \tag{1.1}$$

and its logarithmic potential u_μ is defined as

$$u_\mu(z)=\int_{\mathbb{C}}\log|z-\xi|d\mu(\xi).$$

We note that the integral in (1.1) converges for all z, for which the Newtonian potential $U_{|\mu|}(z)=\int_{\mathbb{C}}\frac{d|\mu|(\xi)}{|\xi-z|}$ of μ is finite, see, e.g., [20, Ch. 2].

In case when $\mu=\mu_p$ is the root-counting measure of a polynomial $p(z)$, we will write $\mathcal{C}_p$ instead of $\mathcal{C}_{\mu_p}$. It follows from Definitions 1 and 2 that the Cauchy transform $\mathcal{C}_p(z)$ of the root-counting measure of a monic polynomial $p(z)$ of degree n coincides with the normalized logarithmic derivative of $p(z)$; i.e.,

$$\mathcal{C}_p(z) = \frac{p'(z)}{np(z)} = \int_{\mathbb{C}} \frac{d\mu_p(\xi)}{z-\xi}, \tag{1.2}$$

and its logarithmic potential $u_p(z)$ is given by the formula:

$$u_p(z) = \frac{1}{n}\log|p(z)| = \int_{\mathbb{C}} \log|z-\xi| d\mu_p(\xi). \tag{1.3}$$

Let $\{p_n(z)\}$ be a sequence of Jacobi polynomials $p_n(z) = P_n^{(\alpha_n,\beta_n)}(z)$ and let $\{\mu_n\}$ be the corresponding sequence of their root-counting measures. The main question we are going to address in this paper is the following:

Problem 1. Assuming that the sequence $\{\mu_n\}$ weakly converges to a measure μ compactly supported in $\mathbb{C}$, what can be said about properties of the support of the measure μ and about its Cauchy transform $\mathcal{C}_\mu$?

Regarding the Cauchy transform $\mathcal{C}_\mu$, our main result in this direction is the following theorem.

Theorem 1. *Suppose that a sequence $\{p_n(z)\}$ of Jacobi polynomials*

$$p_n(z) = P_n^{(\alpha_n,\beta_n)}(z)$$

satisfies conditions:

(a) *the limits $A = \lim_{n\to\infty} \frac{\alpha_n}{n}$ and $B = \lim_{n\to\infty} \frac{\beta_n}{n}$ exist, and $1+A+B \neq 0$;*
(b) *the sequence $\{\mu_n\}$ of the root-counting measures converges weakly to a probability measure μ, which is compactly supported in $\mathbb{C}$.*

Then the Cauchy transform $\mathcal{C}_\mu$ of the limit measure μ satisfies almost everywhere in $\mathbb{C}$ the quadratic equation:

$$(1-z^2)\mathcal{C}_\mu^2 - ((A+B)z + A - B)\mathcal{C}_\mu + A + B + 1 = 0. \tag{1.4}$$

The proof of Theorem 1 given in Section 2 consists of several steps. Our arguments in Section 2 are similar to the arguments used in a number of earlier papers on root asymptotics of orthogonal polynomials.

Equation (1.4) of Theorem 1 implies that the support of the limit measure μ has a remarkable structure described by Theorem 2 below. And this is exactly the point where quadratic differentials, which are the second main theme of this paper, enter into the play.

Theorem 2. *In notation of Theorem 1, the support of μ consists of finitely many trajectories of the quadratic differential*

$$Q(z)\,dz^2 = -\frac{(A+B+2)^2z^2 + 2(A^2-B^2)z + (A-B)^2 - 4(A+B+1)}{(z-1)^2(z+1)^2}\,dz^2$$

and their end points.

Thus, to understand geometrical structure of the support of μ we have to study geometry of critical trajectories, or more generally critical geodesics of the quadratic differential $Q(z)\,dz^2$ of Theorem 1. We will consider a slightly more general family of quadratic differentials $Q(z;a,b,c)\,dz^2$ depending on three complex parameters $a,b,c \in \mathbb{C}$, $a \neq 0$, where

$$Q(z;a,b,c)\,dz^2 = \frac{az^2+bz+c}{(z-1)^2(z+1)^2}\,dz^2. \tag{1.5}$$

It is well known that quadratic differentials appear in many areas of mathematics and mathematical physics such as moduli spaces of curves, univalent functions, asymptotic theory of linear ordinary differential equations, spectral theory of Schrödinger equations, orthogonal polynomials, etc. Postponing necessary definitions and basic properties of quadratic differentials till Section 3, we recall here that any meromorphic quadratic differential $Q(z)\,dz^2$ defines the so-called Q-*metric* and therefore it defines Q-*geodesics* in appropriate classes of curves. Motivated by the fact that the family of quadratic differentials (1.5) naturally appears in the study of the root asymptotics for sequences of Jacobi polynomials and is one of very few examples allowing detailed and explicit investigation in terms of its coefficients, we will consider the following two basic questions:

1) How many simple critical Q-geodesics may exist for a quadratic differential $Q(z)\,dz^2$ of the form (1.5)?
2) For given $a,b,c \in \mathbb{C}$, $a \neq 0$, describe topology of all simple critical Q-geodesics.

A complete description of topological structure of trajectories of quadratic differentials (1.5) which, in particular, answers questions 1) and 2), is given by lengthy Theorem 5 stated in Section 9.

The rest of the paper consists of two parts and is structured as follows. The first part, which is the area of expertise of the first author, includes Sections 2, 4, and 5. Section 2 contains the proof of Theorem 1 and related results. The material presented in Section 4 is mostly borrowed from a recent paper [12] of the first author. It contains some general results connecting signed measures, whose Cauchy transforms satisfy quadratic equations, and related quadratic differentials in $\mathbb{C}$. In particular, these results imply Theorem 2 as a special case. In Section 5, we formulate a number of general conjectures about the type of convergence of root-counting measures of polynomial solutions of a special class of linear differential equations with polynomial coefficients, which includes Riemann's differential equation.

Remaining sections constitute the second part, which is the area of expertise of the second author. In Section 3, we recall basic information about quadratic differentials, their critical trajectories and geodesics. This information is needed for presentation of our results in Sections 6–10. In Section 6, we describe possible domain configurations for the quadratic differentials (1.5). Then, in Section 7, we describe possible topological types of the structure of critical trajectories of quadratic differentials of the form (1.5). Finally in Sections 8–10, we identify sets of parameters corresponding to each topological type. The latter allows us to answer some related questions.

We note here that our main proofs presented in Sections 6–10 are geometrical based on general facts of the theory of quadratic differentials. Thus, our methods can be easily adapted to study trajectory structure of many quadratic differentials other then quadratic differential (1.5).

Section 11 is our Figures Zoo, it contains many figures illustrating our results presented in Sections 6–10.

2. Proof of Theorem 1

To settle Theorem 1 we will need several auxiliary statements. Lemma 1 below can be found as Theorem 7.6 of [3] and apparently was originally proven by F. Riesz.

Lemma 1. *If a sequence $\{\mu_n\}$ of Borel probability measures in $\mathbb{C}$ weakly converges to a probability measure μ with a compact support, then the sequence $\{\mathcal{C}_{\mu_n}(z)\}$ of its Cauchy transforms converges to $\mathcal{C}_\mu(z)$ in L^1_{loc}. Moreover there exists a subsequence of $\{\mathcal{C}_{\mu_n}(z)\}$ which converges to $\mathcal{C}_\mu(z)$ pointwise almost everywhere.*

The next result is recently obtained by the first author jointly with R. Bøgvad and D. Khavinsion, see Theorem 1 of [13] and has an independent interest.

Proposition 2. *Let $\{p_m\}$ be any sequence of polynomials satisfying the following conditions:*

1. $n_m := \deg p_m \to \infty$ *as* $m \to \infty$,
2. *almost all roots of all p_m lie in a bounded convex open $\Omega \subset \mathbb{C}$ when $n \to \infty$. (More exactly, if In_m denotes the number of roots of p_m counted with multiplicities which are located in Ω, then $\lim_{m\to\infty} \frac{In_m}{n_m} = 1$), then for any $\epsilon > 0$,*
$$\lim_{m\to\infty} \frac{In'_m(\epsilon)}{n_m} = 1,$$
where $In'_m(\epsilon)$ is the number of roots of p'_m counted with multiplicities which are located inside $\Omega(\epsilon)$, the latter set being the ϵ-neighborhood of Ω in $\mathbb{C}$.

The next statement is a strengthening of Lemma 8 of [5] based on Proposition 2.

Lemma 2. *Let $\{p_m\}$ be any sequence of polynomials satisfying the following conditions:*

1. $n_m := \deg p_m \to \infty$ *as* $m \to \infty$,
2. *the sequence $\{\mu_m\}$ (resp. $\{\mu'_m\}$) of the root-counting measures of $\{p_m\}$ (resp. $\{p'_m\}$) weakly converges to compactly supported measures μ (resp μ').*
Then u and u' satisfy the inequality $u \geq u'$ with equality on the unbounded component of $\mathbb{C} \setminus supp(\mu)$. Here u (resp. u') is the logarithmic potential of the limiting measure μ (resp. μ').

Proof. Without loss of generality, we can assume that all p_m are monic. Let K be a compact convex set containing almost all the zeros of the sequences $\{p_m\}$ and $\{p'_m\}$, i.e., $\lim_{m\to\infty} \frac{In_m(K)}{n_m} = \lim_{m\to\infty} \frac{In'_m(K)}{n_m} = 1$. By (1.3) we have

$$u(z) = \lim_{m\to\infty} \frac{1}{n_m} \log |p_m(z)|$$

and

$$u'(z) = \lim_{m\to\infty} \frac{1}{n_m - 1} \log \left| \frac{p'_m(z)}{n_m} \right| = \lim_{m\to\infty} \frac{1}{n_m} \log \left| \frac{p'_m(z)}{n_m} \right|$$

with convergence in L^1_{loc}. Hence by (1.2),

$$u'(z) - u(z) = \lim_{m\to\infty} \frac{1}{n_m} \log \left| \frac{p'_m(z)}{n_m p_m(z)} \right| = \lim_{m\to\infty} \frac{1}{n_m} \log \left| \int \frac{d\mu_m(\zeta)}{z-\zeta} \right|. \tag{2.1}$$

Now, if ϕ is a positive compactly supported test function, then

$$\begin{aligned} \int \phi(z)(u'(z) - u(z))\, dA(z) &= \lim_{m\to\infty} \frac{1}{n_m} \int \phi(z) \log \left| \int \frac{d\mu_m(\zeta)}{z-\zeta} \right| dA(z) \\ &\le \lim_{m\to\infty} \frac{1}{n_m} \int \phi(z) \int \frac{d\mu_m(\zeta)}{|z-\zeta|}\, dA(z) \\ &= \lim_{m\to\infty} \frac{1}{n_m} \iint \frac{\phi(z)\, dA(z)}{|z-\zeta|}\, d\mu_m(\zeta) \end{aligned} \tag{2.2}$$

where dA denotes Lebesgue measure in the complex plane. Since $1/|z|$ is locally integrable, the function $\int \phi(z)|z-\zeta|^{-1}\, dA(z)$ is continuous, and hence bounded by a constant M for all z in K. Since asymptotically almost all zeros of $\{p_m\}$ belong to K, the last expression in (2.2) tends to 0 when $m \to \infty$. This proves that $u' \le u$.

In the complement of $\operatorname{supp}\mu$, u is harmonic and u' is subharmonic, hence $u' - u$ is a negative subharmonic function. Moreover, in the complement of $\operatorname{supp}\mu$, $p'_m/(n_m p_m)$ converges to the Cauchy transform $\mathcal{C}(z)$ of μ a.e. in $\mathbb{C}$. Since $\mathcal{C}(z)$ is a nonconstant holomorphic function in the unbounded component of $\mathbb{C} \setminus \operatorname{supp}\mu$, it follows from (2.1) that $u' - u \equiv 0$ there. □

Notice that Lemma 2 implies the following interesting fact.

Corollary 1. *In notation of Lemma* 2*, if* $\operatorname{supp}\mu$ *has Lebesgue area 0 and the complement* $\mathbb{C} \setminus \operatorname{supp}\mu$ *is path-connected, then* $\mu = \mu'$*. In particular, in this case the whole sequence* $\{\mu'_m\}$ *weakly converges to* μ*.*

In general, however $\mu \neq \mu'$ as shown by a trivial example of the sequence $\{z^n - 1\}_{n=1}^{\infty}$. Also even if $\mu = \lim_{m\to\infty} \mu_n$ exists the limit $\lim_{m\to\infty} \mu'_n$ does not have to exist for the whole sequence. An example of this kind is the sequence $\{p_n(z)\}$ where $p_{2l}(z) = z^{2l} - 1$ and $p_{2l+1}(z) = z^{2l+1} - z$, $l = 1, 2, \ldots$.

Luckily, the latter phenomenon can never occur for sequences of Jacobi polynomials, see Proposition 3 below. (Apparently it cannot occur for a much more general class of polynomial sequences introduced in § 5.)

Lemma 3. *If the sequence $\{\mu_n\}$ of the root-counting measures of a sequence of Jacobi polynomials $\{p_n(z)\} = \{P_n^{(\alpha_n,\beta_n)}(z)\}$ weakly converges to a measure μ compactly supported in $\mathbb{C}$, and the sequence $\{\mu'_n\}$ of the root-counting measures of a sequence $\{p'_n(z)\}$ weakly converges to a measure μ' compactly supported in $\mathbb{C}$, then one of the following alternatives holds:*

(i) *the sequences $\left\{\frac{\alpha_n+\beta_n}{n}\right\}$ and $\left\{\frac{\beta_n-\alpha_n}{n}\right\}$ (and, therefore, the sequences $\{\frac{\alpha_n}{n}\}$ and $\left\{\frac{\beta_n}{n}\right\}$) are bounded;*

(ii) *the sequence $\left\{\frac{\alpha_n+\beta_n}{n}\right\}$ is unbounded and the sequence $\left\{\frac{\beta_n-\alpha_n}{n}\right\}$ is bounded, in which case $\{\mu_n\} \to \delta_0$ where δ_0 is the unit point mass at $z = 0$ (or, equivalently, $\mathcal{C}_{\delta_0}(z) = 1/z$);*

(iii) *both sets $\left\{\frac{\alpha_n+\beta_n}{n}\right\}$ and $\left\{\frac{\beta_n-\alpha_n}{n}\right\}$ are unbounded, in which case, there exists at least one $\kappa \in \mathbb{C}$ and a subsequence $\{n_m\}$ such that $\lim_{m\to\infty} \frac{\beta_{n_m}-\alpha_{n_m}}{\alpha_{n_m}+\beta_{n_m}} = \kappa$ and $\{\mu_{n_m}\} \to \delta_\kappa$, where δ_κ is the unit point mass at $z = \kappa$ (or, equivalently, $\mathcal{C}_{\delta_\kappa}(z) = 1/(z-\kappa)$).*

Proof. Indeed, assume that the alternative (i) does not hold. Then there is a subsequence $\{n_m\}$ such that at least one of $\left|\frac{\alpha_{n_m}+\beta_{n_m}}{n_m}\right|$, $\left|\frac{\beta_{n_m}-\alpha_{n_m}}{n_m}\right|$ is unbounded along this subsequence. By our assumptions $\mu_n \to \mu$ and $\mu'_n \to \mu'$ weakly. Hence, by Lemma 1, there exists a subsequence of indices along which $\mathcal{C}_{\mu_n} := \frac{p'_n}{np_n}$ pointwise converges to $\mathcal{C}_\mu$ and $\mathcal{C}_{\mu'_n} := \frac{p''_n}{(n-1)p'_n}$ pointwise converges to $\mathcal{C}_{\mu'}$ a.e. in $\mathbb{C}$. Consider the sequence of differential equations satisfied by $\{p_n\}$ and divided termwise by $n(n-1)p_n$:

$$\begin{aligned}(1-z^2)\frac{p''_n}{(n-1)p'_n}\cdot\frac{p'_n}{np_n} + \left(\frac{(\beta_n-\alpha_n)-(\alpha_n+\beta_n+2)z}{n-1}\right)\frac{p'_n}{np_n}\\ + \frac{n+\alpha_n+\beta_n+1}{n-1} = 0.\end{aligned} \tag{2.3}$$

If for a subsequence of indices, $\left|\frac{\beta_n-\alpha_n}{n}\right| \to \infty$ while $\left|\frac{\alpha_n+\beta_n}{n}\right|$ stays bounded, then the Cauchy transform $\mathcal{C}_\mu$ of the limiting (along this subsequence) measure μ must vanish identically in order for (2.3) to hold in the limit $n \to \infty$. But $\mathcal{C}_\mu \equiv 0$ is obviously impossible.

On the other hand, if for a subsequence of indices, $\left|\frac{\alpha_n+\beta_n}{n}\right| \to \infty$ while $\left|\frac{\beta_n-\alpha_n}{n}\right|$ stays bounded, then the limit of (2.3) when $n \to \infty$ coincides with $-z\mathcal{C}_\mu + 1 = 0 \Leftrightarrow \mathcal{C}_\mu = \frac{1}{z}$ implying $\mu = \delta_0$. Thus in Case (ii), the sequence $\{\mu_n\}$ converges to δ_0.

Now assume, that or a subsequence of indices, both $\left|\frac{\alpha_n+\beta_n}{n}\right|$ and $\left|\frac{\beta_n-\alpha_n}{n}\right|$ tend to ∞. Then dividing (2.3) by $\frac{\alpha_n+\beta_n}{n}$ and letting $n \to \infty$, we conclude that the sequence $\left\{\frac{\beta_n-\alpha_n}{\alpha_n+\beta_n}\right\}$ must be bounded. Therefore there exists its subsequence which converges to some $\kappa\in\mathbb{C}$. Taking the limit along this subsequence, we obtain

$$(z-\kappa)\mathcal{C}_\mu=1.$$

This is true for all z, for which the Cauchy transform converges, i.e., almost everywhere outside the support of μ. Using the main results of [7, 8] claiming that the support of μ consists of piecewise smooth compact curves and/or isolated points together with the fact that $\mathcal{C}_\mu$ must have a discontinuity along every curve in its support, we conclude that the support of μ is the point $z=\kappa$. Thus in Case (iii), the sequence $\{\mu_{n_m}\}$ converges to δ_κ. □

The next statement provides more information about Case (i) of Lemma 3.

Proposition 3. *Assume that the sequence $\{\mu_n\}$ of the root-counting measures for a sequence of Jacobi polynomials $\{p_n(z)=P_n^{(\alpha_n,\beta_n)}(z)\}$ weakly converges to a compactly supported measure μ in $\mathbb{C}$. Assume additionally that* $\lim_{n\to\infty}\frac{\alpha_n}{n}=A$ *and* $\lim_{n\to\infty}\frac{\beta_n}{n}=B$ *with $1+A+B\neq 0$. Then, for any positive integer j, the sequence $\{\mu_n^{(j)}\}$ of the root-counting measures for the sequence $\{p_n^{(j)}(z)\}$ of the jth derivatives converges to the same measure μ.*

Proof. Observe that if an arbitrary polynomial sequence $\{p_m\}$ of increasing degrees has almost all roots in a convex bounded set $\Omega\subset\mathbb{C}$, then, by Proposition 2, almost all roots of $\{p_m'\}$ are in Ω_ϵ, for any $\epsilon>0$. Therefore, if the sequence $\{\mu_m\}$ of the root-counting measures of $\{p_m\}$ weakly converges to a compactly supported measure μ, then there exists at least one weakly converging subsequence of $\{\mu_m'\}$. Additionally, by the Gauss–Lucas Theorem, the support of its limiting measure belongs to the (closure of the) convex hull of the support of μ. Thus the weak convergence of $\{\mu_m\}$ implies the existence of a weakly converging subsequence $\{\mu_{n_m}'\}$.

Proposition 3 is obvious in Cases (ii) and (iii) of Lemma 3. Let us concentrate on the remaining Case (i). Our assumptions imply that along a subsequence of the sequence $\left\{\frac{p_n'}{np_n}\right\}$ of Cauchy transforms of polynomials p_n converges pointwise almost everywhere. We first show that the above sequence $\left\{\frac{p_n'}{np_n}\right\}$ cannot converge to 0 on a set of positive measure.

Indeed, the differential equation satisfied by p_n after its division by $n(n-1)p_n$ is given by (2.3). Since the sequences $\left\{\frac{\alpha_n+\beta_n}{n}\right\}$ and $\left\{\frac{\beta_n-\alpha_n}{n}\right\}$ converge and $1+A+B\neq 0$, equation (2.3) shows that $\frac{p_n'}{np_n}$ cannot converge to 0 on a set of positive measure. Analogously, we see that $\frac{p_n''}{(n-1)p_n'}$ cannot converge to 0 on a set of positive measure either. Indeed, differentiating (2.3), we get that p_n' satisfies the equation

$$(1-z^2)p_n'''+((\beta_n-\alpha_n)-(\alpha_n+\beta_n+4)z)p_n''+(n(n+\alpha_n+\beta_n+1)+(\alpha_n+\beta_n+2))p_n'=0.$$

Using the same analysis as for p_n, we can conclude that the limit $\frac{p_n''}{n(n-1)p_n}$ along a subsequence exists pointwise and is non-vanishing almost everywhere.

Denote the logarithmic potentials of the root-counting measures associated to p_n and p_n' by u_n and u_n' respectively. Denote their limits by u and u' (where u' apriori is a limit only along some subsequence). With a slight abuse of notation, the following holds

$$|u - u'| = \lim_{n\to\infty} |u_n - u_n'| = \lim_{n\to\infty} \frac{1}{n} \log \left| \frac{p_n''}{n(n-1)p_n} \right| = 0$$

due to the above claim about $\frac{p_n''}{n(n-1)p_n}$. But since $u \geq u'$ by Lemma 2, we see that $u = u'$ and, in particular u' exists as a limit over the whole sequence. Hence the asymptotic root-counting measures of $\{p_n\}$ and $\{p_n'\}$ actually coincide. Similar arguments apply to higher derivatives of the sequence $\{p_n\}$. □

Proof of Theorem 1. The polynomial $p_n(z) = P_n^{(\alpha_n,\beta_n)}(z)$ satisfies equation (2.3). By Proposition 3 we know that, under the assumptions of Theorem 1, if $\left\{\frac{p_n'}{np_n}\right\}$ converges to $\mathcal{C}_\mu$ a.e. in $\mathbb{C}$, then the sequence $\left\{\frac{p_n''}{np_n'}\right\}$ also converges to the same $\mathcal{C}_\mu$ a.e. in $\mathbb{C}$. Therefore, the expression $\frac{p_n''}{n^2p_n} = \frac{p_n''p_n'}{n^2p_np_n'}$ converges to $\mathcal{C}_\mu^2$ a.e. in $\mathbb{C}$. Thus $\mathcal{C}_\mu$ (which is well defined a.e. in $\mathbb{C}$) should satisfy the equation

$$(1-z^2)\mathcal{C}_\mu^2 - ((A+B)z + A - B)\mathcal{C}_\mu + A + B + 1 = 0,$$

where $A = \lim_{n\to\infty} \frac{\alpha_n}{n}$ and $B = \lim_{n\to\infty} \frac{\beta_n}{n}$. □

Remark 1. Apparently the condition that the sequences $\left\{\frac{\alpha_n}{n}\right\}$ and $\left\{\frac{\beta_n}{n}\right\}$ are bounded should be enough for the conclusion of Theorem 1. (The existence of the limits $\lim \frac{\alpha_n}{n}$ and $\lim \frac{\beta_n}{n}$ should follow automatically with some weak additional restriction.) Indeed, since the sequences $\left\{\frac{\alpha_n}{n}\right\}$ and $\left\{\frac{\beta_n}{n}\right\}$ are bounded, we can find at least one subsequence $\{n_m\}$ of indices along which both sequences of quotients converge. Assume that we have two possible distinct (pairs of) limits (A_1, B_1) and (A_2, B_2) along different subsequences. But then the same complex-analytic function $\mathcal{C}_\mu(z)$ should satisfy a.e. two different algebraic equations of the form (1.4) which is impossible at least for generic (A_1, B_1) and (A_2, B_2).

3. Preliminaries on quadratic differentials

In this section, we recall some definitions and results of the theory of quadratic differentials on the complex sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Most of these results remain true for quadratic differentials defined on any compact Riemann surface. But for the purposes of this paper, we will focus on results concerning the domain structure and properties of geodesics of quadratic differentials defined on $\overline{\mathbb{C}}$. For more information on quadratic differentials in general, the interested reader may consult classical monographs of Jenkins [22] and Strebel [34] and papers [31] and [32].

A quadratic differential on a domain $D \subset \overline{\mathbb{C}}$ is a differential form $Q(z)\,dz^2$ with meromorphic $Q(z)$ and with conformal transformation rule

$$Q_1(\zeta)\,d\zeta^2 = Q(\varphi(z))\,(\varphi'(z))^2\;dz^2, \tag{3.1}$$

where $\zeta = \varphi(z)$ is a conformal map from D onto a domain $G \subset \overline{\mathbb{C}}$. Then zeros and poles of $Q(z)$ are critical points of $Q(z)\,dz^2$, in particular, zeros and simple poles are finite critical points of $Q(z)\,dz^2$. Below we will use the following notations. By H_p, C, and H we denote, respectively, the set of all poles, set of all finite critical points, and set of all infinite critical points of $Q(z)\,dz^2$. Also, we will use the following notations: $\mathbb{C}' = \overline{\mathbb{C}} \setminus H$, $\mathbb{C}'' = \overline{\mathbb{C}} \setminus H_p$, $\mathbb{C}''' = \overline{\mathbb{C}} \setminus (C \cup H)$.

A trajectory (respectively, orthogonal trajectory) of $Q(z)\,dz^2$ is a closed analytic Jordan curve or maximal open analytic arc $\gamma \subset D$ such that

$$Q(z)\,dz^2 > 0 \quad \text{along } \gamma \qquad (\text{respectively}, Q(z)\,dz^2 < 0 \quad \text{along } \gamma).$$

A trajectory γ is called *critical* if at least one of its end points is a finite critical point of $Q(z)\,dz^2$. By a closed critical trajectory we understand a critical trajectory together with its end points z_1 and z_2 (not necessarily distinct), assuming that these end points exist.

Let $\overline{\Phi}$ denote the closure of the set of points of all critical trajectories of $Q(z)\,dz^2$. Then, by Jenkins' Basic Structure Theorem [22, Theorem 3.5], the set $\overline{\mathbb{C}} \setminus \overline{\Phi}$ consists of a finite number of *circle, ring, strip and end domains*. The collection of all these domains together with so-called *density domains* constitute the so-called *domain configuration* of $Q(z)\,dz^2$. Here, we give definitions of circle domains and strip domains only; these two types will appear in our classification of possible domain configurations in Section 5. Figures 1–4 show several domain configurations with circle and strip domains. For the definitions of other domains, we refer to [22, Ch. 3].

We recall that a *circle domain* of $Q(z)\,dz^2$ is a simply connected domain D with the following properties:

1) D contains exactly one critical point z_0, which is a second-order pole,
2) the domain $D \setminus \{z_0\}$ is swept out by trajectories of $Q(z)\,dz^2$ each of which is a Jordan curve separating z_0 from the boundary ∂D,
3) ∂D contains at least one finite critical point.

Similarly, a strip domain of $Q(z)\,dz^2$ is a simply connected domain D with the following properties:

1) D contains no critical points of $Q(z)\,dz^2$,
2) ∂D contains exactly two boundary points z_1 and z_2 belonging to the set H (these boundary points may be situated at the same point of $\overline{\mathbb{C}}$),
3) the points z_1 and z_2 divide ∂D into two boundary arcs each of which contains at least one finite critical point,

4) D is swept out by trajectories of $Q(z)\,dz^2$ each of which is a Jordan arc connecting points z_1 and z_2.

As we mentioned in the Introduction, every quadratic differential $Q(z)dz^2$ defines the so-called (singular) Q-metric with the differential element $|Q(z)|^{1/2}\,|dz|$. If γ is a rectifiable arc in D then its Q-length is defined by

$$|\gamma|_Q = \int_\gamma |Q(z)|^{1/2}\,|dz|.$$

According to their Q-lengthes, trajectories of $Q(z)\,dz^2$ can be of two types. A trajectory γ is called *finite* if its Q-length is finite, otherwise γ is called *infinite*. In particular, a critical trajectory γ is finite if and only if it has two end points each of which is a finite critical point.

An important property of quadratic differentials is that transformation rule (8.1) respects trajectories and orthogonal trajectories and their Q-lengthes, as well as it respects critical points together with their multiplicities and trajectory structure nearby.

Definition 3. A locally rectifiable (in the spherical metric) curve $\gamma \subset \mathbb{C}'$ is called *a Q-geodesic* if it is locally shortest in the Q-metric.

Next, given a quadratic differential $Q(z)\,dz^2$, we will discuss geodesics in homotopic classes. For any two points $z_1, z_2 \in \mathbb{C}'$, let $\mathcal{H}^J = \mathcal{H}^J(z_1, z_2)$ denote the set of all homotopic classes H of Jordan arcs $\gamma \subset \mathbb{C}'$ joining z_1 and z_2. Here the letter J stands for "Jordan". It is well known that there is a countable number of such homotopic classes. Thus, we may write $\mathcal{H}^J = \{H_k^J\}_{k=1}^\infty$.

Every class H_k^J can be extended to a larger class H_k by adding non-Jordan continuous curves γ joining z_1 and z_2, each of which is homotopic on $\mathbb{C}'$ to some curve $\gamma_0 \in H_k^J$ in the following sense.

There is a continuous function $\varphi(t,\tau)$ from the square $I^2 := [0,1] \times [0,1]$ to $\mathbb{C}'$ such that

1) $\varphi(0,\tau) = z_1$, $\varphi(1,\tau) = z_2$ for all $0 \le \tau \le 1$,
2) $\gamma_0 = \{z = \varphi(t,0) : 0 \le t \le 1\}$,
3) $\gamma = \gamma_1 = \{z = \varphi(t,1) : 0 \le t \le 1\}$,
4) For every fixed $\tau, 0 < \tau < 1$, the curve $\gamma_\tau = \{z = \varphi(t,\tau) : 0 \le t \le 1\}$ is in the class H_k^J.

The following proposition is a special case of a well-known result about geodesics, see, e.g., [34, Theorem 18.2.1].

Proposition 4. *For every k, there is a unique curve $\gamma' \in H_k$, called Q-*geodesic *in H_k, such that $|\gamma'|_Q < |\gamma|_Q$ for all $\gamma \in H_k$, $\gamma \ne \gamma'$. This geodesic is not necessarily a Jordan arc.*

A Q-geodesic from z_1 to z_2 is called *simple* if $z_1 \ne z_2$ and γ is a Jordan arc on $\mathbb{C}'''$ joining z_1 and z_2. A Q-geodesic is called *critical* if both its end points belong to the set of finite critical points of $Q(z)\,dz^2$.

Proposition 5. *Let $Q(z)\,dz^2$ be a quadratic differential on $\overline{\mathbb{C}}$. Then for any two points $z_1, z_2 \in \mathbb{C}'$ and every continuous rectifiable curve γ on $\mathbb{C}'''$ joining the points z_1 and z_2 there is a unique shortest curve γ_0 belonging to the homotopic class of γ.*

Furthermore, γ_0 is a geodesic in this class.

Definition 4. Let $z_0 \in \mathbb{C}'$. A geodesic ray from z_0 is a maximal simple rectifiable arc $\gamma : [0,1) \to \mathbb{C}''' \cup \{z_0\}$ with $\gamma(0) = z_0$ such that for every t, $0 < t < 1$, the arc $\gamma((0,1))$ is a geodesic from z_0 to $z = \gamma(t)$.

Lemma 4. *Let D be a circle domain of $Q(z)\,dz^2$ centered at z_0 and let $\gamma_a : [0,1) \to \mathbb{C}''' \cup \{a\}$ be a geodesic ray from $a \in \partial D$ such that $\gamma_a([0,t_0]) \subset \overline{D}$ for some $t_0 > 0$.*

Then either γ_a enters into D through the point a and then approaches to z_0 staying in D or γ_a is an arc of some critical trajectory $\gamma \subset \partial D$.

Lemma 5. *Let a be a second-order pole of $Q(z)\,dz^2$ and let Γ be the homotopic class of closed curves on $\mathbb{C}''$ separating a from $H_p \setminus \{a\}$. Then there is exactly one real θ_0, $0 \le \theta_0 < 2\pi$, such that the quadratic differential $e^{i\theta_0}Q(z)\,dz^2$ has a circle domain, say D_0, centered at a. Furthermore, the boundary ∂D_0 is the only critical Q-geodesic (non-Jordan in general) in the class Γ.*

In particular, Γ may contain at most one critical geodesic loop.

We will need some simple mapping properties of the canonical mapping related to the quadratic differential $Q(z)\,dz^2$, which is defined by

$$F(z) = \int_{z_0} \sqrt{Q(z)}\,dz$$

with some $z_0 \in \overline{\mathbb{C}}$ and some fixed branch of the radical. A simply connected domain D without critical points of $Q(z)\,dz^2$ is called a Q-rectangle if the boundary of D consists of two arcs of trajectories of $Q(z)\,dz^2$ separated by two arcs of orthogonal trajectories of this quadratic differential. As well a canonical mapping $F(z)$ maps any Q-rectangle conformally onto a geometrical rectangle in the plane with two sides parallel to the horizontal axis.

4. Cauchy transforms satisfying quadratic equations and quadratic differentials

Below we relate the question for which triples of polynomials (P, Q, R) the equation

$$P(z)\mathcal{C}^2 + Q(z)\mathcal{C} + R(z) = 0, \tag{4.1}$$

with $\deg P = n+2$, $\deg Q \le n+1$, $\deg R \le n$ admits a compactly supported signed measure μ whose Cauchy transform satisfies (4.1) almost everywhere in $\mathbb{C}$ to a certain problem about rational quadratic differentials. We call such measure μ a *motherbody measure* for (4.1).

For a given quadratic differential Ψ on a compact surface $\mathcal{R}$, denote by $K_\Psi \subset \mathcal{R}$ the union of all its critical trajectories and critical points. (In general, K_Ψ can be very complicated. In particular, it can be dense in some subdomains of $\mathcal{R}$.) We denote by $DK_\Psi \subseteq K_\Psi$ (the closure of) the set of finite critical trajectories of (4.2). (One can show that DK_Ψ is an imbedded (multi)graph in $\mathcal{R}$. Here by a *multigraph* on a surface we mean a graph with possibly multiple edges and loops.) Finally, denote by $DK_\Psi^0 \subseteq DK_\Psi$ the subgraph of DK_Ψ consisting of (the closure of) the set of finite critical trajectories whose both ends are zeros of Ψ.

A non-critical trajectory $\gamma_{z_0}(t)$ of a meromorphic Ψ is called *closed* if $\exists\, T > 0$ such that $\gamma_{z_0}(t+T) = \gamma_{z_0}(t)$ for all $t \in \mathbb{R}$. The least such T is called the *period* of γ_{z_0}. A quadratic differential Ψ on a compact Riemann surface $\mathcal{R}$ without boundary is called *Strebel* if the set of its closed trajectories covers $\mathcal{R}$ up to a set of Lebesgue measure zero.

Going back to Cauchy transforms, we formulate the following necessary condition of the existence of a motherbody measure for (4.1).

Proposition 6. *Assume that equation* (4.1) *admits a signed motherbody measure* μ. *Denote by* $D(z) = Q^2(z) - 4P(z)R(z)$ *the discriminant of equation* (4.1). *Then the following two conditions hold:*

(i) *any connected smooth curve in the support of* μ *coincides with a horizontal trajectory of the quadratic differential*
$$\Theta = -\frac{D(z)}{P^2(z)}dz^2 = \frac{4P(z)R(z) - Q^2(z)}{P^2(z)}dz^2. \tag{4.2}$$

(ii) *the support of* μ *includes all branching points of* (4.1).

Remark. Observe that if $P(z)$ and $Q(z)$ are coprime, the set of all branching points coincides with the set of all zeros of $D(z)$. In particular, in this case part (ii) of Proposition 6 implies that the set DK_Θ^0 for the differential Θ should contain all zeros of $D(z)$.

Remark. Proposition 6 applied to quadratic differential $Q(z)\,dz^2$ of Theorem 1 implies Theorem 2.

Proof. The fact that every curve in $\operatorname{supp}(\mu)$ should coincide with some horizontal trajectory of (4.2) is well known and follows from the Plemelj–Sokhotsky's formula. It is based on the local observation that if a real measure $\mu = \frac{1}{\pi}\frac{\partial \mathcal{C}}{\partial \bar{z}}$ is supported on a smooth curve γ, then the tangent to γ at any point $z_0 \in \gamma$ should be perpendicular to $\overline{\mathcal{C}_1(z_0)} - \overline{\mathcal{C}_2(z_0)}$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are the one-sided limits of $\mathcal{C}$ when $z \to z_0$, see, e.g., [5]. (Here $^-$ stands for the usual complex conjugation.) Solutions of (4.1) are given by
$$\mathcal{C}_{1,2} = \frac{-Q(z) \pm \sqrt{Q^2(z) - 4P(z)R(z)}}{2P(z)},$$
their difference being
$$\mathcal{C}_1 - \mathcal{C}_2 = \frac{\sqrt{Q^2(z) - 4P(z)R(z)}}{P(z)}.$$

Since the tangent line to the support of the real motherbody measure μ satisfying (4.1) at its arbitrary smooth point z_0, is orthogonal to $\overline{\mathcal{C}_1(z_0)}-\overline{\mathcal{C}_2(z_0)}$, it is exactly given by the condition $\frac{4P(z_0)R(z_0)-Q^2(z_0)}{P^2(z_0)}dz^2>0$. The latter condition defines the horizontal trajectory of Θ at z_0.

Finally the observation that supp μ should contain all branching points of (4.1) follows immediately from the fact that $\mathcal{C}_\mu$ is a well-defined univalued function in $\mathbb{C}\setminus$ supp μ. □

In many special cases statements similar to Proposition 6 can be found in the literature, see, e.g., recent [1] and references therein.

Proposition 6 allows us, under mild nondegeneracy assumptions, to formulate necessary and sufficient conditions for the existence of a motherbody measure for (4.1) which however are difficult to verify. Namely, let $\Gamma\subset\mathbb{CP}^1\times\mathbb{CP}^1$ with affine coordinates $(\mathcal{C},z)$ be the algebraic curve given by (the projectivization of) equation (4.1). Γ has bidegree $(2,n{+}2)$ and is hyperelliptic. Let $\pi_z:\Gamma\to\mathbb{C}$ be the projection of Γ on the z-plane $\mathbb{CP}^1$ along the $\mathcal{C}$-coordinate. From (4.1) we observe that π_z induces a branched double covering of $\mathbb{CP}^1$ by Γ. If $P(z)$ and $Q(z)$ are coprime and if $\deg D(z)=2n+2$, the set of all branching points of $\pi_z:\Gamma\to\mathbb{CP}^1$ coincides with the set of all zeros of $D(z)$. (If $\deg D(z)<2n+2$, then ∞ is also a branching pont of π_z of multiplicity $2n+2-\deg D(z)$.) We need the following lemma.

