CONTRIBUTIONS TO DIRECTED ALGEBRAIC TOPOLOGY

– with inspirations from concurrency theory

Martin Raussen
Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7G
DK - 9220 Aalborg Øst
E-mail: `raussen@math.aau.dk`

Contents

Preface

This doctoral thesis is concerned with contributions to the research field *Directed Algebraic Topology*. It is based on seventeen of the author's published research papers ([**1**] to [**17**]) in this area from the period 1998 – 2013.

Directed Algebraic Topology is a quite new research discipline. It tries to modify and twist methodology from "classical" Algebraic Topology to a situation where paths, in general, are no-longer reversible. So far, its main motivation comes from the theory of concurrent processes in theoretical computer science. As our main example, we use a geometric/combinatorial abstract model for concurrent computations. We try to find answers to questions concerned with the space of executions for such a model.

While concurrency has served as a prime motivation, the field has also arisen purely mathematical and computational interest. I have to admit that I personally am most passionate about these mathematical aspects; but I am very happy to see that some of them can, in the hands of others, be transformed into actually useful running algorithms.

My thanks go to friends, collaborators, and colleagues around the world. Particular thanks go to my colleague and companion from Aalborg University in this endeavour, Lisbeth Fajstrup. Furthermore, I wish to thank explicitly our long-time friends and collaborators from the LIST laboratory of the CEA at Saclay/Paris in France, first of all Éric Goubault, Emmanuel Haucourt and Samuel Mimram. All the good colleagues in the ACAT network of the ESF have to be mentioned here, as well.

It is my pleasure to thank my colleagues at the Department of Mathematical Sciences at Aalborg University for providing a friendly, supportive and inspiring professional environment.

Finally, I would like to express my gratitude to my family for love and support throughout many years.

Aalborg, September 2013 Martin Raussen

Thesis papers

This thesis is based on the following seventeen publications, in chronological order:

[1] L. Fajstrup, É. Goubault, and M. Raussen. Detecting Deadlocks in Concurrent Systems. In D. Sangiorgi and R. de Simone, editors, *CONCUR '98; Concurrency Theory*, volume 1466 of *Lect. Notes Comp. Science*, pages 332 – 347, Nice, France, September 1998. Springer-Verlag. 9th Int. Conf., Proceedings.

[2] M. Raussen. On the classification of dipaths in geometric models for concurrency. *Math. Structures Comput. Sci.*, 10(4):427–457, 2000.

[3] É. Goubault and M. Raussen. Dihomotopy as a tool in state space analysis. In S. Rajsbaum, editor, *LATIN 2002: Theoretical Informatics*, volume 2286 of *Lect. Notes Comput. Sci.*, pages 16 – 37, Cancun, Mexico, April 2002. Springer-Verlag.

[4] M. Raussen. State spaces and dipaths up to dihomotopy. *Homology, Homotopy Appl.*, 5(2):257–280, 2003.

[5] L. Fajstrup, É. Goubault, E. Haucourt, and M. Raussen. Components of the fundamental category. *Appl. Categ. Structures*, 12(1):81–108, 2004.

[6] L. Fajstrup, É. Goubault, and M. Raussen. Algebraic Topology and Concurrency. *Theor. Comput. Sci*, 357:241–278, 2006. Revised version of Aalborg University preprint, 1999.

[7] M. Raussen. Deadlocks and dihomotopy in mutual exclusion models. *Theoret. Comput. Sci.*, 365(3):247 – 257, 2006.

[8] R. Wisniewski and M. Raussen. Geometric analysis of nondeterminacy in dynamical systems. *Acta Inform.*, 43(7):501 – 519, 2007.

[9] M. Raussen. Invariants of directed spaces. *Appl. Categ. Structures*, 15:355–386, 2007.

[10] U. Fahrenberg and M. Raussen. Reparametrizations of continuous paths. *J. Homotopy Relat. Struct.*, 2(2):93–117, 2007.

[11] M. Raussen. Reparametrizations with given stop data. *J. Homotopy Relat. Struct.*, 4(1):1–5, 2009.

[12] ———. Trace spaces in a pre-cubical complex. *Topology Appl.*, 156(9):1718 – 1728, 2009.

[13] ———. Simplicial models for trace spaces. *Algebr. Geom. Topol.*, 10:1683 – 1714, 2010.

[14] L. Fajstrup, É. Goubault, E. Haucourt, S. Mimram, and M. Raussen. Trace Spaces: An Efficient New Technique for State-Space Reduction. In H. Seidl, editor, *Programming Languages and Systems*, volume 7211 of *Lect. Notes Comput. Sci.*, pages 274–294. ESOP 2012, Springer-Verlag, 2012.

[15] M. Raussen. Execution spaces for simple higher dimensional automata. *Appl. Algebra Engrg. Comm. Comput.*, 23:59–84, 2012.

[16] ———. Simplicial models for trace spaces II: General higher-dimensional automata. *Algebr. Geom. Topol.*, 12(3):1745–1765, 2012.

[17] M. Raussen and K. Ziemiański. Homology of spaces of directed paths on Euclidean cubical complexes. *J. Homotopy Relat. Struct.*, 18 pages, 2013. DOI 10.1007/s40062-013-0045-4.

A brief guide to the papers

This is no attempt to describe the content of the papers in any detail; we just try to explain connections between them and express our view on their relative importance:

The paper [**1**] was the first of our attempts to use geometrical and combinatorial (rather than topological) reasoning in the investigation of a concurrency problem: the detection of *deadlocks* and of *unsafe regions* for a so-called linear PV program.

The papers [**3, 4**] are precursors for [**5**] that uses a combination of categorical and topological ideas and methods to define and investigate *components* for models of certain concurrent programs.

The paper [**9**] attempts to describe and organize in rather great generality the many path spaces that a directed space comes equipped with – and also associated algebraic invariants – in a functorial manner. It proposes a candidate for the title *directed homotopy equivalence*.

The papers [**10, 11**] investigate a peculiar but surprisingly rich topic, the algebra underpinning *reparametrizations* of (directed) paths.

The papers [**2, 7, 12**] can be considered as precursors to [**13**]. The first of them treats first of all the (untypical) 2-dimensional case, the second analyses an eye-opener example in 3D, whereas the last investigates general topological properties of path and trace spaces.

I personally view the paper [**13**] as the most significant contribution. It shows a way to model path and trace spaces *simplicially* (or combinatorially) and gives, for the first time to my knowledge, an algorithmic way to calculate *algebraic topological invariants* of spaces of directed paths given a decent description of the state space. The implementation of the algorithm from [**13**] and an application/extension to a case with directed loops is the main topic of the paper [**14**].

The methods from [**13**] have afterwards been modified and generalized in the papes [**15, 16**] so that they – at least in principle – can be used to identify spaces of executions of general *Higher Dimensional Automata* up to homotopy equivalence by simplicial complexes.

The last paper [**17**] has its origin in a frustration over the fact that the algorithm designed in [**13**] resulted in an all too large simplicial complex in a quite simple interesting case (a directed torus with a hole). More delicate homotopy theoretical tools helped to overcome this problem in this particular case; they will hopefully show to be useful in greater generality.

The paper [**8**] is a bit of an outlier. It builds on Wisniewski's phd-thesis – supervised by me – that investigates almost flow-lines (with respect to a vector field) as d-paths. It uses, first of all, Morse theoretic tools.

The paper [**6**] is a survey article; a first version had been published as an Aalborg University preprint already in 1999. Though quite out of date and deserving an update, it still gets citations today.

Full list of publications

Mathematical Papers

In *reverse* chronological order, based on Mathematical Reviews and Zentralblatt für Mathematik und ihre Grenzgebiete:

[1] M. Raussen and K. Ziemiański. Homology of spaces of directed paths on Euclidean cubical complexes. *J. Homotopy Relat. Struct.*, 18 pages, 2013. DOI 10.1007/s40062-013-0045-4.

[2] M. Raussen. Simplicial models for trace spaces II: General higher-dimensional automata. *Algebr. Geom. Topol.*, 12(3):1745–1765, 2012.

[3] ———. Execution spaces for simple higher dimensional automata. *Appl. Algebra Engrg. Comm. Comput.*, 23:59–84, 2012.

[4] L. Fajstrup, É. Goubault, E. Haucourt, S. Mimram, and M. Raussen. Trace Spaces: An Efficient New Technique for State-Space Reduction. In H. Seidl, editor, *Programming Languages and Systems*, volume 7211 of *Lect. Notes Comput. Sci.*, pages 274–294. ESOP 2012, Springer-Verlag, 2012.

[5] M. Raussen. Simplicial models for trace spaces. *Algebr. Geom. Topol.*, 10:1683–1714, 2010.

[6] ———. Trace spaces in a pre-cubical complex. *Topology Appl.*, 156(9):1718–1728, 2009.

[7] ———. Reparametrizations with given stop data. *J. Homotopy Relat. Struct.*, 4(1):1–5, 2009.

[8] U. Fahrenberg and M. Raussen. Reparametrizations of continuous paths. *J. Homotopy Relat. Struct.*, 2(2):93–117, 2007.

[9] M. Raussen. Invariants of directed spaces. *Appl. Categ. Struct.*, 15(4):355–386, 2007.

[10] R. Wisniewski and M. Raussen. Geometric analysis of nondeterminacy in dynamical systems. *Acta Inf.*, 43(7):501–519, 2007.

[11] M. Raussen. Deadlocks and dihomotopy in mutual exclusion models. *Theor. Comput. Sci.*, 365(3):247–257, 2006.

[12] L. Fajstrup, M. Raussen, and É. Goubault. Algebraic topology and concurrency. *Theor. Comput. Sci.*, 357(1–3):241–278, 2006.

[13] L. Fajstrup, M. Raussen, É. Goubault, and E. Haucourt. Components of the fundamental category. *Appl. Categ. Struct.*, 12(1):81–108, 2004.

[14] M. Raussen. A second look at normal curvature. *Normat*, 51(2):59–62, 2003.

[15] ———. State spaces and dipaths up to dihomotopy. *Homology Homotopy Appl.*, 5(2):257–280, 2003.

[16] É. Goubault and M. Raussen. Dihomotopy as a tool in state space analysis. Rajsbaum, Sergio (ed.), LATIN 2002: Theoretical informatics. 5th Latin American symposium, Cancun, Mexico, April 3–6, 2002. Proceedings. Berlin: Springer. Lect. Notes Comput. Sci. 2286, 16-37 (2002).

[17] M. Raussen. Symmetries on manifolds, deformations and rational homotopy: A survey. Bak, Anthony (ed.) et al., Current trends in transformation groups. Dedicated to the memory of Professor Katsuo Kawakubo. Dordrecht: Kluwer Academic Publishers. K-Monogr. Math. 7, 167-179 (2002).

[18] ———. On the classification of dipaths in geometric models for concurrency. *Math. Struct. Comput. Sci.*, 10(4):427–457, 2000.

[19] M. Raussen. Cyclic group actions on manifolds from deformations of rational homotopy types. *Math. Ann.*, 312(4):737–760, 1998.

[20] L. Fajstrup, É. Goubault, and M. Raussen. Detecting deadlocks in concurrent systems. In *CONCUR'98: concurrency theory (Nice)*, volume 1466 of *Lecture Notes in Comput. Sci.*, pages 332–347. Springer, Berlin, 1998.

[21] E. Laitinen and M. Raussen. Homotopy types of locally linear representation forms. *Manuscr. Math.*, 88(1):33–52, 1995.

[22] M. Raussen. Circle actions on rational homology manifolds and deformations of rational homotopy types. *Trans. Am. Math. Soc.*, 347(1):137–153, 1995.

[23] ———. Rational cohomology and homotopy of spaces with circle action. Algebraic topology, Proc. Conf., S. Feliu de Guíxols/Spain 1990, Lect. Notes Math. 1509, 313-325 (1992)., 1992.

[24] I. Madsen and M. Raussen. Locally linear representation forms. *Osaka J. Math.*, 27(3):567–591, 1990.

[25] ———. Smooth and locally linear G homotopy representations. Algebraic topology, Proc. Conf., Göttingen/Ger. 1984, Lect. Notes Math. 1172, 130-156 (1985).

[26] J.M. Møller and M. Raussen. Rational homotopy of spaces of maps into spheres and complex projective spaces. *Trans. Am. Math. Soc.*, 292:721–732, 1985.

[27] P. Löffler and M. Raussen. Symmetrien von Mannigfaltigkeiten und rationale Homotopietheorie. *Math. Ann.*, 271:549–576, 1985.

[28] M. Raussen. Non-orthogonalisable vector fields on spheres. *Proc. Edinb. Math. Soc., II. Ser.*, 27:275–281, 1984.

[29] M. Raussen and R. Wiegmann. *Liftings, homotopy liftings and localization applied to vector field and immersion theory*, volume 6 of *Osnabrücker Schriften zur Mathematik. Reihe M: Mathematische Manuskripte [Osnabrück Publications in Mathematics. Series M: Mathematical Manuscripts]*. Universität Osnabrück Fachbereich Mathematik, Osnabrück, 1984.

[30] M. Raussen. Some invariants for vector field problems. Topology, Proc. spec. Semin., Vol. 4, México 1982, 163-185 (1982).

[31] ———. Symmetries on simply-connected manifolds. Topology, Proc. spec. Semin., Vol. 4, Mexico 1982, 117-133 (1982).

[32] ———. Liftings and homotopy liftings into fibre bundles: vector fields, unstable vector bundles, immersions. Topology, Proc. spec. Semin., Vol. 2, México 1981,

133-150 (1981).

[33] ______. *Liftungen und Homotopieliftungen in Faserbündel: Vektorfelder, instabile Vektorraumbündel, Immersionen. (Dissertation).* PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Georg-August-Universität zu Göttingen. XIII, 113 S., 1981.

[34] M. Raussen and L. Smith. A geometric interpretation of sphere bundle boundaries and generalized J- homomorphisms with an application to a diagram of I. M. James. *Q. J. Math., Oxf. II. Ser.*, 30:113–117, 1979.

[35] M. Raussen. Hurewicz isomorphism and Whitehead theorems in pro-categories. *Arch. Math.*, 30:153–164, 1978.

Interviews and articles in journals of mathematical societies

In *reverse* chronological order:

[36] M. Raussen and Chr. Skau. Interview with Abel laureate Pierre Deligne. *Eur. Math. Soc. Newsl.*, 89:15–23, 2013.

[37] ———. Interview with Abel laureate Endre Szemerédi. *Eur. Math. Soc. Newsl.*, 85:39–48, 2012. Reprinted in *Notices Am. Math. Soc.*, 60(2):221–231, 2012.

[38] M. Raussen. Computational topology [book review]. *Eur. Math. Soc. Newsl.*, 82:58–59, 2011.

[39] M. Raussen and Chr. Skau. Interview with Abel laureate John Milnor. *Eur. Math. Soc. Newsl.*, 81:31–40, 2011. Reprinted in *Notices Am. Math. Soc.*, 59(3):400–408, 2012.

[40] ———. Interview with Abel laureate John Tate. *Eur. Math. Soc. Newsl.*, 77:41–48, 2010. Reprinted in *Notices Am. Math. Soc.*, 58(3):444–452, 2011.

[41] ———. Interview with Abel laureate Mikhail Gromov. *Eur. Math. Soc. Newsl.*, 73:19–30, 2009. Reprinted in *Notices Am. Math. Soc.*, 57(3):391–403, 2010.

[42] M. Raussen. Interview with Jacques Dixmier. *Eur. Math. Soc. Newsl.*, 72:34–41, 2009.

[43] ———. Invitation to topological robotics [book review]. *Eur. Math. Soc. Newsl.*, 72:46–47, 2009.

[44] M. Raussen and Chr. Skau. Interview with John G. Thompson and Jacques Tits. *Eur. Math. Soc. Newsl.*, 69:31–38, 2008. Reprinted in *Notices Am. Math. Soc.*, 56(4):471–478, 2009.

[45] M. Raussen and A. Valette. An interview with Beno Eckmann. *Eur. Math. Soc. Newsl.*, 66:31–37, 2007. Reprinted in math.ch/100 (Eds: B. Colbois, C. Riedtmann, V. Schroeder), EMS publishing house (2010), 389–401.

[46] M. Raussen and Chr. Skau. Interview with Srinivasa Varadhan. *Eur. Math. Soc. Newsl.*, 65:33–40, 2007. Reprinted in *Notices Am. Math. Soc.*, 55(2): 238–246, 2008.

[47] U. Persson and M. Raussen. Two mathematicians forced to resign at Uppsala University, Sweden. *Eur. Math. Soc. Newsl.*, 64:15, 2007.

[48] M. Raussen and Chr. Skau. Interview with Abel Prize recipient Lennart Carleson. *Eur. Math. Soc. Newsl.*, 61:31–36, 2006. Reprinted in *Notices Am. Math. Soc.*, 54(2):223–229, 2007 and *Mitt. Dtsch. Math.-Ver.*, 14(4):206–212, 2006.

[49] ———. Interview with Peter D. Lax. *Eur. Math. Soc. Newsl.*, 57:24–31, 2005. Reprinted in *Notices Am. Math. Soc.*, 53(2):223–229, 2006.

[50] ———. Interview with Michael Atiyah and Isadore Singer. *Eur. Math. Soc. Newsl.*, 53:24–30, 2004. Reprinted in *Notices Am. Math. Soc.*, 52(2):225–233, 2005 and *Mitt. Dtsch. Math. Ver.*, 12(4): 272–281, 2004.

[51] ———. Interview with Jean-Pierre Serre. *Eur. Math. Soc. Newsl.*, 49:18–20, 2003. Reprinted in *Notices Am. Math. Soc.*, 51(2):210–214, 2004 and Nieuw Archief Wiskunde 5(1): 38–41, 2004.

Moreover, interviews with Danish mathematicians Ebbe Thue Poulsen, Bent Fuglede, Tobias Colding and Ib Madsen and a few articles were published in the journal *Matilde* of the Danish Mathematical Society (in Danish).

Interviews with the Abel prize recipients were broadcasted by the 2nd chain of Norwegian TV (kunnskabpskanalen) and are archived on the Abel Prize website.

CHAPTER 1

Introduction

1.1. Motivations. Background

1.1.1. A personal report. During my education as a mathematician, I was primarily trained within differential and algebraic topology. This is clearly visible from the older entries in my list of publications: Until the middle of the 1990s, my research was focused on various aspects of algebraic topology, often on problems concerning group actions on manifolds, some of them quite technical at the end of the day.