Lemma 6. *If $P(z)$ and $Q(z)$ are coprime, then at each pole of* (4.1)*, i.e., at each zero of $P(z)$, only one of two branches of Γ goes to ∞. Additionally the residue of this branch at this zero equals that of $-\frac{Q(z)}{P(z)}$.*

Proof. Indeed if $P(z)$ and $Q(z)$ are coprime, then no zero z_0 of $P(z)$ can be a branching point of (4.1) since $D(z_0)\neq 0$. Therefore only one of two branches of Γ goes to ∞ at z_0. More exactly, the branch $\mathcal{C}_1=\frac{-Q(z)+\sqrt{Q^2(z)-4P(z)R(z)}}{2P(z)}$ attains a finite value at z_0 while the branch $\mathcal{C}_2=\frac{-Q(z)-\sqrt{Q^2(z)-4P(z)R(z)}}{2P(z)}$ goes to ∞ where we use the agreement that $\lim_{z\to z_0}\sqrt{Q^2-4P(z)R(z)}=Q(z_0)$. Now consider the residue of the branch $\mathcal{C}_2$ at z_0. Since residues depend continuously on the coefficients $(P(z),Q(z),R(z))$ it suffices to consider only the case when z_0 is a simple zero of $P(z)$. Further if z_0 is a simple zero of $P(z)$, then

$$Res(\mathcal{C}_2,z_0)=\frac{-2Q(z_0)}{2P'(z_0)}=Res\left(-\frac{Q(z)}{P(z)},z_0\right),$$

which completes the proof. □

By Proposition 6 (besides the obvious condition that (4.1) has a real branch near ∞ with the asymptotics $\frac{\alpha}{z}$ for some $\alpha\in\mathbb{R}$) the necessary condition for (4.1) to admit a motherbody measure is that the set DK^0_Θ for the differential (4.2) contains all branching points of (4.1), i.e., all zeros of $D(z)$. Consider $\Gamma_{cut}:=\Gamma\setminus\pi_z^{-1}(DK^0_\Theta)$. Since DK^0_Θ contains all branching points of π_z, Γ_{cut} consists of some number of open sheets, each projecting diffeomorphically on its image in $\mathbb{CP}^1 \setminus DK^0_\Theta$. (The number of sheets in Γ_{cut} equals to twice the number of connected components in $\mathbb{C}\setminus DK^0_\Theta$.) Observe that since we have chosen a real branch of (4.1) at infinity with the asymptotics $\frac{\alpha}{z}$, we have a marked point $p_{br} \in \Gamma$ over ∞. If we additionally assume that $\deg D(z) = 2n+2$, then ∞ is not a branching point of π_z and therefore $p_{br} \in \Gamma_{cut}$.

Lemma 7. *If* $\deg D(z) = 2n+2$, *then any choice of a spanning (multi)subgraph* $G \subset DK^0_\Theta$ *with no isolated vertices induces the unique choice of the section* S_G *of* Γ *over* $\mathbb{CP}^1 \setminus G$ *which:*

a) *contains* p_{br}*;*
b) *is discontinuous at any point of* G*;* c) *is projected by* π_z *diffeomorphically onto* $\mathbb{CP}^1 \setminus G$.

Here by a spanning subgraph we mean a subgraph containing all the vertices of the ambient graph. By a section of Γ over $\mathbb{CP}^1 \setminus G$ we mean a choice of one of two possible values of Γ at each point in $\mathbb{CP}^1 \setminus G$. After these clarifications the proof is evident.

Observe that the section S_G might attain the value ∞ at some points, i.e., contain some poles of (4.1). Denote the set of poles of S_G by Poles_G. Now we can formulate our necessary and sufficient conditions.

Theorem 3. *Assume that the following conditions are valid:*

(i) *equation* (4.1) *has a real branch near* ∞ *with the asymptotic behavior* $\frac{\alpha}{z}$ *for some* $\alpha \in \mathbb{R}$*;*
(ii) $P(z)$ *and* $Q(z)$ *are coprime, and the discriminant* $D(z) = Q^2(z) - 4P(z)R(z)$ *of equation* (4.1) *has degree* $2n+2$*;*
(iii) *the set* DK^0_Θ *for the quadratic differential* Θ *given by* (4.2) *contains all zeros of* $D(z)$*;*
(iv) Θ *has no closed horizontal trajectories.*

Then (4.1) *admits a real motherbody measure if and only if there exists a spanning (multi)subgraph* $G \subseteq DK^0_\Theta$ *with no isolated vertices, such that all poles in* Poles_g *are simple and all their residues are real, see notation above.*

Proof. Indeed assume that (4.1) satisfying (ii) admits a real motherbody measure μ. Assumption (i) is obviously neccesary for the existence of a real motherbody measure and the necessity of assumption (iii) follows from Proposition 6 if (ii) is satisfied. The support of μ consists of a finite number of curves and possibly a finite number of isolated points. Since each curve in the support of μ is a trajectory of Θ and Θ has no closed trajectories, then the whole support of μ consists of finite critical trajectories of Θ connecting its zeros, i.e., belongs to DK^0_Θ. Moreover the support of μ should contain sufficiently many finite critical trajectories of Θ such that they include all the branching points of (4.1). By (ii) these are exactly all zeros of $D(z)$. Therefore the union of finite critical trajectories of Θ belonging to the support of μ is a spanning (multi)graph of DK^0_Θ without isolated vertices. The isolated points in the support of μ are necessarily the poles of (4.1). Observe that the Cauchy transform of any (complex-valued) measure can only have simple poles (as opposed to the Cauchy transform of a more general distribution). Since μ is real the residue of its Cauchy transform at each pole must be real as well. Therefore the existence of a real motherbody under the assumptions (i)–(iv) implies the existence of a spanning (multi)graph G with the above properties. The converse is also immediate. □

Remark. Observe that if (i) is valid, then assumptions (ii) and (iv) are generically satisfied. Notice however that (iv) is violated in the special case when $Q(z)$ is absent. Additionally, if (iv) is satisfied, then the number of possible motherbody measures is finite. On the other hand, it is the assumption (iii) which imposes severe additional restrictions on admissible triples $(P(z), Q(z), R(z))$. At the moment the authors have no information about possible cardinalities of the sets Poles_G introduced above. Thus it is difficult to estimate the number of conditions required for (4.1) to admit a motherbody measure. Theorem 3 however leads to the following sufficient condition for the existence of a real motherbody measure for (4.1).

Corollary 2. *If, additionally to assumptions* (i)–(iii) *of Theorem* 3, *one assumes that all roots of* $P(z)$ *are simple and all residues of* $\frac{Q(z)}{P(z)}$ *are real, then* (4.1) *admits a real motherbody measure.*

Proof. Indeed if all roots of $P(z)$ are simple and all residues of $\frac{Q(z)}{P(z)}$ are real, then all poles of (4.1) are simple with real residues. In this case for any choice of a spanning (multi)subgraph G of DK_Θ^0, there exists a real motherbody measure whose support coincides with G plus possibly some poles of (4.1). Observe that if all roots of $P(z)$ are simple and all residues of $\frac{Q(z)}{P(z)}$ are real one can omit assumption (iv). In case when Θ has no closed trajectories, then all possible real motherbody measures are in a bijective correspondence with all spanning (multi)subgraphs of DK_Θ^0 without isolated vertices. In the opposite case such measures are in a bijective correspondence with the unions of a spanning (multi)subgraph of DK_Θ^0 and an arbitrary (possibly empty) finite collection of closed trajectories. □

5. Does weak convergence of Jacobi polynomials imply stronger forms of convergence?

Observe that, if one considers an arbitrary sequence $\{s_n(z)\},\ n = 0, 1, \dots$ of monic univariate polynomials of increasing degrees, then even if the sequence $\{\theta_n\}$ of their root-counting measures weakly converges to some limiting probability measure Θ with compact support in $\mathbb{C}$, in general, it is not true that the roots of s_n stay on some finite distance from $\operatorname{supp}\Theta$ for all n simultaneously. Similarly nothing can be said in general about the weak convergence of the sequence $\{\theta'_n\}$ of the root-counting measures of $\{s'_n(z)\}$. However we have already seen that the situation with sequences of Jacobi polynomials seems to be different, compare Proposition 3.

In the present appendix we formulate a general conjecture (and give some evidence of its validity) about sequences of Jacobi polynomials as well as sequences of more general polynomial solutions of a special class of linear differentials equations which includes Riemann's differential equation.

Consider a linear ordinary differential operator

$$\mathfrak{d}(z) = \sum_{i=1}^{k} Q_j(z) \frac{d^j}{dz^j} \tag{5.1}$$

with polynomial coefficients. We say that (5.1) is *exactly solvable* if a) $\deg Q_j \le j$, for all $j = 1, \dots, k$; b) there exists at least one value j_0 such that $\deg Q_{j_0}(z) = j_0$. We say that an exactly solvable operator (5.1) is *non-degenerate* if $\deg Q_k = k$.

Observe that any exactly solvable operator $\mathfrak{d}(z)$ has a unique (up to a constant factor) eigenpolynomial of any sufficiently large degree, see, e.g., [5]. Fixing an arbitrary monic polynomial $Q_k(z)$ of degree k, consider the family $\mathcal{F}_{Q_k}$ of all exactly solvable operators of the form (5.1) whose leading term is $Q_k(z)\frac{d^k}{dz^k}$. ($\mathcal{F}_{Q_k}$ is a complex affine space of dimension $\binom{k+1}{2} - 1$.) Given a sequence $\{\mathfrak{d}_n(z)\}$ of exactly solvable operators from $\mathcal{F}_{Q_k}$ of the form

$$\mathfrak{d}_n(z) = Q_k(z)\frac{d^k}{dz^k} + \sum_{i=1}^{k-1} Q_{j,n}(z) \frac{d^j}{dz^j},$$

we say that this sequence has a *moderate growth* if, for each $j = 1, \dots, k-1$, the sequence of polynomials $\left\{\frac{Q_{j,n}(z)}{n^{k-j}}\right\}$ has all bounded coefficients. (Recall that $\forall n$, $\deg Q_{j,n} \le j$.)

Conjecture 1. [1] *For any sequence $\{\mathfrak{d}_n(z)\}$ of exactly solvable operators of moderate growth, the union of all roots of all the eigenpolynomials of all $\mathfrak{d}_n(z)$ is bounded in $\mathbb{C}$.*

Now take a sequence $\{s_n(z)\}$, $\deg s_n = n$ of polynomial eigenfunctions of the sequence of operators $\mathfrak{d}_n(z) \in \mathcal{F}_{Q_k}$. (Observe that, in general, we have a different exactly solvable operator for each eigenpolynomial but with the same leading term.)

Conjecture 2. *In the above notation, assume that $\{\mathfrak{d}_n(z)\}$ is a sequence of exactly solvable operators of moderate growth and that $\{s_n(z)\}$ is the sequence of their eigenpolynomials (i.e., $s_n(z)$ is the eigenpolynomial of $\mathfrak{d}_n(z)$ of degree n) such that:*

a) *the limits $\widetilde{Q}_j(z) := \lim_{n\to\infty} \frac{1}{n^{k-j}} Q_{j,n}(z)$, $j = 1, \dots, k-1$ exist;*
b) *the sequence $\{\theta_n\}$ of the root-counting measures of $\{s_n(z)\}$ weakly converges to a compactly supported probability measure Θ in $\mathbb{C}$,*

then

[1] Conjecture 1 was disproved in a very recent article [14].

(i) *the Cauchy transform* $\mathcal{C}_\Theta$ *of* Θ *satisfies a.e. in* $\mathbb{C}$ *the algebraic equation*

$$Q_k(z)\left(\frac{\mathcal{C}_\Theta}{\gamma}\right)^k + \sum_{j=1}^{k-1} \widetilde{Q}_j(z)\left(\frac{\mathcal{C}_\Theta}{\gamma}\right)^j = 1, \tag{5.2}$$

where $\gamma = \lim_{n\to\infty} \frac{\sqrt[k]{\lambda_n}}{n}$, λ_n *being the eigenvalue of* $s_n(z)$.

(ii) *for any positive* $\epsilon > 0$, *there exist* n_ϵ *such that, for* $n \geq n_\epsilon$, *all roots of all eigenpolynomials* $s_n(z)$ *are located within* ϵ*-neighborhood of* $\operatorname{supp}\Theta$, *i.e., the weak convergence of* $\theta_n \to \Theta$ *implies a stronger form of this convergence.*

Certain cases of Part (i) of the above Conjecture are settled in [5] and [9] and a version of Part (ii) is discussed in an unpublished preprint [11].

Now we present some partial confirmation of the above conjectures. Consider the family of linear differential operators of second order depending on parameter λ and given by

$$T_\lambda = Q_2(z)\frac{d^2}{dz^2} + \big(Q_1(z)\lambda + P_1(z)\big)\frac{d}{dz} + (\lambda^2 + p\lambda + q)Q_0, \tag{5.3}$$

where $Q_2(z)$ is a quadratic polynomial in z, $Q_1(z)$ and $P_1(z)$ are polynomials in z of degree at most 1, and Q_0 is a non-vanishing constant. (Observe that our use of parameter λ here is the same as of the parameter γ in the latter Conjecture.)

Denote $Q_i(z) = \sum_{j=0}^{i} q_{ji}z^j$, $i = 0, 1, 2$ and put $P_1 = p_{11}z + p_{01}$. The quadratic polynomial

$$q_{22} + q_{11}t + q_{00}t^2 \tag{5.4}$$

is called the *characteristic polynomial* of T_λ. Here $q_{22} \neq 0$ and $q_{00} = Q_0 \neq 0$.

Definition 5. We say that the family T_λ has a *generic type* if the roots of (5.4) have distinct arguments (and in particular 0 is not a root of (5.4) which is guaranteed by $q_{22} \neq 0$ together with $q_{00} \neq 0$), comp. [9].

Below we will denote the roots of characteristic polynomial (5.4) by α_1 and α_2. Thus T_λ has a generic type if and only if $\arg\alpha_1 \neq \arg\alpha_2$.

Lemma 8. *Equation* (5.4) *has two roots with the same arguments if and only if* $q_{22}q_{00} = \rho q_{11}^2$, *where* $0 \leq \rho \leq \frac{1}{4}$.

Proof. Straightforward calculation, see Example 1 of [10]. □

Lemma 9. *In the above notation, for a family* T_λ *of generic type, there exists a positive integer* N *such that, for any integer* $n \geq N$, *there exist two eigenvalues* $\lambda_{1,n}$ *and* $\lambda_{2,n}$ *such that the differential equation*

$$T_\lambda(y) = 0 \tag{5.5}$$

has a polynomial solution of degree n. *Moreover,* $\lim_{n\to\infty} \frac{\lambda_{i,n}}{n} = \alpha_i$ *where* α_1, α_2 *are the roots of the characteristic polynomial of* T_λ.

Proof. Observe that for any $\lambda \in \mathbb{C}$, the operator T_λ acts on each linear space Pol_n of all polynomials of degree at most n, $n = 0, 1, 2, \ldots$, and its matrix presentation $(c_{ij})_{i,j=0}^n$ in the standard monomial basis $(1, z, z^2, \ldots, z^n)$ of Pol_n is an upper-triangular matrix with diagonal entries

$$c_{jj} = j(j-1)q_{22} + jq_{11} + q + (jq_{11} + p)\lambda + q_{00}\lambda^2.$$

Therefore, for any given non-negative integer n, we have a (unique) polynomial solution of (5.5) of degree n if and only if $c_{nn} = 0$ but $c_{jj} \neq 0$ for $0 \leq j < n$. The asymptotic formula for $\lambda_{i,n}$ follows from the form of the equation $c_{nn} = 0$. The genericity assumption that the equations

$$n(n-1)q_{22} + nq_{11} + q + (nq_{11} + p)\lambda + q_{00}\lambda^2 = 0$$

and

$$j(j-1)q_{22} + jq_{11} + q + (jq_{11} + p)\lambda + q_{00}\lambda^2 = 0$$

should not have a common root, for $0 \leq j < n$ and n sufficiently large, is clearly satisfied if we assume that the characteristic equation does not have two roots with the same argument. $\square$

We can now prove the following stronger result.

Proposition 7. *For a general type family of differential operators T_λ of the form* (5.3)*, all roots of all polynomial solutions of $T_\lambda(p) = 0$, $\lambda \in \mathbb{C}$ are located in some compact set $K \subset \mathbb{C}$.*

Proof. Since T_λ is assumed to be of general type, one gets $Q_0 \neq 0$. Therefore, without loss of generality we can assume that $Q_0 = 1$ in (5.5). Let $\{p_n\}, \deg(p_n) = n$ be a sequence of eigenpolynomials for (5.5), and assume that $\lim_{n\to\infty} \frac{\lambda_n}{n} = \alpha$. (By Lemma 9, α equals either α_1 or α_2.) Define $w_n = \frac{p_n'}{\lambda_n p_n}$ and notice that $p_n = e^{\lambda_n \int w_n dz}$. We then have

$$p_n' = \lambda_n w_n p_n; \; p_n'' = (\lambda_n^2 w_n^2 + \lambda_n w_n')p_n.$$

Substituting these expressions in (5.5), we obtain:

$$p_n(Q_2(z)(\lambda_n^2 w_n^2(z) + \lambda_n w_n'(z)) + \lambda_n^2 Q_1(z) w_n(z) + P_1(z)\lambda_n w_n(z) + \lambda_n^2 + p\lambda_n + q = 0.$$

For each fixed n, near $z = \infty$ we can conclude that

$$Q_2(z)(\lambda_n^2 w_n^2(z) + \lambda_n w_n'(z)) + \lambda_n^2 Q_1(z) w_n(z) + P_1(z)\lambda_n w_n(z) + \lambda_n^2 + p\lambda_n + q = 0.$$

This relation defines a rational function w_n near infinity. We will show that the sequence $\{w_n\}$ converges uniformly to an analytic function w in a sufficiently small disc around ∞. Moreover w does not vanish identically. Proposition 7 will immediately follow from this claim. Introducing $t = \frac{1}{z}$, one obtains

$$\widetilde{Q}_2\left(\left(\frac{w_n}{t}\right)^2 - \frac{1}{\lambda_n} w_n'\right) + \widetilde{Q}_1\left(\frac{w_n}{t}\right) + \frac{1}{\lambda_n}\widetilde{P}_1\left(\frac{w_n}{t}\right) + 1 + \frac{p}{\lambda_n} + \frac{q}{\lambda_n^2} = 0,$$

where $\widetilde{Q}_2(t) := t^2 Q_2(1/t)$, $\widetilde{Q}_1(t) := tQ_1(1/t)$ and $\widetilde{P}_1(t) := tP_1(1/t)$. Expand $w_n = c_1 t + c_2 t^2 + \cdots$ in a power series around ∞, i.e., around $t = 0$. (By a slight abuse of notation, we temporarily disregard the fact that the coefficients c_k depend on n until we make their proper estimate.) Set $(w_n/t)^2 = b_0 + b_1 t + \cdots$. Then

$$b_k = c_1 c_{k+1} + c_2 c_k + \cdots + c_k c_2 + c_{k+1} c_1.$$

Finally, introduce $\epsilon_n = 1/\lambda_n$. Using these notations we obtain the following system of recurrence relations for the coefficients c_k:

$$q_{22}c_1^2 + (q_{11} - \epsilon_n q_{22} + \epsilon_n p_{11})c_1 + 1 + \epsilon_n p + \epsilon_n^2 q = 0,$$
$$q_{22}(b_1 - 2\epsilon_n c_2) + q_{12}(b_0 - \epsilon_n c_1) + (q_{11} + \epsilon_n p_{11})c_2 + (q_{01} + \epsilon_n p_{01})c_1 = 0,$$
$$q_{22}(b_2 - 3\epsilon_n c_3) + q_{12}(b_1 - 2\epsilon_n c_2) + q_{02}(b_0 - \epsilon_n c_1)$$
$$+ (q_{11} + \epsilon_n p_{11})c_3 + (q_{01} + \epsilon_n p_{01})c_2 = 0,$$

and, more generally,

$$q_{22}(b_k - (k+1)\epsilon_n c_{k+1}) + q_{12}(b_{k-1} - k\epsilon_n c_k) + q_{02}(b_{k-2} - (k-1)\epsilon_n c_{k-1})$$
$$+ (q_{11} + \epsilon_n p_{11})c_{k+1} + (q_{01} + \epsilon_n p_{01})c_k = 0 \quad \text{for} \quad k \geq 2.$$

Therefore, for any given n, we get 2 possible values for $c_1(n)$, which tend to the roots of $q_{22}t^2 + q_{11}t + 1 = 0$ as $n \to \infty$. Notice that $c_1(n) \to \frac{1}{\alpha}$ as $n \to \infty$. Choosing one of two possible values for c_1, we uniquely determine the remaining coefficients (as rational functions of the previously calculated coefficients). Introducing $\tilde{b}_k = b_k - 2c_1 c_{k+1}$, we can observe that $\tilde{b}_k$ is independent of c_{k+1} and we obtain the following explicit formulas:

$$c_2 = -\frac{q_{12}(c_1^2 - \epsilon_n c_1) + (q_{01} + \epsilon_n p_{01})c_1}{(2c_1 - 2\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11}},$$

$$c_3 = -\frac{q_{22}\tilde{b}_2 + q_{12}(b_1 - 2\epsilon_n c_2) + q_{02}(b_0 - \epsilon_n c_1) + (q_{01} + \epsilon_n p_{01})c_2}{(2c_1 - 3\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11}},$$

and more generally,

$$c_k = -\frac{q_{22}\tilde{b}_{k-1} + q_{12}(b_{k-2} - (k-1)\epsilon_n c_{k-1})}{(2c_1 - k\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11}}$$
$$+ \frac{q_{02}(b_{k-2} - (k-3)\epsilon_n c_{k-3}) + (q_{01} + \epsilon_n p_{01})c_{k-1}}{(2c_1 - k\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11}}.$$

We will now include the dependence of c_k on n and show that the coefficients $c_k(n)$ are majorated by the coefficients of a convergent power series independent of n. First we show that the denominators in these recurrence relations are bounded from below. Notice that under our assumption, the rational functions w_n exist and have a power series expansion near $z = \infty$ with coefficients given by the above recurrence relations. Therefore the denominators in these recurrences do not vanish. Notice also that $\epsilon_n \simeq \frac{c_1(n)}{n}$ asymptotically. For fixed k, it is therefore clear that the limits

$$\lim_{n\to\infty} (2c_1(n) - k\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11} = \lim_{n\to\infty} 2c_1(n)q_{22} + q_{11}$$

vanish if and only if the characteristic polynomial (5.4) has a double root. We must however find a uniform bound for $c_k(n)$ valid for all k simultaneously. Indeed, there might exist a subsequence $I \subset \mathbb{N}$ of k_n such that

$$\lim_{n\in I;n\to\infty}(2c_1(n) - k_n\epsilon_n)q_{22} + q_{11} + \epsilon_n p_{11} = 0. \tag{5.6}$$

But this implies, using the asymptotics of $c_1(n)$ and ϵ_n, the existence of a real number r such that $\frac{1-r}{\alpha} = -\frac{q_{22}}{2q_{11}}$ which is clearly impossible if the characteristic equation does not have two roots with the same argument. Thus we have established a positive lower bound for the absolute value of the denominators in the recurrence relations for the coefficients c_k. The latter circumstance gives us a possibility of majorizing the coefficients $c_k(n)$ independently of k and n. Namely, if there is a unbounded sequence $k_n\epsilon_n$, then we can factor it out from the rational functions in the recurrence. The existence of the sequence mentioned above follow from an elementary lemma stated below, which we leave without a proof. Thus, Proposition 7 is now settled. □

Lemma 10. *Consider a recurrence relation $c_{m+1} = P_m(c_1, \ldots, c_m)$ where each P_m is a polynomial and assume that $d_{m+1} = Q_m(d_1, \ldots, d_m)$ is a similar recurrence relation whose polynomials have all positive coefficients. If the polynomials under consideration satisfy the inequalities*

$$|P_m(z_1, \ldots, z_m)| \le Q_m(|z_1|, \ldots, |z_m|),$$

then the power series $\sum c_i z^i$ is dominated by the series $\sum d_i z^i$ whenever $d_1 \ge |c_1|$.

6. Domain configurations of normalized quadratic differentials

Let $Q(z; a, b, c)\, dz^2$ be a quadratic differential of the form (1.5). Multiplying $Q(z; a, b, c)\, dz^2$ by a non-zero constant $A \in \mathbb{C}$, we rescale the corresponding Q-metric $|Q|^{1/2}\,|dz|$ by a positive constant $|A|^{1/2}$. Hence $A\,Q(z; a, b, c)\, dz^2$ has the same geodesics as the quadratic differential $Q(z; a, b, c)\, dz^2$ has. Obviously, multiplication does not affect the homotopic classes. Thus, while studying geodesics of the quadratic differential $Q(z; a, b, c)\, dz^2$, we may assume without loss of generality that it has the form

$$Q(z)\, dz^2 = -\frac{(z-p_1)(z-p_2)}{(z-1)^2(z+1)^2}\, dz^2. \tag{6.1}$$

In Sections 6–9, we will work with the generic case; i.e., we assume that

$$p_1 \ne \pm 1, \quad p_2 \ne \pm 1, \quad p_1 \ne p_2, \tag{6.2}$$

unless otherwise is mentioned. Some typical configurations in the limit (or non-generic) cases are shown in Figures 5a–5g. Expanding $Q(z)$ into Laurent series at $z = \infty$, we obtain

$$Q(z) = -\frac{1}{z^2} + \text{higher degrees of } z \qquad \text{as } z \to \infty. \tag{6.3}$$

Since the leading coefficient in the series expansion (6.3) is real and negative it follows that $Q(z)\,dz^2$ has a circle domain D_∞ centered at $z=\infty$. The boundary $L_\infty = \partial D_\infty$ of D_∞ consists of a finite number of critical trajectories of the quadratic differential $Q(z)\,dz^2$ and therefore L_∞ contains at least one of the zeros p_1 and p_2 of $Q(z)\,dz^2$.

Next, we will discuss possible trajectory structures of $Q(z)\,dz^2$ on the complement $D_0 = \mathbb{C}\setminus\overline{D}_\infty$. As we have mentioned in Section 3, according to the Basic Structure Theorem, [22, Theorem 3.5], the domain configuration of a quadratic differential $Q(z)\,dz^2$ on $\overline{\mathbb{C}}$, which will be denoted by $\mathcal{D}_Q$, may include circle domains, ring domains, strip domains, end domains, and density domains. For the quadratic differential (6.1), by the Three Pole Theorem [22, Theorem 3.6], there are no density domains in its domain configuration $\mathcal{D}_Q$. In addition, since $Q(z)\,dz^2$ has only three poles of order two each, the domain configuration $\mathcal{D}_Q$ does not contain end domains and may contain at most three circle domains centered at $z=\infty$, $z=-1$, and $z=1$.

We note here that $\mathcal{D}_Q$ may have strip domains (also called *bilaterals*) with vertices at the double poles $z=-1$ and $z=1$ but $\mathcal{D}_Q$ does not have ring domains. Indeed, if there were a ring domain $\widehat{D}\subset D_0$ with boundary components l_1 and l_2 then, by the Basic Structure Theorem, each component must contain a zero of $Q(z)\,dz^2$. In particular, $p_1\neq p_2$ in this case. Suppose that l_1 contains a zero p_1 and that $p_1\in L_\infty$. Then L_∞ contains a critical trajectory γ', which has both its end points at p_1. There is one more critical trajectory γ'', which has one of its end points at p_1. This trajectory γ'' is either lies on the boundary of the circle domain D_∞ or it lies on the boundary of the ring domain $\widehat{D}$. Therefore the second end point of γ'' must be at a zero of $Q(z)\,dz^2$. Since the only remaining zero is p_2, which lies on the boundary component l_2 not intersecting l_1, we obtain a contradiction with our assumption. The latter shows that $\mathcal{D}_Q$ does not have ring domains.

Next, we will classify topological types of domain configurations according to the number of circle domains in $\mathcal{D}_Q$. The first digit in our further classifications stands for the section where this classification is introduced. The second and further digits will denote the case under consideration.

6.1. Assume first that $\mathcal{D}_Q$ contains three circle domains $D_\infty\ni\infty$, $D_{-1}\ni -1$, and $D_1\ni 1$. Then, of course, there are no strip domains in $\mathcal{D}_Q$. In this case, the domains D_∞, D_{-1}, D_1 constitute an extremal configuration of the Jenkins extremal problem for the weighted sum of reduced moduli with appropriate choice of positive weights α_∞, α_{-1}, and α_1; see, for example, [34], [31], [32]. More precisely, the problem is to find all possible configurations realizing the following maximum:

$$\max\ \left(\alpha_\infty^2 m(B_\infty,\infty)+\alpha_{-1}^2 m(B_{-1},-1)+\alpha_1^2 m(B_1,1)\right) \tag{6.4}$$

over all triples of non-overlapping simply connected domains $B_\infty\ni\infty$, $B_{-1}\ni -1$, and $B_1\ni 1$. Here, $m(B,z_0)$ stands for the reduced module of a simply connected domain B with respect to the point $z_0\in B$; see [22, p.24].

Since the extremal configuration of problem (6.4) is unique it follows that the domains D_∞, D_{-1}, and D_1 are symmetric with respect to the real axis. In particular, the zeros p_1 and p_2 are either both real or they are complex conjugates of each other. Of course, this symmetry property of zeros can be derived directly from the fact that the leading coefficient of the Laurent expansion of $Q(z)$ at each its pole is negative in the case under consideration. We have three essentially different possible positions for the zeros:

(a) $-1 < p_2 < p_1 < 1$,
(b) $1 < p_2 < p_1$ or $p_1 < p_2 < -1$,
(c) $p_1 = \overline{p}_2 = p$, where $\Im p > 0$.

We note here that in the case when $-1 < p_2 < 1$ and, in addition, $p_1 > 1$ or $p_1 < -1$ the domain configuration $\mathcal{D}_Q$ must contain a strip domain.

Case **(a)**. The trajectory structure of $Q(z)\,dz^2$ corresponding to this case is shown in Figure 1a. There are three critical trajectories: γ_{-1}, which is on the boundary of D_{-1} and has both its end points at $z = p_2$; γ_1, which is on the boundary of D_1 and has both its end points at $z = p_1$, and γ_0, which is the segment $[p_2, p_1]$.

Case **(b)**. An example of a domain configuration for the case $1 < p_2 < p_1$ is shown in Figure 1b. The boundary of D_1 consists of a single critical trajectory γ_1 having both end points at p_2. The boundary of D_{-1} consists of critical trajectories γ_∞, γ_1, and γ_0, which is the segment $[p_2, p_1]$. In the case $p_1 < p_2 < -1$, the domain configuration is similar.

Case **(c)**. Since the domain configuration is symmetric, p_1 and p_2 both belong to the boundary of D_∞. Furthermore, there are three critical trajectories: γ_{-1}, which joins p_1 and p_2 and intersects the real axis at some point $d_{-1} < -1$, γ_1, which joins p_1 and p_2 and intersects the real axis at some point $d_1 > 1$, and γ^0, which joins p_1 and p_2 and intersects the real axis at some point d_0, $-1 < d_0 < 1$. In this case, $\gamma_1 \cup \gamma_0 \subset \partial D_1$, $\gamma_{-1} \cup \gamma_0 \subset \partial D_{-1}$. An example of a domain configuration of this type is shown in Figure 1c.

6.2. Next we consider the case when $\mathcal{D}_Q$ has exactly two circle domains. Suppose that these domains are $D_\infty \ni \infty$ and $D_{-1} \ni -1$. In this case it is not difficult to see that L_∞ contains exactly one zero. Indeed, if $p_1, p_2 \in L_\infty$, then L_∞ must contain one or two critical trajectories joining p_1 and p_2. Suppose that L_∞ contains one such trajectory, call it γ_0. Since $p_1, p_2 \in L_\infty$ the boundary of D_∞ must contain a trajectory γ_1, which has both its end points at p_1 and a trajectory γ_{-1}, which has both its end points at p_2. Thus, $\gamma_1 \cup \{p_1\}$ and $\gamma_{-1} \cup \{p_2\}$ each surrounds a simply connected domain, which must contain a critical point of $Q(z)\,dz^2$. This implies that $z = -1$ and $z = 1$ are centers of circle domains of $Q(z)\,dz^2$, which is the case considered in part **6.1(a)**.

If L_∞ contains two critical trajectories joining p_1 and p_2, then there are critical trajectories γ' having one of its end points at p_1 and γ'' having one of its end points at p_2. If $\gamma' = \gamma''$, then $D_0 \setminus \gamma'$ consists of two simply connected domains, which in this case must be circle domains of $Q(z)\,dz^2$ as it is shown in Figure 1c.

If $\gamma' \neq \gamma''$, then each of these trajectories must have its second end point at one of the poles $z = -1$ or $z = 1$. Moreover, if γ' has an end point at $z = -1$ then γ'' must have its end point at $z = 1$. Thus, there is no second circle domain of $Q(z)\,dz^2$ in this case. Instead, there is one circle domain D_∞ and a strip domain, call it G_2, as it shown in Figures 3a–3e.

Now, let p_1 be the only zero of $Q(z)\,dz^2$ lying on L_∞. Then L_∞ consists of a single critical trajectory of $Q(z)\,dz^2$, call it γ_∞, together with its end points, each of which is at p_1. There is one more critical trajectory, call it γ_1^+, that has one of its end points at p_1. Then the second end point of γ_1^+ is either at the point p_2 or at the second-order pole at $z = 1$.

If γ_1^+ terminates at p_2, then there is one more critical trajectory, call it γ_2, having one of its end points at p_2. Since D_{-1} is a circle domain and ∂D_{-1} contains at least one zero of $Q(z)\,dz^2$ it follows that γ_2 belongs to the boundary of D_{-1}. Since γ_2 lies on the boundary of D_{-1} it have to terminate at a finite critical point of $Q(z)\,dz^2$ and the only possibility for this is that γ_2 terminates at p_2. In this case, γ_∞, γ_1^+, and γ_2 divide $\overline{\mathbb{C}}$ into three circle domains, the case which was already discussed in part **6.1(b)**.

Suppose that γ_1^+ joins the points $z = p_1$ and $z = 1$. Then $\mathcal{D}_Q$ contains a strip domain G_1. Since $z = 1$ is the only second-order pole of $Q(z)\,dz^2$, which has a non-negative non-zero leading coefficient, the strip domain G_1 has both its vertices at the point $z = 1$. Furthermore, one side of G_1 consists of two critical trajectories γ_∞ and γ_1^+. Therefore there is a critical trajectory, call it γ_1^- of $Q(z)\,dz^2$ lying on ∂G_1, which joins $z = 1$ and $z = p_2$. Now, the remaining possibility is that the boundary of D_{-1} consists of a single critical trajectory γ_{-1}, which has both its end points at p_2. Then G_1 is the only strip domain in $\mathcal{D}_Q$ and the second side of G_1 consists of the critical trajectories γ_1^- and γ_{-1}. Two examples of a domain configuration of this type, symmetric and non-symmetric, are shown in Figure 2a and Figure 2b.

6.3. Finally, we consider the case when D_∞ is the only circle domain of $Q(z)\,dz^2$. We consider two possibilities.

Case **(a)**. Suppose that both zeros p_1 and p_2 belong to the boundary of D_∞. As we have found in part **6.2** above, the domain configuration in this case consists of the circle domain D_∞ and the strip domain G_2. The boundary of D_∞ consists of two critical trajectories γ_∞^+ and γ_∞^- and their end points, while the boundary of G_2 consists of the trajectories γ_∞^+, γ_∞^-, γ_1, and γ_{-1} and their end points, as it is shown in Figures 3a–3c.

Case **(b)**. Suppose that the boundary L_∞ of D_∞ contains only one zero p_1. Then there is a critical trajectory γ_∞ having both its end points at p_1 such that $L_\infty = \gamma_\infty \cup \{p_1\}$. Since p_1 is a simple zero of $Q(z)\,dz^2$ there is one more critical trajectory having one of its end points at p_1. The second end point of this trajectory is either at the pole $z = 1$, or at the pole $z = -1$, or at the zero $z = p_2$. Depending on which of these possibilities is realized, this trajectory will be denoted by γ_1, or γ_{-1}, or γ_0, respectively. Thus, we have two essentially different subcases.

Case **(b1)**. Suppose that there is a critical trajectory γ_0 joining the zeros p_1 and p_2. Then there are two critical trajectories, call them γ_1 and γ_{-1}, each of which has one of its end point at p_2. We note that $\gamma_1 \neq \gamma_{-1}$. Indeed, if $\gamma_1 = \gamma_{-1}$, then the closed curve $\gamma_1 \cup \{p_2\}$ must enclose a bounded circle domain of $Q(z)\, dz^2$, which does not exist. Furthermore, γ_1 and γ_{-1} both cannot have their second end points at the same pole at $z = 1$ or $z = -1$. If this occurs then again γ_1 and γ_{-1} will enclose a simply connected domain having a single pole of order 2 on its boundary, which is not possible. The remaining possibility is that one of these critical trajectories, let assume that γ_1, joins the zero $z = p_2$ and the pole at $z = 1$ while γ_{-1} joins $z = p_2$ and $z = -1$.

In this case the domain configuration $\mathcal{D}_Q$ consists of the circle domain D_∞ and the strip domain G_2; see Figure 3d and Figure 3e. The boundary of G_2 consists of two sides, call them l_1 and l_2. The side l_1 is the set of boundary points of G_2 traversed by the point z moving along γ_1 from $z = 1$ to $z = p_2$ and then along γ_{-1} from the point $z = p_2$ to $z = -1$. The side l_2 is the set of boundary points of G_2 traversed by the point z moving along γ_1 from $z = 1$ to $z = p_2$, then along γ_0 from $z = p_2$ to $z = p_1$, then along γ_∞ from $z = p_1$ to the same point $z = p_1$, then along γ_0 from $z = p_1$ to $z = p_2$, and finally along γ_{-1} from $z = p_2$ to $z = -1$.

Case **(b2)**. Suppose that there is a critical trajectory γ_1 joining the zero p_1 and the pole $z = 1$. Then there is a strip domain, call it G_1, which has both its vertices at the pole $z = 1$ and has the critical trajectories γ_1 and γ_∞ on one of its sides, call it l_1^1. More precisely, the side l_1^1 is the set of boundary points of G_1 traversed by the point z moving along γ_1 from $z = 1$ to $z = p_1$, then along γ_∞ from $z = p_1$ to the same point $z = p_1$, and then along γ_1 from $z = p_1$ to $z = 1$.

Let l_1^2 denote the second side of G_1. Since a side of a strip domain always has a finite critical point it follows that l_1^2 contains two critical trajectories, call them γ_0^+ and γ_0^-, which join the pole $z = 1$ with zero $z = p_2$. There is one critical trajectory of $Q(z)\, dz^2$, call it γ_{-1}, which has one of its end points at $z = p_2$. Since $z = -1$ is a second-order pole, which is not the center of a circle domain, there should be at least one critical trajectory of $Q(z)\, dz^2$ approaching $z = -1$ at least in one direction. Since the end points of all critical trajectories, except γ_{-1}, are already identified and they are not at $z = -1$, the remaining possibility is that γ_{-1} has its second end point at $z = -1$. In this case there is one more strip domain, call it G_2, which has vertices at the poles $z = 1$ and $z = -1$ and sides l_2^1 and l_2^2. Two examples of configurations with one circle domain and two strip domains, symmetric and non-symmetric, are shown in Figure 4a and Figure 4b. Now we can identify all sides of G_1 and G_2. The side l_1^2 is the set of boundary points of G_1 traversed by the point z moving along γ_0^+ from $z = 1$ to $z = p_2$ and then along γ_0^- from $z = p_2$ to $z = 1$. The side l_2^1 is the set of boundary points of G_2 traversed by the point z moving along γ_0^+ from $z = 1$ to $z = p_2$ and then along γ_{-1} from $z = p_2$ to $z = -1$. Finally, the side l_2^2 is the set of boundary points of G_2 traversed by the point z moving along γ_0^- from $z = 1$ to $z = p_2$ and then along γ_{-1} from $z = p_2$ to $z = -1$; see Figure 4a and Figure 4b.

Case **(b3)**. In the case when there is a critical trajectory joining the zero p_1 and the pole $z=-1$, the domain configuration is similar to one described above, we just have to switch the roles of the poles at $z=1$ and $z=-1$.

Remark 2. We have described above all possible configurations in the generic case; i.e., under conditions (6.2). The remaining special cases can be obtained from the generic case as limit cases when $p_2 \to -1$, when $p_2 \to p_1$; etc. In the case $p_1 = p_2$, possible configurations are shown in Figures 5a–5c.

In the case when $p_2 = -1$, $p_1 \neq \pm 1$, possible configurations are shown in Figures 5d–5g.

In the case when $p_1 = p_2 = 1$, the limit position of critical trajectories is just a circle centers at $z=-1$ with radius 2configuration and in the case when $p_1 = 1$, $p_2 = -1$ there is one critical trajectory which is an open interval from $z=-1$ to $z=1$.

7. How parameters determine the type of domain configuration

Our goal in this section is to identify the ranges of the parameters p_1 and p_2 corresponding to topological types discussed in Section 6. For a fixed p_1 with $\Im p_1 \neq 0$, we will define four regions of the parameter p_2. These regions and their boundary arcs will correspond to domain configurations with specific properties; see Figure 6.

It will be useful to introduce the following notation. For $a \in \mathbb{C}$ with $\Im a \neq 0$, by $L(a)$ and $H(a)$ we denote, respectively, an ellipse and hyperbola with foci at $z=1$ and $z=-1$, which pass through the point $z=a$. If $\Im a \neq 0$, then the set $\mathbb{C} \setminus (L(a) \cup H(a))$ consists of four connected components, which will be denoted by $E_1^+(a)$, $E_1^-(a)$, $E_{-1}^+(a)$, and $E_{-1}^-(a)$. We assume here that $1 \in E_1^+(a)$, $-1 \in E_{-1}^+(a)$, $E_1^-(a) \cap \mathbb{R}_+ \neq \emptyset$, and $E_{-1}^-(a) \cap \mathbb{R}_- \neq \emptyset$. Furthermore, assuming that $\Im a \neq \emptyset$, we define the following open arcs: $L^+(a) = (L(a) \cap \partial E_1^+(a)) \setminus \{a, \bar{a}\}$, $L^-(a) = (L(a) \cap \partial E_{-1}^+(a)) \setminus \{a, \bar{a}\}$, $H^+(a) = (H(a) \cap \partial E_1^+(a)) \setminus \{a, \bar{a}\}$, $H^-(a) = (H(a) \cap \partial E_1^-(a)) \setminus \{a, \bar{a}\}$. Let $l_1(a)$ and $l_{-1}(a)$ be straight lines passing through the points 1 and $\bar{a}$ and -1 and $\bar{a}$, respectively. Let $l_1^+(a)$ and $l_{-1}^+(a)$ be open rays issuing from the points $z=1$ and $z=-1$, respectively, which pass through the point $z=\overline{a}$ and let $l_1^-(a)$ and $l_{-1}^-(a)$ be their complementary rays. The line $l_1(a)$ divides $\mathbb{C}$ into two half-planes, we call them P_1 and P_2 and enumerate such that $P_1 \ni 2$. Similarly, the line $l_{-1}(a)$ divides $\mathbb{C}$ into two half-planes P_3 and P_4, where $P_3 \ni -2$.

Before we state the main result of this section, we recall the reader that the local structure of trajectories near a pole z_0 is completely determined by the leading coefficient of the Laurent expansion of $Q(z)$ at z_0, see [22, Ch. 3]. In particular, for the quadratic differential $Q(z)\,dz^2$ defined by (6.1) we have

$$Q(z) = -\frac{1}{4}\frac{C_1}{(z-1)^2} + \text{higher degrees of } (z-1) \quad \text{as } z \to 1 \tag{7.1}$$

and

$$Q(z) = -\frac{1}{4}\frac{C_{-1}}{(z+1)^2} + \text{higher degrees of } (z+1) \quad \text{as } z \to -1.$$

Then, assuming that $p_1 \neq \pm 1$, $p_2 \neq \pm 1$, we find

$$C_1 = (p_1 - 1)(p_2 - 1) \neq 0 \quad \text{and} \quad C_{-1} = (p_1 + 1)(p_2 + 1) \neq 0. \tag{7.2}$$

A complete description of sets of pairs p_1, p_2 with $\Im p_1 > 0$ corresponding to all possible types of domain configurations discussed in Section 6 is given by the following theorem.

Theorem 4. *Let p_1 with $\Im p_1 > 0$ be fixed. Then the following holds.*

7.A. *The types of domain configurations $\mathcal{D}_Q$ correspond to the following sets of the parameter p_2.*

- **(1)** *If $p_2 = \bar{p}_1$, then the domain configuration $\mathcal{D}_Q$ is of the type* **6.1(c)**.
- **(2)** *If $p_2 \in l_1^+(p_1)\setminus\{\bar{p}_1\}$, then $\mathcal{D}_Q$ has the type* **6.2** *with circle domains $D_\infty \ni \infty$ and $D_1 \ni 1$. Furthermore, if $p_2 \in l_1^+(p_1) \cap E_1^+(p_1)$, then $p_1 \in \partial D_\infty$ and if $p_2 \in l_1^+(p_1) \cap E_{-1}^-(p_1)$, then $p_2 \in \partial D_\infty$.*

  If $p_2 \in l_{-1}^+(p_1) \setminus \{\bar{p}_1\}$, then $\mathcal{D}_Q$ has the type **6.2** *with circle domains $D_\infty \ni \infty$ and $D_{-1} \ni -1$. Furthermore, if $p_2 \in l_{-1}^+(p_1) \cap E_{-1}^+(p_1)$, then $p_1 \in \partial D_\infty$ and if $p_2 \in l_{-1}^+(p_1) \cap E_{-1}^-(p_1)$, then $p_2 \in \partial D_\infty$.*
- **(3a)** *If $p_2 \in L(a) \setminus \{p_1, \bar{p}_1\}$, then the domain configuration $\mathcal{D}_Q$ has type* **6.3(a)**. *Furthermore, if $p_2 \in L^+(p_1)$, then there is a critical trajectory having one end point at p_2, which in other direction approaches the pole $z = 1$. Similarly, if $p_2 \in L^-(p_1)$, then there is a critical trajectory having one end point at p_2, which in other direction approaches the pole $z = -1$.*
- **(3b1)** *If $p_2 \in H(p_1) \setminus \{p_1, \bar{p}_1\}$, then $\mathcal{D}_Q$ has type* **6.3(b1)**. *Furthermore, if $p_2 \in H^+(p_1)$, then there is a critical trajectory having both end points at p_1. If $p_2 \in H^-(p_1)$, then there is a critical trajectory having both end points at p_2.*
- **(3b2)** *In all remaining cases, i.e., if $p_2 \notin L(p_1) \cup H(p_1) \cup l_1^+(p_1) \cup l_{-1}^+(p_1) \cup \{-1, 1\}$, the domain configuration $\mathcal{D}_Q$ belongs to type* **6.3(b2)**. *Furthermore, if $p_2 \in (E_1^+(p_1) \cup E_{-1}^+(p_1)) \setminus (l_1^+(p_1) \cup l_{-1}^+(p_1) \cup \{-1, 1\})$, then $p_1 \in \partial D_\infty$ and if $p_2 \in (E_1^-(p_1) \cup E_{-1}^-(p_1)) \setminus (l_1^+(p_1) \cup l_{-1}^+(p_1))$, then $p_2 \in \partial D_\infty$.*

  In addition, if $p_2 \in E_1^+(p_1)\setminus(l_1^+(p_1)\cup\{1\})$, then the pole $z = 1$ attracts only one critical trajectory of the quadratic differential (6.1), *which has its second end point at $z = p_2$ and if $p_2 \in E_{-1}^-(p_1) \setminus (l_1^+(p_1))$, then the pole $z = 1$ attracts only one critical trajectory of the quadratic differential* (6.1), *which has its second end point at $z = p_1$. If $p_2 \in E_{-1}^+(p_1)\setminus(l_{-1}^+(p_1)\cup\{-1\})$, then the pole $z = -1$ attracts only one critical trajectory of the quadratic differential* (6.1), *which has its second end point at $z = p_2$ and if $p_2 \in E_1^-(p_1)\setminus(l_{-1}^+(p_1))$, then the pole $z = -1$ attracts only one critical trajectory of the quadratic differential* (6.1), *which has its second end point at $z = p_1$.*

7.B. *The local behavior of the trajectories near the poles $z = 1$ and $z = -1$ is controlled by the position of the zero p_2 with respect to the lines $l_1(p_1)$ and $l_{-1}(p_1)$. Precisely, we have the following possibilities.*

(1) *If $p_2 \in l_1^-(p_1)$ or, respectively, $p_2 \in l_{-1}^-(p_1)$, then $Q(z)\,dz^2$ has radial structure of trajectories near the pole $z = 1$ or, respectively, near the pole $z = -1$.*

(2) *If $p_2 \in P_1$ or, respectively, $p_2 \in P_2$, then the trajectories of $Q(z)\,dz^2$ approaching the pole $z = 1$ spiral counterclockwise or, respectively, clockwise.*

If $p_2 \in P_3$ or, respectively, $p_2 \in P_4$, then the trajectories of $Q(z)\,dz^2$ approaching the pole $z = -1$ spiral counterclockwise or, respectively, clockwise.

Proof. **7.A(1).** We have shown in Section 6 that a domain configuration $\mathcal{D}_Q$ of the type **6.1(c)** occurs if and only if $p_2 = \bar{p}_1$. Thus, we have to consider cases **7.A(2)** and **7.A(3)**. We first prove statements about positions of zeros p_1 and p_2 for each of these cases. Then we will turn to statements about critical trajectories.

7.A(2). A domain configuration $\mathcal{D}_Q$ contains exactly two circle domains centered at $z = \infty$ and $z = -1$ if and only if $C_{-1} > 0$ and C_1 is not a positive real number. This is equivalent to the following conditions:

$$\arg(p_1 + 1) = -\arg(p_2 + 1) \qquad \mathrm{mod}\ (2\pi), \tag{7.3}$$

$$\arg(p_1 - 1) \neq -\arg(p_2 - 1) \qquad \mathrm{mod}\ (2\pi). \tag{7.4}$$

Geometrically, equations (7.3) and (7.4) mean that the points p_1 and p_2 lie on the rays issuing from the pole $z = -1$, which are symmetric to each other with respect to the real axis. Furthermore, each ray contains one of these points and $p_1 \neq \bar{p}_2$.

Assuming (7.3), (7.4), we claim that $p_1 \in \partial D_\infty$ if and only if $|p_2+1| < |p_1+1|$. First we prove that the claim is true for all p_2 sufficiently close to $z = -1$ if p_1 is fixed. Arguing by contradiction, suppose that there is a sequence $s_k \to -1$ such that $\arg(s_k+1) = -\arg(p_1+1)$ and $p_1 \in \partial D_{-1}^k$, $s_k \in \partial D_\infty^k$ for all $k = 1, 2, \ldots$ Here $D_{-1}^k \ni -1$ and $D_\infty^k \ni \infty$ denote the corresponding circle domains of the quadratic differential

$$Q_k(z)\,dz^2 = -\frac{(z-p_1)(z-s_k)}{(z-1)^2(z+1)^2}\,dz^2. \tag{7.5}$$

Changing variables in (7.5) via $z = (s_k+1)\zeta - 1$ and then dividing the resulting quadratic differential by $\delta_k = |s_k+1|$, we obtain the following quadratic differential:

$$\widehat{Q}_k(\zeta)\,d\zeta^2 = \frac{\zeta-1}{\zeta^2}\,\frac{|1+p_1| - \delta_k^{-1}(s_k+1)^2\zeta}{(2-(s_k+1)\zeta)^2}\,d\zeta^2. \tag{7.6}$$

We note that the trajectories of $Q_k(z)\,dz^2$ correspond under the mapping $z = (s_k+1)\zeta - 1$ to the trajectories of the quadratic differential $\widehat{Q}_k(\zeta)\,d\zeta^2$. Thus, $\widehat{Q}_k(\zeta)\,d\zeta^2$ has two circle domains $\widehat{D}_{k,\infty} \ni \infty$ and $\widehat{D}_{k,0} \ni 0$. The zeros of $\widehat{Q}_k(\zeta)\,d\zeta^2$ are at the points

$$\zeta_k' = 1 \in \partial\widehat{D}_{k,\infty}, \quad \zeta_k'' = \delta_k|1+p_1|(s_k+1)^{-2} \in \partial\widehat{D}_{k,0}. \tag{7.7}$$

From (7.6), we find that

$$\widehat{Q}_k(\zeta)\, d\zeta^2 \to \widehat{Q}(\zeta)\, d\zeta^2 := \frac{|1+p_1|}{4} \frac{\zeta - 1}{\zeta^2}\, d\zeta^2, \tag{7.8}$$

where convergence is uniform on compact subsets of $\mathbb{C} \setminus \{0\}$. Since

$$\widehat{Q}(\zeta) = -(|1+p_1|/4)\zeta^{-2} + \cdots \qquad \text{as } \zeta \to 0$$

the quadratic differential $\widehat{Q}(\zeta)\, d\zeta^2$ has a circle domain $\widehat{D}$ centered at $\zeta = 0$. Let $\widehat{\gamma}$ be a trajectory of $\widehat{Q}(\zeta)\, d\zeta^2$ lying in $\widehat{D}$ and let $\hat{\gamma}_k$ be an arbitrary trajectory of $\widehat{Q}_k(\zeta)\, d\zeta^2$ lying in the circle domain $\widehat{D}_{k,0}$. Since $\hat{\gamma}_k$ is a $\widehat{Q}_k$-geodesic in its class and by (7.8) we have

$$|\hat{\gamma}_k|_{\widehat{Q}_k} \le |\widehat{\gamma}|_{\widehat{Q}_k} \to |\widehat{\gamma}|_{\widehat{Q}} = |1+p_1|^{1/2} \quad \text{as } k \to \infty. \tag{7.9}$$

On the other hand, conditions (7.7) imply that for every $R > 1$ there is k_0 such that for every $k \ge k_0$ there is an arc τ_k joining the circles $\{\zeta :\, |\zeta| = 1\}$ and $\{\zeta :\, |\zeta| = R\}$, which lies on regular trajectory of the quadratic differential $\widehat{Q}_k(\zeta)\, d\zeta^2$ lying in the circle domain $\widehat{D}_{k,0}$. Then, using (7.6), we conclude that there is a constant $C > 0$ independent on R and k such that

$$|\hat{\gamma}_k|_{\widehat{Q}_k} \ge |\tau_k|_{\widehat{Q}_k} = \int_{\tau_k} \left|\widehat{Q}_k(\zeta)\right|^{1/2} |d\zeta| \ge C \int_1^R \frac{\sqrt{|\zeta| - 1}}{|\zeta|}\, d|\zeta|$$

for all $k \ge k_0$. Since $\int_1^R x^{-1}\sqrt{x-1}\, dx \to \infty$ as $R \to \infty$, the latter equation contradicts equation (7.9). Thus, we have proved that if p_1 is fixed and p_2 is sufficiently close to $z = -1$ then $p_1 \in \partial D_\infty$ and $p_2 \in \partial D_{-1}$.

Now, we fix p_1 with $\Im p_1 \ne 0$ and consider the set A consisting of all points p_2' on the ray $r = \{z :\, \arg(z+1) = -\arg(p_1+1)\}$ such that $p_1 \in \partial D_\infty(p_1, p_2)$ and $p_2 \in \partial D_{-1}(p_1, p_2)$ for all $p_2 \in r$ such that $|p_2 + 1| < |p_2' + 1|$. Here $D_\infty(p_1, p_2)$ and $D_{-1}(p_1, p_2)$ are corresponding circle domains of the quadratic differential (6.1). Our argument above shows that $A \ne \emptyset$. Let $p_2^m \in r$ be such that

$$|p_2^m + 1| = sup_{p_2 \in A}\, |p_2 + 1|.$$

Consider the quadratic differential $Q(z; p_1, p_2^m)\, dz^2$ of the form (6.1) with p_2 replaced by p_2^m. Let $D_\infty(p_1, p_2^m) \ni \infty$ and $D_{-1}(p_1, p_2^m) \ni -1$ be the corresponding circle domains of $Q(z; p_1, p_2^m)\, dz^2$. Since the quadratic differential (6.1) depends continuously on the parameters p_1 and p_2, it is not difficult to show, using our definition of p_2^m, that both zeros of $Q(z; p_1, p_2^m)\, dz^2$ belong to the boundary of each of the domains $D_{-1}(p_1, p_2^m)$ and $D_\infty(p_1, p_2^m)$. But, as we have shown in part **6.2** of Section 6, in this case the domain configuration of $Q(z; p_1, p_2^m)\, dz^2$ must consist of three circle domains. Therefore, as we have shown in part **6.1** of Section 6, we must have $p_1^m = \bar{p}_1$.

Thus, we have shown that $p_2 \in \partial D_{-1}$ if p_1 and p_2 satisfy (7.3) and $|p_2 + 1| < |p_1 + 1|$. The Möbius map $w = \frac{3-z}{1+z}$ interchanges the poles $z = \infty$ and $z = -1$ of the quadratic differential (6.1) and does not change the type of its domain configuration. Therefore, our argument shows also that $p_1 \in \partial D_\infty$ if $|p_2 + 1| <$

$|p_1+1|$. This completes the proof of our claim that $p_1 \in \partial D_\infty$ if and only if $|p_2+1| < |p_1+1|$.

Similarly, if $Q(z)\,dz^2$ has exactly two circle domains $D_\infty \ni \infty$ and $D_1 \ni 1$, then $p_2 \in \partial D_1$ and $p_1 \in \partial D_\infty$ if and only if

$$\arg(p_1-1) = -\arg(p_2-1) \mod 2\pi \quad \text{and} \quad |p_2-1| < |p_1-1|.$$

7.A(3). In this part, we will discuss cases **6.3(a)**, **6.3(b1)**, and **6.3(b2)** discussed in Section 6. A domain configuration $\mathcal{D}_Q$ contains exactly one circle domains centered at $z=\infty$ if and only if neither C_1 or C_{-1} is a positive real number. As we have found in Section 6, in this case there exist one or two strip domains G_1 and G_2 having their vertices at the poles $z=1$ and $z=-1$. In what follows, we will use the notion of the *normalized height* h of a strip domain G, which is defined as

$$h = \frac{1}{2\pi}\Im \int_\gamma \sqrt{Q(z)}\,dz > 0,$$

where the integral is taken over any rectifiable arc $\gamma \subset G$ connecting the sides of G.

The sum of *normalized heights* in the Q-metric of the strip domains, which have a vertex at the pole $z=1$ or at the pole $z=-1$ can be found using integration over circles $\{z:\ |z-1|=r\}$ and $\{z:\ |z+1|=r\}$ of radius r, $0<r<1$, as follows:

$$h_+ = \frac{1}{2\pi}\Im \int_{|z-1|=r} \sqrt{Q(z)}\,dz = \frac{1}{2}\Im\sqrt{C_1} = \frac{1}{2}\Im\sqrt{(p_1-1)(p_2-1)} \tag{7.10}$$

if $z=1$ and

$$h_- = \frac{1}{2\pi}\Im \int_{|z+1|=r} \sqrt{Q(z)}\,dz = \frac{1}{2}\Im\sqrt{C_{-1}} = \frac{1}{2}\Im\sqrt{(p_1+1)(p_2+1)} \tag{7.11}$$

if $z=-1$. The branches of the radicals in (7.10) and (7.11) are chosen such that $h_+ \ge 0$, $h_- \ge 0$. Also, we assume here that if a strip domain has both vertices at the same pole then its height is counted twice.

Comparing h_+ and h_-, we find three possibilities:

1) If $h_+ = h_-$, then the domain configuration $\mathcal{D}_Q$ has only one strip domain G_2. This is the case discussed in parts **6.3(a)** and **6.3(b1)** in Section 6.
2) The case $h_+ > h_-$ corresponds to the configuration with two strip domains G_1 and G_2 discussed in part **6.3(b2)** in Section 6. In this case, the normalized heights h_1 and h_2 of the strip domains G_1 and G_2 can be calculated as follows:

$$h_1 = \frac{1}{2}\left(h_+ - h_-\right), \quad h_2 = h_-. \tag{7.12}$$

3) The case $h_+ < h_-1$ corresponds to the configuration with two strip domains mentioned in part **6.3(b3)** in Section 6.

Next, we will identify pairs p_1, p_2, which correspond to each of the cases **6.3(a)**, **6.3(b1)**, and **6.3(b2)**. The domain configuration $\mathcal{D}_Q$ has exactly one strip domain if and only if $h_+ = h_-$. Now, (7.10) and (7.11) imply that the latter equation is equivalent to the following equation:

$$\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(\bar{p}_1-1)(\bar{p}_2-1)}\right)^2$$
$$= \left(\sqrt{(p_1+1)(p_2+1)} - \sqrt{(\bar{p}_1+1)(\bar{p}_2+1)}\right)^2.$$

Simplifying this equation, we conclude that $h_+ = h_-$ if and only if p_1 and p_2 satisfy the following equation:

$$p_1 + \bar{p}_1 + p_2 + \bar{p}_2 + |p_1 - 1||p_2 - 1| - |p_1 + 1||p_2 + 1| = 0 \tag{7.13}$$

We claim that for a fixed p_1 with $\Im p_1 \neq 0$, the pair p_1, p_2 satisfies equation (7.13) if and only if $p_2 \in L(p_1)$ or $p_2 \in H(p_1)$. Indeed, $p_2 \in L(p_1)$ if and only if

$$|p_1 - 1| + |p_1 + 1| = |p_2 - 1| + |p_2 + 1|. \tag{7.14}$$

Similarly, $p_2 \in H(p_1)$ if and only if

$$|p_1 - 1| - |p_1 + 1| = |p_2 - 1| - |p_2 + 1|. \tag{7.15}$$

Multiplying equations (7.14) and (7.15), after simplification we again obtain equation (7.13). Therefore, $p_2 \in L(p_1)$ or $p_2 \in H(p_1)$ if and only if the pair p_1, p_2 satisfy equation (7.13). Thus, $\mathcal{D}_Q$ has only one strip domain if and only if $p_2 \in L(p_1) \setminus \{p_1, \bar{p}_1\}$ or $p_2 \in H(p_2) \setminus \{p_1, \bar{p}_1\}$. This proves the first parts of statements **6.3(a)** and **6.3(b1)**.

Now, we will prove that $p_1 \in \partial D_\infty$ for all $p_2 \in E^+_{-1}(p_1)$. First, we claim that $p_1 \in \partial D_\infty$ for all p_2 sufficiently close to -1. Arguing by contradiction, suppose that there is a sequence $s_k \to -1$ such that $s_k \in \partial D^k_\infty$ for all $k = 1, 2, \ldots$ Here $D^k_\infty \ni \infty$ denotes the corresponding circle domain of the quadratic differential $Q_k(z)\, dz^2$ having the form (7.5). From (7.5) we find that

$$Q_k(z)\, dz^2 \to \widehat{Q}(z)\, dz^2 := -\frac{z - p_1}{(z+1)(z-1)^2}\, dz^2,$$

where convergence is uniform on compact subsets of $\mathbb{C} \setminus \{-1, 1\}$. Since the residue of $\widehat{Q}(z)$ at $z = \infty$ equals 1, the quadratic differential $\widehat{Q}(z)\, dz^2$ has a circle domain $\widehat{D}_\infty \ni \infty$ and if $\gamma \subset \widehat{D}_\infty$ is a closed trajectory of $\widehat{Q}(z)\, dz^2$, then $|\gamma|_{\widehat{Q}} = 2\pi$.

Let us show that the boundary of $\widehat{D}_\infty$ consists of a single critical trajectory $\widehat{\gamma}_\infty$ of $\widehat{Q}(z)\, dz^2$, which has both its end points at $z = p_1$. Indeed, $\partial \widehat{D}_\infty$ consists of a finite number of critical trajectories of $\widehat{Q}(z)\, dz^2$, which have their end points at finite critical points. Therefore, if $-1 \in \partial \widehat{D}_\infty$, then $\partial \widehat{D}_\infty$ contains a critical trajectory, call it $\widehat{\gamma}_1$, which joins $z = -1$ and $z = p_1$. Some notations used in this part of the proof are shown in Figure 7a. This figure shows the limit configuration, which is, in fact, impossible as we explain below. In this case, $\partial \widehat{D}_\infty$ must contain a second critical trajectory, call it $\widehat{\gamma}_2$, which has both its end points at $z = p_1$. This implies that $z = 1$ is the only pole of $\widehat{Q}(z)\, dz^2$ lying in a simply connected domain, call it $\widehat{D}_1$, which is bounded by critical trajectories. Hence, $\widehat{D}_1$ must be a circle domain of $\widehat{Q}(z)\, dz^2$. Furthermore, the domain configuration $\mathcal{D}_{\widehat{Q}}$ consists of two circle domains $\widehat{D}_1$, $\widehat{D}_\infty$, which in this case must be the extremal domains of Jenkins module problem on the following maximum of the sum of reduced moduli:

$$m(B_\infty,\infty)+t^2m(B_1,1)\quad \text{with some fixed } t>0,$$

where the maximum is taken over all pairs of simply connected non-overlapping domains $B_\infty \ni \infty$ and $B_1 \ni 1$. It is well known that such a pair of extremal domains is unique; see for example, [31]. Therefore, $\widehat{D}_1$ and $\widehat{D}_\infty$ must be symmetric with respect to the real line (as is shown, for instance, in Figure 5d), which is not the case since $\widehat{Q}(z)\,dz^2$ has only one zero p_1 with $\Im p_1>0$.

Thus, $\partial\widehat{D}_\infty=\widehat{\gamma}_\infty\cup\{p_1\}$ and $z=-1$ lies in the domain complementary to the closure of $\widehat{D}_\infty$. Figure 7b illustrates notations used further on in this part of the proof.

Let $\tilde{\gamma}_{-1}$ denote the $\widehat{Q}$-geodesic in the class of all curves having their end points at $z=-1$, which separate the points $z=1$ and $z=p_1$ from $z=\infty$. Since $-1\notin\partial\widehat{D}_\infty$ it follows that

$$|\tilde{\gamma}_{-1}|_{\widehat{Q}}>|\widehat{\gamma}_\infty|_{\widehat{Q}}=2\pi. \tag{7.16}$$

Let $\varepsilon>0$ be such that

$$\varepsilon<\frac{1}{4}\left(|\tilde{\gamma}_{-1}|_{\widehat{Q}}-2\pi\right). \tag{7.17}$$

Let $r>0$ be sufficiently small such that

$$|[-1,-1+re^{i\theta}]|_{\widehat{Q}}<\varepsilon/8\quad \text{for all } 0\le\theta<2\pi. \tag{7.18}$$

Now let $\tilde{\gamma}_r$ be the shortest in the $\widehat{Q}$-metric among all arcs having their end points on the circle $C_r(-1)=\{z:\ |z+1|=r\}$ and separating the points $z=1$ and $z=p_1$ from the point $z=\infty$ in the exterior of the circle $C_r(-1)$. It is not difficult to show that there is at least one such curve $\tilde{\gamma}_r$. It follows from (7.18) that

$$|\tilde{\gamma}_r|_{\widehat{Q}}>|\tilde{\gamma}_{-1}|_{\widehat{Q}}-\varepsilon/4. \tag{7.19}$$

Since $s_k\to-1$, $s_k\in\partial D_\infty^k$, and $p_1\notin D_\infty^k$, it follows that for every sufficiently large k there is a regular trajectory $\gamma(k)$ of $Q_k(z)\,dz^2$ intersecting the circle $C_r(-1)$ and such that the arc $\gamma'(k)=\gamma(k)\setminus\{z:\ |z+1|\le r\}$ separates the points $z=1$ and $z=p_1$ from $z=\infty$ in the exterior of $C_r(-1)$. Since $|\gamma(k)|_{Q_k}=2\pi$ for all k and since every quadratic differential $Q_k(z)\,dz^2$ has second-order poles at $z=1$ and $z=\infty$ it follows from (7.5) that there is $r_0>0$ small enough such that $\gamma'(k)$ lies on the compact set $K_0=\{z:\ |z|\le 1/r_0\}\setminus(\{z:\ |z-1|<r_0\}\cup\{z:\ |z+1|<r\})$ for all k sufficiently large. We note also that $Q_k(z)\to\widehat{Q}(z)$ uniformly on K_0. This implies, in particular, that for all k the Euclidean lengthes of $\gamma'(k)$ are bounded by the same constant and that

$$|\gamma'(k)|_{Q_k}\ge|\gamma'(k)|_{\widehat{Q}}-\varepsilon/4 \tag{7.20}$$

for all k sufficiently large.

Combining (7.16)–(7.20), we obtain the following relations:

$$2\pi = |\gamma(k)|_{Q_k} \geq |\gamma'(k)|_{Q_k} \geq |\gamma'(k)|_{\widehat{Q}} - \varepsilon/4 \geq |\tilde{\gamma}_r|_{\widehat{Q}} - \varepsilon/4$$
$$> |\tilde{\gamma}_{-1}|_{\widehat{Q}} - \varepsilon/2 > |\tilde{\gamma}_{-1}|_{\widehat{Q}} - \tfrac{1}{2}\big(|\tilde{\gamma}_{-1}|_{\widehat{Q}} - 2\pi\big) = \tfrac{1}{2}\big(|\tilde{\gamma}_r|_{\widehat{Q}} + 2\pi\big) > 2\pi,$$

which, of course, is absurd. Thus, $p_2 \in \partial D_\infty$ for all p_2 sufficiently close to -1.

Let $\Delta \neq \emptyset$ be the set of all $p_2 \in E^+_{-1}(p_1)$ such that $p_1 \in \partial D_\infty$. To prove that $\Delta = E^+_{-1}(p_1) \setminus \{-1\}$, it is sufficient to show that Δ is closed and open in $E^+_{-1}(p_1)$. Arguing by contradiction, we suppose that there is a sequence of poles $s_k := p_2^k \in E^+_{-1}(p_1)$, $k = 1, 2, \ldots$, such that $s_k \to s_0 := p_2^0 \in E^+_{-1}(p_1)$ and $p_1 \in \partial D^k_\infty$ for all $k = 1, 2, \ldots$ but $p_1 \notin \partial D^0_\infty$. In this part of the proof, the index $k = 0, 1, 2, \ldots$, used in the notations D^k_∞, $\tilde{\gamma}_k$, etc., will denote domains, trajectories, and other objects corresponding to the quadratic differential $Q_k(z)\,dz^2$ defined by (7.5). Since ∂D^0_∞ contains a critical point and $p_1 \notin \partial D^0_\infty$, we must have $p_2^0 \in \partial D^0_\infty$.

Figure 7c illustrates some notations used in this part of the proof. In this case, the boundary ∂D^0_∞ consists of a single critical trajectory γ^0_∞ and its end points, each of which is at $z = p_2^0$. In addition, there is a critical trajectory of infinite Q^0-length, called it $\hat{\gamma}$, which has one end point at p_2^0 and which approaches to the pole $z = -1$ or the pole $z = 1$ in the other direction. Let P_0 be a point on $\hat{\gamma}$ such that the Q^0-length of the arc $\hat{\gamma}_0$ of $\hat{\gamma}$ joining p_2^0 and P_0 equals L, where $L > 0$ is sufficiently large. For $\delta > 0$ sufficiently small, let $\gamma_1^\perp$ and $\gamma_2^\perp$ denote disjoint open arcs on the orthogonal trajectory of $Q^0(z)\,dz^2$ passing through P_0 such that each of $\gamma_1^\perp$ and $\gamma_2^\perp$ has one end point at P_0 and each of them has Q^0-length equal to δ. If δ is small enough, then there is an arc of a trajectory of $Q^0(z)\,dz^2$, call it $\tilde{\gamma}$, which connects the second end point of $\gamma_1^\perp$ with the second end point of $\gamma_2^\perp$. Now, let $D(\delta)$ be the domain, the boundary of which consists of the arcs γ^0_∞, $\hat{\gamma}_0$, $\gamma_1^\perp$, $\gamma_2^\perp$, and their end points. In the terminology explained in Section 3, the domain $D(\delta)$ is a Q^0-rectangle of Q^o-height δ.

If $\delta > 0$ is sufficiently small, then p_1 belong to the bounded component of $\mathbb{C} \setminus \overline{D(\delta)}$. Let $\tilde{\gamma}_1$ be the arc of a trajectory of $Q^0(z)\,dz^2$, which divide $D(\delta)$ into two Q^0-rectangles, each of which has the Q^0-height equal to $\delta/2$. Since $p_1 \in \partial D_k$ for all k and p_1 belongs to the bounded component of $\mathbb{C} \setminus \overline{D(\delta)}$, it follows that, for each $k = 1, 2, \ldots$, there is a closed trajectory $\hat{\gamma}_k$ of $Q_k(z)\,dz^2$ lying in D^k_∞, which intersects $\tilde{\gamma}_1$ at some point $\tilde{z}_k \in D(\delta)$.

Since $Q_k(z) \to Q^0(z)$ it follows that, for all sufficiently large k, the trajectory $\hat{\gamma}_k$ has an arc $\tilde{\gamma}_k$ such that $\tilde{\gamma}_k \subset D(\delta)$ and $\tilde{\gamma}_k$ has one end point on each of the arcs $\gamma_1^\perp$ and $\gamma_2^\perp$.

Now, since $Q_k(z) \to Q^0(z)$ uniformly on $\overline{D(\delta)}$ it follows that

$$|\hat{\gamma}_k|_{Q_k} \geq |\tilde{\gamma}_k|_{Q_k} \to |\tilde{\gamma}_1|_{Q^0} = |\gamma^0_\infty|_{Q^0} + 2|\hat{\gamma}_0|_{Q^0} = 2\pi + 2L,$$

contradicting to the fact that $|\hat{\gamma}_k|_{Q_k} = 2\pi$. The latter fact follows from the assumption that $\hat{\gamma}_k$ is a closed trajectory of $Q_k(z)\,dz^2$, which lies in a circle domain D^k_∞.

Thus, we have proved that Δ is closed in $E^+_{-1}(p_1)$. A similar argument can be used to show that Δ is open in $E^+_{-1}(p_1)$. The difference is that to construct a domain $D(\delta)$, we now use an arc $\tilde{\gamma}_1$ of a critical trajectory $\hat{\gamma}_1$, which has one of its end points at the pole p_1 and not at the pole p_1^0 as we had in the previous case.

Therefore, we have proved that if $p_2 \in E_{-1}^+(p_1)$, then $p_1 \in \partial D_\infty$. The same argument can be used to prove that if $p_2 \in E_1^+(p_1)$, then $p_1 \in \partial D_\infty$.

Finally, if $p_2 \in E_1^-(p_1)$ or $p_2 \in E_{-1}^-(p_1)$, then we can switch the roles of the poles p_1 and p_2 in our previous proof and conclude that $p_2 \in \partial D_\infty$ in these cases. This proves the first part of statement **6.3(b2)**.

Now, possible positions of zeros p_1 and p_2 on boundaries of the corresponding circle and strip domains are determined for all cases. Next, we will discuss limiting behavior of critical trajectories. We will give a proof for the most general case when the domain configuration consists of a circle domain D_∞ and strip domains G_1 and G_2. In all other cases proofs are similar.

Let Δ denote the set of pairs (p_1, p_2), for which the limiting behavior of critical trajectories is shown in Fig. 4a or in more general case in Fig. 4b. That is when γ_1 joins $p_1 \in \partial D_\infty \cap \partial G_1$ and $z = 1$, γ_{-1} joins $p_2 \in \partial G_1 \cap \partial G_2$ and $z = -1$, and γ_0^+ and γ_0^- each joins p_2 and $z = 1$. First, we note that Δ is not empty since $(p_1, p_2) \in \Delta$ when $p_1 > 1$ and $-p_1 < p_2 < -1$. In this case the intervals $(p_2, -1)$ and $(1, p_1)$ represent critical trajectories γ_1 and γ_{-1} and critical trajectories γ_0^+ and γ_0^- connect a zero at p_2 with a pole at $z = 1$; see Fig. 4a.

We claim that Δ is open. To prove this claim, suppose that $(p_1^0, p_2^0) \in \Delta$ and that $(p_1^k, p_2^k) \to (p_1^0, p_2^0)$ as $k \to \infty$, $k = 1, 2, \ldots$ Fix $\varepsilon >$ small enough and consider the arc $\gamma_1^0(\varepsilon) = \gamma_1^0 \setminus \{z : |z - 1| < \varepsilon\}$ of the critical trajectory γ_1^0, which goes from p_1^0 to the pole $z = 1$. Since $(p_1^k, p_2^k) \to (p_1^0, p_2^0)$ it follows that for all k sufficiently big there is a critical trajectory γ_1^k having one point at p_1^k which has a subarc $\gamma_1^k(\varepsilon)$ which lies in the $\varepsilon/10$-neighborhood of the arc $\gamma_1^0(\varepsilon)$. In particular, eventually, $\gamma_1^k(\varepsilon)$ enters the disk $\{z : |z - 1| < \varepsilon\}$. Therefore, it follows from the standard continuity argument and Lemma 4 that γ_1^k approaches the pole $z = 1$. The same argument works for all other critical trajectories of the quadratic differential (6.1) with $p_1 = p_1^k$, $p_2 = p_2^k$. Thus, we have proved that Δ is open.