Being based at Aalborg University with an emphasis on engineering and applied sciences, I felt after all quite alone – even after another topologist, Dr. Lisbeth Fajstrup, had been appointed in 1992. This feeling and a combination of encouragement and pressure from leading people at the department led to a look for alternative research directions, not too far from our experiences. Such an alternative became apparent when the two of us participated in a weeklong workshop at the Isaac Newton Institute in Cambridge in late 1995 under the title *New Connections between Mathematics and Computer Science*, cf Gunawardena [**Gun96**].

This workshop was held in a very nice, open and fruitful atmosphere with a variety of stimulating talks, including very famous speakers like M. Gromov and S. Smale. But there were other talks that turned out to be more decisive for our work. Apart from lectures by John Baez (*n-categories in logic, topology and physics*), Yves Lafont (*Homological methods and word problems*), I would like to emphasize in particular

- Éric Goubault, *Scheduling problems and homotopy theory*
- Sergio Rajsbaum, *On the decidability of a distributed decision task*

that introduced us for the first time to the possibility of applying topological methods for purposes in concurrency theory and in distributed systems theory.

1.1.2. Concurrency and distributed computing. To say it very briefly[1], *concurrency* in computer science means a property of systems in which several computations are executing simultaneously, and potentially interacting with each other. Concurrent systems open up for faster algorithms, but the mumber of possible execution paths (schedules) in such a system can be extremely large and the resulting outcome may be indeterminate. Methods are sought to identify schedules that do produce results (not ending in a deadlock) and that lead (by construction) to correct – or at least tolerable – results.

[1]A comprehensive survey over a wealth of models in concurrency is given in a chapter of the Handbook of Logic in Computer Science by Winskel and Nielsen [**WN95**].

Distributed systems, studied in *distributed computing* consist of multiple autonomous computers communicating through a computer network in order to achieve a common goal. Typically, a problem is divided into many tasks, each of which is solved by one or more computers, communicating with each other by a variety of protocols. What sorts of problems can be solved in distributed architectures – possibly assuming that a number of participating computers may fail to work, without notice to the others? An intriguing introduction to topological methods had at the time just appeared (cf Herlihy and Rajsbaum [**HR95**], in particular; moreover Herlihy and Shavit [**HS99**]).

The two talks mentioned above advocated that these two disciplines, certainly related to each other, but with slightly different goals to achieve, may benefit from a perspective from combinatorial/algebraic topology and showed indications and some results in that direction. These prospects made an impression on us. We began to think, in particular, about how to detect *deadlocks* in semaphore models, cf Chapter 2 of this thesis. We made soon personal connections with Éric Goubault; at the time employed at the ENS in Paris, now a professor at CEA Saclay and at Ecole Polytechnique. This encounter started a very fruitful collaboration that has been ongoing ever since.

It turned out that we would try to take inspiration from methods in combinatorial and algebraic topology that we knew; but it was not possible to apply those directly, before "twisting" them. That twist consisted in taking *directedness* properties serious. No longer are all continuous paths allowed, only *directed* paths, reflecting that the time flow in the execution of a schedule is not reversible. This fact makes it more difficult to exhibit suitable algebraic topological invariants describing phenomena of interest. At least, one has to get involved with *categories* instead of groups. These were the first indications for a need for a systematic investigation of methods for and properties of *Directed Algebraic Topology*. It should be mentioned in passing that other topological methods (order topologies etc) had been applied previously in Computer Science, in particular in *domain theory*.

1.2. Collaboration

While algebraic topology for a long time had the reputation of an exclusive and very pure mathematical discipline, more and more areas of application have popped up during the last fifteen to twenty years, cf also Chapter 8.

1.2.1. Workshops, conferences, networks. One of the first intitiatives to collect researchers with an interest in applying methods from algebraic topology to problems in Computer Science after the conference at the Newton Institute in 1995, was our own series of modest workshops called GETCO (Geometric and Topological Methods in Computer Science): the first of those was held in 1999 at Aalborg University and followed up by a series of similar workshops lasting between a day and a week; several times attached to conferences of the CONCUR or DISC communities in Computer Science.

At a much larger scale, I would like to mention the very inspiring conference series ATMCS (Algebraic Topological Methods in Computer Science), that has been organized five times, lastly in 2012 under the title Applied and computational topology. Moreover workshop series at Schloss Dagstuhl, Germany and dedicated conferences at MSRI, Berkeley, USA, Oberwolfach, Germany, the Fields Institute, Toronto, Canada, and BIRS, Banff, Canada.

On the European arena, collaboration on applied aspects of algebraic topology gained force in the recent Research Networking Programme ACAT – Applied and Computational Algebraic Topology – in the framework of the European Science Foundation. This programme that lasts from 2011 to 2015, has some funds to support conferences, workshops and summer schools within the field; moreover it gives grants to visits between collaborating partners. I am the chair of the steering committee of that network.

1.2.2. Acknowledgements. It would have been impossible to achieve substantial progress without a network of people who have been interested in the research line taken and with whom I have had the pleasure to collaborate for a while or also on an almost permanent basis. It is impossible to mention them all, but I need to give special thanks to

- Lisbeth Fajstrup, colleague at the Department of Mathematical Sciences at Aalborg University, a long term partner and a coauthor in this endeavour; giving inspiration and – very important – ready to listen whatever I had on the agenda
- Éric Goubault, CEA/Saclay and Ecole Polytechnique, France, who would explain many times to us many of the Computer Science aspects of our work; moreover an important source of inspiration and coauthor
- his colleagues Emmanuel Haucourt and Samuel Mimram who joined a little later
- my Ph.D.-students Ulrich Fahrenberg, Rafael Wisniewski and John-Josef Leth for inspiring discussion during an extended period
- Maurice Herlihy who inspired us many times with his well-planned and funny talks on decision problems in distributed computing and who was my host during a three weeks visit at the Computer Science Department at Brown University, Providence, RI, USA, in 2000
- Marco Grandis (Genova University) for deep interest in the subject and for writing the first book Grandis **[Gra09]** on Directed Algebraic Topology
- Krisztof Ziemiański (Warsaw University) for recent collaboration.

Thanks are also due to the referees of my work in the area who have often given me indications how to make my drafts more consistent and/or more readable.

1.3. Not more than a survey!

This thesis can only give a quick guide through the material described in much more detail in the articles that are submitted together with it (cf the list of thesis papers right in the beginning). It is written in retrospect; only some highlights are dealt with and deviations from the main route are deliberately kept obscure.

Most of the work is not technically deep or sophisticated. Beginning with a new research interest meant that many concepts had to be developed if not from scratch, then from only a few basic definitions. The only exception is the development of combinatorial models for trace spaces, cf Chapter 7, that nevertheless, from a mathematical perspective, stands on the shoulders of well-developed techniques.

For proofs of results, the reader is referred to the original papers. Only a few particularly important proofs that have a particular impact on this story have been detailed in this thesis. Emphasis is put on more recent work – that of course uses insights from previous articles. The selection of topics has mainly been made from a *mathematical* perspective.

CHAPTER 2

Topology and order for semaphore models

2.1. Semaphore models

2.1.1. History. First notions and examples. The following notes on the (pre)-history of the subject are mainly drawn from Goubault [**Gou00**]; they describe the initial motivation for our and for related work.

An option for scheduling the access of several processes to shared resources is by *semaphores*: Each resource is provided with a semaphore. A schedule has to obey to the rule that, at any given time, only one (in the case of mutual exclusion) or at most a fixed number k of processes – called the *arity* of that resource – can acquire a lock to a given resource. It has to relinquish the lock having finished working with the resource. It is well-known (even in daily life) that "bad" schedules can lead to deadlock states from which the combined execution has no way to proceed.

Semaphore methods can be given a description that has an inherently geometric flavour: The so-called "progress graph" was first introduced in the literature in Shoshani and Coffman [**SC70**] and Coffman etal. [**CES71**]. The famous Dutch computer scientist Edsger W. Dijkstra [**Dij68**] had given an abstract semantics for handling mutual exclusion: Each (deterministic sequential) process Q_i gives rise to a sequence $R^1a_i^1 \dots R^{n_i}a_i^{n_i}$ with $R^j = P, V$ and a_j one of the shared objects; P (*prolaag* in Dutch; procure?) means acquiring a lock, V (*verhogen* in Dutch; vacate?)[1] means relinquishing it again.

Simple test examples can be found in the early literature [**SC70**] and [**CES71**]; they are attributed there to Dijkstra. The following two examples must suffice here: The first (the "*Swiss flag*") consists of two processes $T_1 = Pa.Pb.Vb.Va$ and $T_2 = Pb.Pa.Va.Vb$ competing for resources a and b and gives rise to the two dimensional progress graph of Figure 1.

The second (of which the first is a special case) is known under the name *Dining Philosophers*: Here n philosophers $T_1, \dots, T_n$ at a round dining table compete for n resources (forks) $a_1, \dots, a_n$ according to the schedules $T_i = Pa_i.Pa_{i+1}.Va_i.Va_{i+1}$ (with $a_{n+1} = a_1$) giving rise to an n-dimensional progress graph. It is obvious that trouble (a deadlock) arises when all philophers start to pick up their left forks at the same time.

For a general linear progress graph, each individual process executes linearily on a (time) unit interval $I = [0, 1]$ on which the P and V actions are marked as an ordered sequence. If there are n processes involved, the state space X consists of a hypercube I^n (each point has n coordinates, one for every process) from which a *forbidden region F*

[1]These terminology explanations come from the Wikipedia page on semaphores.

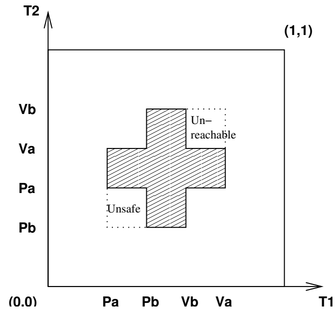


FIGURE 2.1. The Swiss flag as example of a 2D progress graph

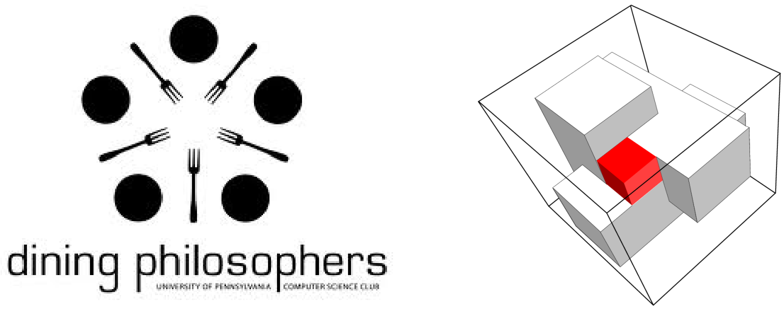


FIGURE 2.2. Left: Dining philosophers. Right: Forbidden and unsafe regions for the three dining philosophers protocol

has been deleted: $X = I^n \setminus F$. The forbidden region consists of points for which more than one (generally more than the arity k) coordinates are situated inbetween a Pa_i and a Va_i with the *same* a_i. It is easy to see that the forbidden region in the semaphore case is a union of (open) higher dimensional *isothetic* rectangles – with facets parallel to the coordinate hyperplanes.

A joint schedule in a progress graph corresponds to a path p in $\mathbf{R}^n$ joining the compound start state, the lowest vertex $\mathbf{0} \in I^n$ to the compound end state, the upmost vertex $\mathbf{1} \in I^n$. The projections of every such path to one of the axes (describing the execution of the corresponding process with the locking instructions on its way) has to be *non-decreasing* since an execution does not run backwards in time. Moreover, such a path may not enter the forbidden region $F \subset I^n$.

The characteristic features of such a *dipath* **[FGR06]** – di for directed – $p : I \to I^n$, are hence that

(1) $p(t)$ avoids the forbidden region F for all t;
(2) all projections $p_i : I \to I$ are non-decreasing.

We will often also fix start and end points.

Deadlocks (originally called "deadly embrace" by Dijkstra) occur in many schedules; these are states (points) from which no *directed* path can proceed without immediately entering the forbidden region F. Furthermore, there may occur "unsafe regions" (a dipath entering the unsafe region cannot reach a final state without entering the forbidden region, cf. Figure 2.1 and Figure 2.2) and "unreachable regions" (that no dipath starting from the compound start state can ever enter). Deadlocks, unsafe regions and unreachable regions had, to a certain extent, already been analysed and described in this context (including interesting test examples) in many articles, see eg Lipski and Papadimitiou [**LP81**] or Carson and Reyolds [**CR87**].

If the processes (on the axes of the progress graph) run without or with restricted coordination, many possible schedules - and in a progress graph even infinitely many – will arise as dipaths. The following is an important insight indicating that topology might have a role to play: Two schedules will yield the same result (of a joint calculation or whatever) if the two respective dipaths can be connected to each other by a *one-parameter deformation* of *di*paths (avoiding the forbidden region along that deformation); the reason for this will be explained in detail in Section 3.2. In topology, one-parameter deformations are called homotopies; we call one-parameter deformations that respect directedness *dihomotopies*. Loosely speaking, dihomotopic dipaths aquire their locks and relinquish them in the same order, at least for semaphores of arity one. In the paper Fajstrup etal. [**FGR06**], an example of a simple PV-program is given (with simple calculation steps between accesses to pieces of shared memory) showing that results in general will be different for dipaths that are *not* dihomotopic. The schedules in this example are in fact two dipaths circumvening the forbidden region in Figure 1 in two different senses.

2.1.2. Discrete versus continuous. This gives room for speculation. At first sight, it may seem strange to replace a large discrete state space (as is common in concurrency theory, e.g., a graph, that is contained in the product of the directed graphs describing the actions for each of the processes) by an *infinite* state space. That directed graph can be considered as the 1-skeleton of a subdivided progress graph (cf. Section 3.1.4). Higher-dimensional rectangles in the progress graph describe additional information, i.e., higher coordination (independence relations) between actions of the individual processes.

For realistic examples, the state space is built either as a discrete product of general *directed graphs* (with branchings, mergings and loops; not only linear graphs) or as the topological product of their *geometric realizations*, in both cases after deleting a forbidden region. This situation is more complicated than the linear one and will be dealt with later.

Anyway, when the number of states and/or the number of processes increases, a discrete product will suffer from (combinatorial) state space explosion. Moreover it is quite difficult to determine which of the directed paths (even much bigger in number) in such a product graph are equivalent to each other and which not.

Deformations (homotopies) are well-studied in algebraic topology, and we attempt to use methods from this area to reason and to do calculations concerning the space of directed paths (up to homotopy) in the model. The "detour" through the continuous models seemingly allows a quicker (or more comprehensible) path to a determination of equivalence classes of execution paths.

2.1.3. Example of an application: data base theory. A nice example where topological reasoning gives rise to insights was originally described by J. Gunawardena in [**Gun94**] and later made more rigorous in Fajstrup etal. [**FGR06**]: In data base engineering, one uses often *2-phase locked protocols* with the following defining property: Each PV-protocol for each of the processes needs to acquire *all* locks before these are *all* relinquished (possibly in a different order):$P \dots P.V \dots V$.

It can then be shown that every schedule using this strategy is *serializable*, i.e., equivalent to a serial schedule characterized by the property: One process at a time! The schedule is a concatenation of the schedules of the individual processes (in some order); no interleaving takes place.

This is important since the results of serial schedules can be checked and understood quite easily. The geometric/topological picture that corresponds to a 2-phase locked protocol yields a progress graph in which the forbidden region has a *center region* from which one may deform all dipaths inductively to dipaths on the *1-skeleton* of the hypercube I^n. One has to be careful to make sure that the deformation is through *di*paths at any time – this is the contribution of Fajstrup etal. [**FGR06**]. But still the argument given (and certainly the intuition behind it) seems to be far easier to comprehend than arguments of a merely discrete type.

2.2. Detection of deadlocks and of unsafe regions

First joint work describing how to find deadlocks and unsafe directions in a progress graph can be found in the preprint Fajstrup and Raussen [**FR96**]. This work was completed in collaboration with Éric Goubault (Fajstrup etal. [**FGR98b**, **FGR98a**]) with the description of a running implementation of the detection algorithm.

2.2.1. A combination of combinatorial and geometric insight leads to an algorithm. How can one find *deadlock states* in the n-dimensional state space $X = I^n \setminus F$ with the combinatorial input given by the hyperrectangles R^i making up the forbidden region F? Those hyperrectangles are products of intervals $R^i = \prod_{j=1}^{m}]a^i_j, b^i_j[\subset I^n$ (between a P-action at a^i_j of the ith process and the corresponding V-action at b^i_j). If n such hyperrectangles intersect (generically), then the lowest corner of the intersection hyperrectangle is a deadlock state! Whichever process progresses, it will have to enter one of the hyperrectangles. This corner point is coordinatewise given by the *maximal* coordinates of the lower corners a^i_j of the contributing hyperrectangles.

A simple illustrating example consists of the walls and the ceiling of a rectangular room making up parts of the forbidden region: No way from the lower corner of the intersection! The corners obtained in this way are the only deadlock points in the interior of I^n. Deadlocks on the boundary of I^n can be found by the same mechanism after having extended the given hyperrectangles touching the boundary and added hyperrectangles representing the boundary; for details see Fajstrup etal. [**FGR98b**, **FGR98a**].

It is even more important to describe the *unsafe regions* that schedules had better avoid. It turned out, that these can be determined step by step using similar ideas. For the first step, note that *below* a deadlock state, there is a hyperrectangle whose lower corner has as coordinates the *second largest* among all the lower corner coordinates a^i_j of the participating hyperrectangles. It is easy to convince yourself that dipaths in this hyperrectangle cannot escape without entering the forbidden region.

In subsequent steps, one adds the unsafe hyperrectangles found so far to the forbidden region; this may then give rise to *additional* unsafe hyperrectangles. It is not difficult to see (and it is shown in Fajstrup etal. [**FGR98a**]) that the algorithm thus described ends after finitely many steps with a complete description of the unsafe regions.

Although the algorithm can be described entirely in discrete terms without any use of topological machinery, it is certain that we would not have found it (and it would be hard to explain it) without geometric thinking and intuition.

2.2.2. Implementation issues. The algorithm for detection of deadlocks and unsafe regions was implemented by Éric Goubault as a C-program. It had essentially to manipulate intersections and unions of isothetic n-rectangles from a given list and to implement the effect of adding an additional n-rectangle. In a first step, the initial list of forbidden n-rectangles is established, the second step works out an array of intersection rectangles (including as a special case those with deadlocks at corner points), and the third adds pieces of the unsafe regions, recursively.