Same argument can be applied to show that all other sets of points (p_1, p_2) responsible for different types of limiting behavior of critical trajectories mentioned in part **6.3(b2)** of Theorem 4 are also nonempty and open. The latter implies that each of these sets must coincide with some connected component of the set $\mathbb{C} \setminus (L(p_1) \cup H(p_1))$. This proves the desired statement in the case under consideration.

7.B. The local behavior of trajectories near second-order poles at $z = 1$ and $z = -1$ is controlled by Laurent coefficients C_1 and C_{-1}, respectively, which are given by formula (7.2). The radial structure near $z = 1$ or near $z = -1$ occurs if and only if $C_1 < 0$ or $C_{-1} < 0$, respectively. The latter inequalities are equivalent to the following relations:

$$\arg(p_1 - 1) = -\arg(p_2 - 1) + \pi \tag{7.21}$$

or

$$\arg(p_1 + 1) = -\arg(p_2 + 1) + \pi. \tag{7.22}$$

Now, statement **(1)** about radial behavior follows from (7.21) and (7.22).

Next, trajectories of $Q(z)\,dz^2$ approaching the pole $z = 1$ spiral clockwise if and only if $0 < \arg C_1 < \pi$. The latter is equivalent to the inequalities:

$$-\arg p_1 - 1 < \arg(p_2 - 1) < -\arg(p_1 - 1) + \pi,$$

which imply the desired statement for the case when trajectories of $Q(z)\,dz^2$ approaching $z = 1$ spiral clockwise. In the remaining cases the proof is similar.

The proof of Theorem 4 is now complete. □

Remark 3. The case when $\Im p_1 = 0$ but $\Im p_2 \neq 0$ can be reduced to the case covered by Theorem 4 by changing numeration of zeros. In the remaining case when $\Im p_1 = 0$ and $\Im p_2 = 0$, the domain configurations are rather simple; they are symmetric with respect to the real axis as it is shown in Figures 1a, 1b, 2a, 3a, and some other figures.

8. Identifying simple critical geodesics and critical loops

Topological information obtained in Section 6 is sufficient to identify all critical geodesics and all critical geodesic loops of the quadratic differential (6.1) in all cases. In particular, we can identify all simple geodesics.

Cases **6.1(a)** and **6.1(b)**; see Figure 1a and Figure 1b. Let γ be a geodesic joining p_1 and p_2. Since D_∞, D_1, and D_{-1} are simply connected and $p_1 \in \partial D_\infty \cap \partial D_1$ and $p_2 \in \partial D_\infty \cap \partial D_{-1}$ it follows from Lemma 4 that γ does not intersect D_∞, D_1, and D_{-1}. In this case, γ must be composed of a finite numbers of copies of γ_0, a finite number of copies of γ_1, and a finite number of copies of γ_{-1}. Therefore the only simple geodesic joining p_1 and p_2 in this case is the segment $\gamma_0 = [p_2, p_1]$.

In addition, by Lemma 5, γ_1 is the only simple non-degenerate geodesic from the point p_1 to itself and γ_{-1} is the only short geodesic from p_2 to p_2.

Case **6.1(c)**; see Figure 1c. As in the previous case, any geodesic γ joining p_1 and p_2 must be composed of a finite number of copies of γ_0, a finite number of copies of γ_1, and a finite number of copies of γ_{-1}. Thus, in this case there exist exactly three simple geodesics joining p_1 and p_2, which are γ_0, γ_1, and γ_{-1}. By Lemma 5, there are no geodesic loops in this case.

Case **6.2**; see Figures 2a, 2b. Suppose that $\mathcal{D}_Q$ consists of circle domains D_∞ and D_{-1} and a strip domain G_1. Let γ be a geodesic joining p_1 and p_2. If γ contains a point $\zeta \in \gamma_{-1}$ or a point $\zeta \in \gamma_\infty$, then it follows from Lemma 4 that γ_{-1} or, respectively, γ_∞ is a subarc of γ. Thus, γ is not simple in these cases.

Suppose now that $\gamma \subset G_1 \cup \gamma_1^+ \cup \gamma_1^-$. Since G_1 is a strip domain the function $w = F(z)$ defined by

$$F(z) = \frac{1}{2\pi} \int_{p_1}^{z} \sqrt{Q(z)}\,dz, \tag{8.1}$$

with an appropriate choice of the radical, maps G_1 conformally and one-to-one onto the horizontal strip S_{h_1}, where $S_h = \{w : 0 < \Im w < h_1\}$, in such a way that the trajectory γ_∞ is mapped onto an interval $(x_1, x_1') \subset \mathbb{R}$ with $x_1 = 0$ and

$x_1' = 1$. Here h_1 is the normalized height of the strip domain G_1 defined by (7.12). Figure 8a and Figure 9a illustrate some notions relevant to Case **6.2**. To simplify notations in our figures, we will use the same notations for Q-geodesics (such as γ_∞, γ_{11}, γ_{12}', etc.) in the z-plane and for their images under the mapping $w = F(z)$ in the w-plane.

The indefinite integral $\Phi(z) = \frac{1}{2\pi}\int \sqrt{Q(z)}\,dz$ can be expressed explicitly in terms of elementary functions as follows:

$$\begin{aligned}\Phi(z) = \tfrac{1}{4\pi i}\Big(&\sqrt{(p_1-1)(p_2-1)}\log(z-1) - \sqrt{(p_1+1)(p_2+1)}\log(z+1)\\ &+ 4\log(\sqrt{z-p_1}+\sqrt{z-p_2})\\ &+ 2\sqrt{(p_1+1)(p_2+1)}\log(\sqrt{(p_1+1)(z-p_2)} - \sqrt{(p_2+1)(z-p_1)})\\ &- 2\sqrt{(p_1-1)(p_2-1)}\log(\sqrt{(p_1-1)(z-p_2)} - \sqrt{(p_2-1)(z-p_1)})\Big).\end{aligned} \tag{8.2}$$

Equation (8.2) can be verified by straightforward differentiation. Alternatively, it can be verified with *Mathematica* or *Maple*. With (8.2) at hands, the function $F(z)$ can be written as

$$F(z) = \Phi(z) - \Phi(p_1), \tag{8.3}$$

where

$$\Phi(p_1) = \frac{1}{4\pi i}\left(2 + \sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\log(p_1-p_2). \tag{8.4}$$

Calculating $\Phi(p_2)$, after some algebra, we find that:

$$F(p_2) = \frac{1}{2} + \frac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right). \tag{8.5}$$

Of course, all branches of the radicals and logarithms in (8.2)–(8.5) have to be appropriately chosen.

To explain more precisely our choice of branches of multi-valued functions in (8.2)–(8.5), we note that the points p_1, p_2 and points of the arcs γ_1^+ and γ_1^- each represents two distinct boundary points of G_1 and therefore every such point has two images under the mapping $F(z)$. These images will be denoted by $x_1(\zeta)$ and $x_1'(\zeta)$ if $\zeta \in \gamma_1^+ \cup \{p_1\}$ and by $x_2(\zeta) + ih_1$ and $x_2'(\zeta) + ih_1$ if $\zeta \in \gamma_1^- \cup \{p_2\}$. We assume here that $x_1(\zeta) < x_1'(\zeta)$ for all $\zeta \in \gamma_1^+ \cup \{p_1\}$ and $x_2(\zeta) < x_2'(\zeta)$ for all $\zeta \in \gamma_1^- \cup \{p_2\}$. In accordance with our notation above, $x_1(p_1) = x_1 = 0$ and $x_1'(p_1) = x_1' = 1$. We also will abbreviate $x_2(p_2)$ and $x_2'(p_2)$ as x_2 and x_2', respectively.

For every $\zeta \in \gamma_1^+$, the segments $[x_1(\zeta), x_1]$ and $[x_1', x_1'(\zeta)]$ are the images of the same arc on γ_1^+. Therefore they have equal lengthes. Similarly, the segments $[x_2(\zeta) + ih_1, x_2 + ih_1]$ and $[x_2' + ih_1, x_2'(\zeta) + ih_1]$ have equal lengthes. Thus, for every $\zeta \in \gamma_1^+$ and every $\zeta \in \gamma_1^-$, we have, respectively:

$$x_1 - x_1(\zeta) = x_1'(\zeta) - x_1' \quad \text{and} \quad x_2 - x_2(\zeta) = x_2'(\zeta) - x_2'. \tag{8.6}$$

We know that the preimage under the mapping $F(z)$ of every straight line segment is a geodesic. This immediately implies that in the case under consideration there exist four simple critical geodesics, which are the following preimages:

$$\begin{aligned} \gamma_{12} &= F^{-1}((x_1, x_2 + ih_1)), \quad \gamma'_{12} = F^{-1}((x_1, x'_2 + ih_1)), \\ \gamma_{21} &= F^{-1}((x'_1, x_2 + ih_1)), \quad \gamma'_{21} = F^{-1}((x'_1, x'_2 + ih_1)). \end{aligned} \tag{8.7}$$

The geodesic loops γ_∞ and γ_{-1} are the following preimeges:

$$\gamma_\infty = F^{-1}((x_1, x_1)), \quad \gamma_{-1} = F^{-1}((x_2 + ih_1, x'_2 + ih_1)). \tag{8.8}$$

We claim that there is no other simple geodesic joining the points p_1 and p_2. Figure 9a illustrates some notation used in the proof of this claim. Suppose that τ is a geodesic ray issuing from p_1 into the region G_1. Let τ_k, $k = 1, \dots, N$, be connected components of the intersection $\tau \cap G_1$ enumerated in their natural order on τ. In particular, τ_1 starts at p_1. We may have finite or infinite number of such components. Thus, N is a finite number or $N = \infty$. Let $l_k = F(\tau_k)$. Since all τ_k lie on the same geodesic it follows that l_k are parallel line intervals in S joining the real axis and the horizontal line L_{h_1}, where $L_h = \{w : \Im w = h\}$. Let v'_k and v''_k be the initial point and terminal point of l_k, respectively. Then $v'_k = e'_k$ and $v''_k = e''_k + ih_1$ with real e'_k and e''_k if k is odd and $v'_k = e'_k + ih_1$, $v''_k = e''_k$ with real e'_k and e''_k if k is even.

The interval l_1 may start at x_1 or at x'_1. To be definite, suppose that $e'_1 = x_1$. For the position of e''_1 we have the following possibilities:

(a) $e''_1 = x_2$ or $e''_1 = x'_2$. In this case, $\tau_1 = \gamma_{12}$ or $\tau_1 = \gamma'_{12}$. Thus we obtain two out of four geodesics in (8.7).

(b) $x_1 < e''_1 < x'_1$. In this case, τ_1 has its end point on γ_{-1}. By Lemma 4, the continuation of τ_1 as a geodesic will stay in D_{-1} and will approach to the pole $z = -1$. Thus, τ is not a geodesic from p_1 to p_2 or a geodesic loop from p_1 to itself in this case.

(c) $e''_1 > x'_2$. Let $d = e''_1 - x'_2$. It follows from (8.6) that $e'_2 = x_2 - d$. Then $e''_2 = x_1 - d$. In general, $e'_{2k-1} = x'_1 + (k-1)d$, $e''_{2k-1} = x'_2 + kd$ for $k = 1, 2, \dots$, and $e'_{2k} = x_2 - kd$, $e''_{2k} = x_1 - kd$ for $k = 1, 2, \dots$. Thus, τ cannot terminate at p_1 or p_2. Instead, τ approaches to the pole at $z = 1$ as a logarithmic spiral.

(d) $e''_1 < x_2$. Let $d_0 = x_2 - e''_1$. Then $e'_2 = x'_2 + d_0$ by (8.6). For the position of e''_2 we have three possibilities.

 (α) $x_1 < e''_2 < x'_1$. In this case by Lemma 4, the continuation of τ_2 as a geodesic ray will stay in D_∞ and will approach to the pole $z = \infty$. Thus, τ is not a geodesic from p_1 to p_2 or a geodesic loop in this case.

 (β) $e''_2 = x'_1$. In this case, τ is a critical geodesic loop $\gamma_{11} = F^{-1}((x_1, v''_1] \cup [v'_2, x'_1))$ from p_1 to itself. We emphasize here, that since the segments l_1 and l_2 are parallel a critical geodesic loop from p_1 to itself occurs if and only if $|\gamma_\infty|_Q = x'_1 - x_1 > x'_2 - x_2 = |\gamma_{-1}|_Q$. If $|\gamma_\infty|_Q < |\gamma_{-1}|_Q$, then there is a critical geodesic loop γ_{22} with end points at p_2.

(γ) $e_2'' > x_1'$. Let $d = e_2'' - x_1'$. Then, as in the case c), we obtain that $e_{2k+1}' = x_1 - kd$, $e_{2k+1}'' = x_2 - d_0 - kd$ for $k = 1, 2, \ldots$, and $e_{2k}' = x_2' + d_0 + kd$, $e_{2k}'' = x_1' + kd$ for $k = 1, 2, \ldots$. Therefore, τ does not terminate at p_1 or p_2. Instead, τ approaches to the pole at $z = 1$ as a logarithmic spiral.

If l_1 has its initial point at x_1', the same argument shows that there are exactly two geodesics joining p_1 and p_2, which are the geodesics γ_{21} and γ_{21}' defined by (8.7).

Combining our findings for Case **6.2**, we conclude that in this case there exist exactly four distinct geodesics joining p_1 and p_2, which are given by (8.7). The geodesic loops γ_∞ and γ_{-1} are given by (8.8). In addition, if $|\gamma_\infty|_Q \ne |\gamma_{-1}|_Q$, then there is exactly one geodesic loop containing the pole $z = 1$ in its interior domain, which has its end points at a zero of $Q(z)\, dz^2$. This loop has the pole $z = 1$ in its interior domain, which does not contain other critical points of $Q(z)\, dz^2$, and has both its end points at p_1 or at p_2, if $|\gamma_\infty|_Q > |\gamma_{-1}|_Q$ or $|\gamma_\infty|_Q < |\gamma_{-1}|_Q$, respectively.

Finally, if $|\gamma_\infty|_Q = |\gamma_{-1}|_Q$, then the geodesics γ_{12} and γ_{21}' together with points $z = p_1$ and p_2 form a boundary of a simply connected bounded domain, which contains the pole $z = 1$ and does not contain other critical points of $Q(z)\, dz^2$. There are no geodesic loops containing $z = 1$ in its interior domain in this case.

The argument based on the construction of parallel segments divergent to ∞, which was used above to prove non-existence of some geodesics, will be used for the same purpose in several other cases considered below. Since the detailed construction is rather lengthy, the detailed exposition will be given for one more case when we have two strip domains. In other cases, we will just refer to this argument (which actually is rather standard, see [34, Ch. IV]) and call it the "proof by construction of divergent geodesic segments".

Case **6.3(a)**; see Figure 8b. In this case, the domain configuration $\mathcal{D}_Q$ consists of a circle domain D_∞ and a strip domain G_2 having its vertices at the poles $z = 1$ and $z = -1$. The function $F(z)$ defined by (8.1) maps G_2 conformally and one-to-one onto the strip S_{h_1} such that the trajectory γ_∞^+ is mapped onto the interval $(x_1, x_2) \subset \mathbb{R}$ with $x_1 = 0$ and some x_2, $0 < x_2 < 1$. The points $z = p_1$ and $z = p_2$ each has two images under the mapping $F(z)$. Let $x_1 = 0$ and $x_1' + ih_1$ with some real x_1' be the images of p_1 and let x_2 and $x_2' + ih_1$ with $x_2' = x_1' + (1 - x_2)$ be the images of p_2. Arguing as in Case **6.2**, one can easily find four distinct simple geodesics joining the points p_1 and p_2. These geodesics are:

$$\gamma_{12} = F^{-1}((x_1, x_2)) = \gamma_\infty^+, \quad \gamma_{12}' = F^{-1}((x_1' + ih_1, x_2' + ih_1)) = \gamma_\infty^-,$$
$$\gamma_{21} = F^{-1}((x_1, x_2' + ih_1)), \quad \gamma_{21}' = F^{-1}((x_2, x_1' + ih_1)).$$

In addition, there are two critical geodesic loops:

$$\gamma_{11} = F^{-1}((x_1, x_1' + ih_1)) \quad \text{and} \quad \gamma_{22} = F^{-1}((x_2, x_2' + ih_1)).$$

It follows from Lemma 5 that there are no other such loops.

Using the proof by construction of divergent geodesic segments as in Case **6.2**, we can show that there are no other simple geodesics joining p_1 and p_2.

Case **6.3(b1)**; see Figure 8c. We still have a circle domain D_∞ and a strip domain G_2. In this case, the function $F(z)$ defined by (8.1) as in Case **6.2** maps G_2 conformally and one-to-one onto S_{h_1} such that γ_∞ is mapped onto the interval $(x_1, x_1') \subset \mathbb{R}$, where $x_1 = 0$ and $x_1' = 1$. The difference is that now the point p_2 represents three boundary points of G_2. Two of them belong to the side l_2 and the third point belongs to the side l_1. Accordingly, there are three images of p_2 under the mapping $F(z)$, which we will denote by $x_2 + ih_1$, x_2', and x_2''. Here x_2 may be any real number while x_2' and x_2'' satisfy the following conditions:

$$x_2' > x_1', \quad x_2'' < x_1, \quad \text{and} \quad x_2' - x_1' = x_1 - x_2''.$$

In this case, there are three short geodesics, which are the following preimages:

$$\gamma_0 = F^{-1}((x_1'', x_1)) = F^{-1}((x_1', x_2'))$$

and

$$\gamma_{12} = F^{-1}((x_1, x_2 + ih_1)), \quad \gamma_{12}' = F^{-1}((x_1', x_2 + ih_1)).$$

In addition, there are three geodesic loops:

$$\gamma_\infty = F^{-1}((x_1, x_1')), \quad \gamma_{22}' = F^{-1}((x_2 + ih_1, x_2')), \quad \gamma_{22}'' = F^{-1}((x_2 + ih_1, x_2'')).$$

Using the proof by construction of divergent segments as above, it is not difficult to show that there are no other simple geodesics joining the points p_1 and p_2.

Case **6.3(b2)**. This is the most general case with many subcases illustrated in Figures 10a–10i. In this case we have a circle domain D_∞ and two strip domains G_1 and G_2. We assume that $\mathcal{D}_Q$ has topological type shown in Figure 4b. In other cases the proof follows same lines. The function $F(z)$ defined by (8.1) maps G_1 conformally and one-to-one onto the strip S_{h_1} such that γ_∞ is mapped onto the interval $(x_1, x_1') \subset \mathbb{R}$, where $x_1 = 0$ and $x_1' = 1$. The point p_2 represents one boundary point of G_1 and two boundary points of G_2. Let $x_2 + ih_1$ be the image of p_2 considered as a boundary point of G_1. Then the trajectory γ_0^+ considered as boundary arc of G_1 is mapped onto the ray $r_1 = \{w = t + ih_1 : t < x_2\}$, while the trajectory γ_0^- is mapped onto the ray $r_2 = \{w = t + ih_1 : t > x_2\}$. The function $F(z)$ can be continued analytically through the trajectory γ_0^+. The continued function (for which we keep our previous notation $F(z)$) maps G_2 conformally and one-to-one onto the strip $S(h_1, h) = \{w : h_1 < \Im w < h\}$ with $h = h_1 + h_2$, where h_1 and h_2 are defined by (7.12). Two boundary points of G_2 situated at p_2 are mapped onto the points $x_2 + ih_1$ and $x_2' + ih$ with some $x_2' \in \mathbb{R}$. Thus, the domain $\widetilde{D} = G_1 \cup G_2 \cup \gamma_0^+$ is mapped by $F(z)$ conformally and one-to-one onto the slit strip $\widehat{S}(h_1, h) = \{w : 0 < \Im w < h\} \setminus \{w = t + ih_1 : t \geq x_2\}$.

We note that every boundary point $\zeta \in \gamma_1 \cup \gamma_{-1} \cup \gamma_0^-$ under the mapping $F(z)$ has two images $w_1(\zeta)$ and $w_2(\zeta)$, which satisfy the following conditions similar to conditions (8.6):

$$x_1 - w_1(\zeta) = w_2(\zeta) - x_1' > 0 \quad \text{if } \zeta \in \gamma_1, \tag{8.9}$$

$$w_1(\zeta) = u_1(\zeta) + ih, \quad w_2(\zeta) = u_2(\zeta) + ih_1, \tag{8.10}$$

where $x_2' - u_1(\zeta) = u_2(\zeta) - x_2 > 0$ if $\zeta \in \gamma_0^-$, and

$$w_1(\zeta) = u_1(\zeta) + ih, \quad w_2(\zeta) = u_2(\zeta) + ih_1,$$

where $u_1(\zeta) - x_2' = u_2(\zeta) - x_2 > 0$ if $\zeta \in \gamma_{-1}$.

Consider four straight lines P_k, $k = 1, 2, 3, 4$, where P_2 passes through x_1' and $x_2 + ih_1$, P_3 passes through x_1 and $x_2 + ih_1$, P_1 passes through x_1 and is parallel to P_2, and P_4 passes through x_1' and is parallel to P_3. Let $u_k + ih$ denote the point of intersection of P_k and the horizontal line $L(h)$, where $L(m)$ stands for the line $\{w : \Im w = m\}$. Then the points $u_k + ih$, $k = 1, 2, 3, 4$, are ordered in the positive direction on $L(h)$; see Figure 10a.

Next, we consider five possible positions for x_2', which correspond to "non-degenerate" cases and four positions corresponding to "degenerate" cases. Figures 10a–10i illustrate our constructions of critical geodesics and critical geodesic loops in all these cases. First, we will work with non-degenerate cases, which are cases (a), (c), (e), (g), and (i) and after that we will briefly mention degenerate cases (b), (d), (f), and (h).

(a) $x_2' < u_1$. Then the slit strip S_1 contains four intervals: $(x_1, x_2 + ih_1)$, $(x_1', x_2 + ih_1)$, $(x_1, x_2' + ih)$, and $(x_1', x_2' + ih)$. Therefore the preimages of these intervals under the mapping $F(z)$ provide four distinct geodesics joining the points p_1 and p_2:

$$\begin{aligned} \gamma_{12} &= F^{-1}((x_1, x_2 + ih_1)), & \gamma_{12}' &= F^{-1}((x_1', x_2 + ih_1)), \\ \gamma_{21} &= F^{-1}((x_1, x_2' + ih)), & \gamma_{21}' &= F^{-1}((x_1', x_2' + ih)). \end{aligned} \tag{8.11}$$

In addition, there are two critical geodesic loops:

$$\gamma_\infty = F^{-1}((x_1, x_1')) \quad \text{and} \quad \gamma_{22} = F^{-1}((x_2 + ih_1, x_2' + ih)). \tag{8.12}$$

The curve $\gamma_{22} \cup \{p_2\}$ bounds a simply connected domain, call it D_{-1}, which contains the trajectory γ_2 and the pole $z = -1$.

One more critical geodesic loop can be found as follows. Let P_5 be the line through $x_2' + ih$ that is parallel to P_1 and let u_5' be the point of intersection of P_5 with the real axis. It follows from elementary geometry that there exists a point u_5, $u_5' < u_5 < x_1$ such that the line segments $[x_2' + ih, u_5]$ and $[u_6, x_2 + ih_1]$ with $u_6 = x_1' + x_1 - u_5$ are parallel to each other. Therefore, it follows from equation (8.9) that the preimage $\gamma_{22}' = F^{-1}((x_2' + ih, u_5] \cup [u_6, x_2 + ih_1))$ is a geodesic loop from p_2 to p_2 containing the pole $z = 1$ in its interior domain.

We claim that there no other simple critical geodesics in this case. The proof is by the method of construction of divergent geodesic segments. An example of such construction for the case under consideration is shown in Figure 9b.

Suppose that τ is a geodesic ray issuing from p_1 into the region $\widetilde{G}$. Let τ_k, $k = 1, \dots, N$, where N is a finite integer or $N = \infty$, be connected component of $\tau \cap \widetilde{G}$ enumerated in the natural order on τ. Let $l_k = F(\tau_k)$ and let e'_k and e''_k be the initial and terminal points of l_k, respectively.

The interval l_1 may start at x_1 or at x'_1. To be definite, assume that $e'_1 = x_1$. Then for e''_1 we have the following cases:

(α) $e''_1 = x'_2 - d_1 + ih$ with some $d_1 > 0$,
(β) $e''_1 = x'_2 + d_1 + ih$ with some $d_1 > 0$,
(γ) $e''_1 = x_2 + d_1 + ih_1$ with some $d_1 > 0$.

We give a proof for the case α). In two other case the proof is similar. By (8.10), $e'_2 = x_2 + d_1 + ih_1$ and $e''_2 > x'_1$. Let $d = e''_2 - x'_1$. Continuing, we find the following expressions for the end points of the segments l_k:

$$e'_{2k-1} = x_1 + (k-1)d, \qquad e''_{2k-1} = x'_2 + d_1 + (k-1)d + ih,$$
$$e'_{2k} = x_2 + d_1 + (k-1)d + ih_1, \qquad e''_{2k} = x'_1 + kd.$$

Thus, in this case τ cannot terminate at p_2. Instead, it approaches to the pole $z = 1$ as a logarithmic spiral.

(c) $u_1 < x'_2 < u_2$. In this case we still have geodesics (8.11) and loops (8.12). The only difference is that we cannot construct the loop γ'_{22} as in part (a). Instead, we can construct a loop γ'_{11} from p_1 to p_1. Indeed, using elementary geometry, we easily find that there is a point $u_7 + ih$ with $u_7 < x'_2$ such that the segments $[x_1, u_7 + ih]$ and $[u_8 + ih_1, x'_1]$ with $u_8 = x_2 + x'_2 - u_7$ are parallel. Therefore using (8.10), we conclude that $\gamma'_{11} = F^{-1}((x_1, u_7 + ih] \cup [u_8 + ih_1, x_1))$ is a critical geodesic loop.

(e) $u_2 < x'_2 < u_3$. We still have geodesics γ_{12}, γ'_{12}, and γ_{21} given by (8.11) and the loops γ_∞, γ_{22}, and γ'_{11} as in the case c). But the geodesic γ'_{21} in (8.11) should be replaced with a geodesic constructed as follows. From elementary geometry we find that there is $u_9 > x_2$ such that the segments $[x'_1, u_9 + ih_1]$ and $[u_{10} + ih, x_2 + ih_1]$ with $u_{10} = x'_2 - u_9 + x_2$ are parallel. Using (8.10), we conclude that the arc $\gamma'_{21} = F^{-1}((x'_1, u_9 + ih_1] \cup [u_{10} + ih, x_2 + ih_1))$ is a geodesic from p_1 to p_2.

(g) $u_3 < x'_2 < u_4$. The geodesics γ_{12}, γ'_{12}, and γ'_{21} and all three critical geodesic loops can be constructed as in part (e). The geodesic γ_{21} in this case can be constructed as follows. Using elementary geometry one can find that there is $u_{11} > x_2$ such that the segments $[x_1, u_{11} + ih_1]$ and $[u_{12} + ih, x_2 + ih_1]$ with $u_{12} = x'_2 + x_2 - u_{11}$ are parallel. Using (8.10) we conclude that the arc $\gamma_{21} = F^{-1}((x_1, u_{11} + ih_1] \cup [u_{12} + ih, x_2 + ih_1))$ is a geodesic from p_1 to p_2.

(i) $x'_2 > u_4$. The geodesics from p_1 to p_2 can be constructed as in case (g). Of course, we still have loops (8.12). The third geodesic critical loop can be obtained as follows. For $u_{13} < x_1 = 0$, let l^1 be the line segment joining the real axis and the line $L(h)$, which has its initial point at $z = u_{13}$ and passes through $z = x_2 + i$. Let $z = u_{14} + ih$ be the terminal point of l^1 on $L(h)$. We consider only those values of u_{13}, for which $u_{14} < x'_2$. Let $d = x'_2 - u_{14}$ and let l^2 be a line segment joining the real axis and $L(h_1)$, which is parallel to

l^1 and has its initial point at $u_{15} = x_1' + d$. Let $z = u_{16} + ih_1$ be the terminal point of l^2 on $L(h_1)$. It follows from elementary geometry that we can find a unique value of u_{13} such that for this value $u_{16} - x_2 = x_2' - u_{14}$.

It follows from our construction and from the identification properties (8.9) and (8.10) that the preimage

$$\gamma_{22}' = F^{-1}([u_{13}, x_2 + ih_1) \cup (x_2 + ih_1, u_{14} + ih] \cup [u_{15}, u_{16} + ih_1])$$

is a geodesic loop from the point p_2 to itself. In addition, this loop contains the pole $z = 1$ in its interior, which does not contain other critical points.

Now we consider four "degenerate" cases.

(b) If $x_2' = u_1$, then we still have critical geodesics (8.11) and critical geodesic loops (8.12). But there is no critical geodesic loop separating the pole $z = 1$ from other critical points. Instead, the boundary of a simply connected domain having $z = 1$ inside and bounded by critical geodesics will consist of geodesics γ_{12}' and γ_{22}.

(d) If $x_2' = u_2$, then we have all critical geodesic loops and geodesics γ_{12}, γ_{12}', and γ_{21} as in the case $u_1 < x_2' < u_2$ but instead of geodesic γ_{21}' we have a non-simple geodesic, which is the union $\gamma_{12}' \cup \gamma_{22}$.

(f) If $x_2' = u_3$, then we have all critical geodesic loops and geodesics γ_{12}, γ_{12}', and γ_{21}' as in the case $u_2 < x_2' < u_3$ but instead of geodesic γ_{21} we have a non-simple geodesic, which is the union $\gamma_{12} \cup \gamma_{22}$.

(h) If $x_2' = u_4$, then we have all geodesics and loops γ_∞, γ_{22} constructed as in the case $u_3 < x_2' < u_4$ but instead of the loop γ_{11}' we will have non-simple critical geodesic separating the pole $z = 1$ from all other critical points. This non-simple critical geodesic is the union $\gamma_{12} \cup \gamma_{21}'$.

Using the proof by construction of divergent geodesic segments one can show that in all cases considered above there are no any other critical geodesics or critical geodesic loops.

Quadratic differentials defined by formula (6.1) depend on four real parameters which are real parts and imaginary parts of zeroes p_1 and p_2. As the reader may noticed in the generic case configurations shown in Figures 10 also depend on four real parameters which are x_2, x_2', h_1, and h. This is not a coincidence; in fact, the set of pairs (p_1, p_2) is in a one-to-one correspondence with the set of these diagrams. To explain how this one-to-one correspondence works, we will show three basic steps. To be definite, we assume that the domain configuration consists of a circle domain D_∞ and strip domains G_1 and G_2. Thus, we will consider diagrams shown in Figures 10.

- As we described above, for any given p_1 and p_2, the function $F(z)$ defined by (8.1) maps G_1 and G_2 onto horizontal strips shown in Figures 10. Furthermore, for fixed p_1 and p_2, the values of the parameters x_2, x_2', h_1, and h are uniquely defined via function $F(z)$.

- To prove that different pairs (p_1, p_2) define different diagrams, we argue by contradiction. Suppose that mappings $F_1(z)$ and $F_2(z)$ constructed by formula (8.1) for distinct pairs (p_1^1, p_2^1) and (p_1^2, p_2^2) produce identical diagrams of the form shown in Figures 10. Then the composition $\varphi = F_1^{-1} \circ F_2$ is well defined and defines a one-to-one meromorphic mapping from $\overline{\mathbb{C}}$ onto itself. Since $\varphi(1) = 1$, $\varphi(-1) = -1$, and $\varphi(\infty) = \infty$ we conclude that φ is the identity mapping. Thus, $\varphi(z) \equiv z$ and therefore $p_1^1 = p_1^2$ and $p_2^1 = p_2^2$.
- Now, we want to show that every diagram of the form shown in Figures 10a–10i corresponds via a mapping defined by formula (8.1) to a quadratic differential of the form (6.1) with some p_1 and p_2.

  To show this, we will construct a compact Riemann surface $\mathcal{R}$ using identification of appropriate edges of the diagram. For more general quadratic differentials, similar construction was used in [32].

  To be definite, we will give detailed construction for the diagram shown in Figure 10a. In all other cases constructions of an appropriate Riemann surface follow same lines. Consider a domain Ω defined by

$$\begin{aligned}\Omega = \quad & \{w : x_1 < \Re w < x_1', \ \Im w \le 0\} \cup \\ & \{w : 0 < \Im w < h\} \setminus \{w = t + ih_1 : t \ge x_2\}.\end{aligned}$$

Thus, Ω is a slit horizontal strip shown in Figure 10a with a vertical half-strip $\{w : x_1 < \Re w < x_1', \ \Im w \le 0\}$ attached to this horizontal strip along the interval (x_1, x_1'); see Figure 11. To construct a Riemann surface $\mathcal{R}$ mentioned above, we identify boundary points of Ω as follows:

$$\begin{aligned} iy &\simeq 1 + iy && \text{for } y \le 0, \\ -x &\simeq 1 + x && \text{for } x \ge 0, \\ x + x_2 + i(h_1 - 0) &\simeq -x + x_2' + ih && \text{for } x \ge 0, \\ x + x_2 + i(h_1 + 0) &\simeq x + x_2' + ih && \text{for } x \ge 0. \end{aligned} \tag{8.13}$$

After identifying points by rules (8.13), we obtain a surface, which is homeomorphic to a complex sphere $\overline{\mathbb{C}}$ punctured at three points. These punctures correspond boundary points of Ω situated at ∞. One puncture corresponds to the point of $\partial\Omega$, we call it b_1, which is accessible along the path $\{z = \frac{1}{2} + it\}$ as $t \to -\infty$. Second puncture corresponds to a point b_2 in $\partial\Omega$, which is accessible along the path $\{z = t + i\frac{h_1+h}{2}\}$ as $t \to \infty$. The third puncture corresponds to two boundary points of Ω; one of them, we call it b_3^1, is accessible along the path $\{z = t + ih_1\}$ as $t \to -\infty$ and the other one, we call it b_3^2, is accessible along the path $\{z = t + \frac{h_1}{2}\}$ as $t \to \infty$. Adding these three punctures, we obtain a compact surface $\mathcal{R}$ which is homeomorphic to a sphere $\overline{\mathbb{C}}$.

Next, we introduce a complex structure on $\mathcal{R}$ as follows. Every point of $\mathcal{R}$ corresponding to a point of Ω inherits its complex structure from Ω as a subset of $\mathbb{C}$. A point of $\mathcal{R}$ corresponding to iy inherits its complex structure from two half-disks $\{z : |z - iy| < \varepsilon, -\pi/2 \le \arg(z - iy) \le \pi/2\}$ and $\{z : |z - (1 + iy)| < \varepsilon, \pi/2 \le \arg(z - iy) \le 3\pi/2\}$. Similarly, every point of $\mathcal{R}$ corresponding to a finite boundary point of Ω, except those which corresponds to the points x_1, and $x_2 + ih_1$, inherits its complex structure from the corresponding boundary half-disks.

Now we assign complex charts for five remaining special points. For a point $x_1 \simeq x_1'$ a complex chart can be assigned as follows:

$$\zeta = \begin{cases} (w-1)^{\frac{2}{3}} & \text{if } |w-1| < \varepsilon,\ 0 \le \arg w \le \frac{3\pi}{2}, \\ (-w)^{\frac{2}{3}} & \text{if } |w| < \varepsilon,\ -\frac{\pi}{2} \le \arg w \le \pi, \end{cases} \tag{8.14}$$

where the branches of the radicals are taken such that $\zeta(w) > 0$ when w is real such that $w > 1$ or $w < 0$.

Similarly, to assign a complex chart to a point $x_2 + ih_1 \simeq x_2' + ih$, we use the following mapping:

$$\zeta = \begin{cases} (w-(x_2+ih_1))^{\frac{2}{3}} & \text{if } |w-(x_2+ih_1)| < \varepsilon, \\ & \quad 0 \le \arg(w-(x_2+ih_1)) \le 2\pi, \\ (w-(x_2'+ih))^{\frac{2}{3}} & \text{if } |w-(x_2'+ih)| < \varepsilon, \\ & \quad \pi \le \arg(w-(x_2'+ih)) \le 2\pi, \end{cases} \tag{8.15}$$

with appropriate branches of the radicals.

To a point of $\mathcal{R}$ corresponding to an infinite boundary point b_1, a complex chart can be assigned via the function

$$\zeta = \exp(-2\pi i w) \quad \text{for } w \text{ such that } 0 \le \Re w \le 1,\ \Im w < 0, \tag{8.16}$$

which maps the half-strip $\{w : 0 \le \Re w \le 1, \Im w < 0\}$ onto the unit disc punctured at $\zeta = 0$. This mapping respects the first identification rule in (8.13) and the origin $\zeta = 0$ represents the point b_1.

To assign a complex chart to a puncture corresponding to a pair of boundary points b_3^1 and b_3^2, we will work with horizontal half-strips H_3^1 and H_3^2 defined as follows. The boundary of H_3^1 consists of two horizontal rays $\{w : w = t : t \ge u_6\}$ and $\{w = t + ih_1 : t \ge x_2\}$ and a line segment $[u_6, x_2 + ih_1]$; the boundary of H_3^2 consists of two horizontal rays $\{w : w = t : t \le u_5\}$ and $\{w = t + ih : t \le x_2'\}$ and a line segment $[u_5, x_2' + ih]$. To construct a required chart, we rotate the half-strip H_3^1 by angle π with respect to the point $w = 1/2$ and then we glue the result to the half-strip H_3^2 along the interval $(-\infty, u_5)$. As a result, we obtain a wider half-strip $\widetilde{H}_3$ the boundary of which consists of horizontal rays $\{w = t + ih : t < x_2'\}$ and $\{w = t - ih_1 : t < 1 - x_2\}$ and a line segment $[1 - x_2 - ih_1, x_2' + ih]$. After that we map an obtained wider half-strip $\widetilde{H}_3$ conformally onto the unit disk in such a way that horizontal rays are mapped onto appropriate logarithmic spirals. The conformal mapping just described can be expressed explicitly in the following form:

$$\zeta = \begin{cases} \exp(2\pi i C_3(1-u_5-w)) & \text{if } w \in H_3^1, \\ \exp(2\pi i C_3 w) & \text{if } w \in H_3^2, \end{cases} \tag{8.17}$$

where

$$C_3 = \frac{(x_2 + x_2' - 1) - i(h + h_1)}{|(x_2 + x_2' - 1) - i(h + h_1)|^2}.$$

In a similar way we can assign a complex chart to the puncture corresponding to the boundary point b_2. In this case, we use the following mapping from the horizontal half-strip H_2, the boundary of which consists of the rays $\{w = t + ih_1 : t \ge x_2\}$ and $\{w = t + ih : t \ge x_2'\}$ and a line segment $[x_2 + ih_1, x_2' + ih]$, onto the unit disk:

$$\zeta = \exp(-2\pi i C_2(w - (x_2 + ih_1))) \quad \text{for } w \in H_2, \tag{8.18}$$

where

$$C_2 = \frac{(x_2' - x_2) - i(h - h_1)}{|(x_2' - x_2) - i(h - h_1)|^2}.$$

Now, our compact surface $\mathcal{R}$ with conformal structure introduced above is conformally equivalent to the Riemann sphere $\overline{\mathbb{C}}$. Let $\Phi(w)$ be a conformal mapping from $\mathcal{R}$ onto $\overline{\mathbb{C}}$ uniquely determined by conditions

$$\Phi(b_1) = \infty, \quad \Phi(b_2) = 1, \quad \Phi(b_3^1) = \Phi(b_3^2) = -1.$$

Next, we consider a quadratic differential $\mathcal{Q}(w)\,dw^2$ on $\mathcal{R}$ defined by

$$\mathcal{Q}(w)\,dw^2 = 1 \cdot dw^2 \tag{8.19}$$

if w is finite and $w \ne x_1$ and $w \ne x_2 + ih_1$. This quadratic differential can be extended to the points $w = x_1$ and $w = x_2 + ih_1$ as a quadratic differential having simple zeroes at these points in terms of the local parameters defined by formulas (8.14) and (8.15), respectively.