The total complexity of the algorithm reflects the geometric complexity of the forbidden region, i.e., the number of intersections of n-rectangles in the forbidden region. The latter express the degree of synchronization of the processes. In most test cases, the implementation gave very competitive results compared to other approaches; in particular, in high dimensions, with many participating processes; some details are given in Fajstrup etal. [**FGR98a**].

These first tools have since been extended into a multi-purpose abstract interpretation based static analyzer ALCOOL by our partners at CEA/LIST, cf eg Goubault and Haucourt [**GH05**] and Fajstrup etal [**FGH$^+$12**].

2.3. Outlook and discussion

As mentioned earlier, the state space X of aconcurrent program will, in general, be modeled as (topological) product of directed graphs from which a subcomplex corresponding to the forbidden region F has to be deleted: $X = (\Gamma_1 \times \cdots \times \Gamma_n) \setminus F$. The individual graphs and thus also the state space X may contain *directed loops*. Still the forbidden region can be understood as a union of generalized hyperrectangles (with identifications on the boundary, giving rise to cylinders or tori). The detection of deadlock points as corners of intersections of n-rectangles is essentially unchanged, but the detection of unsafe points is more tricky: States that seem to be unsafe might be able to escape after several "rounds". The effects of this delooping have been described and investigated in Fajstrup [**Faj00**] and Fajstrup and Sokołowski [**FS00**]. They were the original motivation for extending the theory of coverings from algebraic topology to so-called dicoverings, cf Fajstrup [**Faj03**] which turned out to be much more sophisticated.

In another direction, it turned out that the description of deadlocks and unsafe (and, in an analogous manner, also of unreachable) regions is a helpful step in the classification of dipaths in progress graphs up to dihomotopy in later work (cf Raussen [**Rau00**, **Rau06**, **Rau10**]); this will be taken up in subsequent sections.

CHAPTER 3

Directed spaces and directed homotopy

3.1. Directed spaces

3.1.1. Partially ordered spaces. Apparently the first attempt to combine order and topology in a systematic way can be found in Nachbin's monography [**Nac65**]. The following definition *connecting topology and order* is particularly important:

DEFINITION 3.1.1. Let X denote a topological space and let $\leq\subset X\times X$ denote a partial order (reflexive, transitive and anti-symmetric) on X. The pair $(X,\leq)$ is called a *partially ordered space* (or *po-space* for short) if $\leq$ is a *closed* subset of $X\times X$.

In particular, if x_n, y_n denote sequences in X with $x_n \leq y_n, n\in \mathbf{N}$, with limits $x = \lim x_n, y=\lim y_n$, then $x\leq y$. For example, the standard (coordinatewise) partial order $\leq$ on $\mathbf{R}^n$ makes $(\mathbf{R}^n,\leq)$ a po-space. A closed subspace (like the Swiss flag example from Figure 1) inherits a po-structure.

Another stimulating monography involving order notions is Penrose's [**Pen72**] that deals with questions in relativity theory from the viewpoint of differential geometry. A space-time is seen as a 4=(3+1)-dimensional manifold with a Lorentzian metric of index 1. In particular, the tangent bundle contains a "bundle of cones" consisting of *causal* resp. *time-like* tangent vectors. One studies then properties of *causal* and of *time-like* curves (with causal, resp. time-like tangent vectors all along) in combination with the differential geometry of the underlying manifold, eg with the aim to investigate properties of black holes. Several notions from Penrose's book [**Pen72**] have been useful in connection with our work although they often had to be modified in order to fit for our purposes.

In both cases mentioned above, one considers most often not *all* curves in the topological space; only curves that play together well with order properties are relevant. In a *po-space*, a directed path $p: I=[0,1]\to X$ has to preserve orders, i.e., has to satisfy:

$$t_1\leq t_2 \Rightarrow p(t_1)\leq p(t_2). \tag{3.1}$$

As soon as one considers spaces that contain interesting (from the point of view of order) *loops*, po-spaces are too rigid: Property (3.1) cannot be satisfied for any partial order $\leq$ along a non-constant loop p (with $p(0)=p(1)$) since one obtains for any $x=p(t)\neq p(0)$: $p(0)\leq x\leq p(1)=p(0)$ contradicting anti-symmetry. On the other hand, state spaces with loops occur naturally in applications, since most relevant programs contain loops.

There are several reasonable approaches to widen the definition of a po-space:

3.1.2. Lpo-spaces. The one originally taken by us was an approach using charts (compare the definition of manifolds) as in the following definition:

DEFINITION 3.1.2. (Fajstrup etal. [**FGR06**])

(1) A covering $\mathcal{U}$ of a topological space X by open sets U with partial orders $\leq_U$ is an atlas of an *lpo-space* (locally partially ordered) if, for every $x \in X$, there is a non-empty neighbourhood $W(x)$ such that
$$y \leq_U z \Leftrightarrow y \leq_V z \text{ for all } U, V \in \mathcal{U}, y, z \in U \cap V \cap W(x).$$
(2) Two atlas $\mathcal{U}$ and $\mathcal{V}$ define equivalent lpo-structures if their union $\mathcal{U} \cup \mathcal{V}$ defines an lpo-structure in X.

For example, (counter-clockwise) rotations cannot be used to give the circle $X = S^1$ a po-structure, but they give rise to a perfectly well-defined lpo-structure. Moreover, one may ask a *dimap* to preserve partial local orders (cf Fajstrup etal. [**FGR06**]) and in particular investigate *dipaths* and *diloops* in lpo-spaces as dimaps with domain the ordered interval $\vec{I}$ and the lpo-space $\vec{S}^1$.

Directed paths are the essential object of our study – they correspond to the execution paths in the model; the (partial) order on the state space is secondary. This makes the study of lpo-spaces by themselves a doubtful goal. In fact, for a space with a partial order, one may define a (coarser) partial order by:
$$x \preceq y \Leftrightarrow \exists \text{ a dipath } p : I \to X \text{ such that } x = p(0), y = p(1).$$
Then, x and y are no longer related unless they are so via a directed path. For example, the region under the Swiss flag (Figure 1) becomes then completely unrelated to the region above it. On the other hand, it is not always clear that the new partial order $\preceq$ thus obtained is closed.

A more serious drawback of lpo-spaces arises from categorial considerations. An investigation of (finite) limits and colimits shows that these need not always exist; if they exist they need not be the limits or colimits that arise in the category of topological spaces (under the forgetful functor). This is maybe not that surprising; it parallels the bad behaviour of the category of manifolds under limits and colimits. There have been attempts to reconcile lpo-spaces with model categories (cf Bubenik and Worytkiewicz [**BW06**]); but those seemingly remained without further practical applications.

3.1.3. Streams. These categorical problems have been overcome by introducing the more flexible *streams* in the work of Krishnan [**Kri09**]. Roughly speaking, a stream is a topological space with consistent *pre*orders on the open sets, a so-called *circulation*. Consistency means that the preorder on an open set is equal to the transitive closure of the preorders on its open subsets. It is shown in Krishnan [**Kri09**] that streams and stream maps (preserving the local preorders) form a complete and cocomplete category.

Lpo-spaces are more special than streams: Local anti-symmetry is *not* required for streams. In particular, streams may have vortices, ie they may allow for arbitrarily small directed loops.

3.1.4. Cubical complexes. Higher-Dimensional Automata. Higher-Dimensional Automata (HDA) were originally introduced and studied by Pratt [**Pra91**] and van Glabbeek [**vG91**] as combinatorial models extending the progress graphs from Section 2.1.1. Roughly speaking, the concurrent parallel execution of one step taken by each of n individual processes is modelled by an n-box or n-cube $\square_n$ if it is independent of orderings among the processes (and perhaps also subdivisions). An n-cube can thus be seen as the state space for all possible interleavings of the $n!$ directed paths on its one-skeleton. The presence of such a cube in a complex indicates that the order of the partial executions on the 1-skeleton (or even of partial interleavings) is insignificant. On the other hand, if nothing is known about such independence relations, only the 1-skeleton of such a box appears in the HDA.

Knowledge about *partial independence* can be encoded by considering a *subcomplex* of the n-box containg its 1-skeleton. A subcube $\square_k \subset \square_n$ is included if k-processes execute independently as long as the others have come to a halt. The partial ordering on an k-box (equivalent to $\vec{I}^k$) yields a natural directed path structure.

Several such (sub-)boxes (perhaps of varying dimensions; the number of participating processes and the nature of independence relations may vary) can be glued together to yield a Higher-Dimensional Automaton (HDA). Usually, at least the 1-skeleton of such an HDA is equipped with labels, as a generaliation of transition systems in classical concurrency theory. In the following, we shall abstract away from the use of such labels; this is in a sense justified by results of Srba [**Srb01**] showing – for transition systems – that a labelled transition system can be replaced by an equivalent unlabelled one without loosing expressivity. Let us also mention that van Glabbeek [**vG06**] later showed that Higher Dimensional Automata compete favourably with other widely used models for concurrent computing, eg Petri nets.

The definition of an HDA and discussions about their properties came closer to well-known mathematics when Éric Goubault found out that the underlying combinatorics and topology is that of a pre-cubical set (also called $\square$-set, cf. Fajstrup [**Faj05**], in analogy with the term Δ-set from Rourke and Sanderson [**RS71**] for a simplicial set without degeneracies).

Pre-cubical sets had been previously investigated in detail by Brown and Higgins in [**BH81a**, **BH81b**]. Cubical complexes are already present underlying cubical singular homology in Serre's thesis [**Ser51**] from 1951. For modern homotopy theory references involving cubical sets see Jardine [**Jar02**] and Grandis and Mauri [**GM03**].

In the following section, we shall use $\square_n$ as an abbreviation for the n-cube $I^n = [0,1]^n$ with the product topology.

DEFINITION 3.1.3. (1) A $\square$-set or pre-cubical set M is a family of disjoint sets $\{M_n | n \geq 0\}$ with face maps $\partial_i^k : M_n \to M_{n-1}, \quad n > 0, \quad 1 \leq i \leq n, \quad k = 0, 1,$ satisfying the pre-cubical relations $\partial_i^k \partial_j^l = \partial_{j-1}^l \partial_i^k$ for $i < j$.

(2) A pre-cubical set M is called *non-self-linked* (cf Fajstrup etal. [**FGR06**]) if, for all n, $x \in M_n$ and $0 < i \leq n$, the $2^i \binom{n}{i}$ iterated faces
$\partial_{l_1}^{k_1} \cdots \partial_{l_i}^{k_i} x \in M_{n-i}, \quad k_i = 0, 1, \quad 1 \leq l_1 < \cdots < l_i \leq n,$ are all different.

(3) The geometric realization $|M|$ of a pre-cubical set M is given as the quotient space $|M| = (\coprod_n M_n \times \Box_n)_{/\equiv}$ (a cubical complex) under the equivalence relation induced from

$$(\partial_i^k(x), t) \equiv (x, \delta_i^k(t)), \quad x \in M_{n+1}, \quad t = (t_1, \dots t_n) \in \Box_n,\ 1 \le i \le n,\ k = 0,1$$
$$\text{with} \quad \delta_i^k(t) = (t_1, \dots t_{i-1}, k, t_{i+1}, \dots, t_n).$$

In a *non-self-linked* pre-cubical set, the map $\Box_n \simeq \Box_n \times e \to |M|$ is *injective* for every n-cell $e \in M_n$. In particular, every element $m \in |M|$ in the image of this map has *uniquely* determined coordinates in $\Box_n$, cf Fajstrup etal **[FGR06]**. Moreover, every element $x \in |M|$ has a unique *carrier* cell $e(x) \in M_n, n \ge 0$, such that x comes from an element in the *interior* $\Box_n^o$ under the restriction of the quotient map to $\Box_n \times e(x)$.

We will make use of particular open sets in $|M|$, the *open stars* of *vertices* in M_0. The open star $St(x, M)$ of $x \in M_0$ consists of the union of the interiors of all cells of which x is a vertex. It was shown in Fajstrup etal **[FGR06]**, that every such open star inherits a *consistent* partial order from the partial orders on individual cells given by their identification with $\Box_n \subset \mathbf{R}^n$ and hence that

PROPOSITION 3.1.4. *(*Fajstrup etal. [**FGR06**, Theorem 6.23]*) The po-structure on the cells of a $\Box$-set extends to an lpo-structure on its geometric realization if M is non-selflinked.*

In particular, we know which paths in $|M|$ are to be considered as *directed*; for a down-to-earth description cf Raussen **[Rau09b]**. The geometric realizations of these pre-cubical sets are arguably the most important class of examples of lpo-spaces for applications. They are not as special as they might look at a first glance: Fajstrup **[Faj06]** showed that every triangulable space can be realized as a cubical complex. If this cubical complex is free of immersed cubical Möbius bands, then there are consistent choices of directions; if this is not the case, one subdivision suffices to establish a compatible local partial order.

3.1.5. D-paths, d-spaces, d-$\mathcal{TOP}$. A very general approach to directed spaces that plays together well with established techniques in homotopy theory was suggested by Marco Grandis. He launched in Grandis **[Gra03a]** the idea to take the directed paths (d-paths) as *defining* element of the structure of d-spaces:

As customary, the *concatenation* of two paths $p, q : I \to X$ in a topological space X is defined by $(p * q)(t) = \begin{cases} p(t) & t \le 0.5 \\ q(2t-1) & t \ge 0.5 \end{cases}$.

DEFINITION 3.1.5. (Grandis **[Gra03a]**) Let X denote a topological space and let $\vec{P}(X) \subset X^I := \{p : I \to X \mid p \text{ continuous}\}$ – the subset of *d-paths*. The pair $(X, \vec{P}(X))$ is called a *d-space* if

- $\vec{P}(X)$ contains every *constant* path $p_x(t) = x, t \in I;\ x \in I$;
- The concatenation of two d-paths is a d-path: $p, q \in \vec{P}(X) \Rightarrow p * q \in \vec{P}(X)$;
- $p \in \vec{P}(X), \alpha \in I^I$ a *non-decreasing* reparametrization $\Rightarrow p \circ \alpha \in \vec{P}(X)$.

Remark that only *non-decreasing* reparametrizations are part of the structure. Thus, in general the reverse $\bar{p}$ of a d-path p given by $\bar{p}(t) = p(1-t)$ is not a d-path. On the other hand, a sub-d-path of a d-path is d again. Note two extreme cases:

- $\vec{P}(X)$ consists solely of constant paths.
- $\vec{P}(X) = X^I$ consists of *all* paths.

While the di-paths in an lpo-space X provide it with a d-space structure, the last example above shows that not every d-space arises from an lpo-space.

In particular, antisymmetry in an lpo-space forbids the existence of small (non-constant) loops; but vortices can perfectly arise in a d-space. For example, the d-paths in the plane might consist of paths rotating counterclockwise around the origin. For homotopy theory purposes, it is an advantage that one can give not only the *cylinder* but also the *cone* of a d-space the structure of a d-space; cf. the work of Grandis in **[Gra03a, Gra02, Gra09]**.

A continuous map $f : X \to Y$ is called a d-*map* if it preserves d-paths, i.e., if $f(\vec{P}(X)) \subseteq \vec{P}(Y)$. We can consider the subcategory d-$\mathcal{TOP} \subset \mathcal{TOP}$ (with d-maps as the morphisms) of the category of topological spaces. It is not difficult to see that the category d-$\mathcal{TOP}$ thus arising has all limits and colimits, cf Grandis **[Gra09]**; in particular, there is an obvious notion of product of d-spaces.

A correspondence between the category of streams and that of d-spaces has been described and investigated by Ziemiański **[Zie12a]**. They are related by adjoint functors leading to isomorphisms of categories of "good" streams and of "good" d-spaces. These are very close to the saturated directed spaces of Hirschowitz etal **[HHH13]**.

3.2. Dihomotopy and d-homotopy

3.2.1. Definitions. It was mentioned in the introduction, Section 2.1.1, that two directed execution paths that can be deformed into each other along a one parameter deformation will yield the same result; a combinatorial explanation will be given in Section 3.2.2 below. It needs an adapted notion of *homotopy* to seriously involve methods of algebraic topology with directed spaces.

To this end, we need first to describe two directed intervals, both with $I = [0, 1]$ as the underlying topological space:

I**:** with $\vec{P}(I)$ consisting of the constant paths only;
$\vec{I}$**:** with $\vec{P}(\vec{I})$ consisting of all *non-decreasing* continuous paths $\varphi : I \to I$.

D- and dihomotopies, cf Grandis and Fajstrup etal **[Gra03a, FGR06]**, from a d-space X to a d-space Y are homotopies that preserve certain d-space structures:

DEFINITION 3.2.1. (1) A *dihomotopy* is a d-map $H : X \times I \to Y$. The d-maps H_0, H_1 are called *dihomotopic*: $H_0 \simeq H_1$.
(2) A *d-homotopy* is a d-map $H : X \times \vec{I} \to Y$ establishing a relation $H_0 \preceq H_1$ (from H_0 to H_1).
(3) Two d-maps $f, g : X \to Y$ are called *d-homotopic* if there is a finite sequence $f \preceq f_1 \succeq f_2 \preceq f_3 \ldots f_n = g$ of "zig-zag" d-homotopies connecting them.

Obiously, d-homotopy implies dihomotopy.

There are also pointed and relative versions of these definitions. A particularly relevant case concerns directed homotopy of directed paths ($X = I$) with given end points.

3.2.2. Motivation. Here is how *directed homotopy* relates to the combinatorial framework of HDAs: The 1-skeleton of an HDA can be viewed as a (directed) graph. Directed paths on the 1-skeleton model executions that are locally sequential. Two paths p, q on the 1-skeleton with the same end-points are *elementarily equivalent* if there are decompositions $p = p_- * p_0 * p_+, q = p_- * q_0 * p_+$ (same prefix and same postfix) and p_0, q_0 are directed paths (with the same end points) on the 1-skeleton of the same cube within the HDA[1]; this defines a reflexive and symmetric relation on all such directed paths. By construction of the HDA (cf Section 3.2.1), the paths p_0 and q_0 will always yield the same result for a computation along the respective schedules; hence, so do p and q.

Combinatorial dihomotopy is the equivalence relation obtained by taking the transitive closure of elementary equivalence defined above. It is a direct consequence that executions (computations) that correspond to combinatorially dihomotopic directed paths on the 1-skeleton (and also what may arise from local interleavings) will yield the *same* result; non-dihomotopic paths *may* have different results.

General d-homotopy and dihomotopy are infinitesimal versions of combinatorial homotopy of directed paths (in a subdivided model space). For this motivation, it does not play a role which of the two relations one considers: Any two paths p, q on a k-cube from the bottom to the top are both d-homotopic and dihomotopic: Consider the path $p \vee q : I \to \square_k,\ (p \vee q)(t) = p(t) \vee q(t)$. Here $\vee$ denotes the componentwise maximum. Both p and q are connected to $p \vee q$ by a linear d-homotopy: $p, q \preceq p \vee q$, and hence $p \preceq p \vee q \succeq q$.

3.2.3. Dihomotopy versus homotopy of directed paths. Two directed paths that are dihomotopic (d-homotopic) are certainly homotopic - forgetting about order(s) on the deformation interval. The reverse is in general *not* true. Simple examples are the "room with three barriers", cf Fajstrup et al. **[FGR06]**, or a 3D-progress graph in the form of a cube from which 4 hyperrectangles (forming pairs of corner wedges) have been removed, cf Raussen **[Rau06]** and Figure 3.1.

In both cases, it is easy to convince yourself that certain d-paths are homotopic but not dihomotopic to each other. A formal proof is more difficult; it can best be achieved by abstract methods discussed in Chapter 7.

REMARK 3.2.2. For HDA (cubical complexes) *of dimension 2*, it was shown in Raussen **[Rau10]** that homotopic d-paths are dihomotopic. Hence 2-dimensional complexes are very special and not very useful as background for intuition about results on HDA of higher dimensions.

---

[1] Just for this purpose, one might restrict to 2-dimensional complexes, or to the 2-skeleton of the complex. But that restriction would often make algorithms less efficient!

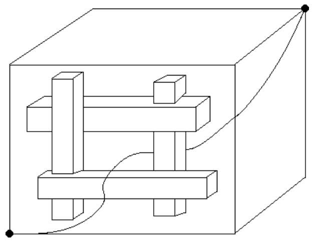


FIGURE 3.1. A d-path that is homotopic but not dihomotopic to a d-path on the boundary

Another essential difference between dihomotopy and homotopy concerns *cancellation*: A homotopy $p_1 * p_2 \simeq p_1 * p_3$ of paths rel end point implies the existence of a homotopy $p_2 \simeq p_3$. A similar statement is *not* true for dihomotopy. An easy counterexample is found on the 2D-complex that is the boundary of a 3D cube with the lower facet removed. D-paths from bottom to top are all dihomotopic, whereas there are non-dihomotopic paths on the boundary of the removed facet.

3.2.4. Dihomotopy versus d-homotopy. Of course, d-homotopy of d-maps implies dihomotopy. It seems to be an open question in which generality the reverse holds. Fajstrup [**Faj05**] showed that d-paths on the 1-skeleton of a cubical complex that are dihomotopic are necessarily also d-homotopic. In greater generality, a directed version of a simplicial/cubical approximation theorem has recently been proved by Krishnan [**Kri13**].

Dihomotopy does not always imply d-homotopy, not even for directed paths. Here is a simple (but non-cubical) example due to Fajstrup: Consider the directed suspension of an interval I with (only) *constant* d-paths; the d-paths on that suspension from the bottom to the top point are contained in subspaces $\{a\} \times I$ that a d-homotopy (but not a dihomotopy) needs to respect.

3.3. A first invariant: The fundamental category of a d-space

In classical algebraic topology, the space of all paths $P(X)(x_0; x_1)$ in X from x_0 to x_1 is homotopy equivalent to the loop space $\Omega(X)(x_0)$ if just x_0 and x_1 belong to the same path component. The easy construction of a homotopy equivalence makes essential use of the *reverse* of a path connecting x_0 and x_1. A similar statement is wrong for d-paths in directed spaces; one needs to take end points (or more generally, end sets) as essential parts of any algebraic structure capturing directed homotopy.

On the level of directed paths, appropriate structures are given by the *path category* $\vec{P}(X)$ and the *fundamental category* $\vec{\pi}_1(X)$ of a directed space X.

DEFINITION 3.3.1. (Grandis [**Gra03a, Gra09**]) Both the path category and the fundamental category of a directed space X have the elements of X as objects. The morphisms from x_0 to x_1 are given by

$\vec{P}(X)(x_0, x_1)$**:** the space of d-paths from x_0 to x_1 (with the topology inherited from the compact-open topology on the space of all paths)

$\vec{\pi}_1(X)(x_0, x_1)$: the set of dihomotopy (d-homotopy) classes of d-paths from x_0 to x_1.

Composition of morphisms in both categories is given by concatenation (of paths, respectively their di/d-homotopy classes). The path category is enriched in the category $\mathcal{T}op$.

In Chapter 6, we will also consider the trace category $\vec{T}(X)$ of X whose morphism are spaces of traces (reparametrization equivalence classes) of d-paths between given end points (with the quotient topology; also $\vec{T}(X)$ is $\mathcal{T}op$-enriched). There are obvious forgetful functors $\vec{P}(X) \to \vec{T}(X) \to \vec{\pi}_1(X)$.

The most important result at this level – making calculations possible at least in relatively easy cases – is a van Kampen type result for fundamental categories. Note that a d-subspace $X_i \subset X$ has d-paths $\vec{P}(X_i) = \vec{P}(X) \cap P(X_i)$.

THEOREM 3.3.2. (Grandis [**Gra03a**]) *Let X denote a d-space with two d-subspaces X_1 and X_2; let $X_0 = X_1 \cap X_2$. If $X = intX_1 \cup intX_2$, then the following diagram of fundamental categories (induced by inclusions) is a pushout diagram in $\mathcal{C}at$:*

$$\begin{array}{ccc} \vec{\pi}_1(X_0) & \longrightarrow & \vec{\pi}_1(X_1) \\ \downarrow & & \downarrow \\ \vec{\pi}_1(X_2) & \longrightarrow & \vec{\pi}_1(X) \end{array}$$

The proof (in Grandis [**Gra03a**, **Gra09**]) – along the lines of the proof of the classical van Kampen theorem – makes essential use of "zig-zag" d-homotopies. Hence, the morphisms in the fundamental category have to be considered as d-homotopy classes of d-paths.

3.4. Outlook and Discussion

It depends very much on the aim of a study which framework one should choose. If one wishes to study homotopy theoretic properties, then the categories of d-spaces (Grandis) or of streams (Krishnan) are to prefer; the categories of flows (Gaucher), cf Chapter 8, has similar aims.

Cubical complexes have the huge advantage of combining topological, combinatorial and directed structures in a very natural way. This is why concrete results and calculations have mainly been established for such spaces, cf Chapter 7. Simplicial complexes, cf Ziemiański [**Zie12b**] might also become useful. Lpo-spaces filter out some of the general properties of cubical complexes (local antisymmetry and the non-existence of vortices). It is true that they do not form well-behaved categories – but neither do differentiable manifolds!

CHAPTER 4

Localization techniques and components

Given a d-space X. As explained in Chapter 3, it comes equipped with (topological) path spaces $\vec{P}(X)(x_0, x_1)$ and their sets of components $\vec{\pi}_1(X)(x_0, x_1)$ forming the morphisms of the path category, resp. the fundamental category of X. It is relevant to ask how sensitive these path spaces and their components are with respect to *variations of the end points* x_0 and x_1. If homotopy types (or some algebraic invariants) only change at certain thresholds, one may compress the representation of these categories without losing information.

4.1. Weakly invertible morphisms and wide subcategories

4.1.1. Problem and aim: State space explosion problem and components. Discrete models (transition systems etc.) in concurrency theory share the property that the number of discrete states grows very fast with respect to the length (or rather complexity) of each individual program and also with respect to the number of processors. This is known as the *state space explosion problem*. At first sight, this problem does not get any better by replacing discrete state spaces with infinite state spaces, as with Higher Dimensional Automata. On the other hand, the state spaces (in the form of cubical complexes) are well-structured, and hence, one may hope that they can be decomposed in such a way that the homotopy types of path spaces (or at least their components) only change along some very specific "cut-locuses"; this would mean, that the essential information in path or fundamental categories can be compressed to quotient categories of a much smaller size. It was the aim of the series of papers Goubault and Raussen [**GR02**], Raussen [**Rau03**], Fajstrup et al [**FGHR04**], later extended by Goubault and Haucourt [**GH07**], to achieve such a compression of information in the *fundamental category* of a d-space by using a *calculus of fractions*.

4.1.2. Weakly invertible morphisms. To achieve this, one identifies in the fundamental category $\vec{\pi}_1(X)$ a system (in fact a wide subcategory) $\mathcal{W}$ of so-called *weakly invertible* morphisms. A morphism (d-path class) $\sigma \in \vec{\pi}_1(X)(x, y)$ is called *weakly invertible*[1] if (pre and post-) compositions with σ in form of the maps

$$\begin{array}{rcl} \sigma^\sharp : \vec{\pi}_1(X)(y,z) & \to & \vec{\pi}_1(X)(x,z); \\ \alpha & \mapsto & \sigma * \alpha \end{array} \qquad \begin{array}{rcl} \sigma_\sharp : \vec{\pi}_1(X)(v,x) & \to & \vec{\pi}_1(X)(v,y) \\ \beta & \mapsto & \beta * \sigma \end{array}$$

are *bijections* – as soon as the domains are non-empty.

[1] Goubault and Haucourt [**GH07**] use the term Yoneda morphism instead.

REMARK 4.1.1.
(1) The notions discussed apply to general categories, not only to the fundamental category.
(2) Weakly invertible morphisms can be interpreted to be those that do not contribute to a "decision" when concatenated with other morphisms. The aim is to take a quotient with respect to a subcategory of the weakly invertible morphisms, as large as possible.
(3) In Section 5.1, we will formulate other versions of weak invertibility and also generalize the localization techniques below.

4.1.3. lr and pure subcategories. In order to apply localization techniques and to arrive at quotient categories with nice properties, we need further assumptions to a wide subcategory $\Sigma \subseteq \mathcal{W}$:

DEFINITION 4.1.2. (Gabriel and Zisman [**GZ67**], Borceux [**Bor94**]):

(1) Σ satisifies the lr-*extension* properties (admits a left/right calculus of fractions) if every diagram of morphisms in $\vec{\pi}_1(X)$

$$\begin{array}{ccc} x' & \overset{\gamma'}{\dashrightarrow} & y' \\ {\scriptstyle\sigma}\uparrow & & \uparrow{\scriptstyle\sigma'} \\ x & \xrightarrow{\gamma} & y \end{array} \qquad\qquad \begin{array}{ccc} x' & \xrightarrow{\gamma} & y' \\ {\scriptstyle\sigma'}\uparrow & & \uparrow{\scriptstyle\sigma} \\ x & \overset{\gamma'}{\dashrightarrow} & y \end{array}$$

with $\sigma \in \Sigma$ can be filled in with $\sigma' \in \Sigma$ (and γ' in $\vec{\pi}_1(X)$);

(2) Σ is *pure* if it only allows decompositions within Σ, ie $\sigma_1 * \sigma_2 \in \Sigma$ implies $\sigma_1, \sigma_2 \in \Sigma$.

REMARK 4.1.3. Applied to weakly invertible morphisms, the condition in Definition 4.1.2(1) tells you that a "non-decision" at the start point can be reflected by a "non-decision" at the end point and vice versa. Pureness means that a decision cannot be converted to a non-decision by pre-or post-composition.

4.2. Localization, categories of fractions and component categories

4.2.1. The setup. The general idea, initially formulated in Goubault and Raussen [**GR02**] and in Raussen [**Rau03**], is to formally invert a subcategory of the weakly invertible morphsisms (turning those into isomorphisms) and to consider a quotient category (turning isomorphisms into identities).

In general, given a wide subcategory $\Sigma \subset \mathcal{C}$ of a category $\mathcal{C}$ containing the $\mathcal{C}$-isomorphisms, one may consider the category "of fractions" $\mathcal{C}[\Sigma^{-1}]$ in which a formal inverse σ^{-1} has been added to every morphism $\sigma \in \Sigma$. It is universal with respect to functors that send all Σ-morphisms into isomorphisms. To make the exposition here as simple as possible, we work with a pure wide subcategory Σ satisfying the lr-extension property right away:

It is easy to see that all morphisms of $\mathcal{C}[\Sigma^{-1}]$ then have a description of the form $\sigma^{-1} \circ \alpha$, resp. $\beta \circ \sigma^{-1}$ with $\sigma \in \Sigma$. A morphism of the form $\sigma_1^{-1} \circ \sigma_2$, resp. $\sigma_1 \circ \sigma_2^{-1}$ is invertible in $\mathcal{C}[\Sigma^{-1}]$; such a morphism is called a Σ-zig-zag.

Two objects x, y of $\mathcal{C}$ are called Σ-*equivalent* ($x \simeq_\Sigma y$) if there exists a Σ-zig-zag-morphism between them. The equivalence classes with respect to that relation are called the Σ-*components* of $\mathcal{C}$; they are the (usual) path components *with respect to the* Σ-*zig-zag morphisms*. Moreover, we generate an equivalence relation on the morphisms of $\mathcal{C}[\Sigma^{-1}]$ by requiring that $\tau \simeq \tau \circ \sigma, \tau \simeq \sigma \circ \tau$ whenever $\sigma \in \Sigma$ and the composition is defined.

The *component category* $\pi_0(\mathcal{C};\Sigma)$ of the category $\mathcal{C}$ with respect to Σ has as objects the Σ-components; the morphisms from $[x]$ to $[y]$ are the equivalence classes of morphisms in $\bigcup_{x' \simeq x, y' \simeq y} \mathcal{C}[\Sigma^{-1}]$. Two morphisms in $\pi_0(\mathcal{C};\Sigma)$ that are represented by $\tau_i \in \mathcal{C}(x_i, y_i), 1 \leq i \leq 2$, with $y_1 \simeq_\Sigma x_2$ can be composed by inserting any Σ-zig-zag-morphism connecting y_1 and x_2, cf. Fajstrup etal. **[FGHR04]** for details.

Taking equivalence classes results in a quotient functor $q_\Sigma : \mathcal{C} \to \pi_0(\mathcal{C};\Sigma)$.

4.2.2. Properties of component categories. The most important properties of components and the quotient function q_Σ shown in Fajstrup etal. **[FGHR04]** can be summarized as follows:

PROPOSITION 4.2.1. (Fajstrup et al. [**FGHR04**, Proposition 2 – 7]) *Let $\mathcal{C}$ denote a category that has only identities as endomorphisms. Let $\Sigma \subset \mathcal{C}$ denote a pure wide subcategory of weakly invertible morphisms satisfying the lr-extension properties,[2] cf Definition 4.1.2.*
Let x, y, z denote objects in $\mathcal{C}$.

(1) *$\Sigma(x, y)$ is either empty or it consists of a single morphism (in the latter case x, y are Σ-equivalent).*
(2) *If x and y are Σ-equivalent and $f \in \mathcal{C}(x, z), g \in \mathcal{C}(z, y)$, then $f, g \in \Sigma$ (and hence z is Σ-equivalent to x and y).*
(3) *Let $x, y \in C$ for a component $C \subseteq ob(\mathcal{C})$. Every morphism $\tau' \in \mathcal{C}(x', y')$ with $x' \in C$ (resp. $y' \in C$) is Σ-equivalent to a morphism $\tau \in \mathcal{C}(x, -)$ (resp. $\tau \in \mathcal{C}(-, y)$).*
(4) *For every object $x \in C$, every morphism in $\pi_0(\mathcal{C};\Sigma)(C, D)$ has a lift to a morphism in $\mathcal{C}(x, y)$ for some $y \in D$.*
(5) *If, moreover, $\pi_0(\mathcal{C};\Sigma)(C, D)$ is finite, then there exists $y \in D$ such that the quotient map $q_C : \mathcal{C}(x, y) \to \pi_0(\mathcal{C};\Sigma)(C, D)$ is a bijection.*
(6) *Every isomorphism in $\pi_0(\mathcal{C};\Sigma)$ is an endomorphism.*
(7) *If $\tau_1 \circ \tau_2 \in \pi_0(\mathcal{C};\Sigma)(C, C)$ is an isomorphism, then the $\tau_i, 1 \leq i \leq 2$, are isomorphisms.*
(8) *$\pi_0(\mathcal{C};\Sigma)$ has only identities as endomorphisms.*

REMARK 4.2.2. The lifting properties (3) and (4) show the usefulness of the construction to yield information concerning the original category. Properties (6) and (7) show that it is impossible to return to a component that has been left.

Under reasonable additional assumptions, cf Fajstrup etal [**FGHR04**, Section 5.3], we were able to show that the component category $\pi_0(\mathcal{C};\Sigma)$ has desirable properties. In particular, cf [**FGHR04**, Proposition 9], one can define the concept of neighbouring $\Sigma-$ components C_1, C_2; these allow *precisely one* morphism in $\pi_0(\mathcal{C};\Sigma)(C_1, C_2)$.

[2] Not all conditions are needed in all statements below. For details, consult Fajstrup et al. [**FGHR04**].

4.2.3. Further developments. How to choose a convenient large subcategory of the weakly invertible morphisms, that is both pure and that satisfies the extension properties? It is not difficult to see that the subcategories of a given category that satisfy the extension properties form a lattice with a maximal element (Fajstrup etal **[FGHR04**, Lemma 5]); this is the largest subcategory satisfying the extension properties. But there seems to be no way to find a maximal *pure* subcategory of a given system of morphisms, in general; let alone one that satisfies extension properties.

4.3. Outlook and discussion

This problem has been overcome by Goubault and Haucourt [**GH07**] by strengthening the extension properties from Definition 4.1.2:

DEFINITION 4.3.1. (Goubault and Haucourt [**GH07**]) A wide subcategory $\Sigma \subset \mathcal{C}$ satisfies the *strong* extension properties if the diagrams in Definition 4.1.2 can be filled in so that they yield pushout, resp. pullback squares in $\mathcal{C}$.

REMARK 4.3.2. This means that the pushout (resp. pullback) is universal with respect to all other fillouts. The intuitive idea is that the pushout arises from a $\vee-$, resp. $\wedge$-operation in a lattice.

As a result, Goubault and Haucourt show that the strong extension properties *imply* pureness, in the following sense:

PROPOSITION 4.3.3. (Goubault and Haucourt [**GH07**]) *Let $\mathcal{B}$ denote a wide subcategory of $\mathcal{C}$ and suppose that the pair $(\mathcal{C}, Iso(\mathcal{C}))$ is pure.*

(1) *If $(\mathcal{C}, \mathcal{B})$ satisfies the strong extension properties, then $(\mathcal{C}, \mathcal{B})$ is pure.*
(2) *The family of all wide subcategories $Iso(\mathcal{C}) \subseteq \mathcal{D} \subseteq \mathcal{B} \subseteq \mathcal{C}$ such that $(\mathcal{C}, \mathcal{D})$ satisfies the strong extension properties is a* complete lattice*; in particular, there is a wide subcategory $\Sigma_{\mathcal{B}} \subseteq \mathcal{B}$ such that $(\mathcal{C}, \Sigma_{\mathcal{B}})$ satisfies the strong extension properties and such that $\mathcal{D} \subseteq \Sigma_{\mathcal{B}}$ for all $\mathcal{D}$ above.*

Hence, this maximal subcategory satisfying the strong extension properties is a good candidate for the definition of components in *loopfree* categories.