Similarly, using local parameters defined by formulas (8.16), (8.17), and (8.18), we can extend quadratic differential (8.19) to the points of $\mathcal{R}$ corresponding to the infinite boundary points of Ω situated at b_1 b_2, and $b_3^1 \simeq b_3^2$, respectively.

We note that the horizontal strips $\{w : 0 < \Im w < h_1\}$ and $\{w : h_1 < \Im w < h\}$ are strip domains of the quadratic differential (8.19), while the half-strip $\{w : 0 \le \Re w \le 1, \Im w < 0\}$, which boundary points are identified by the first rule in (8.13), defines a circle domain of this quadratic differential.

Now, when the quadratic differential (8.19) have been extended to a quadratic differential defined on the whole Riemann surface $\mathcal{R}$, we may use conformal mapping $z = \Phi(w)$ to transplant this quadratic differential to get a quadratic differential $\widehat{Q}(z)\,dz^2$ defined on $\overline{\mathbb{C}}$. Since critical points of a quadratic differential are invariant under conformal mapping, it follows that $\widehat{Q}(z)\,dz^2$ has second-order poles at the points $z = \infty$, $z = 1$ and $z = -1$ and it has simple zeroes at the images $\Phi(x_1)$ and $\Phi(x_2 + ih_1)$ of the points $w = x_1$ and $w = x_2 + ih_1$.

Furthermore, the pole $z = \infty$ belongs to a circle domain of $\widehat{Q}(z)\,dz^2$ and every trajectory in this circle domain has length 1. Using the above information, we conclude that $\widehat{Q}(z)\,dz^2 = \frac{1}{4\pi^2} Q(z)\,dz^2$, where $Q(z)\,dz^2$ is given by formula (6.1) with $p_1 = \Phi(x_1)$ and $p_2 = \Phi(x_2 + ih_1)$.

Combining our observations made in this section, we conclude the following:

Every quadratic differential of the form (6.1) *having two strip domains generates a diagram of the type shown in Figures* 10a–10i *and every diagram of this type corresponds to one and only one quadratic differential with two strip domains in its domain configuration of the form* (6.1).

9. How parameters count critical geodesics and critical loops

In Section 8, we described Q-geodesics corresponding to the quadratic differential (6.1) in terms of Euclidean geodesics in the w-plane. In this section, we explain how this information can be used to find the number of short geodesics and geodesic loops for each pair of zeros p_1 and p_2.

To be definite, we will work with the case **6.3(b2)** of Theorem 4 assuming that

$$\Im p_1 > 0, \quad \text{and } p_2 \in E^{+}_{-1}(p_1). \tag{9.1}$$

In all other cases, the number of short geodesics and geodesic loops can be found similarly.

Under conditions (9.1), the domain configuration of the quadratic differential (6.1) consists of domains D_∞, G_1, and G_2 as it is shown in Figure 4a and Figure 4b and possible configurations of images of G_1 and G_2 under the mapping (8.1) are shown in Figures 10a–10i.

Let $\varepsilon > 0$ be sufficiently small and let dz^{+}_{ε} denote a tangent vector to the trajectory of the quadratic differential (6.1) at $z = 1+\varepsilon$, which can be found from the equation $Q(z)\,dz^2 > 0$. Using (7.1) and (7.2), we find that

$$\arg(dz^{+}_{\varepsilon}) = \frac{\pi}{2} - \frac{1}{2}\arg C_1 + o(1) = \frac{\pi}{2} - \frac{1}{2}\arg((p_1-1)(p_2-1)) + o(1), \tag{9.2}$$

where $o(1) \to 0$ as $\varepsilon \to 0$. We assume here that $-\frac{\pi}{2} \le \arg(dz^{+}_{\varepsilon}) \le \frac{\pi}{2}$.

If $1+\varepsilon \in \gamma_1$ then the tangent vector dz^{+}_{ε} corresponds to the direction on γ_1 from $z = 1$ to $z = p_1$. Let $\alpha^{+}_{\varepsilon} = \alpha^{+} + o(1)$, where α^{+} is a constant such that $0 \le \alpha^{+} \le \pi$, denote the angle formed at the point $1+\varepsilon \in \gamma_1$ by dz^{+}_{ε} and the vector $\overrightarrow{v} = -i$, which is tangent to the circle $\{z : |z-1| = \varepsilon\}$ at $z = 1+\varepsilon$. It follows from (9.2) that

$$\alpha^{+} = \pi - \frac{1}{2}\arg C_1 = \pi - \frac{1}{2}\arg((p_1-1)(p_2-1)). \tag{9.3}$$

Similarly, if dz^{-}_{ε} denote the tangent vector to the trajectory of the quadratic differential (6.1) at $z = -1+\varepsilon$, then

$$\arg(dz^{-}_{\varepsilon}) = \frac{\pi}{2} - \frac{1}{2}\arg C_{-1} + o(1) = \frac{\pi}{2} - \frac{1}{2}\arg((p_1+1)(p_2+1)) + o(1). \tag{9.4}$$

Suppose that $1+\varepsilon \in \gamma_{-1}$ and that d^{-}_{ε} shows direction on γ_{-1} from $z = -1$ to $z = p_2$. As before we can find constant α^{-}, $0 \le \alpha^{-} \le \pi$, such that the angle formed at $z = -1+\varepsilon \in \gamma_{-1}$ by the vectors dz_ε^+ and $\overrightarrow{v} = -i$ is equal to $\alpha^- + o(1)$, where $o(1) \to 0$ as $\varepsilon \to 0$ and

$$\alpha^- = \pi - \frac{1}{2}\arg C_{-1} = \pi - \frac{1}{2}\arg((p_1+1)(p_2+1)). \tag{9.5}$$

To relate angles α^+ and α^- to geometric characteristics of diagrams in Figures 10a–10i, we recall that geodesics are conformally invariant and that for small $\varepsilon > 0$ a geodesic loop γ_ε^+ which passes through the point $z = 1+\varepsilon$ and surrounds the pole $z = 1$ is an infinitesimal circle. Therefore the angle formed by the vector dz_ε^+ and the tangent vector to γ_ε^+ at $z = 1+\varepsilon$ equals $\alpha^+ + o(1)$.

Similarly, the angle formed by the vector dz_ε^- and the tangent vector to the corresponding geodesic loop $\gamma_\varepsilon^- \ni -1+\varepsilon$ surrounding the pole at $z = -1$ is equal to $\alpha^- + o(1)$.

Since geodesics are conformally invariant and since conformal mappings preserve angles, we conclude that trajectories of the quadratic differential $\mathcal{Q}(w)\,dw^2$ defined in Section 8 (see formula (8.19)) form angles of opening α^+ or α^- with the images of the corresponding geodesic loops γ_ε^+ or γ_ε^-, respectively. Since the metric defined by the quadratic differential (8.19) is Euclidean, it follows that the corresponding images of geodesic loops are line segments joining pairs of points identified by relations (8.13).

Using this observation and identification rule $-x + x_2' + ih \simeq x + x_2 + ih_1$, we conclude that the segment $[x_2 + ih_1, x_2' + ih]$ forms an angle $\pi - \alpha^-$ with the positive real axis; i.e.,

$$\pi - \alpha^- = \arg((x_2' - x_2) + i(h - h_1)). \tag{9.6}$$

To find an equation for the angle α^+, we will use the half-strip $\widetilde{H}_3$ constructed at the end of Section 8, which is related to a conformal mapping defined by formula (8.17). In this case, $\pi - \alpha^+$ is equal to the angle formed by the segment $[1 - x_2 - ih_1, x_2' + ih]$ with the positive real axis; i.e.,

$$\pi - \alpha^+ = \arg((x_2 + x_2' - 1) + i(h + h_1)). \tag{9.7}$$

Equating the right-hand sides of equations (9.3) and (9.4) to the right-hand sides of equations (9.7) and (9.6), respectively, we obtain two equations, which relate parameters x_2, x_2', h_1, and h. Combining this with equations (7.10)–(7.12), we obtain the following system of four equations:

$$\begin{aligned}
\arg((x_2 + x_2' - 1) + i(h + h_1)) &= \tfrac{1}{2}\arg((p_1-1)(p_2-1))\\
\arg((x_2' - x_2) + i(h - h_1)) &= \tfrac{1}{2}\arg((p_1+1)(p_2+1))\\
h_1 &= \tfrac{1}{4}\Im\Big(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\Big)\\
h &= \tfrac{1}{4}\Im\Big(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\Big).
\end{aligned}$$

This system of equations can be solved to obtain the following:

$$\begin{aligned} x_2 + ih_1 &= \tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right), \\ x_2' + ih &= \tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right). \end{aligned} \tag{9.8}$$

Now, when the points x_2+ih_1 and $x_2'+ih$ are determined, we can give explicit conditions on the zeros p_1 and p_2 which correspond to all subcases (a)–(i) of the case **6.3(b2)** discussed in Section 8.

Theorem 5. *Suppose that zeros p_1 and p_2 satisfy conditions* (9.1). *Then the number of short geodesics and geodesic loops and their topology are determined by the following inequalities, which corresponds to the subcases* (a)–(i) *of Case* **6.3(b2)** *described in Section* 8 *and shown in Figures* 10a–10i*:*

Case (a) *with four short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned} 0 &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\right) \\ &< \arg\left(\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right)\right) < \pi. \end{aligned}$$

Case (b) *with four short geodesics and two critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned} 0 &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\right) \\ &= \arg\left(\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right)\right) < \pi. \end{aligned}$$

Case (c) *with four short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned} 0 &< \arg\left(\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right)\right) \\ &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\right) < \pi, \end{aligned}$$

$$\begin{aligned} 0 &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\right) \\ &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right)\right) < \pi. \end{aligned}$$

Case (d) *with three short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned} 0 &< \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} - \sqrt{(p_1+1)(p_2+1)}\right)\right) \\ &= \arg\left(-\tfrac{1}{2} + \tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)} + \sqrt{(p_1+1)(p_2+1)}\right)\right) < \pi. \end{aligned}$$

Case (e) *with four short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned}0 &< \arg\left(-\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &< \arg\left(-\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi,\\ 0 &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi.\end{aligned}$$

Case (f) *with three short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned}0 &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &= \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi.\end{aligned}$$

Case (g) *with four short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned}0 &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi,\\ 0 &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &< \arg\left(-\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi.\end{aligned}$$

Case (h) *with four short geodesics and two critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned}0 &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &= \arg\left(-\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi.\end{aligned}$$

Case (i) *with four short geodesics and three critical geodesic loops occurs if the following conditions are satisfied:*

$$\begin{aligned}0 &< \arg\left(-\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}+\sqrt{(p_1+1)(p_2+1)}\right)\right)\\ &< \arg\left(\tfrac{1}{2}+\tfrac{1}{4}\left(\sqrt{(p_1-1)(p_2-1)}-\sqrt{(p_1+1)(p_2+1)}\right)\right)<\pi.\end{aligned}$$

10. Some related questions

Our results presented in Sections 6–9 provide complete information concerning critical trajectories and Q-geodesic of the quadratic differential (6.1). This allows us to answer many related questions. As an example, we will discuss three questions originated in the study of limiting distributions of zeros of Jacobi polynomials.

Below, we suppose that $p_1, p_2 \in \mathbb{C}$ are fixed. Then we consider the family of quadratic differentials $Q_s(z)\,dz^2$ depending on the real parameter s, $0 \le s < 2\pi$, such that

$$Q_s(z)\,dz^2 := e^{-is}Q(z)\,dz^2 = -e^{-is}\frac{(z-p_1)(z-p_2)}{(z-1)^2(z+1)^2}\,dz^2. \tag{10.1}$$

1) For how many values of s, $0 \le s < 2\pi$, the quadratic differential $Q_s(z)\,dz^2$ has a trajectory loop with end points at p_1 and for how many values of s $Q_s(z)\,dz^2$ has a trajectory loop with end points at p_2?
2) For how many values of s, $0 \le s < 2\pi$, the corresponding quadratic differential $Q_s(z)\,dz^2$ has a short critical trajectory?
3) How we can find the values of s, $0 \le s < 2\pi$, mentioned in questions stated above?

To answer these questions we need two simple facts:

(a) First, we note that γ is a short trajectory loop or, respectively, a short critical trajectory for the quadratic differential (10.1) with some s if and only if γ is a short geodesic loop or, respectively, a short geodesic joining points p_1 and p_2 for the quadratic differential (6.1). Thus, the numbers of values s in question 1) and question 2), respectively, are bounded by the number of short geodesic loops and the number of short geodesics, respectively. In the most general case with one circle domain and two strip domains, these short geodesic loops and short geodesics were described in Theorem 5 and their images under the canonical mapping were shown in Figures 10a–10i. Of course, one value of s can correspond to more than one short geodesic loop and more than one short geodesic.

(b) To find the values of s in question 3), we use the following observation. If l is a straight line segment in the image domain Ω forming an angle α, $0 \le \alpha < \pi$, with the direction of the positive real axis, then l is an image under the canonical mapping (8.1) of an arc of a trajectory of the quadratic differential (10.1) with

$$s = 2\alpha. \tag{10.2}$$

We will use (10.2) to find values of s which turn short geodesic loops and short geodesics into short trajectory loops and short trajectories, respectively. It is convenient to introduce notations α_∞, α_{12}, α'_{12}, α_{22}, α'_{22}, α''_{22}, and so on, to denote the angles formed by corresponding geodesics γ_∞, γ_{12}, γ'_{12}, γ_{22}, γ'_{22}, γ''_{22}, and so on (considered in the w-plane) with the positive direction of the real axis. Furthermore, we will use notations $\mathcal{A}(6.1)$, $\mathcal{A}(6.1(\mathrm{a}))$, $\mathcal{A}(6.2)$, $\mathcal{A}(6.3(\mathrm{a}))$, $\mathcal{A}(6.3(\mathrm{b1}))$, $\mathcal{A}(6.3(\mathrm{b2})(\mathrm{a}))$, and so on, to denote the sets of all angles introduced above in the cases under consideration; i.e., in the cases **6.1**, **6.2**, **6.3(a)**, **6.3($\mathbf{b_1}$)**, **6.3(b2)**(a), and so on.

Now, we are ready to answer questions stated above. We proceed with two steps. First, we identify the type of domain configuration $\mathcal{D}_Q$. This will provide us with the first portion of necessary information. We recall that in general there are at most three geodesic loops centered at $z = \infty$, $z = 1$, and $z = -1$. Thus, the maximal number of values s in question 1) is at most three. Then we identify which of the schemes corresponds to the parameters p_1, p_2 (in the most general case these schemes are shown in Figures 10a–10i). This will provide us with the remaining portion of necessary information.

- Suppose that $\mathcal{D}_Q$ has type **6.1**. Then we already have three circle domains and therefore $s = 0$ is the only value for which $Q_s z)\,dz^2$ may have short trajectory loops. In case **6.1(a)**, we have short trajectory loops centered at $z = 1$ and $z = -1$ and no other such loops. In case **6.1(b)** with $1 < p_2 < p_1$ (respectively with $p_1 < p_2 < -1$), we have short trajectory loops centered at $z = \infty$ and $z = 1$ (respectively, at $z = \infty$ and $z = -1$). In case **6.1(c)**, there are no short geodesic loops.

As concerns short critical trajectories for domain configuration of type **6.1**, again $s = 0$ is the only value for which there are such trajectories. This follows from the fact discussed in Section 8 that in case **6.1** there are no other simple geodesics joining p_1 and p_2. In cases **6.1(a)** and **6.1(b)**, there is a single short critical trajectory which is the interval $\gamma_0 = (p_2, p_1)$. In case **6.1(c)**, there are three short critical trajectories which are arcs γ_0, γ_1, and γ_{-1} shown in Figure 1c.

- Next, we consider the case when $\mathcal{D}_Q$ has type **6.2**. For $s = 0$, we have two short trajectory loops. As before, we assume that these loops surround points $z = -1$ and $z = \infty$. In other cases discussion is similar, we just have to switch roles of the poles of the quadratic differential (10.1).

In this case, $\mathcal{A}(6.2) = \{0, \alpha_{11}, \alpha_{12}, \alpha'_{12}, \alpha_{21}, \alpha'_{21}\}$. One more value of s, for which we may have a short trajectory loop (centered at $z = 1$) may occur for $s = 2\alpha_{11} = -\arg((1 - p_1)(1 - p_2))$. If $|\gamma_\infty|_Q > |\gamma_{-1}|_Q$ then we will have a short geodesic loop from p_1 to p_1. This loop corresponds to a geodesic γ_{11} in Figure 8a. If $|\gamma_\infty|_Q < |\gamma_{-1}|_Q$, then we will have a similar short geodesic loop from p_2 to p_2. In the case $|\gamma_\infty|_Q = |\gamma_{-1}|_Q$, we have $\alpha_{11} = \alpha_{12} = \alpha'_{21}$. In this case, we do not have the third short geodesic loop. Instead, we have two short critical trajectories joining p_1 and p_2.

By (10.2), the value of s, which corresponds to the third loop (if it exists) is equal to $2\alpha_{11}$. As concerns values of s corresponding to short critical trajectories, in case **6.2** with $|\gamma_\infty|_Q \neq |\gamma_{-1}|_Q$ we have four such values. These values are $2\alpha_{12}$, $2\alpha'_{12}$, $2\alpha_{21}$, and $2\alpha'_{21}$ (see Fig. 8a).

If $|\gamma_\infty|_Q = |\gamma_{-1}|_Q$, then there are three values of s, which produce short geodesics from p_1 to p_2. Two of these values, $s = 2\alpha'_{12}$ and $s = 2\alpha_{21}$, generate one short critical trajectory each. The third value $s = 2\alpha_{12}$ generates two short critical trajectories.

• Turning to the most general case **6.3**, we will give detailed account for subcases **6.3(b1)** and **6.3(b2)**(i), in all other subcases consideration is similar.

First, we consider the subcase **6.3(b1)** when the domain configuration $\mathcal{D}_Q$ consists of one circle domain and one strip domain; see Figures 3a–3e. In this case, $\mathcal{A}(6.3(b1)) = \{0, \alpha'_{22}, \alpha''_{22}, \alpha_{12}, \alpha'_{12}\}$. The value $s = 0$ generates one short trajectory loop and one short trajectory. The values $s = 2\alpha'_{22}$ and $s = 2\alpha''_{22}$ generate one short trajectory loop each and the values $s = 2\alpha_{12}$ and $s = 2\alpha'_{12}$ generate one short trajectory each.

Let us consider case **6.3(b2)**(i) shown in Figure 10i. We have $\mathcal{A}(6.3(b2)(i)) = \{0, \alpha_{22}, \alpha'_{22}, \alpha_{12}, \alpha'_{12}, \alpha_{21}, \alpha'_{21}\}$ where all angles are distinct. The values $s = 0$, $s = 2\alpha_{22}$, and $s = 2\alpha'_{22}$ generate short trajectory loops γ_∞, γ_{22}, and γ''_{22}, respectively. Remaining values $s = 2\alpha_{12}$, $s = 2\alpha'_{12}$, $s = 2\alpha_{21}$, $s = 2\alpha'_{21}$ generate short trajectories γ_{12}, γ'_{12}, γ_{21}, and γ'_{21}, respectively.

Finally, we note that position of points x_1, x'_1, $x_2 + ih_1$, and $x'_2 + ih$ are given explicitly; see formulas (9.8). Using these formulas one can find explicit expressions for all angles α_{12}, α'_{12}, α_{21}, α'_{21}, and so on, in all possible cases.

11. Figures Zoo

This section contains all our figures. For convenience, we divide the set of all figures in eleven groups.

I. Configurations with three circle domains.

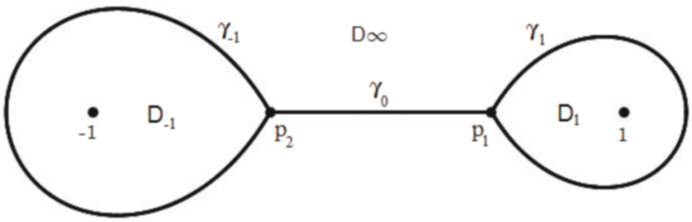


FIG. 1a. Three circle domains. Case **6.1(a)**.

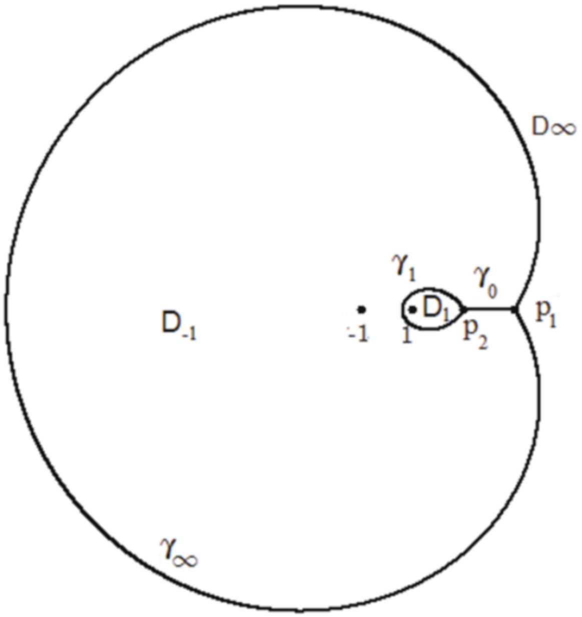


FIG. 1b. Three circle domains. Case **6.1(b)**.

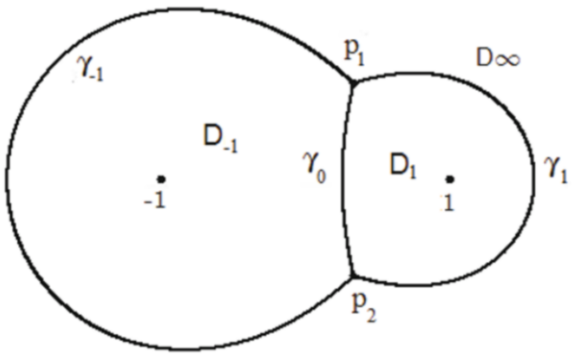


FIG. 1c. Three circle domains. Case **6.1(c)**.

II. Configurations with two circle domains.

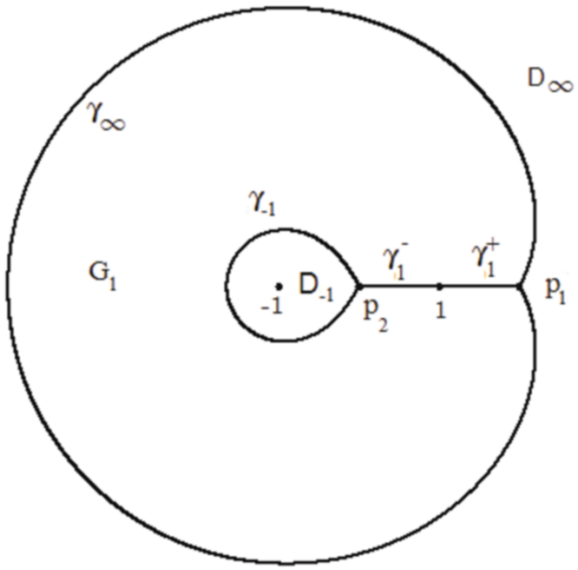


FIG. 2a. Two circle domains. Case **6.2** with symmetric domains.

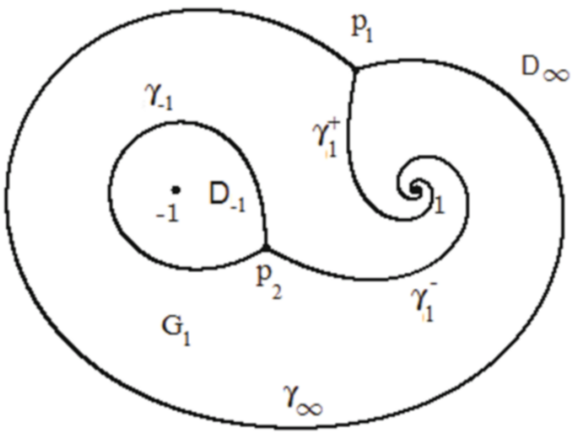


FIG. 2b. Two circle domains. Case **6.2** with non-symmetric domains.

III. Configurations with one circle domain and one strip domain.

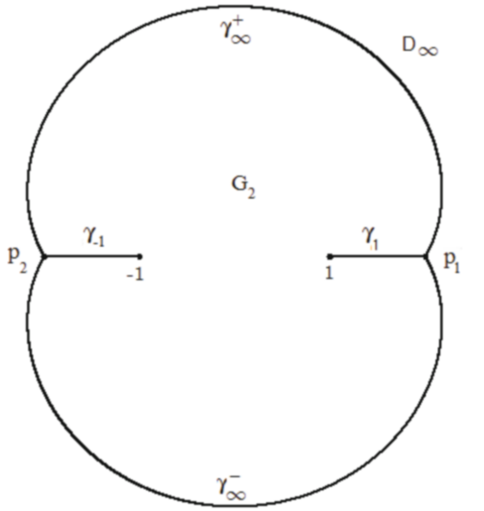


FIG. 3a. One circle domain. Case **6.3(a)** with axial symmetry.

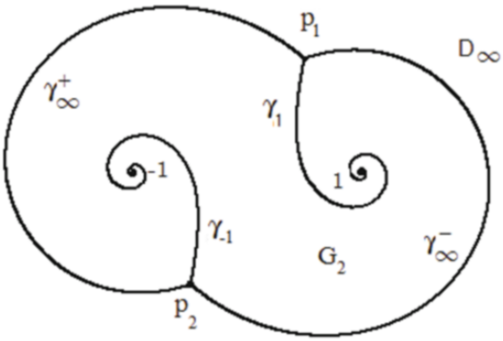


FIG. 3b. One circle domain. Case **6.3(a)** with central symmetry.

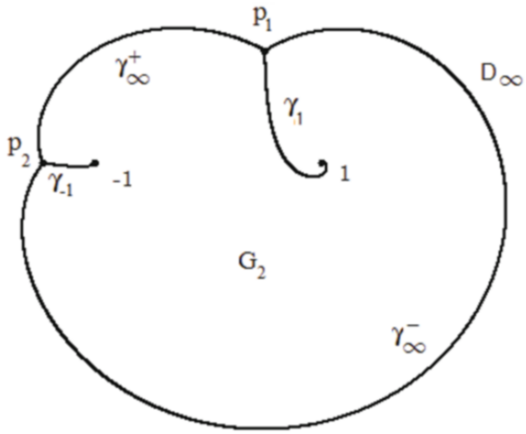


FIG. 3c. One circle domain. Case **6.3(a)** with non-symmetric domains.

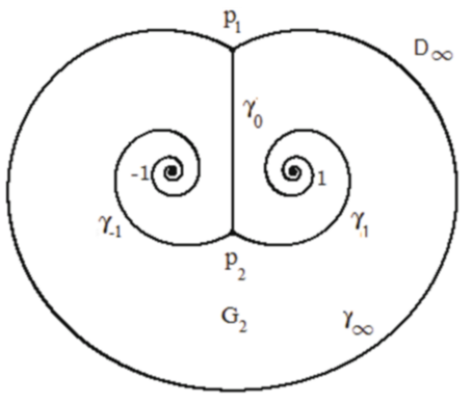


FIG. 3d. One circle domain. Case **6.3(b1)** with symmetric domains.

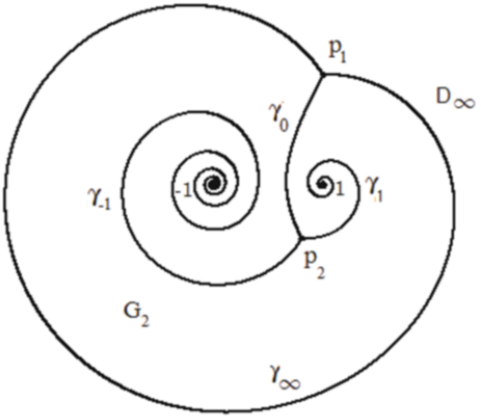


FIG. 3e. One circle domain. Case **6.3(b1)** with non-symmetric domains.

IV. Configurations with one circle domain and two strip domains.

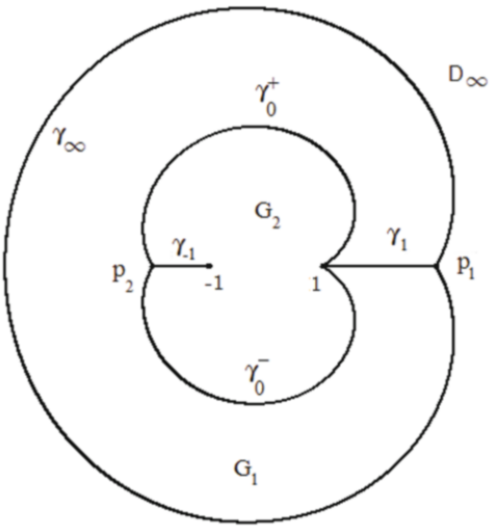


FIG. 4a. One circle domain. Case **6.3(b2)** with symmetric domains.

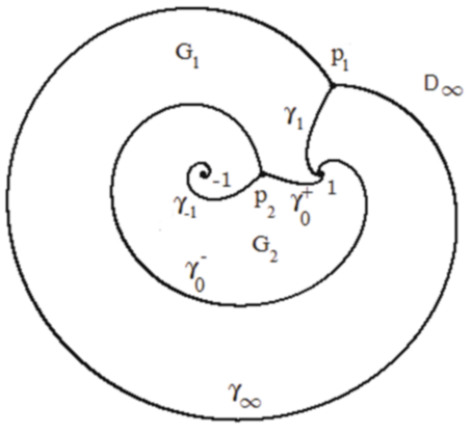


FIG. 4b. One circle domain. Case **6.3(b2)** with non-symmetric domains.

V. Degenerate configurations.

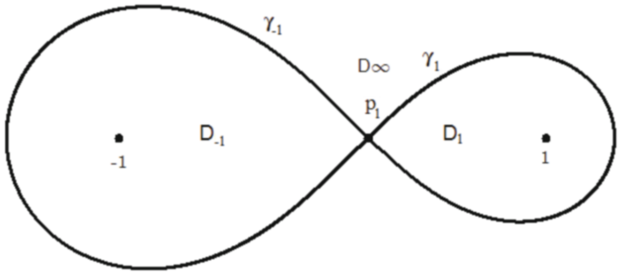


FIG. 5a. Degenerate case with $-1 < p_1 = p_2 < 1$.

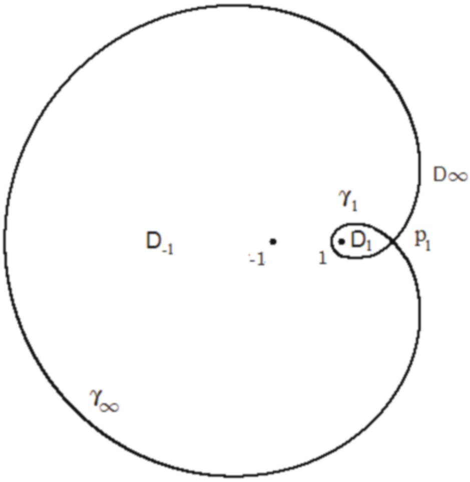


FIG. 5b. Degenerate case with $p_1 = p_2 > 1$.

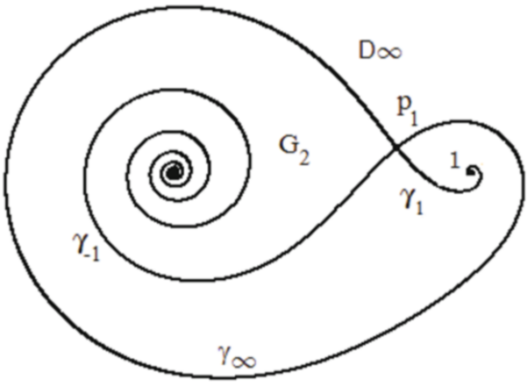


FIG. 5c. Degenerate case with $p_1 = p_2$, $\Im p_1 > 0$.

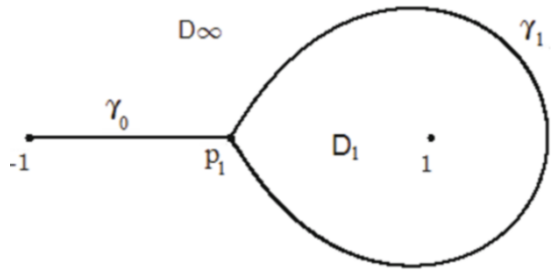


FIG. 5d. Degenerate case with $p_2 = -1$, $-1 < p_1 < 1$.

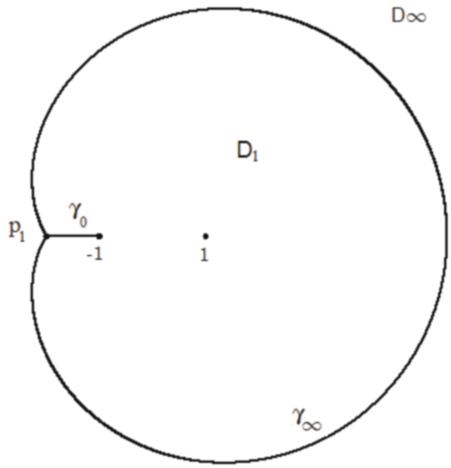


FIG. 5e. Degenerate case with $p_2 = -1$, $p_1 < -1$.

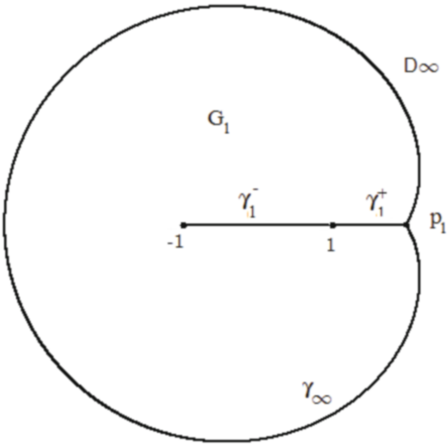


FIG. 5f. Degenerate case with $p_2 = -1$, $p_1 > 1$.

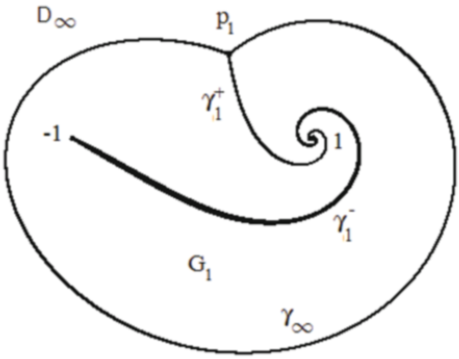


FIG. 5g. Degenerate case with $p_2 = -1$, $\Im p_1 > 0$.

VI. Type regions.

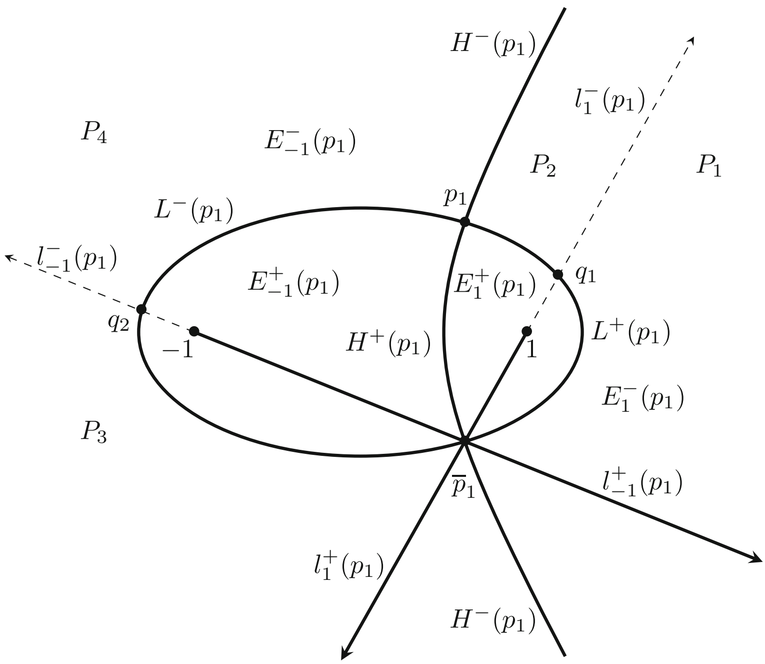


FIG. 6. Type regions.

VII. Figures for the proof of Theorem 4.

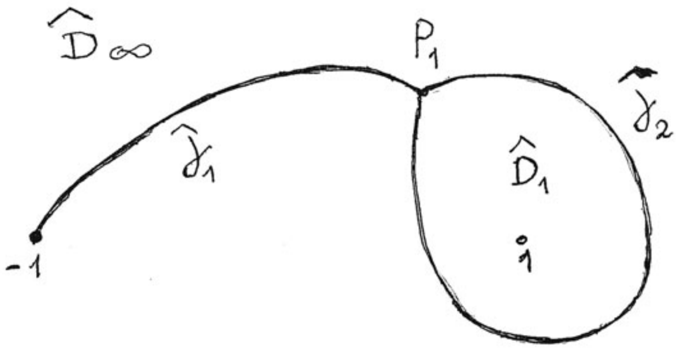


FIG. 7a. Proof of Theorem 4: Impossible limit configuration.