Moreover, also in Goubault and Haucourt [**GH07**], the authors establish an equivalence between the category of fractions $\mathcal{C}[\Sigma^{-1}]$ and the quotient category $\mathcal{C}/_{\Sigma}$ using generalized equivalences that had previously been investigated by Bednarczyk etal [**BBP99**]. They go on to show that the van Kampen theorem for the fundamental category from Grandis [**Gra03a**] infers a similar statment for component categories (with respect to a maximal subcategory of the weakly invertible morphisms satisfying the strong extension properties).

For many applications, it may only be important to distinguish elements of a d-space (and d-paths starting at such elements) with respect to their *future*. In that case, only r-extensions properties are relevant. A framework to handle future components in that direction has been dealt with in Goubault etal [**GHK10**].

CHAPTER 5

A general categorical approach to invariants of directed spaces

5.1. Categorical approaches

How about algebraic topological invariants of a d-space X? Of course, one may define invariants of path spaces $\vec{P}(X)(x,y)$ or their quotient trace spaces $\vec{T}(X)(x,y)$ (up to reparametrization, cf Chapter 6) for given $x, y \in X$ and have them organized by way of various categories related to X itself. In fact, since both source and target play a role, it is more natural to use categories related to the *product* $X \times X$ for indexing purposes; this is certainly necessary for categories that are not acyclic – in the notation of Kozlov **[Koz08]** – or loop-free – in the notation of Haucourt **[Hau06]**. This bookkeeping point of view and also an analysis of associated component categories (with distinctions up to various invariants from algebraic topology) are the main contributions of Raussen **[Rau07]**:

5.1.1. Preorder categories as indexing categories and functors. A d-space X comes equipped with a natural preorder $x \preceq y \Leftrightarrow \vec{P}(X)(x,y) \neq \emptyset$. For all the preorder categories below, the *objects* are the pairs $(x,y) \in X \times X$ with $x \preceq y$.

The *morphisms* in the category $\vec{D}(X)$ are $\vec{D}(X)((x,y),(x',y')) := \vec{T}(X)(x',x) \times \vec{T}(X)(y,y')$ with composition given by pairwise contra-, resp. covariant concatenation. Hence, $\vec{D}(X)$ is a full subcategory of the category $\vec{T}(X)^{op} \times \vec{T}(X)$. Note that every morphism $(\sigma_x, \sigma_y) \in \vec{T}(X)(x',x) \times \vec{T}(X)(y,y')$ decomposes as

$$(\sigma_x, \sigma_y) = (c_{x'}, \sigma_y) \circ (\sigma_x, c_y) = (\sigma_x, c_{y'}) \circ (c_x, \sigma_y)$$

with $c_u \in \vec{T}(X)(u,u)$ the constant trace at $u \in X$.

Trace spaces with varying pairs of end points in X can be organised by the trace space functor $\vec{T}^X : \vec{D}(X) \to Top$ given by $\vec{T}^X(x,y) = \vec{T}(X)(x,y)$ and $\vec{T}^X(\sigma_x, \sigma_y)(\sigma) := \sigma_x \circ \sigma \circ \sigma_y \in \vec{T}(X)(x',y')$ for $\sigma \in \vec{T}(X)(x,y)$. This functor can be viewed as (a restriction of) the *Top*-enriched hom-functor of $\vec{T}(X)$.

A d-map $f : X \to Y$ induces a functor $\vec{D}(f) : \vec{D}(X) \to \vec{D}(Y)$ with $\vec{D}(f)(x,y) = (fx, fy)$ and $\vec{D}(f)(\sigma_x, \sigma_y) = (\vec{T}(f)(\sigma_x), \vec{T}(f)(\sigma_y)) = (f \circ \sigma_x, f \circ \sigma_y)$; moreover, it induces a natural transformation $\vec{T}(f)$ from $\vec{T}^X$ to $\vec{T}^Y$.

A homotopical variant is given by the category $\vec{D}_\pi(X)$ with the same objects as above and with $\vec{D}_\pi(X)((x,y),(x',y')) := \vec{\pi}_1(X)(x',x) \times \vec{\pi}_1(X)(y,y')$. Hence, this category is a full subcategory of the category $\vec{\pi}_1(X)^{op} \times \vec{\pi}_1(X)$, and $\vec{\pi}_1(X)$ denotes the fundamental category (cf. Sect. 3.3). It comes with a functor $\vec{T}^X_\pi : \vec{D}_\pi(X) \to Ho - Top$ into the homotopy category; a d-map $f : X \to Y$ induces a natural transformation

$\vec{T}_\pi(f)$ from $\vec{T}^X_\pi$ to $\vec{T}^Y_\pi$. Together with the (vertical) forgetful functors, we obtain a commutative diagram

$$\begin{array}{ccc} \vec{D}(X) & \xrightarrow{\vec{T}^X} & Top \\ \downarrow & & \downarrow \\ \vec{D}_\pi(X) & \xrightarrow{\vec{T}^X_\pi} & Ho-Top. \end{array} \tag{5.1}$$

The functors $\vec{T}^X$ and $\vec{T}^X_\pi$ may be composed with homology functors into categories of (graded) abelian groups, R-modules or graded rings. In particular, we obtain, for $n \geq 0$, functors $\vec{H}_{n+1}(X) : \vec{D}(X) \to Ab$ with $(x,y) \mapsto H_n(\vec{T}(X)(x,y))$ and $(\sigma_x, \sigma_y)_*$ given by the map induced on n-th homology groups by concatenation with those two traces on trace space $\vec{T}(X)(x,y)$. This functor factors over $\vec{D}_\pi(X)$ by homotopy invariance. In the same spirit, one can define homology with coefficients and cohomology. A d-map $f : X \to Y$ induces a natural transformation $\vec{H}_{n+1}(f) : \vec{H}_{n+1}(X) \to \vec{H}_{n+1}(Y), n \geq 0$.

Composing with the functor $\pi_0 : Top \to Sets$ that associates to a topological space its set of path components, defines a functor $\vec{\Pi}_1 : \vec{D}(X) \to Sets$ with $\vec{\Pi}_1(X)(x,y) = \vec{\pi}_1(X)(x,y)$, the set of morphisms in the fundamental category – with dihomotopy and not d-homotopy, cf Section 3.2.4, as equivalence relation.

If one needs to take care of base points (essential for homotopy groups of the very often non-connected spaces of d-paths), more care is needed. For that purpose, one may use a factorization category (cf Baues [**Bau89**]) of the trace category; for details, cf Raussen [**Rau07**, Section 3.3].

5.1.2. Components with respect to functors. Using functors – like the ones discussed above – on preorder categories, one may define components as subspaces of $X \times X$, for a variety of algebraic topological invariants. In greater generality:

Consider a functor $F : \mathcal{C} \to \mathcal{D}$ between two small categories. A morphism $\sigma \in \mathcal{C}(x,y)$ will be called *F-invertible* if and only if $F(\sigma) \in \mathcal{D}(Fx, Fy)$ is a *$\mathcal{D}$-isomorphism*. Let $\mathcal{C}_F(x,y) \subseteq \mathcal{C}(x,y)$ denote the set of *all* F-invertible morphisms from x to y. The collection of all $\mathcal{C}_F(x,y)$ form a wide subcategory $\mathcal{C}_F$ of $\mathcal{C}$ since the composition of two F-invertible morphisms obviously is F-invertible again; remark that $\mathcal{C}_F(x,y)$ contains the $\mathcal{C}$-isomorphisms.

For example, consider the functor $\vec{T}^X_\pi : \vec{D}_\pi(X) \to Ho-Top$ or the functors $\vec{H}_{n+1}(X) : \mathcal{D}_\pi(X) \to Ab$ from Sect. 5.1.1. A morphism $(\sigma_x, \sigma_y) \in \vec{D}_\pi(X)((x,y),(x',y'))$ is $\vec{T}^X$-invertible if and only if $\vec{T}(X)(\sigma_x, \sigma_y) : \vec{T}(X)(x,y) \to \vec{T}(X)(x',y')$ is a homotopy equivalence; it is $\vec{H}_{n+1}$-invertible if $(\sigma_x, \sigma_y)_* : H_n(\vec{T}(X)(x,y)) \to H_n(\vec{T}(X)(x',y'))$ is an isomorphism.

When $\mathcal{C}$ is the homotopy preorder category $\vec{D}_\pi(X)$ and F one of the functors from Section 5.1.1, it makes sense to apply the framework of component categories from Section 4.2. Reasonable components can then be defined at least in the case when the

$\vec{T}^X$-invertible morphisms satisfy the strong extension properties, cf. Definition 4.3.1. For details, compare Raussen [**Rau07**, Section 4].

5.2. Homotopy flows and d-homotopy equivalences

We discuss a candidate for a definition of the notion *directed homotopy equivalence* and an investigation of its properties:

5.2.1. Introduction. Which requirements should a d-map $f : X \to Y$ satisfy in order to qualify as a directed homotopy equivalence? Obviously, there should be a reverse d-map $g : Y \to X$ such that both $g \circ f$ and $f \circ g$ are d-homotopic to the respective identity maps. But this is not enough: The (d-path) structures on X and Y ought to be homotopically related, i.e., the maps $\vec{T}(f) : \vec{T}(X)(x,y) \to \vec{T}(Y(fx, fy)$ should be ordinary homotopy equivalences – for all x, y with $\vec{T}(X)(x,y) \neq \emptyset$ – and that in a natural way. Moreover note the following:

EXAMPLE 5.2.1. The subspace $L = [0,1] \times \{0\} \cup \{0\} \times [0,1] \subset \vec{R}^2$ – the branch figure "letter L" – is homotopy equivalent to the one point space $O = \{(0,0)\}$ included in it; there is a d-homotopy of the map $c \circ i$ (i: inclusion, c: the (constant) reverse map) to the identity map on L. Moreover, all non-empty trace spaces are contractible. But a branch should not be d-homotopy equivalent to a point!

5.2.2. Homotopy flows. The main tool in the definition below is the notion of a homotopy flow generalising the concept of a flow on a differentiable manifold. This notion will be used as an ingredient in the requirements for a d-homotopy equivalence. Moreover, it is useful in order to generate subcategories of weakly invertible d-paths (cf Section 4.1) and thus to reason about component categories (cf Section 4.2).

DEFINITION 5.2.2. (1) A d-map $H : X \times \vec{I} \to X$ is called a *future homotopy flow* if $H_0 = id_X$ and a *past homotopy flow* if $H_1 = id_X$.
(2) The sets consisting of all future homotopy flows, resp. of all past homotopy flows will be denoted by $\vec{P}_+C(X)$, resp. by $\vec{P}_-C(X)$.

Remark that we do not require that the maps $H(-,t) : X \to X$ are homeomorphisms.

The orbits of a flow have the following counterpart: For every $x \in X$, the map $H_x : \vec{I} \to X,\ t \mapsto H(x,t),$ is a d-path (with $H_x(0) = x$, resp. $H_x(1) = x$). Evaluation at $x \in X$ sends x to H_x and defines maps $H \mapsto H_x$

$$ev^x_+ : \vec{P}_+C_0(X) \to \vec{T}(X)(x,-), \text{ resp. } ev^x_- : \vec{P}_-C_0(X) \to \vec{T}(X)(-,x) \tag{5.2}$$

by $H \mapsto H_x$. Every homotopy flow gives thus rise to a well-organized collection of d-paths.

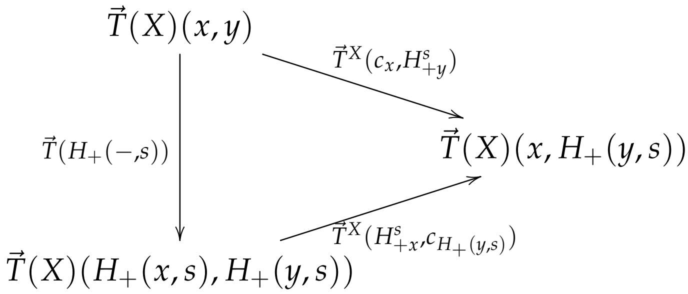
Remark that a maximal element $x_+ \in X$ – the only d-path with source x_+ is constant – will be fixed under a future homotopy flow; likewise a minimal element under a past homotopy flow. Moreover, a branch point like in Example 5.2.1 is fixed under all (future, resp. past) homotopy flows.

Homotopy flows can be concatenated, for future homotopy flows as follows:

$$(H_1 * H_2)(x,t) = \begin{cases} H_1(x,2t), & t \le \frac{1}{2} \\ H_2(H_1(x,1),2t-1), & t \ge \frac{1}{2}, \end{cases}$$

A homotopy flow $H_+ : X \times \vec{I} \to X$ and its restrictions $H_+^s : X \times \vec{I} \to X, H_+^s(x,t) = H_+(x,st)$ induce interesting maps on trace spaces, collected in the homotopy commutative diagram below (similarly for a past homotopy flow H_-):

(5.3)



A homotopy flow on X does not change the topology of trace spaces if it induces homotopy equivalences on associated trace spaces:

DEFINITION 5.2.3. (1) A future homotopy flow $H : X \times \vec{I} \to X$ is called *automorphic* if, for all $x, y \in X$ with $\vec{T}(X)(x,y) \neq \emptyset$ and all $s \in I$, the map $\vec{T}(H_+^s)$, (vertical in (5.3)) is a *homotopy equivalence*.
Similarly for past a past homotopy flow and $\vec{T}(H_-^s)$.

(2) A self-d-map $f : X \to X$ is called a future/past-automorphism if there exists an automorphic future/past homotopy flow connecting f and the identity on X.

In particular, given such an automorphic homotopy flow H_+, resp. H_-, the maps $f = H_+(-,1)$, resp. $g = H_-(-,0)$ induce homotopy equivalences

$$\vec{T}(f) : \vec{T}(X)(x,y) \to \vec{T}(X)(fx, fy) \text{ resp. } \vec{T}(g) : \vec{T}(X)(gx, gy) \to \vec{T}(X)(x,y).$$

Relations to the remaining maps on trace spaces in (5.3) are formulated in the easy

LEMMA 5.2.4. *Let H denote a future/past homotopy flow on X.*

(1) *If all concatenation maps (the skew ones in (5.3)) are homotopy equivalences, then H is automorphic.*
(2) *Let H be automorphic. If one of the concatenation maps (skew in (5.3)) is a homotopy equivalence, then the other is as well.*
(3) *$P_+C(X)$ satisfies the r-extension properties, and $P_-C(X)$ stisfies the l-extension property.*

Weaker properties (concerning maps on trace spaces that induce isomorphisms on homotopy, resp. homology groups) are also formulated in Raussen [**Rau07**, Definition 5.9].

5.2.3. d-homotopy equivalences: Definition and properties. The following is a suggestion for a reasonable definition of a d-homotopy equivalence making sure that trace spaces are related by homotopy equivalences. The conditions should be seen as requirements to a d-homotopy *class* of d-maps from X to Y:

DEFINITION 5.2.5. (1) A d-map $f : X \to Y$ is called a *future d-homotopy equivalence* if there exist d-maps $f_+ : X \to Y, g_+ : Y \to X$ such that f, f_+ are d-homotopic and *automorphic* d-homotopies $H^X : id_X \to g_+ \circ f_+$ on X and $H^Y : id_Y \to f_+ \circ g_+$ on Y (ie $H^X : X \times \vec{I}$ with $H^X_0 = id_X$ and $H^X_1 = g_+ \circ f_+$) etc.

(2) The d-map $f : X \to Y$ is called a *past d-homotopy equivalence* if there exist d-maps $f_- : X \to Y, g_- : Y \to X$ such that f, f_- are d-homotopic and *automorphic* d-homotopies $H^X : g_- \circ f_- \to id_X$ on X and $H^Y : f_- \circ g_- \to id_Y$ on Y.

(3) A d-map f is called a *d-homotopy equivalence* if it is *both* a future and a past d-homotopy equivalence.

Remark that a self-d-homotopy equivalence preserves branch points (like in Example 5.2.1).

d-homotopy equivalences have the desired properties:

PROPOSITION 5.2.6. *The natural transformation $\vec{T}_\pi(f) : \vec{T}^X_\pi \to \vec{T}^Y_\pi$ induced by a (past or future) d-homotopy equivalence $f : X \to Y$ between d-spaces X and Y is an equivalence, i.e., the induced maps $\vec{T}(f)(x,y) : \vec{T}(X)(x,y) \to \vec{T}(Y)(fx, fy)$ are homotopy equivalences for all $x, y \in X$ with $x \preceq y$.*

PROOF. By abuse of notation, we write f, g instead of f^+, g^+, resp. f^-, g^- in the following. In the diagram

$$\vec{T}(X)(x,y) \xrightarrow{\vec{T}(f)} \vec{T}(Y)(fx,fy) \xrightarrow{\vec{T}(g)} \vec{T}(X)(gfx,gfy) \xrightarrow{\vec{T}(f)} \vec{T}(Y)(fgfx,fgfy),$$

$$I : \vec{T}(X)(gfx,gfy) \to \vec{T}(X)(x,y), \qquad J : \vec{T}(Y)(fgfx,fgfy) \to \vec{T}(Y)(fx,fy)$$

let I denote a homotopy inverse to $\vec{T}(g) \circ \vec{T}(f)$ and let J denote a homotopy inverse to $\vec{T}(f) \circ \vec{T}(g)$. Then $\vec{T}(g)$ has a homotopy right inverse $\vec{T}(f) \circ I$ and a homotopy left inverse $J \circ \vec{T}(f)$. By general nonsense, the right homotopy inverse and the left homotopy inverse are homotopic to each other, and thus $\vec{T}(g)$ is a homotopy equivalence. Since $\vec{T}(g \circ f) = \vec{T}(g) \circ \vec{T}(f)$ is a homotopy equivalence by definition, the map $\vec{T}(f)$ is a homotopy equivalence, as well. $\square$

Furthermore, we prove in **[Rau07**, Proposition 6.8] that future and past dihomotopy equivalences behave well under composition:

PROPOSITION 5.2.7. *The composition $g \circ f : X \to Z$ of (future or past) dihomotopy equivalences $X \xrightarrow{f} Y \xrightarrow{g} Z$ is again a (f/p) dihomotopy equivalence.*

5.3. Outlook and discussion

5.3.1. Construction of homotopy flows and of components? While the definition of a d-homotopy equivalence ensures some of the most desirable properties, it seems not that easy to construct the homotopy flows needed in the definition in a systematic way. The fact that they have to preserve branch points (for a more precise definition cf Raussen **[Rau12b]** or Section 7.3.1) and thus also the regions between branch points requires them to be very "stiff". On the other hand, this is perhaps good news for handling components in simple cases, like for the PV-models from Section 2.1 and Chapter 7.

5.3.2. Topology change and relations to multidimensional persistence. It is clear that the discussion of topology change of trace spaces under variation of end points and the identification of components cannot be easy in general. In fact, variation of end points can be thought of as a (double) filtration on the trace spaces: You begin with trace spaces $\vec{T}(X)(x,x)$ – trivial in spaces without directed loops – and let start and end point drift further and further away from each other and observe topology changes – at certain thresholds – on the way. At a first glance, the situation resembles that investigated in the theory of (homological) persistence:

For a one-dimensional filtration with coefficients over a field k, the homology of a filtered space can be analyzed via the ranks of homology maps in the form of so-called barcodes, cf eg Edelsbrunner and Harer **[EH10]**. The theoretical background herefore is based on the classification of modules over the polynomial ring $k[t]$, cf Zomorodain and Carlsson **[ZC05]**.

For a multi-scale filtration, the situation is much more complicated; one needs to analyze modules over a polynomial ring of several variables. In general, discrete invariants like the barcodes are not sufficient; compare Carlsson and Zomorodian **[CZ09]**.

Our case is even more difficult; for at least three reasons:

- The state space has "holes" in cases of interest; a filtration is thus not homogeneous and theoretically understandable by considering a polynomial ring.
- There may and will often be more than one map comparing trace spaces between given pairs of end points (this is the essence of the category $D(X)$ and also $D_\pi(X)$).
- In full generality, one would have to consider a double (multi-scale) filtration, taking into accounts both end points.

CHAPTER 6

Topology of executions: General properties of path and trace spaces

One of the fundamental tasks in directed algebraic topology is to translate information about a directed space (d-space) into information about the *space of directed paths* (d-paths, executions) in that space. To this end, one has to structure the path spaces and

- to define them as (ordinary) *topological* spaces and to study properties of such spaces
- to define and study appropriate subspaces (eg subspaces of d-paths with given end points) and natural maps between such
- to investigate the influence of (weakly increasing) reparametrizations and to study properties of quotient spaces (trace spaces).

We begin here with the last issue:

6.1. Reparametrizations and traces

6.1.1. In ordinary spaces – without directions. Before turning to spaces of d-paths, it turned out to be beneficial to study the space of paths $P(X) = X^I$ from an interval I into a Hausdorff space X in the compact-open topology more closely. This Section reports briefly on the work of Fahrenberg and Raussen **[FR07]** with a minor correction that appeared in Raussen **[Rau09a]**:

In differential geometry, one studies usually (spaces of) *regular* paths in a manifold. Those have never-vanishing speed, can be reparametrized by a diffeomorphism of the unit interval to yield a path parametrized by arc length – with *unit speed*.

The closest analogon for paths in a Hausdorff space leads to

DEFINITION 6.1.1. Let $p : I \to X$ denote a continuous path in a Hausdorff space X.

(1) A path $p : I \to X$ is called *regular* if, for every interval $J \subseteq I$ with $p_{|J}$ constant, J is either a degenerate (one-point) interval $[a, a]$ or $J = I$.
(2) A continuous map $\phi : I \to I$ is called a *reparametrization* if $\phi(0) = 0, \phi(1) = 1$ and if ϕ is *increasing*, i.e. if $s \leq t \in I$ implies $\phi(s) \leq \phi(t)$.
The subspace (and monoid under composition) of all reparametrizations (within the mapping space I^I) is denoted $Rep_+(I)$.
(3) The subspace (and group under composition) of all (increasing) *homeomorphisms* within $Rep(I)$ is called $Homeo_+(I)$.
(4) Two paths $p, q : I \to X$ are called *reparametrization equivalent* if there exist reparametrizations ϕ, ψ such that $p \circ \phi = q \circ \psi$.

It turns out that reparametrization equivalence is in fact an equivalence relation, but the proof of the transitivity property is non-trivial. The equivalence classes with respect to reparametrization equivalence are called *traces* in the space X, and the quotient space $T(X) = X^I/_{Rep_+(I)}$ of $P(X) = X^I$ is called the *trace space* for X. Likewise, one may consider the quotient space $T_R(X) = R(X)/_{Homeo_+(I)}$ of *regular* traces.

Since one can contract paths to their start point, trace spaces as such are not interesting from the homotopy point of view. This changes, when one restricts end points and looks at subspaces of all (regular) paths $R(X)(x_0, x_1) \subset P(X)(x_0, x_1)$ with given end points $x_0, x_1 \in X$. In the non-directed case, one may restrict attention to *loops* ($x_0 = x_1$) in every path-connected component, and it is then not difficult to see using path fibrations – and proved in Fahrenberg and Raussen [**FR07**, Remark 3.10]:

PROPOSITION 6.1.2. *For two elements $x_0, x_1 \in X$ in a Hausdorff space X, the inclusion map $R(X)(x_0, x_1) \hookrightarrow P(X)(x_0, x_1)$ is a weak homotopy equivalence.*

It is natural to ask whether every path is reparametrization equivalent to a regular path. Since a general path can have complicated (Cantor) sets of intervals on which it is constant, this is not clear right away. In fact, reparametrizations of the interval have an interesting algebraic and combinatorial structure investigated in Fahrenberg and Raussen [**FR07**, Section2]. Using these insights, we proved in Theorem 3.6 of that article:

THEOREM 6.1.3. *For every two elements $x_0, x_1 \in X$ of a Hausdorff space X, the map $i : T_R(X)(x_0, x_1) \to T(X)(x_0, x_1)$ is a* homeomorphism.

The proof of this theorem is surprisingly intricate and needs a thorough study of the monoid $Rep_+(I)$ of reparametrizations, a characterization of reparametrization by sets of *stop intervals* and relations between those.

REMARK 6.1.4. It seems to be difficult to prove statements about the quotient maps q and q_R in the diagram

$$\begin{array}{ccc} R(X)(x_0, x_1) & \xrightarrow{\subset} & P(X)(x_0, x_1) \\ \downarrow q_R & & \downarrow q \\ T_R(X)(x_0, x_1) & \xrightarrow[i]{\cong} & T(X)(x_0, x_1) \end{array}$$

in full generality. The map q_R is a quotient map for a free action of the contractible group $Homeo_+(I)$. But it is not clear whether this map is a fibration in general; one cannot apply the Vietoris-Begle theorem (cf Smale [**Sma57**]) either since the fibre $Homeo_+(I)$ is not compact.

If the map q_R is a fibration (with contractible fibers of type $Homeo_+(I)$), then it is a weak homotopy equivalence, and then q is so as well. This is certainly the case when there is a ("unit speed") section of q_R leading to product decompositions $R(X)(x_0, x_1) \cong T_R(X)(x_0, x_1) \times Homeo_+(I)$ and $P(X)(x_0, x_1) \cong T(X)(x_0, x_1) \times Rep_+(I)$. In this case, both q_R and q are actually homotopy equivalences. For a framework where this occurs naturally, cf Section 6.2.

6.1.2. In directed spaces. Under mild extra assumptions, a variant of Theorem 6.1.3 holds for a d-space X, as well. Let $\vec{R}(X) = R(X) \cap \vec{T}(X)$ consist of the regular d-paths. The free action of the contractible group $Homeo_+(X)$ restricts to $\vec{R}(X)$ and yields the quotient space $\vec{T}_R(X) = \vec{R}(X)/_{Homeo_+(I)}$. We need to consider so-called *saturated* d-spaces:

DEFINITION 6.1.5. *(Fahrenberg and Raussen [**FR07**])* A d-space X is called a *saturated* d-space if the underlying topological space is Hausdorff and satisfies the following additional property:

$$p \in \vec{P}(X),\ \varphi \in Rep_+(I) \text{ and } p \circ \varphi \in \vec{P}(X) \ \Rightarrow\ p \in \vec{P}(X).$$

This means, that if p becomes a d-path after a reparametrization, then it has to be a d-path itself. It is easy to saturate a given d-space to yield a d-space with possible additional d-paths; therefore it is no harm to assume that a d-space is saturated right away.

Remark that, unlike in the classical case, the topology of the spaces $\vec{T}(X)(x_0, x_1)$ will usually depend crucially on the choice of end points.

COROLLARY 6.1.6. *For every two elements $x_0, x_1 \in X$ of a saturated d-space X, the map $i : \vec{T}_R(X)(x_0, x_1) \to \vec{T}(X)(x_0, x_1)$ is a* homeomorphism*.*

It is not clear, in general, that the inclusion map $\vec{R}(X)(x_0, x_1) \hookrightarrow \vec{P}(X)(x_0, x_1)$ is a homotopy equivalence. But again, as in Section 6.1.1, if inclusion has a ("unit speed") section, then $\vec{P}(X)(x_0, x_1) \cong \vec{R}(X)(x_0, x_1) \times Rep_+(I)$ and hence inclusion is a homotopy equivalence. Section 6.2 handles a case where this is naturally the case.

REMARK 6.1.7. K. Ziemiański [**Zie12b**, Section 5] has noted that spaces of d-paths and of traces in a d-space are homotopy equivalent if one just has a (length) function $l : \vec{P}(X) \to \mathbf{R}$ that is a homomorphism with respect to concatenation and addition, trivial only on constant dipaths, invariant under reparametrization and, moreover and crucially, *continuous*.

6.2. Trace spaces in cubical complexes

Our aim is to describe additional tools that allow a closer investigation of trace spaces $\vec{T}(X)(x_0, x_1)$ for a convenient d-space X.

Convenience has two aspects:

- There should be enough structure on the d-space to make sure that it is possible to describe an associated trace space in terms of a *finite complex*.
- Important classes of d-spaces arising as models for concurrency should be included.

The cubical complexes and their geometric realizations explained in Section 3.1.4 are candidates for convenient d-spaces, certainly satisfying the second criterion as models for Higher Dimensional Automata. With the d-space structure explained in Section 3.1.4, we shall explain that associated trace spaces enjoy several properties that pave the way for more detailed investigatons. In short, it is shown in Raussen [**Rau09b**], that trace spaces in a cubical complex – under mild additional assumptions – are separable metric spaces which are locally contractible and locally compact. Moreover, it turns out that they have the homotopy type of a CW-complex, and hence determination of topological invariants comes within reach.

6.2.1. Arc length parametrization and consequences. The L_1-arc length of a d-path in a cubical complex was introduced and studied in Raussen [**Rau09b, Rau12b**]. The *signed* L_1-length $l_1^{\pm}(p)$ of a path $p : I \to \square_n$ *within* a cube $\square_n$ is defined as $l_1^{\pm}(p) = \sum_{j=1}^{n} p_j(1) - p_j(0)$. For any path p, that is the concatenation of finitely many paths each of which is contained in a single cube, the signed L_1-length is defined as the sum of the lengths of the pieces. The result is independent of the choice of decomposition – and of the parametrization! Moreover, it is non-negative for every d-path and positive for every non-constant d-path.

This construction can be phrased more elegantly using differential one-forms on a cubical complex (a special case of the PL differential forms introduced by D. Sullivan [**Sul77**] in his approach to rational homotopy theory, or of the closed one-forms on topological spaces by M. Farber [**Far02, Far04**]): On an n-cube $e \simeq \square_n$, consider the particular 1-form $\omega_e = dx_1 + \cdots + dx_n \in \Omega^1(\square_n)$. It is obvious that $\omega_{\partial_i^k e} = (i_i^k)^*\omega_e$ with $i_i^k : |\partial_i^k e| \hookrightarrow |e|$ denoting inclusion. Pasting together, one arrives at a particular (closed!) 1-form ω_X on every pre-cubical set X – the one-form that reduces to ω_e on every cell e in X.

The signed length of a (piecewise differentiable) path γ on X can then be defined as $l_1^{\pm}(\gamma) = \int_0^1 \gamma^*\omega_X$ and extended to continuous paths using uniformly converging sequences of such piecewise differentiable paths. This length

- is invariant under orientation preserving reparametrization;
- changes sign under orientation reversing reparametrization;
- is additive under concatenation and non-negative for d-paths.
- yields a *continuous* map $l_1^{\pm}\colon P(X)(x_0, x_1) \to \mathbf{R}$.

An application of Stokes' theorem shows:

PROPOSITION 6.2.1. *Two paths $p_0, p_1 \in P(X)(x_0, x_1)$ that are homotopic rel end points have the same signed length: $l_1^{\pm}(p_0) = l_1^{\pm}(p_1)$.* □

A more direct proof can be given along the lines of Raussen [**Rau09b**] using the continuous d-map $s : X \to S^1 = \mathbf{R}/\mathbf{Z}$ given by $s(e; x_1, \dots, x_n) = \sum x_i \bmod 1$. We think of S^1 as a pre-cubical set with one vertex and one edge from that vertex to itself. Then the map $l_1^{\pm}(S^1) : P(S^1) \to \mathbf{R}$ coincides with the map that associates to $p \in P(S^1)$ the real number $\hat{p}(1) - \hat{p}(0)$ for an arbitrary lift $\hat{p}$ of p under the exponential map $\exp : \mathbf{R} \to S^1$. It follows from the definition that the arc length $l_1^{\pm}(X)$ factors for an arbitrary cubical complex X:

$$l_1^{\pm}(X) : P(X) \xrightarrow{s_\#} P(S^1) \xrightarrow{l_1^{\pm}(S^1)} \mathbf{R}$$

with the following consequences:

PROPOSITION 6.2.2. (1) *The function* $l_1^{\pm} : P(X) \to \mathbf{R}$ *is continuous.*
(2) $l_1^{\pm}(p) \equiv s(p(1)) - s(p(0)) \bmod 1$ *for* $p \in P(X)$.
(3) $l_1^{\pm}(P(X)(x_0, x_1))$ *is constant* $\bmod 1$ *for every pair of points* $x_0, x_1 \in X$.
(4) $l_1^{\pm}$ *is constant on any connected component, i.e., on any homotopy class of paths in* $P(X)(x_0, x_1)$. *In particular, it induces a map* $l_1^{\pm} : \pi_1(X)(x_0, x_1)$ *into a coset in* $\mathbf{R}$ *mod* $\mathbf{Z}$.

Similar results hold for spaces of *directed paths* $\vec{P}(X)(-,-)$ and of *traces* $\vec{T}(X)(-,-)$ (with *di*homotopy classes corresponding to components).

REMARK 6.2.3. Ordinary Euclidean arc length (for piecewise differentiable functions) is *not* a continuous function on a path space. A family of graphs of oscillating functions with increasing frequency and decaying amplitude may have constant arc length and converge to the graph of a constant function with a smaller arc length.

We will now consider the restriction $l_1 : \vec{P}(X) \to \mathbf{R}_{\geq 0}$ of $l_1^{\pm}$ to spaces of d-paths. We call a d-path parametrization $p : I \to X$ *natural* if $l_1(p|_{[0,t]}) = t \cdot l_1(p)$ for all $t \in I$; such a path moves at "unit speed" with respect to l_1. Naturally parameterized d-paths form the subspace $\vec{N}(X) \subset \vec{P}(X)$.

PROPOSITION 6.2.4. *Let* X *denote a cubical complex;* $x_0, x_1 \in X$.
(1) $\vec{N}(X)(x_0, x_1) \subset \vec{P}(X)(x_0, x_1)$ *is a deformation retract.*
(2) *All maps in the diagram*

$$\begin{array}{ccc} \vec{N}(X)(x_0,x_1) & \hookrightarrow & \vec{P}(X)(x_0,x_1) \\ & \searrow \quad \swarrow & \\ & \vec{T}(X)(x_0,x_1) & \end{array}$$

are homotopy equivalences.

PROOF. A homotopy inverse to inclusion is given by the naturalization map nat : $\vec{P}(X)(x_0, x_1) \to \vec{N}(X)(x_0, x_1)$, $nat(p)(t) = p(l_1^{-1}(t))$ – which is well-defined (!), continuous and homotopic to the identity in $\vec{P}(X)(x_0, x_1)$ – since $Rep_+(I)$ is convex and thus contractible. The map $\vec{N}(X)(x_0, x_1) \to \vec{T}(X)(x_0, x_1)$ is even a homeomorphism. For details, we refer to Raussen [**Rau09b**, Section 2.4]. □

Essentially the same construction wass used in the more general setup of Ziemiański [**Zie12b**].

6.2.2. General properties of trace spaces in a cubical complex. Every connected component of a cubical complex has a topology that is induced from a metric space: A chain connecting two elements $x, y \in X$ in the same component is a sequence $x = x_0, x_1, \ldots, x_n = y$ in X such that two subsequent elements x_i, x_{i+1}, $0 \leq i < n$, are contained in the *same* cell. The L_1-distance $d_1(x, y)$ is then defined as the infimum (in fact the minimum) over all sums $\sum_{i=0}^{n-1} |l_1^{\pm}(p_i)|$ with p_i any path from x_i to x_{i+1} in a common cell – extending over all chains between x and y. It is easy to check that d_1 is a metric.

The compact-open topology on path space $\vec{P}(X)$ is induced from the supremum metric $d_1(p,q) = \max_{t\in I} d(p(t), q(t))$ for $p, q \in \vec{P}(X)$ – for paths within the same component. By Proposition 6.2.4, trace space $\vec{T}(X)$ is homeomorphic to the subspace $\vec{N}(X)$ and can thus be seen as a subspace of this metric subspace.

PROPOSITION 6.2.5 (Raussen [**Rau09b**]). *Given a cubical complex X with elements $x_0, x_1 \in X$.*

(1) *Path space $\vec{P}(X)(x_0, x_1)$ and trace space $\vec{T}(X)(x_0, x_1)$ are metrizable and thus* Hausdorff *and* paracompact*.*
(2) *If X is a finite complex, then these metric spaces are* separable*.*

A space of d-paths is never compact – unless it only contains constant paths. This is so since the space of reparametrizations $Rep_+(I) = \vec{P}(I)(0,1)$ is not compact; it is not even equicontinuous, a necessary condition for compactness by the Arzelà-Ascoli theorem (cf. e.g. [**Dug66, Mun75**]).