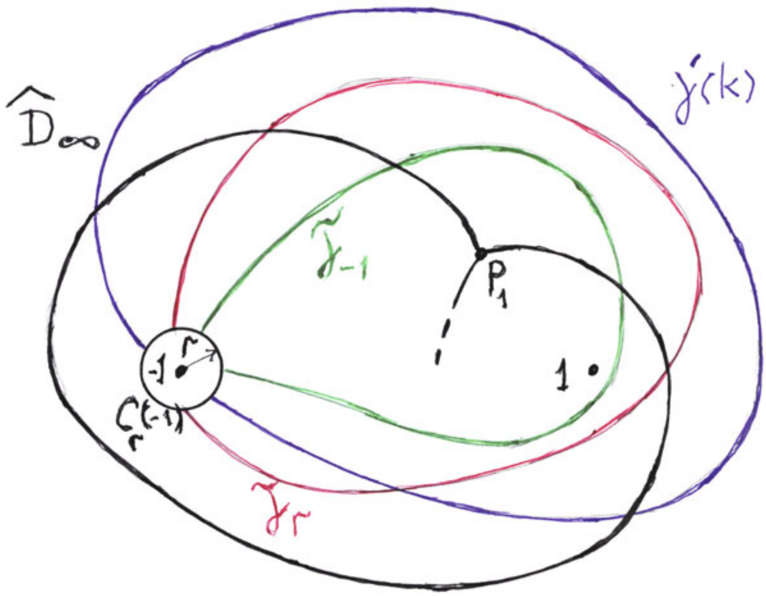


FIG. 7b. Proof of Theorem 4: Limit configuration.

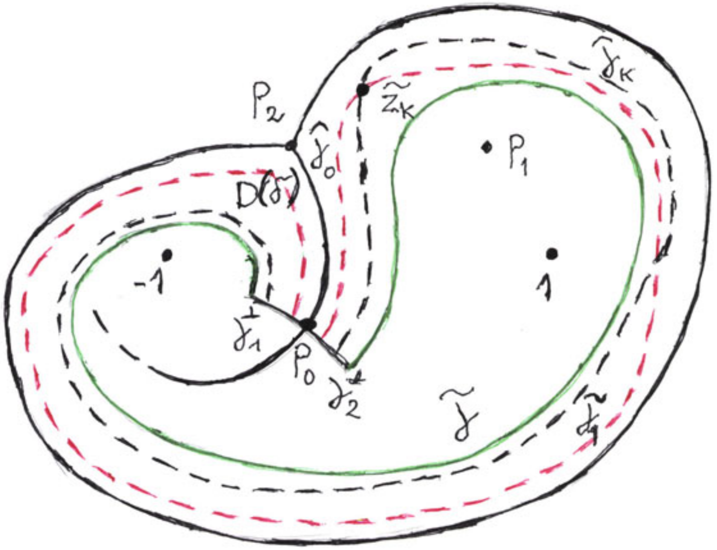


FIG. 7c. Proof of Theorem 4: Q^0-rectangle $D(\delta)$ with trajectories.

VIII. Geodesics and loops in simple cases.

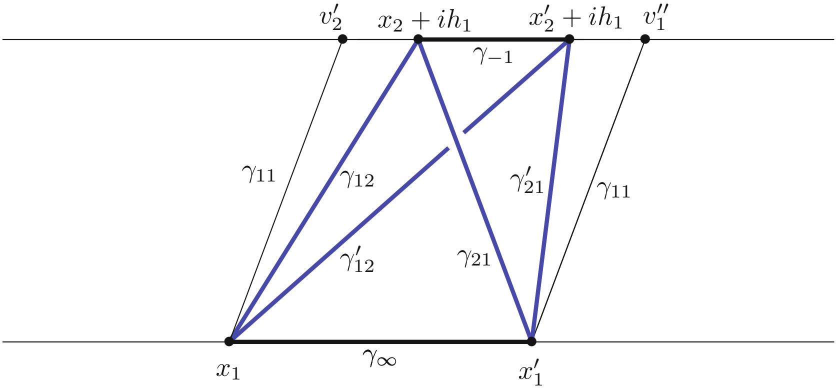


FIG. 8a. Geodesics and loops. Case **6.2**.

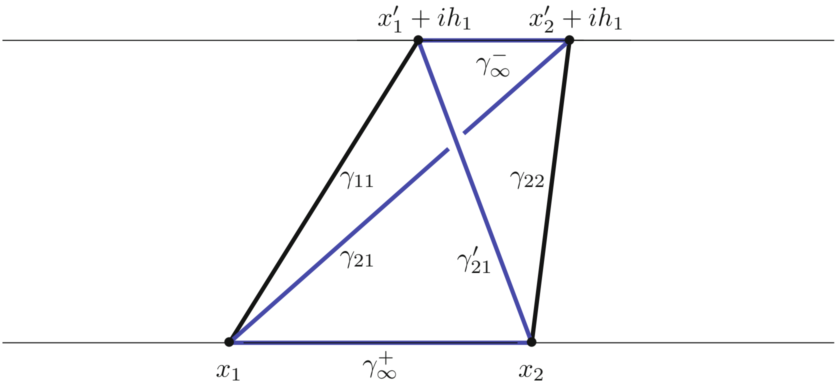


FIG. 8b. Geodesics and loops. Case **6.3(a)**.

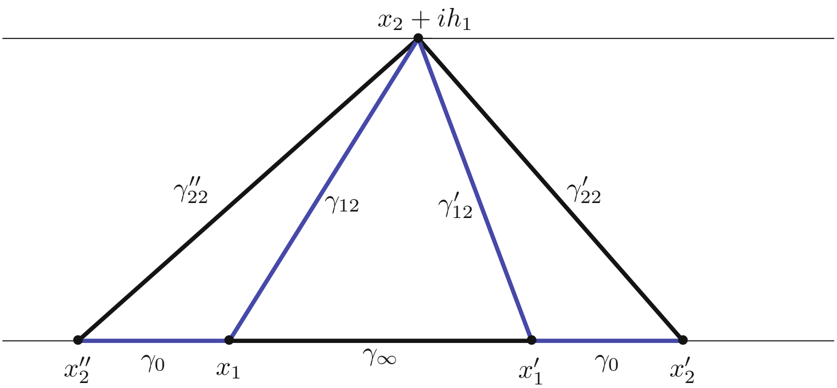


FIG. 8c. Geodesics and loops. Case **6.3(b1)**.

IX. Divergent segments.

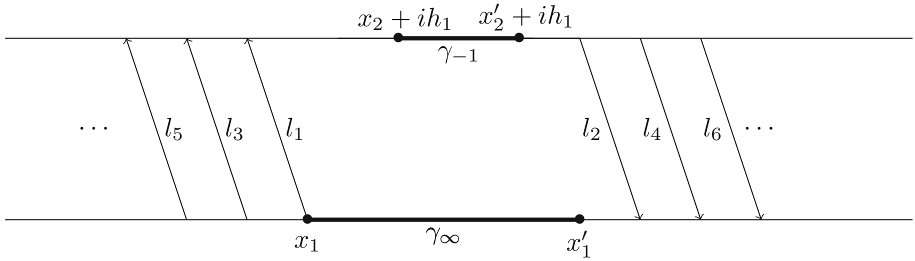


Fig. 9a. Divergent segments. Case **6.2**.

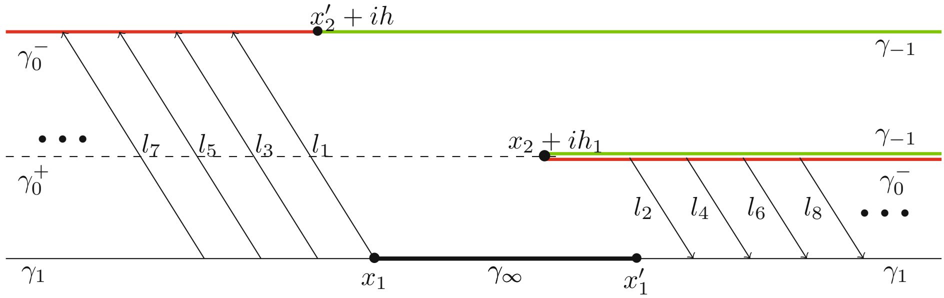


Fig. 9b. Divergent segments. Case **6.3(b2)**.

X. Geodesics and loops in the most general case.

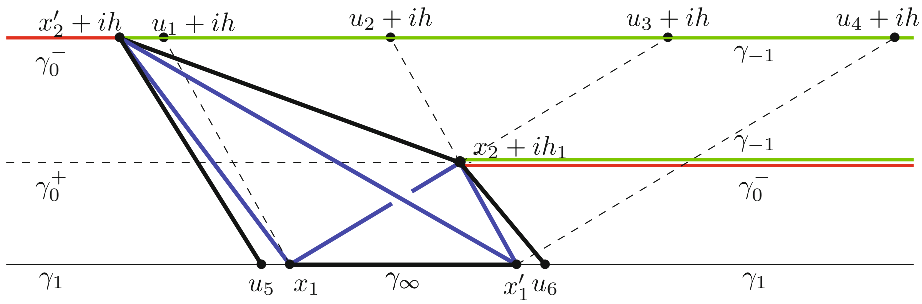


Fig. 10a. Critical geodesics and loops. Case **6.3(b2)**(a).

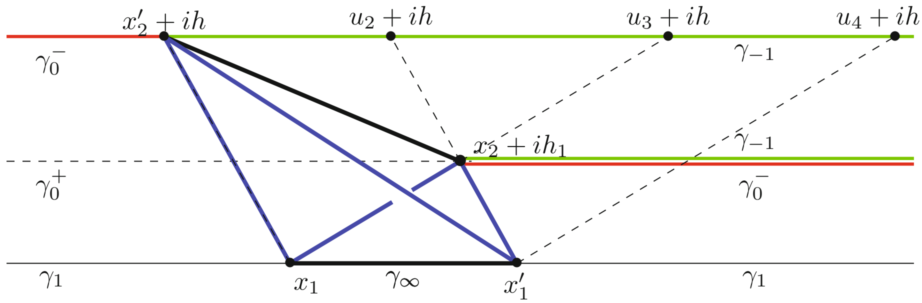


FIG. 10b. Critical geodesics and loops. Case **6.3(b2)**(b).

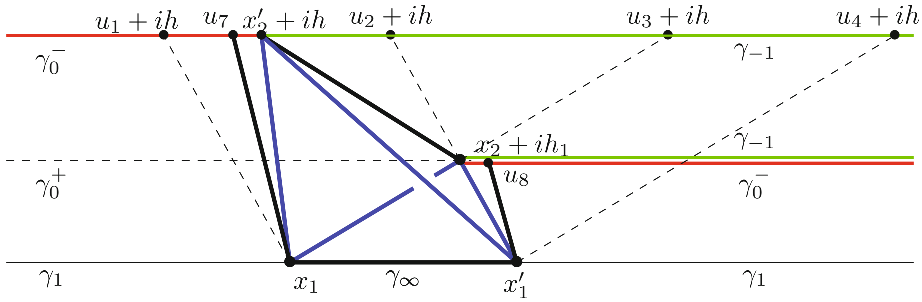


FIG. 10c. Critical geodesics and loops. Case **6.3(b2)**(c).

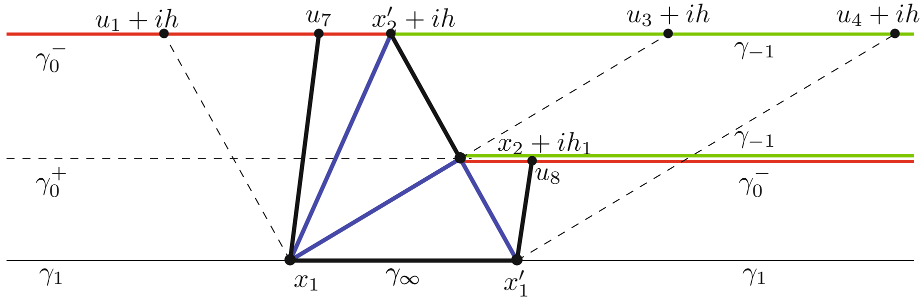


FIG. 10d. Critical geodesics and loops. Case **6.3(b2)**(d).

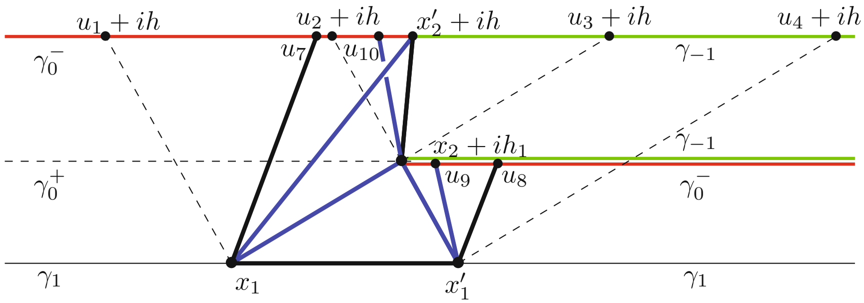


FIG. 10e. Critical geodesics and loops. Case **6.3(b2)**(e).

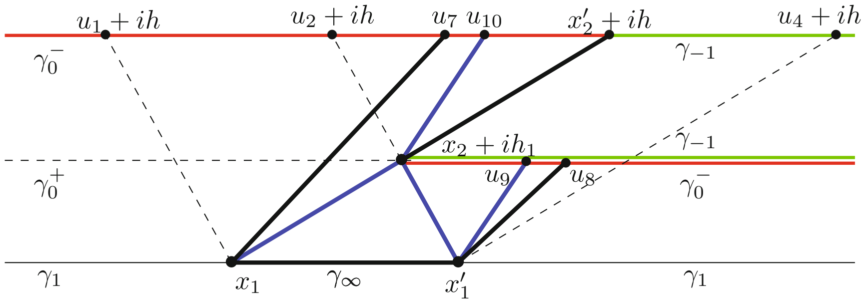


FIG. 10f. Critical geodesics and loops. Case **6.3(b2)**(f).

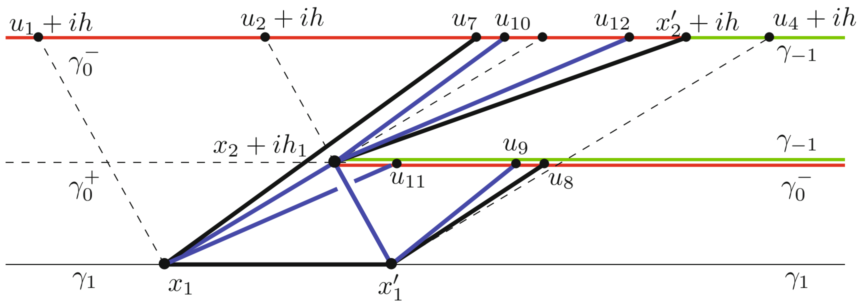


FIG. 10g. Critical geodesics and loops. Case **6.3(b2)**(g).

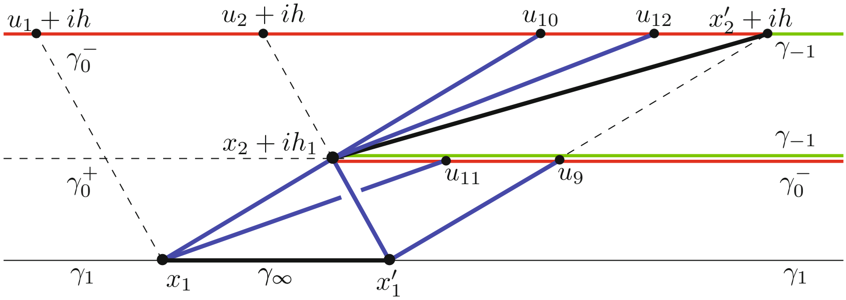


FIG. 10h. Critical geodesics and loops. Case **6.3(b2)**(h).

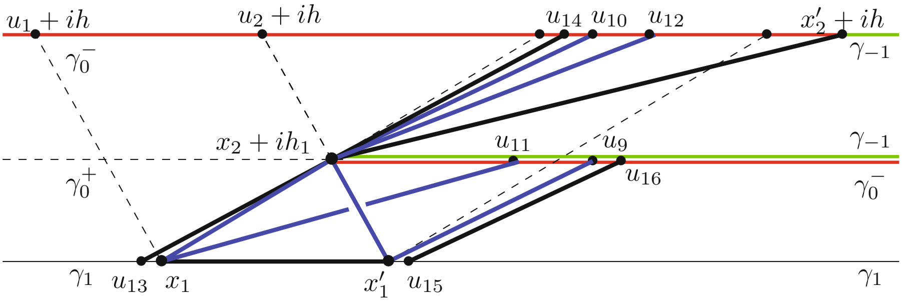


FIG. 10i. Critical geodesics and loops. Case **6.3(b2)**(i).

XI. Identification rules.

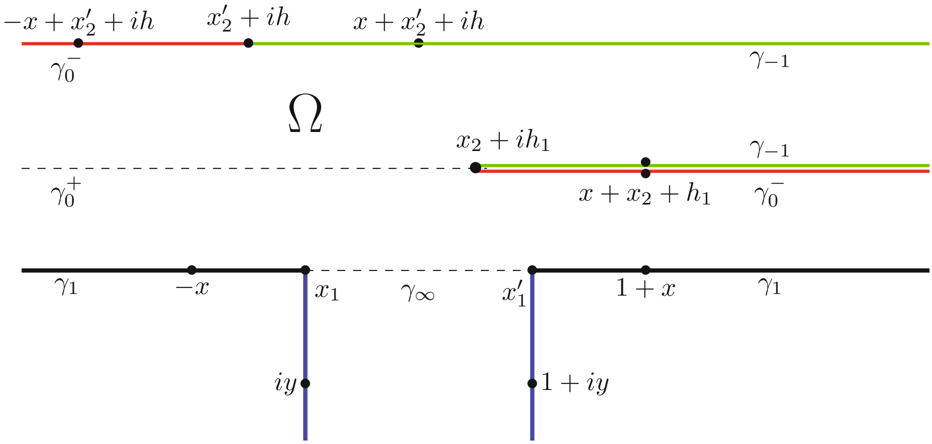


FIG. 11. Domain Ω and identification rules.

Acknowledgement. The authors want to acknowledge the hospitality of the Mittag-Leffler Institute in Spring 2011 where this project was initiated. The first author is also sincerely grateful to R. Bøgvad, A. Kuijlaars, A. Martínez-Finkelshtein, and A. Vasiliev for many useful discussions.

References

[1] M. J. Atia, A. Martínez-Finkelshtein, P. Martínez-González, and F. Thabet, Quadratic differentials and asymptotics of Laguerre polynomials with varying complex coefficients, arXiv: 1311.0372.

[2] M. Abramowitz, I. A. Stegun, eds. (1972), Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, New York: Dover Publications, ISBN 978-0-486-61272-0.

[3] R. G. Bartle, The elements of integration. John Wiley & Sons, Inc., New York-London-Sydney 1966 x+129 pp.

[4] T. Bergkvist, *On asymptotics of polynomial eigenfunctions for exactly solvable differential operators*, J. Approx. Theory 149(2), (2007), 151–187.

[5] T. Bergkvist and H. Rullgård, *On polynomial eigenfunctions for a class of differential operators*, Math. Res. Lett., vol.9, (2002), 153–171.

[6] A. Beurling, Sur la géométrie métrique des surfaces à courbure totale ≤ 0. (French) Comm. Sém. Math. Univ. Lund [Medd. Lunds Univ. Mat. Sem.] 1952, (1952). Tome Supplémentaire, 7–11.

[7] J.-E. Björk, J. Borcea, R. Bøgvad, *Subharmonic Configurations and Algebraic Cauchy Transforms of Probability Measures.* Notions of Positivity and the Geometry of Polynomials Trends in Mathematics 2011, pp. 39–62.

[8] J. Borcea, R.Bøgvad, *Piecewise harmonic subharmonic functions and positive Cauchy transforms.* Pacific J. Math., Vol. 240 (2009), No. 2, 231–265.

[9] J. Borcea, R. Bøgvad and B. Shapiro, Homogenized spectral pencils for exactly solvable operators: asymptotics of polynomial eigenfunctions, Publ. RIMS, vol 45 (2009) 525–568.

[10] J. Borcea, R. Bøgvad, and B. Shapiro, On rational approximation of algebraic functions, Adv. Math. vol 204, issue 2 (2006) 448–480.

[11] J. Borrego Morell, H. Pijeira Cabrera, Strong asymptotic behavior of eigenpolynomials for a class of linear differential operators, preprint, Nov. 15, (2012).

[12] R. Bøgvad, B. Shapiro, On motherbody measures with algebraic Cauchy transform, arXiv:1406.1972, submitted.

[13] R. Bøgvad, D. Khavinson, B. Shapiro, On asymptotic Gauss–Lucas theorem, arXiv:1510.02339, submiited.

[14] R. Bøgvad, Ch. Hägg, A refinement for rational functions of Polya's method to construct Voronoi diagrams, J. Math. Anal. Apple. 452 (2017) 312–334.

[15] M. Charalambides, G. Csordas, The distribution of zeros of a class of Jacobi polynomials. Proc. Amer. Math. Soc. 138 (2010), no. 12, 4345–4357.

[16] C. Díaz Mendoza, R. A. Orive, The Szegö curve and Laguerre polynomials with large negative parameters. J. Math. Anal. Appl. 379 (2011), no. 1, 305–315,

[17] K. A. Driver, P. Duren, Asymptotic zero distribution of hypergeometric polynomials, Numerical Algorithms 21 (1999) 147–156.

[18] K. A. Driver, S. J. Johnston, Asymptotic zero distribution of a class of hypergeometric polynomials. Quaest. Math. 30 (2007), no. 2, 219–230.

[19] P. Duren, B. Guillou, Asymptotic properties of zeros of hypergeometric polynomials, JAT 111 (2001) 329–343.

[20] J. Garnett, *Analytic capacity and measure*. Lecture Notes in Mathematics, Vol. 297. Springer-Verlag, Berlin-New York, 1972. iv+138 pp.

[21] A. A. Gonchar, E. A. Rakhmanov, Equilibrium distributions and the rate of rational approximation of analytic functions, Math.USSR-Sb.62 (2) (1989) 305–348.

[22] J. A. Jenkins, *Univalent functions and conformal mapping*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Neue Folge, Heft 18. Reihe: Moderne Funktionentheorie Springer-Verlag, Berlin-Göttingen-Heidelberg, (1958), vi+169 pp.

[23] A. B. J. Kuijlaars, A. Martínez-Finkelshtein, Strong asymptotics for Jacobi polynomials with varying nonstandard parameters. J. Anal. Math. 94 (2004), 195–234.

[24] A. B. J. Kuijlaars, A. Martínez-Finkelshtein, R. A. Orive, Orthogonality of Jacobi polynomials with general parameters. Electron. Trans. Numer. Anal. 19 (2005), 1–17.

[25] A. Martínez-Finkelshtein, P. Martínez-González and R. A. Orive, Zeros of Jacobi polynomials with varying non-classical parameters, in Special functions (Hong Kong, 1999), 98–113, World Sci. Publishing, River Edge, NJ, 2000.

[26] A. Martínez-Finkelshtein, P. Martínez-González and R. A. Orive, On asymptotic zero distribution of Laguerre and generalized Bessel polynomials with varying parameters, Journal of Computational and Applied Mathematics 133 (2001) 477–487.

[27] G. Pólya, *Sur un théorème de Stieltjes*, C. R. Acad. Sci Paris vol. 155, (1912), 767–769.

[28] W.-Y. Qiu, R. Wong, Global asymptotic expansions of the Laguerre polynomials, a Riemann-Hilbert approach. Numer. Algorithms 49 (2008), no. 1–4, 331–372.

[29] B. Shapiro, Algebro-geometric aspects of Heine–Stieltjes theory, J. London Math. Soc. 83(1) (2011) 36–56.

[30] B. Shapiro, K. Takemura, M. Tater, *On spectral polynomials of the Heun equation. II*, Comm. Math. Phys. 311(2) (2012), 277–300.

[31] A.Yu. Solynin, Modules and extremal metric problems. *Algebra i Analiz* **11** (1999), no. 1, 3–86; English transl. *St. Petersburg Math. J.* **11** (2000), 1–65.

[32] A.Yu. Solynin, Quadratic differentials and weighted graphs on compact surfaces. *Analysis and mathematical physics*, 473–505, Trends Math., Birkhäuser, Basel, 2009.

[33] H. Stahl, Orthogonal polynomials with complex-valued weight function I and II, Constr.Approx.2 (1986) 225–240, 241–251.

[34] K. Strebel, *Quadratic differentials*. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 5. Springer-Verlag, Berlin, (1984), xii+184 pp.

[35] G. Szegö, Orthogonal polynomials. Fourth edition. American Mathematical Society, Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, R.I., 1975. xiii+432 pp.

Boris Shapiro
Department of Mathematics
Stockholm University
SE-106 91 Stockholm, Sweden
e-mail: shapiro@math.su.se

Alexander Solynin
Department of Mathematics and Statistics
Texas Tech. University
Box 41042
Lubbock, TX 79409, USA
e-mail: alex.solynin@math.ttu.edu

Analysis Meets Geometry: A Tribute to Mikael Passare
Trends in Mathematics, 439–466

Interior Eigenvalue Density of Jordan Matrices with Random Perturbations

Johannes Sjöstrand and Martin Vogel

Dedicated to the memory of Mikael Passare

Abstract. We study the eigenvalue distribution of a large Jordan block subject to a small random Gaussian perturbation. A result by E. B. Davies and M. Hager shows that as the dimension of the matrix gets large, with probability close to 1, most of the eigenvalues are close to a circle.

We study the expected eigenvalue density of the perturbed Jordan block in the interior of that circle and give a precise asymptotic description.

Résumé. Nous étudions la distribution de valeurs propres d'un grand bloc de Jordan soumis à une petite perturbation gaussienne aléatoire. Un résultat de E. B. Davies et M. Hager montre que quand la dimension de la matrice devient grande, alors avec probabilité proche de 1, la plupart des valeurs propres sont proches d'un cercle.

Nous étudions la répartitions moyenne des valeurs propres à l'intérieur de ce cercle et nous en donnons une description asymptotique précise.

Mathematics Subject Classification (2010). 47A10, 47B80, 47H40, 47A55.

Keywords. Spectral theory; non-self-adjoint operators; random perturbations.

1. Introduction

In recent years there has been a renewed interest in the spectral theory of non-self-adjoint operators. We begin by recalling that for a closed linear operator $P : D(P) \to \mathcal{H}$ on a complex Hilbert space $\mathcal{H}$ with dense domain $D(P)$, we denote the resolvent set of P by

$$\rho(P) := \{z \in \mathbb{C};\ (P - z) : D(P) \to \mathcal{H} \text{ is a bijection with bounded inverse}\}.$$

For $z \in \rho(P)$ we call $(P - z)^{-1}$ the resolvent of P at z. The spectrum of P is defined as

$$\sigma(P) := \mathbb{C} \backslash \rho(P).$$

In case when P is self-adjoint (or more generally normal) we have a spectral theorem which yields a very good estimate on the norm of the resolvent, i.e.,

$$\|(P-z)^{-1}\| = \frac{1}{\operatorname{dist}(z,\sigma(P))}, \quad z \in \rho(P).$$

However, in the case of non-normal operators the norm of the resolvent can be very large even far away from the spectrum, since generically we only have the lower bound

$$\|(P-z)^{-1}\| \geq \frac{1}{\operatorname{dist}(z,\sigma(P))}, \quad z \in \rho(P).$$

Equivalently, the spectrum of such operators can be highly unstable even under very small perturbations of the operator. Originating from renewed interest in the phenomenon of spectral instability of non-self-adjoint operators in numerical analysis (cf [22, 21]) we possess an excellent tool to describe the region of spectral instability, i.e., the notion of ε-pseudospectrum. For $\varepsilon > 0$ it is defined by

$$\sigma_\varepsilon(P) := \left\{ z \in \rho(P) : \ \|(P-z)^{-1}\| > \frac{1}{\varepsilon} \right\} \cup \sigma(P).$$

The phenomenon of spectral instability of non-self-adjoint operators has become a popular and vital subject of study since it poses a serious difficulty, for example in numerical application when we are interested in determining the eigenvalues of a large non-normal matrix, but it can also be the source of many interesting effects, as emphasized by the works of L. N. Trefethen and M. Embree (eg [21]), E. B. Davies, M. Zworski and many others [3, 4, 6, 25, 5]. In view of this it is very natural to add small random perturbations.

One line of recent research concerns the case of elliptic (pseudo)differential operators subject to small random perturbations, cf. [1, 9, 8, 10, 16, 23], which show that for a very large class of semiclassical non-self-adjoint (pseudo-)differential operators we obtain a probabilistic Weyl law for the eigenvalues in the interior of the range of the principal symbol after adding a tiny random perturbation.

Perturbations of Jordan blocks

In this paper we shall study the spectrum of a random perturbation of the large Jordan block A_0:

$$A_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ . & . & . & . & \dots & . \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} : \mathbb{C}^N \to \mathbb{C}^N, \tag{1.1}$$

whose spectrum is $\sigma(A_0) = \{0\}$. M. Zworski [24] noticed that for every $z \in D(0,1)$, there are associated exponentially accurate quasi-modes when $N \to \infty$. Hence the open unit disc is a region of spectral instability. In $\mathbb{C} \setminus \overline{D(0,1)}$ we have spectral stability (a good resolvent estimate), since $\|A_0\| = 1$. Thus, if $A_\delta = A_0 + \delta Q$ is a small (random) perturbation of A_0 we expect the eigenvalues to move inside a small neighborhood of $\overline{D(0,1)}$.

In the special case when $Qu = (u|e_1)e_N$, where $(e_j)_1^N$ is the canonical basis in $\mathbb{C}^N$, the eigenvalues of A_δ are of the form

$$\delta^{1/N} e^{2\pi i k/N}, \; k \in \mathbb{Z}/N\mathbb{Z},$$

so if we fix $0 < \delta \ll 1$ and let $N \to \infty$, the spectrum "will converge to a uniform distribution on S^1".

E. B. Davies and M. Hager [5] studied random perturbations of A_0: Let $0 < \delta \ll 1$ and consider the following random perturbation of A_0 as in (1.1):

$$A_\delta = A_0 + \delta Q, \quad Q = (q_{j,k}(\omega))_{1\le j,k\le N}, \tag{1.2}$$

where $q_{j,k}(\omega)$ are independent and identically distributed complex random variables, following the complex Gaussian law $\mathcal{N}_{\mathbb{C}}(0,1)$. Recall that a random variable α has complex Gaussian distribution law $\mathcal{N}_{\mathbb{C}}(0,1)$ if

$$\alpha_*(\mathbb{P}(d\omega)) = \pi^{-1} \mathrm{e}^{-\alpha\overline{\alpha}} L(d\alpha)$$

where $L(d\alpha)$ denotes the Lebesgue measure on $\mathbb{C}$ and ω is the random parameter living in a probability space with probability measure $\mathbb{P}$. The above implies that

$$\mathbb{E}[\alpha] = 0, \quad \text{and} \quad \mathbb{E}\left[|\alpha|^2\right] = 1,$$

or in other words $\alpha \sim \mathcal{N}_{\mathbb{C}}(0,1)$ has expectation 0 and variance 1.

Davies and Hager showed that with probability close to 1, most of the eigenvalues are close to a circle:

Theorem 1.1 ([5]). *Let A_δ be as in (1.2). If $0 < \delta \le N^{-7}$, $R = \delta^{1/N}$, $\sigma > 0$, then with probability $\ge 1 - 2N^{-2}$, we have $\sigma(A_\delta) \subset D(0, RN^{3/N})$ and*

$$\#(\sigma(A_\delta) \cap D(0, Re^{-\sigma})) \le \frac{2}{\sigma} + \frac{4}{\sigma} \ln N.$$

The main purpose of this paper is to obtain, for a small coupling constant δ, more information about the distribution of eigenvalues of A_δ in the interior of a disc (cf Theorem 1.2), where the result of Davies and Hager only yields a logarithmic upper bound on the number of eigenvalues.

Theorem 1.2. *Let A_δ be the $N \times N$-matrix in (1.2) and restrict the attention to the parameter range $e^{-N/\mathcal{O}(1)} \le \delta \ll 1$, $N \gg 1$. Let r_0 belong to a parameter range,*

$$\frac{1}{\mathcal{O}(1)} \le r_0 \le 1 - \frac{1}{N},$$

$$\frac{r_0^{N-1} N}{\delta}(1 - r_0)^2 + \delta N^3 \ll 1, \tag{1.3}$$

so that $\delta \ll N^{-3}$. Then, for all $\varphi \in \mathcal{C}_0(D(0, r_0 - 1/N))$

$$\mathbb{E}\left[\mathbb{1}_{B_{\mathbb{C}^{N^2}}(0,C_1 N)}(Q) \sum_{\lambda\in\sigma(A_\delta)} \varphi(\lambda)\right] = \frac{1}{2\pi} \int \varphi(z) \Xi(z) L(dz),$$

where

$$\Xi(z) = \frac{4}{(1-|z|^2)^2}\left(1+\mathcal{O}\left(\frac{|z|^{N-1}N}{\delta}(1-|z|)^2+\delta N^3\right)\right).$$

is a continuous function independent of r_0. $C_1 > 0$ *is a large enough constant* (*satisfying* (4.19)).

A recent result by A. Guionnet, P. Matched Wood and O. Zeitouni [7] implies that when δ is bounded from above by $N^{-\kappa-1/2}$ for some $\kappa > 0$ and from below by some negative power of N, then

$$\frac{1}{N}\sum_{\mu\in\sigma(A_\delta)}\delta(z-\mu) \to \text{the uniform measure on } S^1,$$

weakly in probability.

We conclude the discussion of the results with some comments on Theorem 1.2:

Condition (1.3) is equivalent to $\delta N^3 \ll 1$ and

$$r_0^{N-1}(1-r_0)^2 \ll \frac{\delta}{N}.$$

For this inequality to be satisfied, it is necessary that

$$r_0 < 1 - 2(N+1)^{-1}.$$

For such r_0 the function $[0, r_0] \ni r \mapsto r^{N-1}(1-r)^2$ is increasing, and so inequality (1.3) is preserved if we replace r_0 by $|z| \le r_0$.

The leading contribution to the density $\Xi(z)$ is independent of N and is equal to the Lebesgue density of the volume form induced by the Poincaré metric on the disc $D(0,1)$. This yields a small density of eigenvalues close to the center of the disc $D(0,1)$ which is, however, growing towards the boundary of $D(0,1)$.

A similar result has been obtained by M. Sodin and B. Tsirelson in [20] for the distribution of zeros of a certain class of random analytic functions with domain $D(0,1)$ linking the fact that the density is given by the volume form induced by the Poincaré metric on $D(0,1)$ to its invariance under the action of $SL_2(\mathbb{R})$.

In order to obtain Theorem 1.2, we will study the expected eigenvalue density, adapting the approach of [23]. (For random polynomials and Gaussian analytic functions such results are more classical, see for example [12, 15, 11, 19, 14, 13].)

Organization. In Section 2 we will present some numerical simulations to illustrate the result of Theorem 1.2. In Section 3 we present a general formula for the average density of zeros of a holomorphic function g depending holomorphically on some parameters Q. In Section 4 we will set up an auxiliary Grushin problem yielding an effective function g, as above. Section 5 deals with the appropriate choice of coordinates Q and the calculation of the corresponding Jacobian $J(f)$. Finally, in Section 6 we complete the proof of Theorem 1.2.

2. Numerical simulations

To illustrate the result of Theorem 1.2, we present the following numerical calculations (Figure 1 and 2 which have been obtained using MATLAB) for the eigenvalues of the $N \times N$-matrix in (1.2), where $N = 500$ and the coupling constant δ varies from 10^{-5} to 10^{-2}.

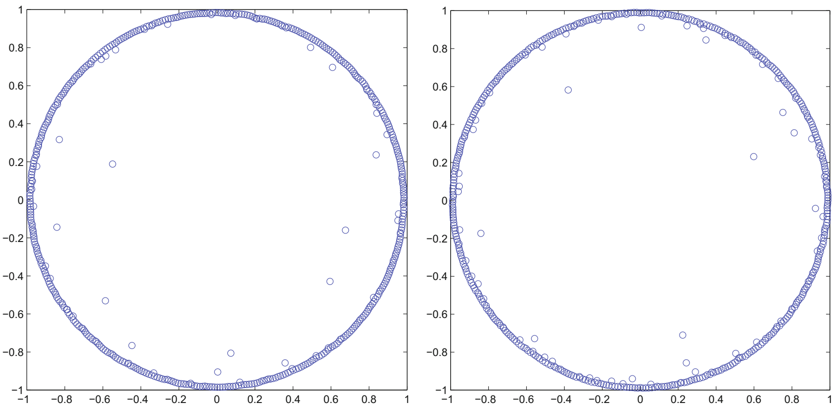


FIGURE 1. On the left-hand side $\delta = 10^{-5}$ and on the right-hand side $\delta = 10^{-4}$.

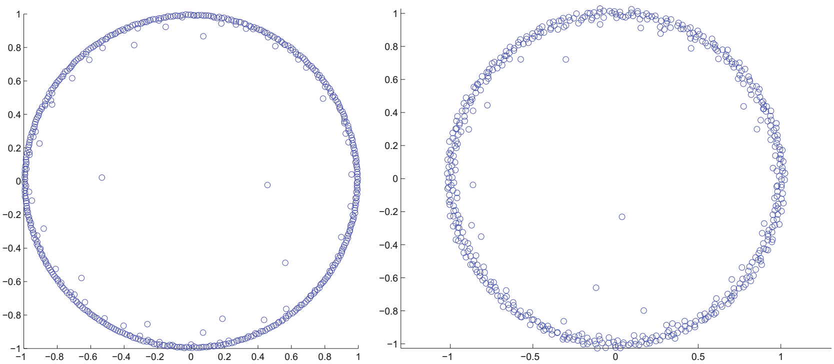


FIGURE 2. On the left-hand side $\delta = 10^{-3}$ and on the right-hand side $\delta = 10^{-2}$.

Note that on the right-hand side of Figure 2 we can see the onset of a different phenomenon: When the perturbation becomes too strong the spectral band will grow larger (for more details on this effect see [2] and [18, Chapter 13]).

3. A general formula

To start with, we shall obtain a general formula (due to [23] in a similar context). Our treatment is slightly different in that we avoid the use of approximations of the delta function and also that we have more holomorphy available.