Trace spaces are in general not compact either. If the d-space X contains a non-trivial loop based at $x_0 \in X$, then the closed subspace $\vec{T}(X)(x_0, x_0)$ has d-paths of infinitely many L_1-arc lengths and thus by Proposition 6.2.2 infinitely many connected components whence it cannot be compact. But compactness results are available if one bounds the L_1-arc lengths of d-paths:

PROPOSITION 6.2.6 (Raussen, [**Rau09b**]). *Let X denote a finite cubical complex.*

(1) *A subset $H \subseteq \vec{T}(X)$ of bounded L_1-arc length is* relatively compact*.*
(2) *For $x_0, x_1 \in X$, every dihomotopy class (connected component) in $\vec{T}(X)(x_0, x_1)$ is compact.*
(3) *Trace space $\vec{T}(X)$ is* locally compact*.*

John Milnor investigated in Milnor [**Mil59**] conditions for spaces that ensure that a variety of mapping spaces have the *homotopy type of a CW-complex*: One may check that these criteria can be applied to spaces of traces in a cubical complex satisfying an extra condition, and we conclude that those spaces have the homotopy type of a CW-complex.

DEFINITION 6.2.7. (Milnor [**Mil59**]) A topological space A is called ELCX (equi locally convex) provided there are

(1) a neighborhood U of the diagonal $\Delta A \subset A \times A$ and a map $\lambda : U \times I \to A$ satisfying
$\lambda(a,b,0) = a, \lambda(a,b,1) = b$ for all $(a,b) \in U$, and
$\lambda(a,a,t) = a$ for all $a \in A, t \in I$;
(2) an open covering of A by sets V_β such that $V_\beta \times V_\beta \subset U$ and
$\lambda(V_\beta \times V_\beta \times I) = V_\beta$.

LEMMA 6.2.8. (a special case of [**Mil59**, Lemma 4])
Every paracompact ELCX space has the homotopy type of a CW-complex. □

In fact, Milnor shows that a paracompact ELCX space is dominated by a simplicial complex and thus (see e.g. Hatcher [**Hat02**, Appendix, Proposition A.11]) homotopy equivalenct to a CW-complex.

PROPOSITION 6.2.9 (Raussen, [**Rau09b**]). *Let X denote a non-self linked cubical complex (cf Definition 3.1.3) and let* $x_0, x_1 \in X$. *Then*

(1) *X is ELCX.*
(2) *The path spaces* $\vec{P}(X)$ *and* $\vec{P}(X)(x_0, x_1)$ *are ELCX.*
(3) *The path and trace spaces* $\vec{P}(X), \vec{P}(X)(x_0, x_1), \vec{T}(X)$ *and* $\vec{T}(X)(x_0, x_1)$ *have the homotopy type of a CW-complex.*

6.3. Outlook and discussion

It is a natural but unfortunately still unachieved project to try to achieve a generalisation of the results on reparametrizations of paths (and d-paths) to mapping spaces with boxes $\square^n$ resp. directed boxes $\vec{\square}^n$ as their domain. A further generalization would concern mapping spaces with domain a cubical complex.

Later work, described in the subsequent Section 7 yields a precise cell structure – as a simplicial complex – for a complex homotopy equivalent to both $\vec{P}(X)(x_0, x_1)$ and $\vec{T}(X)(x_0, x_1)$; in Section 7.1 for a space X arising from a linear semaphore model as introduced in Section 2; then, more generally in Section 7.3, for d-paths in a general cubical complex X.

CHAPTER 7

Homotopy types of trace spaces

In this section, we report on recent work that allows to identify the homotopy types of trace spaces of convenient d-spaces as *explicitly given* (generalized simplicial) complexes. A resulting combinatorial description of the chain complexes associated to these complexes allows to calculate algebraic topological invariants for the trace space under investigation.

It should be admitted right away, though, that the simplicial complexes and chain complexes in question will often be huge. Apart from a few vary easy toy cases, actual calculations need the power of dedicated computer algorithms for homology calculations. Experiments using such software (eg CHomP and CrHom) done by Polish collaborators on top of outputs of the dedicated ALCOOL software of French colleagues have shown limits to the calculations of, say, Betti numbers using this approach – apart from the quintessential β_0, ie the number of components of trace spaces. For an algorithm finding and exploiting the simplicail complexes mentioned above as well as for implementation issues, consult Section 7.1.4.

7.1. Simplicial models for trace spaces associated to semaphore models

It is easiest to describe the general idea for the spaces associated to semaphore models explained in Chapter 2: The state space is of the form $X = I^n \setminus F$ with a forbidden region $F = \bigcup_{i=1}^{l} R^i$ consisting of isothetic hyperrectangles $R^i = \prod_{j=1}^{n}]a_j^i, b_j^i[$. To make notation as easy as possible, we concentrate on a description of trace space $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ between bottom $\mathbf{0}$ and top $\mathbf{1}$ in the hypercube I^n – the generalization to the general case $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ is not very difficult, cf Raussen [**Rau12a**].

The general strategy may be described as follow:

(1) Decompose X into finitely many subspaces $X_k \subset X$ such that the sub-trace spaces $\vec{T}(X_k)(\mathbf{0}, \mathbf{1})$ *cover* the entire trace space $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ and such that all these sub-trace spaces and all intersections of those are either *empty* or *contractible*. The non-empty ones among those intersections form a poset category.

(2) Apply a variant of the nerve lemma (eg Kozlov [**Koz08**, Theorem 15.21]) to identify the homotopy type of $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ with the nerve of that poset category – a finite simplicial complex.

In fact, one can do a bit better by replacing the poset category by a related smaller category $\mathcal{C}(X)(\mathbf{0}, \mathbf{1})$ taking into account the product structure induced from the collection of individual hyperrectangles; the associated complex is then a prodsimplicial complex in the terminology of Kozlov [**Koz08**].

It turns out that the method used to detect deadlocks and unsafe regions described in Section 2 comes in very handy to distinguish those sub-trace spaces that are *empty* from those that are *not*.

7.1.1. Contractible subspaces of trace spaces. For bookkeeping purposes, we consider the set $M_{l,n} = M_{l,n}(\mathbf{Z}/2)$ of all *binary* $l \times n$-matrices – with 2^{ln} elements – and the subset $M^R_{l,n} \subset M_{l,n}$ consisting of the $(2^n - 1)^l$ matrices with the property that *no row vector is a zero vector*. We regard $M_{l,n}$ and $M^R_{l,n}$ as poset categories with the coordinate-wise partial order $\leq$.

For every matrix $M \in M_{l,n}$, we define a subspace X_M of $X = I^n \setminus F$:

DEFINITION 7.1.1.

$$\begin{aligned} X_M &:= \{\mathbf{x} \in X \mid m_{ij} = 1 \Rightarrow x^i_j \leq a^i_j \text{ or } \exists k\colon\ x_k \geq b^i_k\} \\ &= \{\mathbf{x} \in X \mid \forall i : (\forall k\ x_k < b^i_k \Rightarrow (m_{ij} = 1 \Rightarrow x^i_j \leq a^i_j))\}. \end{aligned}$$

REMARK 7.1.2. (1) Interpretation: An execution path in X_M has the property: If $m_{ij} = 1$, then process j will *not* acquire a lock to resource i before at least one of the others has relinquished it.

(2) Matrices in $M^R_{l,n}$ represent areas in which *each* hole is obstructed furthermore in *at least one* direction.

EXAMPLE 7.1.3. (1) Figure 7.1 shows, in each of the two rows, an example of a model space $X = \vec{I}^2 \setminus F$ given as the complement of the forbidden region F consisting of two black squares. The grey-shaded areas show, in both cases, the four subspaces X_M corresponding to the four matrices

$$M = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix};$$

in that order.

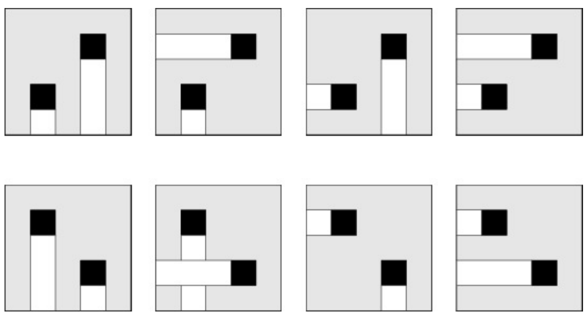


FIGURE 7.1. Two examples of a model space X and of associated subspaces X_M – the grey-shaded areas.

Remark that an empty space of d-paths $\vec{P}(X_M)(\mathbf{0}, \mathbf{1}) = \emptyset$ occurs only in the second row and second column in Figure 7.1. All other spaces $\vec{P}(X_M)(\mathbf{0}, \mathbf{1})$ are non-empty and contractible. This can be used to explain that the trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ has four (contractible) components in the first case and three components in the second.

(2) Figure 7.2 below shows a forbidden region "black box" $\vec{J}^3$ – with $\vec{J} \subset \vec{I}$ an interior open interval – with upper corner $\mathbf{b}$ surrounded by the state space $X = \vec{I}^3\vec{J}^3$. Moreover, you recognize the shaded areas $X_{M_j} \cap (\partial_+ \downarrow \mathbf{b}), 1 \leq j \leq 3$, with $M_1 = [100], M_2 = [010]$ and $M_3 = [001]$.

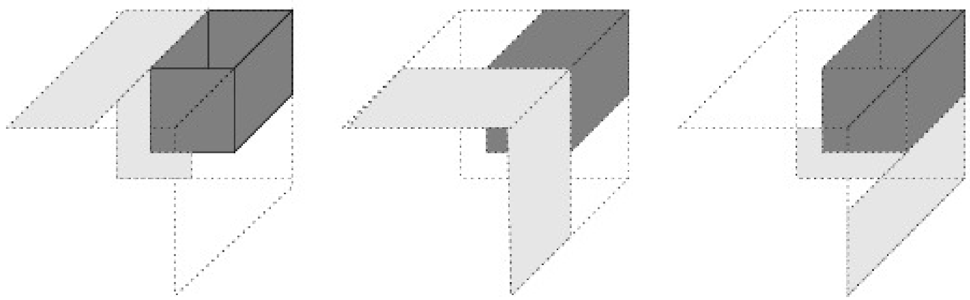


FIGURE 7.2. Intersections of X_{M_j} with the upper boundary $\partial_+ \downarrow \mathbf{b}$ of the black box $\downarrow \mathbf{b}$ with upper corner $\mathbf{b}$.

Remark that every pair of these areas intersect non-trivially, whereas the intersection $X_{[111]} = X_{M_1} \cap X_{M_2} \cap X_{M_3}$ is empty. In particular, $\vec{P}(X_M)(\mathbf{0},\mathbf{1}) = \emptyset$ for $M \in M^R_{1,3}$ if and only if $M = [111]$. The subsequent analysis yields as a consequence that $\vec{T}(X)(\mathbf{0},\mathbf{1}) \simeq \partial\Delta^2 \cong S^1$.

The following constructions depend on an analyis of the binary operation $\vee$ (least upper bound) defined on $\mathbf{R}^n$: $\mathbf{a} \vee \mathbf{b} = (\max(a_1, b_1), \ldots, \max(a_n, b_n))$. This operation restricts to an operation on I^n, but its restriction to $X \times X$ has values in the forbidden region F. This is why we subdivide X into subspaces $X_M, M \in M^R_{l,n}$:

PROPOSITION 7.1.4. (1) *The subspaces $X_M, M \in M^R_{l,n}$ are all* closed *under* $\vee$.

(2) $\vec{T}(X)(\mathbf{0},\mathbf{1}) = \bigcup_{M \in M^R_{l,n}} \vec{T}(X_M)(\mathbf{0},\mathbf{1})$ – *every trace is contained in at least one of the restricted regions X_M.*

(3) *Every trace space $\vec{T}(X_M)(\mathbf{0},\mathbf{1}), M \in M^R_{l,n}$, is* empty *or* contractible.

PROOF. For details, we refer to Raussen **[Rau10]**.

(1) Verify that the inequalities defining X_M are satisfied for $\mathbf{a} \vee \mathbf{b}$, $\mathbf{a}, \mathbf{b} \in X_M$. It is crucial that $M \in M^R_{l,n}$ – the inegalities "guard" *every* hole R^i.

(2) For a given d-path $p \in \vec{P}(X)(\mathbf{0},\mathbf{1})$, let $t_i^+ = \min\{t \mid \exists k : \ p_k(t) = b_k^i\}$ for $1 \leq i \leq l$. Then there exists $j \in [1 : n]$ such that $p_j(t_i^+) \leq a_j^i$ and hence $p_j(t) \leq a_j^i$ for $t \leq t_i^+$; otherwise $p(t) \in R^i$ on a non-empty interval $]t_i^-, t_i^+[$.

(3) If $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ is non-empty, then, for any pair $p, q \in \vec{P}(X_M)(\mathbf{0},\mathbf{1})$, define a one-parameter family $H(p,q) : \vec{P}(X_M)(\mathbf{0},\mathbf{1}) \times I \to \vec{P}(X_M)(\mathbf{0},\mathbf{1})$ by

$$H_t(p,q)(s) := q(s) \vee p(ts), \ t \in I.$$

Remark that $H_0(p,q)(s) = q(s) \vee \mathbf{0} = q(s), H_t(p,q)(0) = \mathbf{0} \vee \mathbf{0} = \mathbf{0}, H_t(p,q)(1) = \mathbf{1} * p(t) = \mathbf{1}$ and that $H_1(p,q)(s) = q(s) \vee p(s)$. Thus $H(p,q)$ defines an increasing d-homotopy (cf Grandis **[Gra03a]**) $q \mapsto p \vee q$ between d-paths within $\vec{P}(X_M)(\mathbf{0},\mathbf{1})$. Likewise, $H(q,p)$ is an increasing d-homotopy $p \mapsto q \vee p = p \vee q$. Their concatenation $G(q,p) = H(p,q) * H^-(q,p)$ (orientations are reversed for the second d-homotopy) is a "zig-zag" d-homotopy from q to p; in particular a path from q to p within $\vec{P}(X_M)(\mathbf{0},\mathbf{1})$.

The map $G(-,-) : \vec{P}(X_M)(\mathbf{0},\mathbf{1}) \times \vec{P}(X_M)(\mathbf{0},\mathbf{1}) \to \vec{P}(X_M)(\mathbf{0},\mathbf{1})^I$ defines a continuous section of the "end path map" $ev_0 \times ev_1 : \vec{P}(X_M)(\mathbf{0},\mathbf{1})^I \to \vec{P}(X_M)(\mathbf{0},\mathbf{1}) \times \vec{P}(X_M)(\mathbf{0},\mathbf{1})$; it associates the d-homotopy $G(p,q)$ to a pair (q,p).

Given an arbitrary $p \in \vec{P}(X_M)(\mathbf{0},\mathbf{1})$, the map $G(-,p)\colon \vec{P}(X_M)(\mathbf{0},\mathbf{1}) \times I \to \vec{P}(X_M)(\mathbf{0},\mathbf{1})$ is a contraction of $\vec{P}(X_M)(\mathbf{0},\mathbf{1})$ to p. By Raussen [**Rau09b**, Proposition 2.16] or Proposition 6.2.4(2), the trace space $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to path space $\vec{P}(X_M)(\mathbf{0},\mathbf{1})$; hence it is also contractible.

□

7.1.2. A variant of the nerve lemma leads to a prodsimplicial complex. In the following, we will work with a restriction of the poset category $M_{l,n}$ of binary matrices from Section 7.1.1: The relevant index category to consider here is the full subposet category $\mathcal{C}(X)(\mathbf{0},\mathbf{1}) \subset M^R_{l,n} \subset M_{l,n}$ consisting of all matrices M such that

$$\vec{T}(X_M)(\mathbf{0},\mathbf{1}) \neq \emptyset. \tag{7.1}$$

This index category gives rise to functors $\mathcal{D}$ and $\mathcal{E}$ into **Top**:

DEFINITION 7.1.5.
- For a non-zero binary vector $\mathbf{m} \in (\mathbf{Z}/2)^n$, let $\Delta(\mathbf{m}) \subseteq \Delta^{n-1}$ denote the simplex spanned by the unit vectors $\mathbf{e}_j \in \mathbf{R}^n$ with $m_j = 1$.
- For $M \in M^R_{l,n}$, let $\Delta(M) = \prod_{i=1}^{l} \Delta(\mathbf{m}_i) \subseteq (\Delta^{n-1})^l$.
- The functor $\mathcal{D} : \mathcal{C}(X)(\mathbf{0},\mathbf{1})^{op} \to$ **Top** associates $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ to the matrix M; the reverse partial order on $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ corresponds to inclusion in **Top**.
- The functor $\mathcal{E} : \mathcal{C}(X)(\mathbf{0},\mathbf{1}) \to$ **Top** restricts from a functor $\mathcal{E}^l_n : M^R_{l,n} \to$ **Top**; it associates to $M \in \mathcal{C}(X)(\mathbf{0},\mathbf{1}) \subseteq M^R_{l,n}$ the simplex product $\Delta(M)$. For this functor, the *original* partial order on $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ corresponds to inclusion in **Top**.

The functor $\mathcal{E}^l_n$ should be considered as a pasting scheme for the product of simplices $(\Delta^{n-1})^l$; the functor $\mathcal{E}$ becomes then a pasting scheme for a sub-prodsimplicial complex (cf Kozlov **[Koz08]**) $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) \subseteq (\Delta^{n-1})^l$ to be explained below.

Regarding the functors $\mathcal{E}$ and $\mathcal{D}$ as pasting schemes, we consider their colimits:

- $\mathrm{colim}(\mathcal{D}) = \vec{T}(X)(\mathbf{0},\mathbf{1})$ – by Proposition 7.1.4;
- $\mathrm{colim}(\mathcal{E}^l_n) = (\Delta^{n-1})^l$;
- $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) := \mathrm{colim}(\mathcal{E}) \subset \mathrm{colim}(\mathcal{E}^l_n) = (\Delta^{n-1})^l$ is a *prodsimplicial* complex consisting of all those products of simplices $\Delta(M)$ that correspond to matrices

$M \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$; in other words, the functor $\mathcal{E}$ is a pasting scheme for a prodsimplicial complex with one simplex product for each $M \in M^R_{l,n}$ giving rise to a *non-empty* trace space $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$.

- As a subcomplex of $(\partial\Delta^{n-1})^l \cong (S^{n-2})^l$, the prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ has at most n^l vertices, and $\dim(\mathbf{T}(X)(\mathbf{0},\mathbf{1})) \leq (n-2)l$.