Let $g(z,Q)$ be a holomorphic function on $\Omega\times W\subset\mathbb{C}\times\mathbb{C}^{N^2}$, where $\Omega\subset\mathbb{C}$, $W\subset\mathbb{C}^{N^2}$ are open bounded and connected. Assume that

$$\text{for every } Q\in W,\ g(\cdot,Q)\not\equiv 0. \tag{3.1}$$

To start with, we also assume that

$$\text{for almost all } Q\in W,\ g(\cdot,Q) \text{ has only simple zeros.} \tag{3.2}$$

Let $\phi\in C_0^\infty(\Omega)$ and let $m\in C_0(W)$. We are interested in

$$K_\phi=\int\Bigg(\sum_{z;\,g(z,Q)=0}\phi(z)\Bigg)m(Q)L(dQ), \tag{3.3}$$

where we frequently identify the Lebesgue measure with a differential form,

$$L(dQ)\simeq(2i)^{-N^2}d\overline{Q}_1\wedge dQ_1\wedge\cdots\wedge d\overline{Q}_{N^2}\wedge dQ_{N^2}=:(2i)^{-N^2}d\overline{Q}\wedge dQ.$$

In (3.3) we count the zeros of $g(\cdot,Q)$ with their multiplicity and notice that the integral is finite: For every compact set $K\subset W$ the number of zeros of $g(\cdot,Q)$ in $\operatorname{supp}\phi$, counted with their multiplicity, is uniformly bounded, for $Q\in K$. This follows from Jensen's formula.

Now assume,

$$g(z,Q)=0\Rightarrow d_Qg\neq 0. \tag{3.4}$$

Then

$$\Sigma:=\{(z,Q)\in\Omega\times W;\, g(z,Q)=0\}$$

is a smooth complex hypersurface in $\Omega\times W$ and from (3.2) we see that

$$K_\phi=\int_\Sigma\phi(z)m(Q)(2i)^{-N^2}d\overline{Q}\wedge dQ, \tag{3.5}$$

where we view $(2i)^{-N^2}d\overline{Q}\wedge dQ$ as a complex (N^2,N^2)-form on $\Omega\times W$, restricted to Σ, which yields a non-negative differential form of maximal degree on Σ.

Before continuing, let us eliminate the assumption (3.2). Without that assumption, the integral in (3.3) is still well defined. It suffices to show (3.5) for all $\phi\in\mathcal{C}_0^\infty(\Omega_0\times W_0)$ when $\Omega_0\times W_0$ is a sufficiently small open neighborhood of any given point $(z_0,Q_0)\in\Omega\times W$. When $g(z_0,Q_0)\neq 0$ or $\partial_z g(z_0,\Omega_0)\neq 0$ we already know that this holds, so we assume that for some $m\geq 2$, $\partial_z^k g(z_0,Q_0)=0$ for $0\leq k\leq m-1$, $\partial_z^m g(z_0,Q_0)\neq 0$.

Put $g_\varepsilon(z,Q)=g(z,Q)+\varepsilon$, $\varepsilon\in\operatorname{neigh}(0,\mathbb{C})$. By Weierstrass' preparation theorem, if Ω_0,W_0 and $r>0$ are small enough,

$$g_\varepsilon(z,Q)=k(z,Q,\varepsilon)p(z,Q,\varepsilon)\quad\text{in } \Omega_0\times W_0\times D(0,r),$$

where k is holomorphic and non-vanishing, and

$$p(z,Q,\varepsilon) = z^m + p_1(Q,\varepsilon)z^{m-1} + \cdots + p_m(Q,\varepsilon).$$

Here, $p_j(Q,\varepsilon)$ are holomorphic, and $p_j(0,0) = 0$.

The discriminant $D(Q,\varepsilon)$ of the polynomial $p(\cdot,Q,\varepsilon)$ is holomorphic on $W_0 \times D(0,r)$. It vanishes precisely when $p(\cdot,Q,\varepsilon)$ – or equivalently $g_\varepsilon(\cdot,Q)$ – has a multiple root in Ω_0.

Now for $0 < |\varepsilon| \ll 1$, the m roots of $g_\varepsilon(\cdot,Q_0)$ are simple, so $D(Q_0,\varepsilon) \neq 0$. Thus, $D(\cdot,\varepsilon)$ is not identically zero, so the zero set of $D(\cdot,\varepsilon)$ in W_0 is of measure 0 (assuming that we have chosen W_0 connected). This means that for $0 < |\varepsilon| \ll 1$, the function $g_\varepsilon(\cdot,Q)$ has only simple roots in Ω for almost all $Q \in W_0$.

Let Σ_ϵ be the zero set of g_ϵ, so that $\Sigma_\epsilon \to \Sigma$ in the natural sense. We have

$$\int \left(\sum_{z;\, g_\epsilon(z,Q)=0} \phi(z) \right) m(Q)(2i)^{-N^2} d\overline{Q} \wedge dQ = \int_{\Sigma_\epsilon} \phi(z) m(Q)(2i)^{-N^2} d\overline{Q} \wedge dQ$$

for $\phi \in \mathcal{C}_0^\infty(\Omega_0 \times W_0)$, when $\epsilon > 0$ is small enough, depending on ϕ, m. Passing to the limit $\epsilon = 0$ we get (3.5) under the assumptions (3.1), (3.4), first for $\phi \in \mathcal{C}_0^\infty(\Omega_0 \times W_0)$, and then by partition of unity for all $\phi \in \mathcal{C}_0^\infty(\Omega \times W)$. Notice that the result remains valid if we replace $m(Q)$ by $m(Q)1_B(Q)$ where B is a ball in W.

Now we strengthen the assumption (3.4) by assuming that we have a non-zero $Z(z) \in \mathbb{C}^{N^2}$ depending smoothly on $z \in \Omega$ (the dependence will actually be holomorphic in the application below) such that

$$g(z,Q) = 0 \Rightarrow \left(\overline{Z}(z) \cdot \partial_Q \right) g(z,Q) \neq 0. \tag{3.6}$$

We have the corresponding orthogonal decomposition

$$Q = Q(\alpha) = \alpha_1 \overline{Z}(z) + \alpha', \quad \alpha' \in \overline{Z}(z)^\perp,\ \alpha_1 \in \mathbb{C},$$

and if we identify unitarily $\overline{Z}(z)^\perp$ with $\mathbb{C}^{N^2-1}$ by means of an orthonormal basis $e_2(z), \ldots, e_{N^2}(z)$, so that $\alpha' = \sum_2^{N^2} \alpha_j e_j(z)$ we get global coordinates α_1, α_2, ..., α_{N^2} on Q-space.

By the implicit function theorem, at least locally near any given point in Σ, we can represent Σ by $\alpha_1 = f(z,\alpha')$, $\alpha' \in \overline{Z}(z)^\perp \simeq \mathbb{C}^{N^2-1}$, where f is smooth. (In the specific situation below, this will be valid globally.) Clearly, since $z, \alpha_2, \ldots, \alpha_{N^2}$ are complex coordinates on Σ, we have on Σ that

$$\frac{1}{(2i)^{N^2}} d\overline{Q} \wedge dQ = J(f) \frac{d\overline{z} \wedge dz}{2i} \wedge \frac{d\overline{\alpha}_2 \wedge d\alpha_2 \wedge ... \wedge d\overline{\alpha}_{N^2} \wedge d\alpha_{N^2}}{(2i)^{N^2-1}}$$

with the convention that

$$J(f) \frac{d\overline{z} \wedge dz}{2i} \geq 0, \quad (2i)^{1-N^2} d\overline{\alpha}_2 \wedge d\alpha_2 \wedge \cdots \wedge d\overline{\alpha}_{N^2} \wedge d\alpha_{N^2} > 0.$$

Thus

$$K_\phi = \int \phi(z) m\left(f(z,\alpha')\overline{Z}(z) + \alpha'\right) J(f)(z,\alpha_2,\dots,\alpha_{N^2}) \\ \times (2i)^{-N^2} d\overline{z}\wedge dz \wedge d\overline{\alpha}_2 \wedge d\alpha_2 \wedge \cdots \wedge d\overline{\alpha}_{N^2} \wedge d\alpha_{N^2}. \tag{3.7}$$

The Jacobian $J(f)$ is invariant under any z-dependent unitary change of variables, $\alpha_2,\dots,\alpha_{N^2} \mapsto \widetilde{\alpha}_2,\dots,\widetilde{\alpha}_{N^2}$, so for the calculation of $J(f)$ at a given point (z_0,α_0'), we are free to choose the most appropriate orthonormal basis $e_2(z),\dots,e_{N^2}(z)$ in $\overline{Z}(z)^\perp$ depending smoothly on z. We write (3.7) as

$$K_\phi = \int \phi(z)\widetilde{\Xi}(z)\frac{d\overline{z}\wedge dz}{2i}, \tag{3.8}$$

where the density $\widetilde{\Xi}(z)$ is given by

$$\widetilde{\Xi}(z) = \int_{\alpha'=\sum_2^{N^2}\alpha_j e_j(z)} m(f(z,\alpha')\overline{Z}(z)+\alpha')J(f)(z,\alpha_2,\dots,\alpha_{N^2}) \\ \times (2i)^{1-N^2} d\overline{\alpha}_2\wedge d\alpha_2\wedge\cdots\wedge d\overline{\alpha}_{N^2}\wedge d\alpha_{N^2}. \tag{3.9}$$

4. Grushin problem for the perturbed Jordan block

4.1. Setting up an auxiliary problem

Following [17], we introduce an auxiliary Grushin problem. Define $R_+ : \mathbb{C}^N \to \mathbb{C}$ by

$$R_+ u = u_1,\ u = (u_1 \ \dots \ u_N)^{\mathrm{t}} \in \mathbb{C}^N. \tag{4.1}$$

Let $R_- : \mathbb{C} \to \mathbb{C}^N$ be defined by

$$R_- u_- = (0\ 0\ \dots\ u_-)^{\mathrm{t}} \in \mathbb{C}^N. \tag{4.2}$$

Here, we identify vectors in $\mathbb{C}^N$ with column matrices. Then for $|z|<1$, the operator

$$\mathcal{A}_0 = \begin{pmatrix} A_0 - z & R_- \\ R_+ & 0 \end{pmatrix} : \mathbb{C}^{N+1} \to \mathbb{C}^{N+1} \tag{4.3}$$

is bijective. In fact, identifying

$$\mathbb{C}^{N+1} \simeq \ell^2([1,2,\dots,N+1]) \simeq \ell^2(\mathbb{Z}/(N+1)\mathbb{Z}),$$

we have $\mathcal{A}_0 = \tau^{-1} - z\Pi_N$, where $\tau u(j) = u(j-1)$ (translation by 1 step to the right) and $\Pi_N u = 1_{[1,N]}u$. Then $\mathcal{A}_0 = \tau^{-1}(1 - z\tau\Pi_N)$, $(\tau\Pi_N)^{N+1} = 0$,

$$\mathcal{A}_0^{-1} = (1 + z\tau\Pi_N + (z\tau\Pi_N)^2 + \cdots + (z\tau\Pi_N)^N)\circ\tau.$$

Write

$$\mathcal{E}_0 := \mathcal{A}_0^{-1} =: \begin{pmatrix} E^0 & E_+^0 \\ E_-^0 & E_{-+}^0 \end{pmatrix}.$$

Then

$$E^0 \simeq \Pi_N(1 + z\tau\Pi_N + \cdots (z\tau\Pi_N)^{N-1})\tau\Pi_N, \tag{4.4}$$

$$E_+^0 = \begin{pmatrix} 1 \\ z \\ \dots \\ z^{N-1} \end{pmatrix}, \ E_-^0 = \begin{pmatrix} z^{N-1} & z^{N-2} & \dots & 1 \end{pmatrix}, \tag{4.5}$$

$$E_{-+}^0 = z^N. \tag{4.6}$$

A quick way to check (4.5), (4.6) is to write $\mathcal{A}_0$ as an $(N+1)\times(N+1)$-matrix where we moved the last line to the top, with the lines labeled from 0 ($\equiv N+1 \mod (N+1)\mathbb{Z}$) to N and the columns from 1 to $N+1$.

Continuing, we see that

$$\|E^0\| \le G(|z|), \ \|E_\pm^0\| \le G(|z|)^{\frac{1}{2}}, \ \|E_{-+}^0\| \le 1, \tag{4.7}$$

where $\|\cdot\|$ denote the natural operator norms and

$$G(|z|) := \min\left(N, \frac{1}{1-|z|}\right) \asymp 1 + |z| + |z|^2 + \ldots + |z|^{N-1}. \tag{4.8}$$

Next, consider the natural Grushin problem for A_δ. If $\delta\|Q\|G(|z|) < 1$, we see that

$$\mathcal{A}_\delta = \begin{pmatrix} A_\delta - z & R_- \\ R_+ & 0 \end{pmatrix} \tag{4.9}$$

is bijective with inverse

$$\mathcal{E}_\delta = \begin{pmatrix} E^\delta & E_+^\delta \\ E_+^\delta & E_{-+}^\delta \end{pmatrix},$$

where

$$\begin{aligned} E^\delta &= E^0 - E^0\delta Q E^0 + E^0(\delta Q E^0)^2 - \ldots = E^0(1+\delta Q E^0)^{-1}, \\ E_+^\delta &= E_+^0 - E^0\delta Q E_+^0 + (E^0\delta Q)^2 E_+^0 - \ldots = (1+E^0\delta Q)^{-1}E_+^0, \\ E_-^\delta &= E_-^0 - E_-^0\delta Q E^0 + E_-^0(\delta Q E^0)^2 - \ldots = E_-^0(1+\delta Q E^0)^{-1}, \\ E_{-+}^\delta &= E_{-+}^0 - E_-^0\delta Q E_+^0 + E_-^0\delta Q E^0 \delta Q E_+^0 - \ldots \\ &= E_{-+}^0 - E_-^0 \delta Q(1+E^0\delta Q)^{-1}E_+^0. \end{aligned} \tag{4.10}$$

We get

$$\begin{aligned} &\|E^\delta\| \le \frac{G(|z|)}{1-\delta\|Q\|G(|z|)}, \ \|E_\pm^\delta\| \le \frac{G(|z|)^{\frac{1}{2}}}{1-\delta\|Q\|G(|z|)}, \\ &|E_{-+}^\delta - E_{-+}^0| \le \frac{\delta\|Q\|G(|z|)}{1-\delta\|Q\|G(|z|)}. \end{aligned} \tag{4.11}$$

Indicating derivatives with respect to δ with dots and omitting sometimes the super/sub-script δ, we have

$$\dot{\mathcal{E}} = -\mathcal{E}\dot{\mathcal{A}}\mathcal{E} = -\begin{pmatrix} EQE & EQE_+ \\ E_-QE & E_-QE_+. \end{pmatrix} \tag{4.12}$$

Integrating this from 0 to δ yields

$$\|E^\delta - E^0\| \le \frac{G(|z|)^2\delta\|Q\|}{(1-\delta\|Q\|G(|z|))^2}, \quad \|E_\pm^\delta - E_\pm^0\| \le \frac{G(|z|)^{\frac{3}{2}}\delta\|Q\|}{(1-\delta\|Q\|G(|z|))^2}. \tag{4.13}$$

We now sharpen the assumption that $\delta\|Q\|G(|z|) < 1$ to

$$\delta\|Q\|G(|z|) < 1/2. \tag{4.14}$$

Then

$$\begin{aligned} &\|E^\delta\| \le 2G(|z|),\ \|E_\pm^\delta\| \le 2G(|z|)^{\frac{1}{2}}, \\ &|E_{-+}^\delta - E_{-+}^0| \le 2\delta\|Q\|G(|z|). \end{aligned} \tag{4.15}$$

Combining this with the identity $\dot{E}_{-+} = -E_-QE_+$ that follows from (4.12), we get

$$\|\dot{E}_{-+} + E_-^0 Q E_+^0\| \le 16G(|z|)^2\delta\|Q\|^2, \tag{4.16}$$

and after integration from 0 to δ,

$$E_{-+}^\delta = E_{-+}^0 - \delta E_-^0 Q E_+^0 + \mathcal{O}(1)G(|z|)^2(\delta\|Q\|)^2. \tag{4.17}$$

Using (4.5), (4.6) we get with $Q = (q_{j,k})$,

$$E_{-+}^\delta = z^N - \delta \sum_{j,k=1}^N q_{j,k} z^{N-j+k-1} + \mathcal{O}(1)G(|z|)^2(\delta\|Q\|)^2, \tag{4.18}$$

still under the assumption (4.14).

4.2. Estimates for the effective Hamiltonian

We now consider the situation of (1.2):

$$A_\delta = A_0 + \delta Q,\ Q = (q_{j,k}(\omega))_{j,k=1}^N,\ q_{j,k}(\omega) \sim \mathcal{N}_{\mathbb{C}}(0,1) \text{ independent.}$$

W. Bordeaux-Montrieux [1] observed the following result.

Proposition 4.1. *There exists a $C_0 > 0$ such that the following holds: Let $X_j \sim \mathcal{N}_{\mathbb{C}}(0,\sigma_j^2)$, $1 \le j \le N < \infty$ be independent complex Gaussian random variables. Put $s_1 = \max \sigma_j^2$. Then, for every $x > 0$, we have*

$$\mathbb{P}\left[\sum_{j=1}^N |X_j|^2 \ge x\right] \le \exp\left(\frac{C_0}{2s_1}\sum_{j=1}^N \sigma_j^2 - \frac{x}{2s_1}\right).$$

According to this result we have

$$P(\|Q\|_{\mathrm{HS}}^2 \ge x) \le \exp\left(\frac{C_0}{2}N^2 - \frac{x}{2}\right)$$

and hence if $C_1 > 0$ is large enough,

$$\|Q\|_{\mathrm{HS}}^2 \le C_1^2 N^2, \text{ with probability } \ge 1 - e^{-N^2}. \tag{4.19}$$

In particular (4.19) holds for the ordinary operator norm of Q. In the following, we often write $|\cdot|$ for the Hilbert–Schmidt norm $\|\cdot\|_{\mathrm{HS}}$ and we shall work under the assumption that $|Q| \le C_1 N$. We let $|z| < 1$ and assume:

$$\delta N G(|z|) \ll 1. \tag{4.20}$$

Then with probability $\ge 1 - e^{-N^2}$, we have (4.14), (4.18) which give for $g(z,Q) := E^{\delta}_{-+}$,

$$g(z,Q) = z^N - \delta(Q|\overline{Z}(z)) + \mathcal{O}(1)(G(|z|)\delta N)^2. \tag{4.21}$$

Here, Z is given by

$$Z = \left(z^{N-j+k-1}\right)_{j,k=1}^{N}. \tag{4.22}$$

A straightforward calculation shows that

$$|Z| = \sum_{0}^{N-1} |z|^{2\nu} = \frac{1-|z|^{2N}}{1-|z|^2} = \frac{1-|z|^N}{1-|z|}\,\frac{1+|z|^N}{1+|z|}, \tag{4.23}$$

and in particular,

$$G(|z|)/\,2 \le |Z| \le G(|z|). \tag{4.24}$$

The middle term in (4.21) is bounded in modulus by $\delta|Q||Z| \le \delta C_1 N G(|z|)$ and we assume that $|z|^N$ is much smaller than this bound:

$$|z|^N \ll \delta C_1 N G(|z|). \tag{4.25}$$

More precisely, we work in a disc $D(0, r_0)$, where

$$r_0^N \le C^{-1}\delta C_1 N G(r_0) \le C^{-2}, \quad r_0 \le 1 - N^{-1} \tag{4.26}$$

and $C \gg 1$. In fact, the first inequality in (4.26) can be written $m(r_0) \le C^{-1}\delta C_1 N$ and $m(r) = r^N(1-r)$ is increasing on $[0, 1-N^{-1}]$ so the inequality is preserved if we replace r_0 by $|z| \le r_0$. Similarly, the second inequality holds after the same replacement since G is increasing.

In view of (4.20), we see that

$$\left(G(|z|)\delta N\right)^2 \ll \delta G(|z|) N$$

is also much smaller than the upper bound on the middle term.

By the Cauchy inequalities,

$$d_Q g = -\delta Z \cdot dQ + \mathcal{O}(1) G(|z|)^2 \delta^2 N. \tag{4.27}$$

The norm of the first term is $\asymp \delta G \gg G^2\delta^2 N$, since $G\delta N \ll 1$. (When applying the Cauchy inequalities, we should shrink the radius $R = C_1 N$ by a factor $\theta < 1$, but we have room for that, if we let C_1 be a little larger than necessary to start with.)

Writing

$$Q = \alpha_1 \overline{Z}(z) + \alpha',\ \ \alpha' \in \overline{Z}(z)^{\perp} \simeq \mathbb{C}^{N^2-1},$$

we identify $g(z,Q)$ with a function $\widetilde{g}(z,\alpha)$ which is holomorphic in α for every fixed z and satisfies

$$\widetilde{g}(z,\alpha) = z^N - \delta|Z(z)|^2\alpha_1 + \mathcal{O}(1)G(|z|)^2\delta^2 N^2, \tag{4.28}$$

while (4.27) gives

$$\partial_{\alpha_1}\widetilde{g}(z,\alpha) = -\delta|Z(z)|^2 + \mathcal{O}(1)G(|z|)^3\delta^2 N, \tag{4.29}$$

and in particular,

$$|\partial_{\alpha_1}\widetilde{g}| \asymp \delta G(|z|)^2.$$

This derivative does not depend on the choice of unitary identification $\overline{Z}^{\perp} \simeq \mathbb{C}^{N^2-1}$. Notice that the remainder in (4.28) is the same as in (4.21) and hence a holomorphic function of (z,Q). In particular it is a holomorphic function of $\alpha_1,\dots,\alpha_{N^2}$ for every fixed z and we can also get (4.29) from this and the Cauchy inequalities. In the same way, we get from (4.28) that

$$\partial_{\alpha_j}\widetilde{g}(z,\alpha) = \mathcal{O}(1)G(|z|)^2\delta^2 N,\ j=2,\dots,N^2. \tag{4.30}$$

The Cauchy inequalities applied to (4.21) give,

$$\partial_z g(z,Q) = Nz^{N-1} - \delta Q\cdot\partial_z Z(z) + \mathcal{O}(1)\frac{(G(|z|)\delta N)^2}{r_0-|z|}. \tag{4.31}$$

Then, for $\widetilde{g}(z,\alpha_1,\alpha') = g(z,\alpha_1\overline{Z}(z)+\alpha')$, $\alpha' = \sum_2^{N^2}\alpha_j e_j$ we shall see that

$$\partial_z\widetilde{g} = Nz^{N-1} - \delta\alpha_1\partial_z\left(|Z|^2\right) + \mathcal{O}(1)\frac{(G\delta N)^2}{r_0-|z|} + \mathcal{O}(1)G^2\delta^2 N\left|\sum_2^{N^2}\alpha_j\partial_z e_j\right|, \tag{4.32}$$

$$\partial_{\overline{z}}\widetilde{g} = -\delta\alpha_1\partial_{\overline{z}}\left(|Z|^2\right) + \mathcal{O}(1)G^2\delta^2 N\left|\alpha_1\overline{\partial_z Z} + \sum_2^{N^2}\alpha_j\partial_{\overline{z}}e_j\right|. \tag{4.33}$$

The leading terms in (4.32), (4.33) can be obtained formally from (4.28) by applying ∂_z, $\partial_{\overline{z}}$ and we also notice that

$$\partial_z|Z|^2 = \overline{Z}\cdot\partial_z Z,\quad \partial_{\overline{z}}|Z|^2 = Z\cdot\overline{\partial_z Z}.$$

However it is not clear how to handle the remainder in (4.28), so we verify (4.32), (4.33), using (4.27), (4.31):

$$\begin{aligned}
\partial_z\widetilde{g} &= \partial_z g + d_Q g\cdot\sum_2^{N^2}\alpha_j\partial_z e_j \\
&= Nz^{N-1} - \delta Q\cdot\partial_z Z + \mathcal{O}(1)\frac{(G\delta N)^2}{r_0-|z|} + (-\delta Z\cdot dQ + \mathcal{O}(1)G^2\delta^2 N)\cdot\sum_2^{N^2}\alpha_j\partial_z e_j \\
&= Nz^{N-1} - \delta\alpha_1\partial_z\left(|Z|^2\right) - \delta\sum_2^{N^2}\alpha_j e_j\cdot\partial_z Z - \delta Z\cdot\sum_2^{N^2}\alpha_j\partial_z e_j \\
&\qquad\qquad + \text{ the remainders in (4.32).}
\end{aligned}$$

The 3d and the 4th terms in the last expression add up to

$$\delta\partial_z\left(\sum\nolimits_2^{N^2}\alpha_j e_j\cdot Z\right) = \delta\partial_z(0) = 0,$$

and we get (4.32).

Similarly,

$$\partial_{\overline{z}}\widetilde{g} = d_Q g \cdot \left(\alpha_1 \overline{\partial_z Z} + \sum_2^{N^2} \alpha_j \partial_{\overline{z}} e_j\right)$$
$$= \left(-\delta Z \cdot dQ + \mathcal{O}(1) G^2 \delta^2 N\right) \cdot \left(\alpha_1 \overline{\partial_z Z} + \sum_2^{N^2} \alpha_j \partial_{\overline{z}} e_j\right).$$

Up to remainders as in (4.33), this is equal to

$$-\delta\alpha_1 Z \cdot \overline{\partial_z Z} - \delta \sum_2^{N^2} \alpha_j Z \cdot \partial_{\overline{z}} e_j = -\delta\alpha_1 \partial_{\overline{z}} \left(|Z|^2\right) - \delta \sum_2^{N^2} \alpha_j \partial_{\overline{z}} \left(Z \cdot e_j\right)$$
$$= -\delta\alpha_1 \partial_{\overline{z}} \left(|Z|^2\right).$$

Here, we know that

$$\begin{aligned} |Z(z)| &= \sum_0^{N-1} (z\overline{z})^\nu =: K(z\overline{z}), \\ \partial_z \left(|Z(z)|^2\right) &= 2KK'\overline{z}, \\ \partial_{\overline{z}} \left(|Z(z)|^2\right) &= 2KK'z. \end{aligned} \tag{4.34}$$

Observe also that $K(t) \asymp G(t)$ and that $G(|z|) \asymp G(|z|^2)$.

The following result implies that $K'(t)$ and $K(t)^2$ are of the same order of magnitude.

Proposition 4.2. *For $k \in \mathbb{N}$, $2 \leq N \in \mathbb{N} \cup \{+\infty\}$, $0 \leq t < 1$, we put*

$$M_{N,k}(t) = \sum_{\nu=1}^{N-1} \nu^k t^\nu, \tag{4.35}$$

so that $K(t) = K_N(t) = M_{N,0}(t) + 1$, $K'(t) \asymp M_{N-1,1}(t) + 1$. For each fixed $k \in \mathbb{N}$, we have uniformly with respect to N, t:

$$M_{\infty,k}(t) \asymp \frac{t}{(1-t)^{k+1}}, \tag{4.36}$$

$$M_{\infty,k}(t) - M_{N,k}(t) \asymp \frac{t^N}{1-t}\left(N + \frac{1}{1-t}\right)^k. \tag{4.37}$$

For all fixed $C > 0$ and $k \in \mathbb{N}$, we have uniformly,

$$M_{N,k}(t) \asymp M_{\infty,k}(t), \text{ for } 0 \leq t \leq 1 - \frac{1}{CN}, \ N \geq 2. \tag{4.38}$$

Notice that under the assumption in (4.38), the estimate (4.37) becomes

$$M_{\infty,k}(t) - M_{N,k}(t) \asymp \frac{t^N N^k}{1-t}.$$

We also see that in any region $1-\mathcal{O}(1)/N \le t < 1$, we have

$$M_{N,k}(t) \asymp N^{k+1},$$

so together with (4.38), (4.36), this shows that

$$M_{N,k}(t) \asymp t \min\left(\frac{1}{1-t}, N\right)^{k+1}. \tag{4.39}$$

Proof. The statements are easy to verify when $0 \le t \le 1-1/\mathcal{O}(1)$ and the N-dependent statements (4.37), (4.38) are clearly true when $N \le \mathcal{O}(1)$. Thus we can assume that $1/2 \le t < 1$ and $N \gg 1$.

Write $t = e^{-s}$ so that $0 < s \le 1/\mathcal{O}(1)$ and notice that $s \asymp 1-t$. For $N \in \mathbb{N}$, we put

$$P_{N,k}(s) = \sum_{\nu=N}^{\infty} \nu^k e^{-\nu s}, \tag{4.40}$$

so that

$$P_{N,k}(s) = \begin{cases} M_{\infty,k}(t) \text{ when } N = 1, \\ M_{\infty,k}(t) - M_{N,k}(t) \text{ when } N \ge 2. \end{cases} \tag{4.41}$$

We regroup the terms in (4.40) into sums with $\asymp 1/s$ terms where $e^{-\nu s}$ has constant order of magnitude:

$$P_{N,k}(s) = \sum_{\mu=1}^{\infty} \Sigma(\mu), \quad \Sigma(\mu) = \sum_{N+\frac{\mu-1}{s} \le \nu < N+\frac{\mu}{s}} \nu^k e^{-\nu s}.$$

Here, since the sum $\Sigma(\mu)$ consists of $\asymp 1/s$ terms of the order $\nu^k e^{-(Ns+\mu)}$,

$$\Sigma(\mu) \asymp e^{-(Ns+\mu)} \sum_{N+\frac{\mu-1}{s} \le \nu < N+\frac{\mu}{s}} \nu^k \asymp e^{-(Ns+\mu)} \frac{(Ns+\mu)^k}{s^{k+1}}.$$

Hence,

$$\begin{aligned} P_{N,k}(s) &\asymp \frac{e^{-Ns}}{s^{k+1}} \sum_{\mu=1}^{\infty} e^{-\mu}(Ns+\mu)^k \\ &\asymp \frac{e^{-Ns}}{s^{k+1}}(Ns+1)^k = \frac{e^{-Ns}}{s}\left(N+\frac{1}{s}\right)^k. \end{aligned}$$

Recalling (4.41) and the fact that $s \asymp 1-t$, $1/2 \le t < 1$, we get (4.36) when $N = 1$ and (4.37) when $N \ge 2$.

It remains to show (4.38) and it suffices to do so for $1/2 \le t \le 1-C/N$, $N \gg 1$ and for $C \ge 1$ sufficiently large but independent of N. Indeed, for $1-C/N \le t \le 1-1/\mathcal{O}(N)$, both $M_{N,k}(t)$ and $M_{\infty,k}(t)$ are $\asymp N^{1+k}$. We can also exclude the case $k = 0$ where we have explicit formulae.

To get the equivalence (4.38) for $1/2 \le t \le 1 - C/N$, $k \ge 1$, it suffices, in view of (4.36), (4.37), to show that for such t and for $N \gg 1$, we have

$$\frac{N^k t^N}{1-t} \le \frac{1}{D} \frac{1}{(1-t)^{k+1}},$$

for any given $D \ge 1$, provided that C is large enough. In other terms, we need

$$t^N (1-t)^k \le \frac{1}{D} N^{-k}, \text{ for } \frac{1}{2} \le t \le 1 - \frac{C}{N},$$

when $C = C(D)$ is large enough and $N \ge N(C) \gg 1$. The left-hand side in this inequality is an increasing function of t on the interval $[0, 1/(1 + k/N)]$. If $t \le 1 - C/N \le 1/(1 + k/N)$ (which is fulfilled when $C \ge 2k$ and $N \gg N(C)$) it is

$$\le \left(1 - \frac{C}{N}\right)^N \left(\frac{C}{N}\right)^k = \left(1 + \mathcal{O}_C\left(\frac{1}{N}\right)\right) e^{-C} C^k N^{-k}.$$

This is $\le N^{-k}/D$ if $C \ge C(D)$, $N \ge N(C)$. □

For simplicity we will restrict the attention to the region

$$|z| \le r_0 - 1/N, \tag{4.42}$$

where $G \asymp (1 - |z|)^{-1}$, $G' \asymp (1 - |z|)^{-2}$.

It follows from the calculation (5.6) below, that

$$|\partial_z Z|^2 = \left(\frac{2}{t}\left(K(t\partial_t)^2 K + (t\partial_t K)^2\right)\right)_{t=|z|^2}.$$

This is $\asymp 1$ for $|z| \le 1/2$ and for $1/2 \le |z| < 1 - 1/N$ it is in view of Proposition 4.2 and the subsequent observation

$$\asymp M_{N,0} M_{N,2} + M_{N,1}^2 \asymp \frac{1}{(1-t)^4}, \ t = |z|^2.$$

In the region (4.42) we get:

$$|Z'(z)| \asymp G(|z|)^2. \tag{4.43}$$

(4.34), (4.42), (4.43) will be used in (4.32), (4.33).

Combining the implicit function theorem and Rouché's theorem with (4.28), we see that for $|\alpha'| < C_1 N$, $\alpha' = \sum_2^N \alpha_j e_j \in \overline{Z}(z)^\perp$, the equation

$$\widetilde{g}(z, \alpha_1, \alpha') = 0 \tag{4.44}$$

has a unique solution

$$\alpha_1 = f(z, \alpha') \in D(0, C_1 N / G(|z|)). \tag{4.45}$$

Here, we also use (4.20), (4.25). Moreover, f satisfies

$$f(z, \alpha') = \frac{z^N}{\delta |Z|^2} + \mathcal{O}(1)\delta N^2 = \mathcal{O}(1)\left(\frac{|z|^N}{\delta G^2} + \delta N^2\right). \tag{4.46}$$

Differentiating the equation (4.44) (where $\alpha_1 = f$) we get

$$\partial_z \widetilde{g} + \partial_{\alpha_1} \widetilde{g} \partial_z f = 0, \ \partial_{\overline{z}} \widetilde{g} + \partial_{\alpha_1} \widetilde{g} \partial_{\overline{z}} f = 0.$$

Hence,

$$\begin{cases} \partial_z f = -\left(\partial_{\alpha_1}\widetilde{g}\right)^{-1} \partial_z \widetilde{g}, \\ \partial_{\overline{z}} f = -\left(\partial_{\alpha_1}\widetilde{g}\right)^{-1} \partial_{\overline{z}} \widetilde{g}. \end{cases} \tag{4.47}$$

Since $\widetilde{g}$ is holomorphic in α_1, α' and in $\alpha_1, \alpha_2, \dots, \alpha_{N^2}$, we see that f is holomorphic in α' and in $\alpha_2, \dots, \alpha_{N^2}$ Applying $\partial_{\alpha_2}, \dots, \partial_{\alpha_{N^2}}$ to (4.44), we get

$$\partial_{\alpha_j} f = -\left(\partial_{\alpha_1}\widetilde{g}\right)^{-1} \partial_{\alpha_j} \widetilde{g},\ 2 \le j \le N^2. \tag{4.48}$$

Combining (4.29) in the form,

$$\partial_{\alpha_1}\widetilde{g}(z,\alpha) = -(1 + \mathcal{O}(G(|z|)\delta N))\delta |Z(z)|^2,$$

(4.30), (4.32), (4.33) with (4.47) and (4.48), we get

$$\begin{aligned} \partial_z f &= \frac{(1+\mathcal{O}(G\delta N))}{\delta |Z(z)|^2} \\ &\quad \times \left(N z^{N-1} - \delta f \partial_z\left(|Z|^2\right) + \mathcal{O}\left(G^2\delta^2 N\right) \left| \sum_2^{N^2} \alpha_j \partial_z e_j \right| + \mathcal{O}(1) \frac{(G\delta N)^2}{r_0 - |z|} \right). \end{aligned} \tag{4.49}$$

$$\begin{aligned} \partial_{\overline{z}} f &= \frac{(1+\mathcal{O}(G\delta N))}{\delta |Z(z)|^2} \\ &\quad \times \left(-\delta f \partial_{\overline{z}}\left(|Z|^2\right) + \mathcal{O}\left(G^2\delta^2 N\right) \left| f\overline{\partial_z Z} + \sum_2^{N^2} \alpha_j \partial_{\overline{z}} e_j \right| \right), \end{aligned} \tag{4.50}$$

$$\partial_{\alpha_j} f = \mathcal{O}(1)\frac{G^2\delta^2 N}{\delta G^2} = \mathcal{O}(\delta N),\ 2 \le j \le N^2. \tag{4.51}$$

From (4.34) and the observation prior to Proposition 4.2 we know that

$$\partial_z\left(|Z|^2\right),\ \partial_{\overline{z}}\left(|Z|^2\right) \asymp G(|z|)^3 |z|.$$

Recall also that $|Z| \asymp G(|z|)$. Using this in (4.49), (4.50), we get

$$\begin{aligned} \partial_z f &= \frac{\mathcal{O}(1)}{\delta G^2} \\ &\quad \times \left(N|z|^{N-1} + \delta |f| G^3 |z| + \mathcal{O}\left(G^2\delta^2 N\right) \left| \sum_2^{N^2} \alpha_j \partial_z e_j \right| + \mathcal{O}(1) \frac{G^2\delta^2 N^2}{r_0 - |z|} \right). \end{aligned} \tag{4.52}$$

5. Choosing appropriate coordinates

The next task will be to choose an orthonormal basis $e_1(z), e_2(z), \dots, e_{N^2}(z)$ in $\mathbb{C}^{N^2}$ with $e_1(z) = |Z(z)|^{-1}\overline{Z}(z)$ such that we get a nice control over $\sum_2^{N^2} \alpha_j \partial_z e_j$, $\sum_2^{N^2} \alpha_j \partial_{\overline{z}} e_j$ and such that

$$dQ_1 \wedge \dots \wedge dQ_{N^2}|_{\alpha_1 = f(z,\alpha')}$$

can be expressed easily up to small errors. Consider a point $z_0 \in D(0, r_0 - N^{-1})$. We shall see below that the vectors $\overline{Z}(z)$, $\overline{\partial_z Z}(z)$ are linearly independent for every $z \in D(0,1)$

Proposition 5.1. *There exists an orthonormal basis* $e_1(z), e_2(z), \dots, e_{N^2}(z)$ *in* $\mathbb{C}^{N^2}$, *depending smoothly on* $z \in \operatorname{neigh}(z_0)$ *such that*

$$e_1(z) = |Z(z)|^{-1}\overline{Z}(z), \tag{5.1}$$

$$\mathbb{C}e_1(z_0) \oplus \mathbb{C}e_2(z_0) = \mathbb{C}\overline{Z}(z_0) \oplus \overline{\partial_z Z}(z_0), \tag{5.2}$$

$$e_j(z) - e_j(z_0) = \mathcal{O}((z-z_0)^2),\ j \geq 3. \tag{5.3}$$

Proof. We choose $e_1(z)$ as in (5.1). Let $e_3(z_0), \dots, e_{N^2}(z_0)$ be an orthonormal basis in $\left(\mathbb{C}\overline{Z}(z_0) \oplus \mathbb{C}\overline{\partial_z Z}(z_0)\right)^{\perp}$. Then we get an orthonormal family $e_3(z), \dots, e_{N^2}(z)$ in $e_1(z)^{\perp}$ in the following way:

Let V_0 be the isometry $\mathbb{C}^{N^2-2} \to \mathbb{C}^{N^2}$, defined by $V_0\nu_j^0 = e_j(z_0)$, $j = 3, \dots, N^2$, where $\nu_3^0, \dots, \nu_{N^2}^0$ is the canonical basis in $\mathbb{C}^{N^2-2}$ with a non-canonical labeling. Let $\pi(z)u = (u|e_1(z))e_1(z)$ be the orthogonal projection onto $\mathbb{C}e_1(z)$. For $z \in \operatorname{neigh}(z_0, \mathbb{C})$, let $V(z) = (1-\pi(z))V_0$. Then $f_j(z) = V(z)\nu_j^0$, $j = 3, \dots, N^2$ form a linearly independent system in $e_1(z)^{\perp}$ and we get an orthonormal system of vectors that span the same hyperplane in $e_1(z)^{\perp}$ by Gram orthonormalization,

$$e_j(z) = V(z)(V^*(z)V(z))^{-\frac{1}{2}}\nu_j^0,\ 3 \leq j \leq N^2.$$

We have

$$V(z)\nu_j^0 = (1-\pi(z))e_j(z_0) = e_j(z_0) - (e_j(z_0)|e_1(z))e_1(z),$$

$$(e_j(z_0)|e_1(z)) = \frac{(e_j(z_0)|\overline{Z}(z))}{|Z(z)|} = \mathcal{O}((z-z_0)^2),$$

since $(e_j(z_0)|\overline{Z}(z)) = e_j(z_0) \cdot Z(z) =: k(z)$ is a holomorphic function of z with $k(z_0) = (e_j(z_0)|\overline{Z}(z_0)) = 0$, $k'(z_0) = (e_j(z_0)|\overline{\partial_z Z}(z_0)) = 0$. Thus, $V(z) = V(z_0) + \mathcal{O}((z-z_0)^2)$ and we conclude that (5.3) holds. Let $e_2(z)$ be a normalized vector in $(e_1(z), e_3(z), e_4(z), \dots, e_{N^2}(z))^{\perp}$ depending smoothly on z. Then $e_1(z)$, $e_2(z)$, $\dots$, $e_{N^2}(z)$ is an orthonormal basis and since $e_3(z_0), \dots, e_{N^2}(z_0)$ are orthogonal to $\overline{Z}(z_0), \overline{\partial Z}(z_0)$ by construction, we get (5.2). □

We can make the following explicit choice:

$$e_2(z) = |f_2(z)|^{-1}f_2(z),\ f_2(z) = \overline{\partial_z Z}(z) - \sum_{j\neq 2}(\overline{\partial_z Z}(z)|e_j(z))e_j(z), \tag{5.4}$$

so that for $z = z_0$,

$$e_2(z_0) = |f_2(z_0)|^{-1}f_2(z_0),\ f_2(z_0) = \overline{\partial_z Z}(z_0) - (\overline{\partial_z Z}(z_0)|e_1(z_0))e_1(z_0). \tag{5.5}$$

We next compute some scalar products and norms with Z and $\partial_z Z$. Recall that $Z(z) = \left(z^{N-j+k-1}\right)_{j,k=1}^{N}$ and that $|Z(z)| = K(|z|^2)$, $K(t) = \sum_0^{N-1} t^{\nu}$.

Repeating basically the same computation, we get

$$z\partial_z Z = \big((N-j+k-1)z^{N-j+k-1}\big)_{j,k=1}^{N},$$

and

$$\begin{aligned}
|z\partial_z Z|^2 &= \sum_{j,k=1}^{N}(N-j+k-1)^2|z|^{2(N-j+k-1)} = \sum_{\nu,\mu=0}^{N-1}(\nu+\mu)^2|z|^{2(\nu+\mu)}\\
&= \sum_0^{N-1}\nu^2|z|^{2\nu}\sum_0^{N-1}|z|^{2\mu} + 2\sum_0^{N-1}\nu|z|^{2\nu}\sum_0^{N-1}\mu|z|^{2\mu} + \sum_0^{N-1}|z|^{2\nu}\sum_0^{N-1}\mu^2|z|^{2\mu}\\
&= 2\left(K(t\partial_t)^2K + (t\partial_t K)^2\right)_{t=|z|^2}.
\end{aligned} \tag{5.6}$$

Similarly,

$$\begin{aligned}
(z\partial_z Z|Z) &= \sum_{j,k=1}^{N}(N-j+k-1)|z|^{2(N-j+k-1)}\\
&= \sum_{\nu=0}^{N-1}\sum_{\mu=1}^{N-1}(\nu+\mu)|z|^{2(\nu+\mu)}\\
&= 2(Kt\partial_t K)_{t=|z|^2}.
\end{aligned}$$

Then, by a straightforward calculation,

$$|\partial_z Z|^2 - \frac{|(\partial_z Z|Z)|^2}{|Z|^2} = \left(\frac{2}{t}\left(K(t\partial_t)^2K - (t\partial_t K)^2\right)\right)_{t=|z|^2} \tag{5.7}$$

Here,

$$\begin{aligned}
\frac{2}{t}\left(K(t\partial_t)^2K - (t\partial_t K)^2\right) &= \frac{2}{t}\sum_0^{N-1}t^\nu\sum_0^{N-1}\nu^2t^\nu - \frac{2}{t}\left(\sum_0^{N-1}\nu t^\nu\right)^2\\
= \sum_{\nu,\mu=0}^{N-1}\left(\nu^2+\mu^2-2\nu\mu\right)t^{\nu+\mu-1} &= \sum_{\nu,\mu=0}^{N-1}(\nu-\mu)^2t^{\nu+\mu-1}\\
&= \sum_{k=0}^{2N-3}a_{k,N}t^k,
\end{aligned}$$

where

$$a_{k,N} = \sum_{\substack{\nu+\mu-1=k\\ 0\le\nu,\mu\le N-1}}(\nu-\mu)^2.$$

We observe that

$$a_{k,N} \le \mathcal{O}(1)(1+k)^3 \text{ uniformly with respect to } N,$$
$$a_{k,N} = a_{k,\infty} \text{ is independent of } N \text{ for } k \le N-2,$$
$$a_{k,\infty} \ge (1+k)^3/\mathcal{O}(1).$$

We conclude that

$$\frac{1}{C}\left(1+M_{N-1,3}\right) \leq \frac{2}{t}\left(K(t\partial_t)^2K-(t\partial_t K)^2\right) \leq C\left(1+M_{2N-2,3}\right)$$

and (4.39) shows that the first and third members are of the same order of magnitude,

$$\asymp 1+M_{N,3}(t) \asymp \min\left(\frac{1}{1-t}, N\right)^4$$

which is $\asymp 1+M_{\infty,3}(t)$, for $0 \leq t \leq 1-1/N$. From this and Proposition 4.2 we get:

Proposition 5.2. *We have*

$$\frac{2}{t}(K(t\partial_t)^2K-(t\partial_t K)^2) \asymp K^4,\ 0 < t \leq 1-1/N, \tag{5.8}$$

where we recall that $K = K_N$ depends on N and that

$$K_N = K_\infty - \frac{t^N}{1-t}.$$

We have

$$\begin{cases} t\partial_t K_N = t\partial_t K_\infty + \mathcal{O}\left(\frac{Nt^N}{1-t}\right), & t \leq 1-\frac{1}{N}, \\ (t\partial_t)^2 K_N = (t\partial_t)^2 K_\infty + \mathcal{O}\left(\frac{N^2t^N}{1-t}\right), & t \leq 1-\frac{1}{N}, \end{cases} \tag{5.9}$$

and it follows that

$$\begin{aligned} &\frac{2}{t}\left(K_N(t\partial_t)^2K_N-(t\partial_t K_N)^2\right) - \frac{2}{t}\left(K_\infty(t\partial_t)^2K_\infty-(t\partial_t K_\infty)^2\right) \\ &= \mathcal{O}\left(\frac{N^2t^N}{(1-t)^2}\right), \end{aligned} \tag{5.10}$$

for $t \leq 1-1/N$.

Proposition 5.2 and (5.7) give

$$|\partial_z Z|^2 - \frac{|(\partial_z Z|Z)|^2}{|Z|^2} \asymp K(|z|^2)^4. \tag{5.11}$$

This implies that $\partial_z Z$, Z are linearly independent.

Assume that

$$|\nabla_z e_1(z)| = \mathcal{O}(m)$$

for some weight $m \geq 1$. We shall see below that this holds when $m = K(|z|^2)$. Then $\|\nabla_z \pi(z)\| = \mathcal{O}(m)$ and hence $\|\nabla_z V\| = \mathcal{O}(m)$. It follows that $\|\nabla_z(V^*(z)V(z))\| = \mathcal{O}(m)$. By standard (Cauchy–Riesz) functional calculus, using also that $\|V(z)^{-1}\| = \mathcal{O}(1)$, we get $\|\nabla_z(V^*(z)V(z))^{-\frac{1}{2}}\| = \mathcal{O}(m)$. Hence $\|\nabla_z U(z)\| = \mathcal{O}(m)$, where $U(z) = V(z)(V^*(z)V(z))^{-1/2}$ is the isometry appearing in the proof of Proposition 5.1. Since $\nabla_z e_j = (\nabla_z U(z))\nu_j^0$, we conclude that

$$\left|\sum_3^{N^2} \alpha_j \nabla_z e_j\right| \leq \mathcal{O}(m)\|\alpha\|_{\mathbb{C}^{N^2-2}}. \tag{5.12}$$

We next show that we can take $m = K(|z|^2)$. We have

$$\nabla_z e_1 = \frac{\nabla_z \overline{Z}}{|Z|} - \frac{\nabla_z |Z|}{|Z|^2}\overline{Z} = \frac{\nabla_z \overline{Z}}{K} - \frac{K' \nabla_z (z\overline{z})}{K^2}\overline{Z}. \tag{5.13}$$

By (5.6),

$$|\partial_z Z| = \left(\frac{2}{t} \left(K(t\partial_t)^2 K + (t\partial_t K)^2 \right) \right)^{\frac{1}{2}}_{t=|z|^2} = \mathcal{O}(K^2).$$

Since Z is holomorphic, this leads to the same estimates for $|\nabla_z Z|$ and $|\nabla_z \overline{Z}|$, and $|\partial_z^2 Z| = \mathcal{O}(K^3)$, for $|z| < 1 - N^{-1}$, by the Cauchy inequalities. Using this in (5.13), we get

$$|\nabla_z e_1| = \mathcal{O}(K). \tag{5.14}$$

Thus we can take $m = K(|z|^2)$ in (5.12). Let f_2 be the vector in (5.4) so that $e_2(z) = |f_2|^{-1} f_2$. Recall that $e_j = U(z)\nu_j^0$, where we now know that $\|\nabla_z U(z)\| = \mathcal{O}(K)$. Write,

$$\nabla_z f_2 = \nabla_z \overline{\partial_z Z} - \sum_{j \neq 2} \left((\nabla_z \overline{\partial_z Z} | e_j) e_j + (\overline{\partial_z Z} | \nabla_z e_j) e_j + (\overline{\partial_z Z} | e_j) \nabla_z e_j \right).$$

Here, $|\nabla_z \overline{\partial_z Z}| = \mathcal{O}(K^3)$, as we have just seen. It is also clear that the term for $j = 1$ in the sum above is $\mathcal{O}(K^3)$. It remains to study $|\mathrm{I} + \mathrm{II} + \mathrm{III}| \leq |\mathrm{I}| + |\mathrm{II}| + |\mathrm{III}|$, where

$$\mathrm{I} = \sum_3^{N^2} (\nabla_z \overline{\partial_z Z} | e_j) e_j,$$

$$\mathrm{II} = \sum_3^{N^2} (\overline{\partial_z Z} | \nabla_z e_j) e_j,$$

$$\mathrm{III} = \sum_3^{N^2} (\overline{\partial_z Z} | e_j) \nabla_z e_j.$$

Here, $|\mathrm{I}| \leq |\nabla_z \overline{\partial_z Z}| = \mathcal{O}(K^3)$ and by (5.12) we have $|\mathrm{III}| \leq \mathcal{O}(K) |\overline{\partial_z Z}| = \mathcal{O}(K^3)$. Further,

$$\begin{aligned} \mathrm{II} &= \sum_3^{N^2} (\overline{\partial_z Z} | (\nabla_z U(z)) \nu_j^0) e_j \\ &= \sum_3^{N^2} ((\nabla_z U(z))^* \overline{\partial_z Z} | \nu_j^0) e_j, \end{aligned}$$

so

$$|\mathrm{II}| = |(\nabla_z U(z))^* \overline{\partial_z Z}| = \mathcal{O}(K) K^2 = \mathcal{O}(K^3).$$

Thus,

$$|\nabla_z f_2| = \mathcal{O}(K^3). \tag{5.15}$$

Recall from (5.5) that for $z = z_0$,

$$f_2 = \overline{\partial_z Z} - (\overline{\partial_z Z}|e_1)e_1,$$

$$|f_2|^2 = |\partial_z Z|^2 - \frac{|(\partial_z Z|Z)|^2}{|Z|^2},$$

so by (5.11),

$$|f_2(z_0)| \asymp K(|z_0|^2)^2,$$

Hence,

$$|f_2(z)| \asymp K^2, \ z \in \operatorname{neigh}(z_0).$$

From this, (5.4) and (5.15), we conclude first that $\nabla_z |f_2| = \mathcal{O}(K^3)$ and then that

$$|\nabla_z e_2| = \mathcal{O}(K). \tag{5.16}$$

This completes the proof of the fact that we can take $m = K$ above. In particular (5.12) holds with $m = K(|z|^2) \asymp G(|z|)$, so

$$\left| \sum_2^{N^2} \alpha_j \partial_z e_j \right| \le \mathcal{O}(1) G|\alpha| \le \mathcal{O}(1) GN, \tag{5.17}$$

where we used the assumption that $|Q| \le C_1 N$ in the last step.

Combining this with (4.52), (4.51), (4.46), (4.34) and the observation prior to Proposition 4.2, we get

$$\begin{aligned} \partial_z f &= \frac{\mathcal{O}(1)}{\delta G^2} \left(N|z|^{N-1} + \delta \left(\frac{|z|^N}{\delta G^2} + \delta N^2 \right) G^3 + G^2 \delta^2 N G N + \frac{G^2 \delta^2 N^2}{r_0 - |z|} \right) \\ &= \mathcal{O}(1) \left(\frac{N|z|^{N-1}}{\delta G^2} + \frac{|z|^N}{\delta G} + G\delta N^2 + \frac{\delta N^2}{r_0 - |z|} \right). \end{aligned}$$

In the last parenthesis the second term is dominated by the first one and the third term is dominated by the fourth. If we recall that $r_0 - |z| \ge 1/N$, we get

$$\partial_z f = \mathcal{O}(1) \left(\frac{N|z|^{N-1}}{\delta G^2} + \delta N^3 \right). \tag{5.18}$$

Similarly, from (4.50), (4.43) we get

$$\begin{aligned} \partial_{\overline{z}} f &= \frac{\mathcal{O}(1)}{\delta G^2} \left(\delta \left(\frac{|z|^N}{\delta G^2} + \delta N^2 \right) G^3 + G^2 \delta^2 N \left(\left(\frac{|z|^N}{\delta G^2} + \delta N^2 \right) G^2 + GN \right) \right) \\ &= \mathcal{O}(1) \left(\frac{|z|^N}{\delta G} + \delta N^2 G + N|z|^N + G^2 \delta^2 N^3 + G\delta N^2 \right). \end{aligned}$$

Using (4.20), we get

$$\partial_{\overline{z}} f = \mathcal{O}(1) \left(\frac{|z|^N}{\delta G} + \delta N^2 G \right), \tag{5.19}$$

see (4.46). This will be used together with the estimates $\partial_{\alpha_j} f = \mathcal{O}(\delta N)$ in (4.51).

Proposition 5.3. *We express Q in the canonical basis in $\mathbb{C}^{N^2}$ or in any other fixed orthonormal basis. Let $e_1(z),\dots,e_{N^2}(z)$ be an orthonormal basis in $\mathbb{C}^{N^2}$ depending smoothly on z and with $e_1(z)=|Z(z)|^{-1}\overline{Z}(z)$, $\mathbb{C}e_1(z)\oplus\mathbb{C}e_2(z)=\mathbb{C}\overline{Z}(z)\oplus\overline{\partial_z Z}(z)$. Write $Q=\alpha_1\overline{Z}(z)+\sum_2^{N^2}\alpha_j e_j(z)$, and recall that the hypersurface*

$$\{(z,Q)\in D(0,r_0-1/N)\times B(0,C_1N);\ E_{-+}^{\delta}(z,Q)=0\}$$

is given by (4.45) *with f as in* (4.46)*. Then the restriction of $dQ\wedge d\overline{Q}$ to this hypersurface, is given by*

$$\begin{aligned} &dQ\wedge d\overline{Q}=J(f)dz\wedge d\overline{z}\wedge d\alpha'\wedge d\overline{\alpha}',\\ &J(f)=-\frac{|\alpha_2|^2}{|Z|^2}\left|\left(e_2|\overline{\partial_z Z}\right)\right|^2+\mathcal{O}(1)\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+|\alpha_2|\delta NG^2\right)^2\\ &\qquad\qquad+\mathcal{O}(1)|\alpha_2|G\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+|\alpha_2|G^2\delta N\right). \end{aligned}\tag{5.20}$$

Here $\alpha'=(\alpha_2,\dots,\alpha_{N^2})$, $d\alpha'\wedge d\overline{\alpha}'=d\alpha_2\wedge d\overline{\alpha}_2\wedge\cdots\wedge d\alpha_{N^2}\wedge d\overline{\alpha}_{N^2}$.

Proof. The differential form $dQ_1\wedge dQ_2\wedge\cdots\wedge dQ_{N^2}$ will change only by a factor of modulus one if we express Q in another fixed orthonormal basis and we will choose for that the basis $e_1(z_0),\dots,e_{N^2}(z_0)$:

$$Q=\sum_1^{N^2}Q_ke_k(z_0),\quad Q_k=(Q|e_k(z_0)).$$

Write

$$Q=\alpha_1\underbrace{\overline{Z}(z)}_{|Z(z)|e_1(z)}+\sum_2^{N^2}\alpha_ke_k(z)$$

and restrict to $\alpha_1=f(z,\alpha_2,\dots,\alpha_{N^2})$, where we sometimes identify $\alpha'\in\overline{Z}(z)^{\perp}$ with $(\alpha_2,\dots,\alpha_{N^2})$:

$$Q_{|\alpha_1=f(z,\alpha')}=f(z,\alpha')\overline{Z}(z)+\sum_2^{N^2}\alpha_ke_k(z).$$

Then,

$$Q_j=f(\overline{Z}(z)|e_j(z_0))+\sum_{k=2}^{N^2}\alpha_k(e_k(z)|e_j(z_0)),$$

$$\begin{aligned} dQ_j=&(d_zf+d_{\alpha'}f)(\overline{Z}(z)|e_j(z_0))+f(d_z\overline{Z}(z)|e_j(z_0))\\ &+\sum_{k=2}^{N^2}\alpha_k(d_ze_k(z)|e_j(z_0))+\sum_{k=2}^{N^2}d\alpha_k(e_k(z)|e_j(z_0)). \end{aligned}$$

Taking $z = z_0$ until further notice, we get with $\alpha' = (\alpha_2, \dots, \alpha_{N^2})$:

$$dQ_j = (d_z f + d_{\alpha'} f)(\overline{Z}|e_j) + f(\overline{\partial_z Z}|e_j)d\overline{z} + \alpha_2(d_z e_2|e_j) + \begin{cases} d\alpha_j, \ j \geq 2, \\ 0, \ j = 1. \end{cases}$$

Here, we used (5.3). The first term to the right is equal to $(d_z f + d_{\alpha'} f)|\overline{Z}|$ when $j = 1$ and it vanishes when $j \geq 2$. The second term vanishes for $j \geq 3$, by (5.2). The third term is equal to $-\alpha_2(e_2|d_z e_j)$ (by differentiation of the identity $(e_2|e_j) = \delta_{2,j}$) and it vanishes for $j \geq 3$ (remember that we take $z = z_0$). Thus, for $z = z_0$:

$$\begin{aligned}
dQ_1 &= |\overline{Z}|(d_z f + d_{\alpha'} f) + f(\overline{\partial_z Z}|e_1)d\overline{z} - \alpha_2(e_2|d_z e_1), \\
dQ_2 &= d\alpha_2 + f(\overline{\partial_z Z}|e_2)d\overline{z} - \alpha_2(e_2|d_z e_2), \\
dQ_j &= d\alpha_j, \ j \geq 3.
\end{aligned}$$

When forming $dQ_1 \wedge d\overline{Q}_1 \wedge \cdots \wedge dQ_{N^2} \wedge d\overline{Q}_{N^2}$ we see that the terms in $d\alpha_j$ for $j \geq 3$ in the expression for dQ_1 will not contribute, so in that expression we can replace $d_{\alpha'} f$ by $\partial_{\alpha_2} f d\alpha_2$. Using (5.18), (5.19), (4.51), (4.46), (4.43) this gives, where "$\equiv$" means equivalence up to terms that do not influence the $2N^2$ form above:

$$\begin{aligned}
dQ_1 \equiv &- \alpha_2(e_2|d_z e_1) + \mathcal{O}(1)\left(\frac{N|z|^{N-1}}{\delta G} + G\delta N^3\right) dz \\
&+ \mathcal{O}(1)\left(\frac{|z|^N}{\delta} + G^2\delta N^2\right) d\overline{z} + \mathcal{O}(\delta N G)d\alpha_2.
\end{aligned}$$

Similarly, using also (5.16),

$$dQ_2 = d\alpha_2 + \mathcal{O}\left(\frac{|z|^N}{\delta} + \delta N^2 G^2 + |\alpha_2|G\right) d\overline{z} + \mathcal{O}\left(|\alpha_2|G\right) dz.$$

When computing $dQ_1 \wedge dQ_2$ we notice that the terms in $dz \wedge d\overline{z}$ will not contribute to the $2N^2$-form $dQ_1 \wedge d\overline{Q}_1 \wedge \cdots \wedge dQ_{N^2} \wedge d\overline{Q}_{N^2}$. We get

$$\begin{aligned}
dQ_1 \wedge dQ_2 \equiv &- \alpha_2(e_2|d_z e_1) \wedge d\alpha_2 \\
&+ \mathcal{O}(1)\left(\frac{N|z|^{N-1}}{\delta G} + G\delta N^3 + |\alpha_2|\delta N G^2\right) dz \wedge d\alpha_2 \\
&+ \mathcal{O}(1)\left(\frac{|z|^N}{\delta} + G^2\delta N^2 + |\alpha_2|\delta N G^2\right) d\overline{z} \wedge d\alpha_2.
\end{aligned} \tag{5.21}$$

Here,

$$\begin{aligned}
(e_2|d_z e_1) &= \left(e_2|d_z\left(|Z|^{-1}\right)\overline{Z}\right) = \left(e_2||Z|^{-1}d_z\overline{Z}\right) + \left(e_2|d_z\left(|Z|^{-1}\right)\overline{Z}\right) \\
&= |Z|^{-1}\left(e_2|\overline{\partial_z Z}d\overline{z}\right) + 0 = |Z|^{-1}(e_2|\overline{\partial_z Z})dz,
\end{aligned}$$

so the first term in (5.21) is equal to

$$-\frac{\alpha_2}{|Z|}(e_2|\overline{\partial_z Z})dz \wedge d\alpha_2 = \mathcal{O}(1)\alpha_2 G dz \wedge d\alpha_2.$$

Notice that $dQ_1 \wedge d\overline{Q}_1 \wedge dQ_2 \wedge d\overline{Q}_2 = -dQ_1 \wedge dQ_2 \wedge d\overline{Q}_1 \wedge d\overline{Q}_2$. From (5.21) and its complex conjugate we get

$$\begin{aligned} &dQ_1 \wedge d\overline{Q}_1 \wedge dQ_2 \wedge d\overline{Q}_2 \\ &\quad \equiv \left(-\frac{|\alpha_2|^2}{|Z|^2} \left|\left(e_2|\overline{\partial_z Z}\right)\right|^2 + \mathcal{O}(1)\left(\frac{N|z|^{N-1}}{\delta G} + G\delta N^3 + |\alpha_2|\delta N G^2\right)^2 \right. \\ &\qquad \left. + \mathcal{O}(1)|\alpha_2|G\left(\frac{N|z|^{N-1}}{\delta G} + G\delta N^3 + |\alpha_2|G^2\delta N\right)\right) dz \wedge d\overline{z} \wedge d\alpha_2 \wedge d\overline{\alpha}_2. \quad \square \end{aligned}$$

6. Proof of Theorem 1.2

Let $Q \in \mathbb{C}^{N^2}$ be an $N \times N$ matrix whose entries are independent random variables $\sim \mathcal{N}_{\mathbb{C}}(0,1)$, so that the corresponding probability measure is

$$\pi^{-N^2} e^{-|Q|^2} (2i)^{-N^2} d\overline{Q}_1 \wedge dQ_1 \wedge \cdots \wedge d\overline{Q}_{N^2} \wedge dQ_{N^2} = \frac{1}{(2\pi i)^{N^2}} e^{-|Q|^2} d\overline{Q} \wedge dQ.$$

We are interested in

$$K_\phi = \mathbb{E}\left(1_{B_{\mathbb{C}^{N^2}}}(0,1) \sum_{\lambda \in \sigma(A_0 + \delta Q)} \phi(\lambda) \right), \quad \phi \in C_0(D(0, r_0 - 1/N), \tag{6.1}$$

which is of the form (3.3) with

$$m(Q) = 1_{B_{\mathbb{C}^{N^2}}}(Q) \pi^{-N^2} e^{-|Q|^2}, \tag{6.2}$$

so we have (3.8), (3.9) with $J(f)$ as in (5.20) and f as in (4.45). More explicitly,

$$\widetilde{\Xi}(z) = \int_{|f|^2|Z(z)|^2 + |\alpha'|^2 \le C_1^2 N^2} \pi^{-N^2} e^{-|f(z,\alpha')|^2 |Z(z)|^2 - |\alpha'|^2} J(f)(z, \alpha') L(d\alpha').$$

By (4.46), (4.20), (4.25):

$$|f| \le \mathcal{O}(1) \frac{N}{G} \left(\frac{|z|^N}{\delta N G} + \delta N G \right) \ll \frac{N}{G}.$$

We now strengthen (4.20), (4.25) to the assumption

$$\frac{|z|^N}{\delta N G} + \delta N G \ll \frac{1}{N}, \text{ for all } z \in D(0, r_0), \tag{6.3}$$

implying that $|f|G \ll 1$, for all $z \in D(0, r_0)$. Equivalently, by the same reasoning as after (4.26), r_0 should satisfy

$$\frac{r_0^N}{\delta N G(r_0)} + \delta N G(r_0) \ll \frac{1}{N}. \tag{6.4}$$

Then

$$e^{-|f(z,\alpha')|^2 |Z(z)|^2} = 1 + \mathcal{O}(1) N^2 \left(\frac{|z|^N}{\delta N G} + \delta N G \right)^2,$$

and using (5.20), we get

$$\begin{aligned}\widetilde{\Xi}(z) = &\left(1+\mathcal{O}(1)N^2\left(\frac{|z|^N}{\delta NG}+\delta NG\right)^2\right)\\ &\times\frac{|(e_2|\overline{\partial_z Z})|^2}{|Z|^2}\int_{|(f|Z|,\alpha')|\leq C_1N}|\alpha_2|^2e^{-|\alpha'|^2}\pi^{1-N^2}L(d\alpha')\\ &+\mathcal{O}(1)\int e^{-|\alpha'|^2}\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+|\alpha_2|\delta NG^2\right)^2\pi^{1-N^2}L(d\alpha')\\ &+\mathcal{O}(1)\int e^{-|\alpha'|^2}|\alpha_2|G\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+|\alpha_2|\delta NG^2\right)\pi^{1-N^2}L(d\alpha').\end{aligned}$$

Since $|f||Z|\ll N$, the first integral is equal to

$$\int_{\mathbb{C}}\frac{1}{\pi}|w|^2e^{-|w|^2}L(dw)+\mathcal{O}\left(e^{-N^2/\mathcal{O}(1)}\right)=1+\mathcal{O}\left(e^{-N^2/\mathcal{O}(1)}\right).$$

The sum of the other two integrals is equal to

$$\begin{aligned}&\mathcal{O}(1)\left(\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+\delta NG^2\right)^2+G\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3+\delta NG^2\right)\right)\\ &=\mathcal{O}(1)\left(\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3\right)^2+G\left(\frac{N|z|^{N-1}}{\delta G}+G\delta N^3\right)\right).\end{aligned}$$

Noticing that

$$\frac{|(e_2|\overline{\partial_z Z})|^2}{|Z|^2}=\mathcal{O}(G^2),$$

we deduce that

$$\begin{aligned}\widetilde{\Xi}(z) =&\frac{|(e_2|\overline{\partial_z Z})|^2}{|Z|^2}\\ &+\mathcal{O}(1)\left(G^2N^2\left(\frac{|z|^{N-1}}{\delta G^2}+\delta N^2\right)^2+G^2N\left(\frac{|z|^{N-1}}{\delta G^2}+\delta N^2\right)\right).\end{aligned}\tag{6.5}$$

We next study the leading term in (6.5), given by

$$\frac{|(\overline{\partial_z Z}|e_2)|^2}{\pi|Z|^2}.\tag{6.6}$$

Since $\overline{\partial_z Z}$ belongs to the span of $e_1=\overline{Z}/|Z|$ and e_2, we have

$$|(\overline{\partial_z Z}|e_2)|^2=|\overline{\partial_z Z}|^2-|(\overline{\partial_z Z}|e_1)|^2,$$

so the leading term (6.6) is

$$\frac{1}{\pi|Z|^2}\left(|\overline{\partial_z Z}|^2-\frac{|(\overline{\partial_z Z}|\overline{Z})|^2}{|Z|^2}\right),$$

which by (5.7) is equal to

$$\frac{2}{\pi t}\left(\frac{(t\partial_t)^2 K}{K} - \frac{(t\partial_t K)^2}{K^2}\right)_{t=|z|^2}. \tag{6.7}$$

Here, $K = K_N(t) = \sum_0^{N-1} t^\nu$ is the function appearing in Proposition 5.2. Let us first compute the limiting quantity obtained by replacing $K = K_N$ in (6.7) by $K_\infty = 1/(1-t)$. Since $\partial_t K_\infty = K_\infty^2$, we get

$$t\partial_t K_\infty = tK_\infty^2, \quad (t\partial_t)^2 K_\infty = tK_\infty^2 + 2t^2 K_\infty^3,$$

and

$$\frac{2}{\pi t}\left(\frac{(t\partial_t)^2 K_\infty}{K_\infty} - \frac{(t\partial_t K_\infty)^2}{K_\infty^2}\right) = \frac{2}{\pi} K_\infty^2 = \frac{2}{\pi}\frac{1}{(1-t)^2}. \tag{6.8}$$

We next approximate the expression (6.7) with (6.8), using (5.10) and the fact that $K = (1+\mathcal{O}(t^N))K_\infty$ (uniformly with respect to N). The expression (6.7) is equal to

$$\begin{aligned} &\frac{2}{\pi t K^2}(K(t\partial_t)^2 K - (t\partial_t K)^2) \\ &= \frac{2(1+\mathcal{O}(t^N))}{\pi t K_\infty^2}\left(K_\infty (t\partial_t)^2 K_\infty - (t\partial_t K_\infty)^2 + \mathcal{O}(N^2 t^N K_\infty^2)\right). \end{aligned}$$

Here,

$$(t\partial_t K_\infty)^2 = \mathcal{O}(t^2 K_\infty^4), \quad K_\infty (t\partial_t)^2 K_\infty = \mathcal{O}(tK_\infty^3 + t^2 K_\infty^4),$$

so the last expression becomes,

$$\frac{2}{\pi t}\left(\frac{(t\partial_t)^2 K_\infty}{K_\infty} - \frac{(t\partial_t K_\infty)^2}{K_\infty^2}\right) + \mathcal{O}(t^N K_\infty + t^{N+1} K_\infty^2 + t^{N-1} N^2),$$

where the first two terms in the remainder are dominated by the last one. We conclude that the difference between the expressions (6.7) and (6.8) is $\mathcal{O}(t^{N-1}N^2)$, and using also (6.5), we get,

$$\begin{aligned} \widetilde{\Xi}(z) =& \frac{2}{\pi(1-|z|^2)^2} + \mathcal{O}(|z|^{2(N-1)}N^2) \\ &+ \mathcal{O}(1)\left(G^2 N^2 \left(\frac{|z|^{N-1}}{\delta G^2} + \delta N^2\right)^2 + G^2 N \left(\frac{|z|^{N-1}}{\delta G^2} + \delta N^2\right)\right). \end{aligned} \tag{6.9}$$

The remainder term can be written

$$\mathcal{O}(G^2)\left(\frac{|z|^{2(N-1)}N^2}{G^2} + \frac{|z|^{2(N-1)}N^2}{\delta^2 G^4} + \delta^2 N^6 + \frac{|z|^{N-1}N}{\delta G^2} + \delta N^3\right).$$

By (6.3), $\frac{1}{\delta G} \gg N^2$, so the second term is

$$\gg \frac{|z|^{2(N-1)}N^2}{G^2} N^4,$$

which is much larger than the first term. We now strengthen (6.3) to

$$\frac{|z|^{N-1}}{\delta G^2} + \delta N^2 \ll \frac{1}{N},$$

or equivalently to

$$\frac{|z|^{N-1}N}{\delta G^2} + \delta N^3 \ll 1. \tag{6.10}$$

Then the remainder in (6.9) becomes

$$\mathcal{O}(G^2)\left(\frac{|z|^{N-1}N}{\delta G^2} + \delta N^3\right),$$

and (6.9) becomes

$$\widetilde{\widetilde{\Xi}}(z) = \frac{2}{\pi(1-|z|^2)^2}\left(1 + \mathcal{O}\left(\frac{|z|^{N-1}N}{\delta G^2} + \delta N^3\right)\right). \tag{6.11}$$

Setting $\widetilde{\widetilde{\Xi}} = \frac{1}{2\pi}\Xi$ concludes the proof of Theorem 1.2.

Acknowledgment

We would like to thank Stéphane Nonnenmacher for his observation on the relation of the density in Theorem 1.2 with the Poincaré metric. The first author was partially supported by the project ANR NOSEVOL 2011 BS 010119 01.

References

[1] W. Bordeaux-Montrieux, *Loi de Weyl presque sûre et résolvent pour des opérateurs différentiels non-autoadjoints, Thèse*, pastel.archives-ouvertes.fr/docs/00/50/12/81/PDF/manuscrit.pdf (2008).

[2] C. Bordenave and M. Capitaine, *Outlier eigenvalues for deformed i.i.d. random matrices*, Communications on Pure and Applied Mathematics **69** (2016), no.11, 2131–2194.

[3] E. B. Davies, *Pseudospectra of Differential Operators*, J. Oper. Th **43** (1997), 243–262.

[4] E. B. Davies, *Pseudo–spectra, the harmonic oscillator and complex resonances*, Proc. of the Royal Soc.of London A **455** (1999), no. 1982, 585–599.

[5] E. B. Davies and M. Hager, *Perturbations of Jordan matrices*, J. Approx. Theory **156** (2009), no. 1, 82–94.

[6] N. Dencker, J. Sjöstrand, and M. Zworski, *Pseudospectra of semiclassical (pseudo-) differential operators*, Communications on Pure and Applied Mathematics **57** (2004), no. 3, 384–415.

[7] A. Guionnet, P. Matchett Wood, and O. Zeitouni, *Convergence of the spectral measure of non-normal matrices*, Proc. AMS **142** (2014), no. 2, 667–679.

[8] M. Hager, *Instabilité Spectrale Semiclassique d'Opérateurs Non-Autoadjoints II*, Annales Henri Poincaré **7** (2006), 1035–1064.

[9] ———, *Instabilité spectrale semiclassique pour des opérateurs non-autoadjoints I: un modèle*, Annales de la faculté des sciences de Toulouse Sér. 6 **15** (2006), no. 2, 243–280.

[10] M. Hager and J. Sjöstrand, *Eigenvalue asymptotics for randomly perturbed non-selfadjoint operators*, Mathematische Annalen **342** (2008), 177–243.

[11] J. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, American Mathematical Society, 2009.

[12] M. Kac, *On the average number of real roots of a random algebraic equation*, Bulletin of the American Mathematical Society **49** (1943), no. 4, 314–320.

[13] B. Shiffman, *Convergence of random zeros on complex manifolds*, Science in China Series A: Mathematics **51** (2008), 707–720.

[14] B. Shiffman and S. Zelditch, *Distribution of Zeros of Random and Quantum Chaotic Sections of Positive Line Bundles*, Communications in Mathematical Physics **200** (1999), 661–683.

[15] ______, *Equilibrium distribution of zeros of random polynomials*, Int. Math. Res. Not. (2003), 25–49.

[16] J. Sjöstrand, *Spectral properties of non-self-adjoint operators*, Notes d'un minicours à Evian les Bains, 8–12 juin (2009), Actes des Journées d'é.d.p. d'Évian 2009 (http://jedp.cedram.org/jedp-bin/fitem?id=JEDP_2009____A1_0 and http://arxiv.org/abs/1002.4844}).

[17] J. Sjöstrand and M. Zworski, *Elementary linear algebra for advanced spectral problems*, Annales de l'Institute Fourier **57** (2007), 2095–2141.

[18] J. Sjöstrand, *Non-self-adjoint differential operators, spectral asymptotics and random perturbations*, current monograph preprint version at http://sjostrand.perso.math.cnrs.fr/161014bok.pdf

[19] M. Sodin, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, Mathematical Research Letters (2000), no. 7, 371–381.

[20] M. Sodin and B. Tsirelson, *Random complex zeroes, I. Asymptotic normality*, Israel Journal of Mathematics **144** (2004), 125–149.

[21] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.

[22] L. N. Trefethen, *Pseudospectra of linear operators*, SIAM Rev. **39** (1997), no. 3, 383–406.

[23] M. Vogel, *The precise shape of the eigenvalue intensity for a class of non-selfadjoint operators under random perturbations*, arxiv:1401.8134v1 [math.SP] (2014).

[24] M. Zworski, *Numerical linear algebra and solvability of partial differential equations,* Comm. Math. Phys. 229(2)(2002), 293-307.

[25] M. Zworski and T. J. Christiansen, *Probabilistic Weyl Laws for Quantized Tori*, Communications in Mathematical Physics **299** (2010).

Johannes Sjöstrand
Institut de Mathématiques de Bourgogne
– UMR 5584 CNRS
Université de Bourgogne
Faculté des Sciences Mirande
9 avenue Alain Savary, BP 47870
F-21078 Dijon Cedex
e-mail: johannes.sjostrand@u-bourgogne.fr

Martin Vogel
Laboratoire de Mathématiques d'Orsay
– UMR 8628 CNRS
Université Paris-Sud
Bâtiment 425
F-91405 Orsay Cedex
e-mail: martin.vogel@math.u-psud.fr