EXAMPLE 7.1.6. Figure 7.3 shows the state space X from Section 3.2.4 with a (slightly compressed) picture of $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) \subset \partial\Delta^2 \times \partial\Delta^2$. It consists of the shaded parts of the torus and the isolated vertex shown four times in the covering (corresponding to the contractible component of the path space with the d-path in the illustration on the left hand side). That latter model space is clearly homotopy equivalent to $S^1 \vee S^1 \sqcup *$.

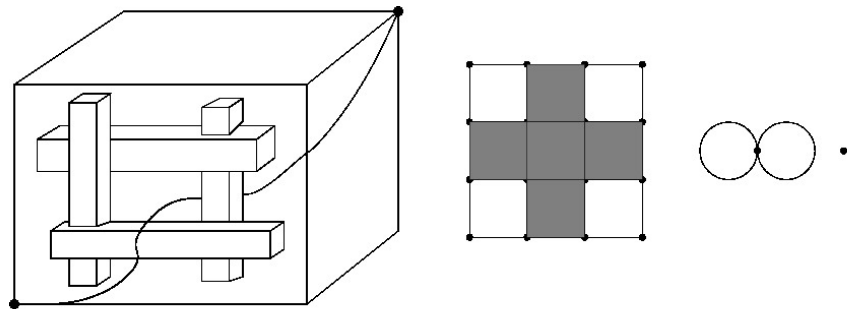


FIGURE 7.3. State space X and prodsimplicial model $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ of the path space

THEOREM 7.1.7. *The trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to the prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) \subset (\partial\Delta^{n-1})^l$ and to the nerve of the category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$; the latter simplicial complex arises as a barycentric subdivision of $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$.*

PROOF. First, we determine the *homotopy* colimits of the functors defining the pasting schemes above. We apply the homotopy lemma (cf eg Kozlov [**Koz08**, Theorem 15.12]) to the natural transformation $\Psi : \mathcal{D} \Rightarrow \mathcal{T}^*$ from $\mathcal{D}$ to the trivial functor $\mathcal{T}^* : \mathcal{C}(X)(\mathbf{0},\mathbf{1})^{op} \to \mathbf{Top}$ which sends every object into the same one-point space. Since the maps corresponding to Ψ are homotopy equivalences at every object M in $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ (from a contractible space $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ – by Proposition 7.1.4(3) – to a point), the map $\operatorname{hocolim}\mathcal{D} \to \operatorname{hocolim}\mathcal{T}^*$ induced by Ψ is a homotopy equivalence by the homotopy lemma. By definition, $\operatorname{hocolim}\mathcal{T}^*$ is the nerve $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$ of the indexing category.

A similar argument shows that also the trivial natural transformation from $\mathcal{E}$ to $\mathcal{T} : \mathcal{C}(X)(\mathbf{0},\mathbf{1}) \to \mathbf{Top}$ induces a homotopy equivalence of homotopy colimites.

Next, we wish to apply the projection lemma (cf Segal [**Seg68**, Proposition 4.1] or Kozlov [**Koz08**, Theorem 15.19]) to the fiber projection maps $\operatorname{hocolim}\mathcal{D} \to \operatorname{colim}\mathcal{D}$ and $\operatorname{hocolim}\mathcal{E} \to \operatorname{colim}\mathcal{E}$. The given cover is not open, but it was shown in Raussen [**Rau10**] how it can be replaced by an open cover of homotopy equivalent spaces.

Hence the projection lemma ensures that these maps are homotopy equivalences. Altogether, the maps discussed above fit to yield a homotopy equivalence

$$\begin{array}{ccccc} \vec{T}(X)(\mathbf{0},\mathbf{1}) = \mathrm{colim}(\mathcal{D}) & \longleftarrow & \mathrm{hocolim}(\mathcal{D}) & \longrightarrow & \mathrm{hocolim}(\mathcal{T}^*)\ , \\ & & & & \updownarrow \\ \mathbf{T}(X)(\mathbf{0},\mathbf{1}) = \mathrm{colim}(\mathcal{E}) & \longleftarrow & \mathrm{hocolim}(\mathcal{E}) & \longrightarrow & \mathrm{hocolim}(\mathcal{T}) \end{array}$$

since the two opposite categories $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ and $\mathcal{C}(X)(\mathbf{0},\mathbf{1})^{op}$ have the same classifying space $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$. In particular, $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ is also homotopy equivalent to the nerve $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$ – which is thus a barycentric subdivision of $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$. □

7.1.3. Determination of dead and of alive matrices. In the following, we call a matrix $M \in M^R_{l,n}$ *alive* it $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) \neq \emptyset$ and *dead* else. It is crucial for an algorithmic description of $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ to find a method that distinguishes dead and alive matrices. It turns out that the determination of dead matrices may be achieved through the method determining deadlocks and unsafe regions from Chapter 2. But first we need to establish an easy order property:

Consider the map $\Psi : M_{l,n} \to \mathbf{Z}/_2,\ \psi(M) = 1 \Leftrightarrow \vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \emptyset$ and the subset $M^C_{l,n} \subset M_{l,n}$ consisting of all matrices with *unit vectors* as columns – minimal *candidates* for dead matrices. Then

PROPOSITION 7.1.8. (1) Ψ *is order-preserving.*
(2) $\Psi(M) = 1 \Leftrightarrow$ *there exists* $N \in M^C_{l,n}$ *with* $\Psi(N) = 1$ *and* $N \leq M$*.*

For the easy proof, we refer to Raussen [**Rau10**, Proposition 4.5].

The first property in Proposition 7.1.8 tells us that $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ is indeed closed under containment and its geometric realization is thus a complex.

By the second property, we can concentrate on determining the subset $D(X)(\mathbf{0},\mathbf{1}) := \{M \in M^C_{l,n} \mid \Psi(M) = 1\} \subset M^C_{l,n}$. It allows to describe $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ as the set of matrices $M \in M^R_{l,n}$ with the property: For every matrix $N \in D(X)(\mathbf{0},\mathbf{1})$, there is a pair $(i,j) \in [1:l] \times [1:n]$ such that $m_{ij} = 0, n_{ij} = 1$.

Since the map Ψ above is order-preserving, it is enough to describe the *maximal* matrices in $\mathcal{C}_{max}(X)(\mathbf{0},\mathbf{1}) \subseteq \mathcal{C}(X)(\mathbf{0},\mathbf{1})$ that are "just alive": replacing just one entry $m_{ij} = 0$ by an entry 1 makes such a matrix greater or equal than a dead matrix.

The crucial idea for the determination of the *dead* matrices in $D(X)(\mathbf{0},\mathbf{1}) \subset M^C_{l,n}$ is a variant of the algorithm determining deadlocks and unsafe regions from Chapter 2: The aim is to describe the subspaces $X_M \subseteq X$ as *complements* of a union of *extended* hyperrectangles of type

$$R^i_j = \prod_{k=1}^{j-1} \tilde{I}^i_k \times I^i_j \times \prod_{k=j+1}^{n} \tilde{I}^i_k,\ 1 \leq i \leq l, 1 \leq j \leq n$$

with $\tilde{I}^i_k = [0, b^i_k[\supset]a^i_k, b^i_k[$. It is then easy to see (Raussen, [**Rau10**, Lemma 4.2]) that $X_M = \vec{I}^n \setminus \bigcup_{m_{ij}=1} R^i_j$.

Furthermore, as soon as n extended hyperrectangles R^i_j, one for each $1 \leq j \leq n$, give rise to a deadlock, the associated unsafe region is a hyperrectangle with **0** as the lowest vertex: the second largest coordinates (cf Section 2.2.1) are all 0 in extended hyperrectangles. In the end, everything boils down to a systematic check of various sets of inequalities between bottom and top coordinates a^i_j and $b^{i'}_{j'}$ of intervals in the product decomposition of the original hyperrectangles, cf Raussen [**Rau10**, Section 4].

7.1.4. Implementation issues. The algorithm described above that determines the poset category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ has been implemented in the ALCOOL tool of our French partners at CEA Saclay, cf Fajstrup etal [**FGH$^+$12**]. That latter paper contains also considerations about how to extend the methods from this section to cubical complexes that arise from the spaces X_M by identifying boundary faces, ie, complexes that are subspaces of products of a torus and a box arising by deletion of forbidden hyperrectangles.

The ALCOOL tool produces (among other deliveries) a description of the maximal alive matrices in $\mathcal{C}_{max}(X)(\mathbf{0},\mathbf{1})$. The boundary operator ∂ of an assoicated chain complex (with $\mathbf{Z}/2$ coefficients) can be implemented by a sum of terms in which exactly one of the digits 1 is replaced by a zero. In this way, it was possible for M. Juda (Krakow) to adapt the homology software created by the Polish group around M. Mrozek (cf eg Kaczynski etal [**KMM04**]) to do homology calculations of trace spaces.

This works well for semaphore protocols of a very moderate size. Unfortunately, when the number l of obstructions grows, the model $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ becomes quickly high-dimensional – although the homological dimension of the trace space is conjectured to be far more limited. The method thus does not yield algorithmically satisfactory results for large l. We work currently on modifications ot the method – that needs new theoretical insights – with the aim to reduce the dimension of the complex and thus to make homology calculations feasible in cases of interest.

7.2. Specific results for mutex semaphores – arity one

In the previous Section 7.1, we assumed for simplicity that none of the forbidden hyperrectangles R^i intersects the boundary ∂I^n. For semaphore models, this is true only for semaphores of arity $n-1$, giving simultaneous access to $n-1$ but not to n of the processors. For semaphores of a lower arity a (cf Section 2.1), one shared object gives rise to a union of $\binom{n}{a+1}$ hyperrectangles each of which contains $n-a-1$ maximal interval factors $[0,1]$ – and thus intersections with the boundary.

This fact, and also a need for investigation of "intermediate" trace spaces $\vec{T}(X)(\mathbf{c},\mathbf{d})$, $\mathbf{0} \leq \mathbf{c} \leq \mathbf{d} \leq 1$, motivated the extension of the method described in Section 7.1 to more general forbidden regions developed in Raussen [**Rau12a**]. We will not try to describe the quite technical development needed in that contribution in general; instead we list some results in the particularly interesting case where all semaphores have arity *one*:

Only one process can access a shared object at any given time. This is the particularly important case of *mutual exclusion* or *mutex* semaphores. In this case, we prove:

PROPOSITION 7.2.1. (Raussen [**Rau12a**, Proposition 7]) *Let $X = I^n \setminus F$ denote the state space corresponding to a collection C of calls to semaphores of arity one. Then the trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to a finite discrete space.*

We knew this already for $n = 2$ – in this case semaphores can only have arity one: The prodsimplicial complexes considered in Section 7.1 are then all subcomplexes of a product of 0-dimensional spheres.

For the state space X corresponding to a program with just a *single call* to one semaphore of arity one, this can be phrased more specifically as follows: X may be decomposed into subspaces X_π, one for every permutation $\pi \in \Sigma_n$. More specifically, we show for this case:

COROLLARY 7.2.2. (Raussen [**Rau12a**, Corollary 2]) *The trace space is a disjoint union $\vec{T}(X)(\mathbf{0},\mathbf{1}) = \bigsqcup_{\pi\in\Sigma_n} \vec{T}(X_\pi)(\mathbf{0},\mathbf{1})$. All $n!$ components $\vec{T}(X_\pi)(\mathbf{0},\mathbf{1})$, $\pi \in \Sigma_n$, are contractible.*

The general situation (more than one call) is far more complex, but it can in the end be translated to the discrete realm using a notion of *compatible permutations*: Every semaphore h can be called several times by a number of processes; each concurrent call c is performed by a subset $J_h \subset [1 : n]$ of (at least two competing) processes; a call c is characterized by a semaphore $h(c)$, by the subset $J_{h(c)}$ and by one out of $r_j(h(c))$ locking intervals on the axes corresponding to $J_{h(c)}$. Every such call $c = (h; m_1(h), \dots, m_{|J_h|}(h)), 1 \le m_j(h) \le r_j(h)$, gives rise to a forbidden region $F(c)$. Such a forbidden region alone would give rise to a trace space homotopy equivalent to a discrete space $\Sigma_{J_{h(c)}} \subset \Sigma_n$ – the stabilizer of $[1 : n] \setminus \Sigma_{J_{h(c)}}$, with cardinality $|J_{h(c)}|!$.

In total, we have to study the complement of a forbidden region $F = \bigcup_{c\in C} F(c)$ with C denoting the set of *all* calls. This suggests to study *collections* of permutations $\boldsymbol{\pi} = (\pi_c)_{c\in C} \in \Sigma = \prod_{c\in C} \Sigma_{J_{h(c)}}$: Consider the set of boundary coordinates $a^i_j, b^i_j \in I, 1 \le j \le n,$ corresponding to *all* cconcurrent calls to the semaphores. For every collection $\boldsymbol{\pi} = (\pi_c)_{c\in C} \in \Sigma = \prod_{c\in C} \Sigma_{J_{h(c)}}$, we consider several order relations on subsets of these real numbers:

- The natural order $\le$, inherited from the reals, on numbers a^i_j, b^i_j with the *same* subscript (direction) j;
- $b^{m_{\pi_c(j)}(h)}_{\pi_c(j)} \preceq a^{m_{\pi_c(j')}(h)}_{\pi_c(j')}$ for $c \in C, j < j' \in J_{h(c)}$ for the *same* call $c = (h; m_1(h), \dots, m_{|J_h|}(h)) \in C$.

DEFINITION 7.2.3. We call the collection $\boldsymbol{\pi} = (\pi_c)_{c\in C} \in \Sigma = \prod_{c\in C} \Sigma_{J_{h(c)}}$ *compatible* if the transitive closure $\sqsubseteq_{\boldsymbol{\pi}}$ of these relations is a *partial order*.

This notion was applied in the proof of

PROPOSITION 7.2.4. (Raussen, [**Rau12a**, Proposition 10]) *Let $X = I^n \setminus F$ denote the state space corresponding to a collection C of calls to semaphores of arity one.*
Then $\overrightarrow{T}(X)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to the discrete space
$\{\boldsymbol{\pi} = (\pi_c)_{c\in C} \in \prod_{c\in C} \Sigma_{J_{h(c)}} \mid \boldsymbol{\pi} \text{ compatible}\} \subseteq \prod_{c\in C} \Sigma_{J_{h(c)}} \subseteq (\Sigma_n)^{|C|}$.

EXAMPLE 7.2.5. Let $X_k \subset I^k$ denote the state space corresponding to the PV-model describing k dining philosophers (each protocol of type $PaPbVaVb$; cf Dijkstra [**Dij71**]) and Section 2.1.1. Then the trace space $\overrightarrow{T}(X_k)(\mathbf{0},\mathbf{1})$ consists of 2^k-2 contractible components: There are 2^k-2 essentially different interleavings of the d-paths corresponding to each individual protocol – indicating who of two neighbouring philosophers uses a fork first. The number 2^k-2 of schedules is, for $k>3$, considerably smaller than the number $k!$ of ordered k-tuples of philosophers. This is due to the fact that several philosophers can serve themselves concurrently for $k>3$. For a detailed analysis, cf Raussen [**Rau12a**, Example 2].

7.3. General Higher Dimensional Automata

In the preceeding sections, methods have only been worked out for semaphore models *without* loops; an important, but restricted family of general Higher Dimensional Automata (HDA). It turned out that some general ideas from this analysis can be applied to the general case, as well. In particular, the general case includes HDA with non-trivial directed *loops*, absolutely essential in the analysis of realistic concurrent programs. It has to be admitted, that the ideas that we describe below are certainly far more difficult to implement; this has not even been tried so far.

The aim is to find a combinatorial model $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ of the trace space $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ for a general non-self-linked (cf Fajstrup et al [**FGR06**], Raussen [**Rau09b**]) cubical complex X. This is done in several steps:

7.3.1. Trace spaces for non-looping cubical complexes. First we study trace spaces for non-self-linked cubical complexes *without* (non-trivial) d-loops and look for a replacement of the subspaces X_M from Section 7.1. These turn out to be (maximal) *non-branching* (sub)-complexes:

A (finite) cubical complex (geometric realization of a pre-cubical set) X will be called *non-branching* if it satisfies the following additional property

- **(NB):** Every vertex $v \in X_0$ is the lower corner vertex of a *unique maximal* cube c_v in X. This maximal cube c_v contains thus all cubes with lower corner vertex v as a (possibly iterated) lower face.

PROPOSITION 7.3.1. *For every pair of elements x_0, x_1 in a* non-branching *cubical complex X, the trace space $\vec{T}(X)(x_0,x_1)$ is either* empty *or* contractible.

Instead of using the least upper bound ($\vee$) operation (essential in the proof of Proposition 7.1.4), we apply the diagonal *directed flow* $F^X\colon X \times \mathbf{R}_{\geq 0} \to X$: Every element $x \in X$ is contained in the interior or the lower boundary of a uniquely determined maximal cube, i.e., the maximal cube c_v of its lowest vertex v. On the interior and the lower faces of such a cube c_v, this flow is locally given by the *diagonal* flow:

$$F_c^X(c;(x_1,\ldots,x_n);t) = (c; x_1+t,\ldots,x_n+t) \text{ for } 0 \leq t \leq 1 - \max_{1\leq i\leq n} x_i. \tag{7.2}$$

On a maximal vertex v_1 with $c = c_{v_1} = v_1$ (a deadlock), F_c^X is defined to be constant in the variable t for $0 \leq t$.

Note that property (NB) is essential: necessary and sufficient for pasting diagonal flows together from flows on individual cubes. Diagonal flows on intersecting different *maximal* cubes do not fit together on intersections of their lower boundaries. For a different description using a diagonal 1-form ω as in Section 6.2.1, we refer to Raussen [**Rau12b**].

One may now construct a cover of a general cubical complex X by maximal subspaces satisfying property (NB) and identify $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ with the nerve of that covering; cf Raussen [**Rau12b**, Theorem 4.2]. To determine such maximal (NB) subspaces algorithmically, one investigates the branch points in the 0-skeleton X_0 of the complex X – with several maximal cubes having that branch point as lowest vertex – and associated branches, one for every such maximal cube.

Simple examples show that these smaller branch subspaces may have (secondary) branch points. Hence, one has to iterate the construction. One ends up with a poset category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ the objects of which are given by so-called coherent and complete sequences of first and higher order branch points and associated choices of branch cubes. This category can be realized as a colimit $\mathbf{T}(\mathbf{x}_0,\mathbf{x}_1)$ of spaces each of which is a product of products of simplices and of cones on such products. In analogy with Theorem 7.1.7, one obtains

THEOREM 7.3.2. *For a cubical complex X without non-trivial loops and points $\mathbf{x}_0,\mathbf{x}_1 \in X$, the trace space $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is homotopy equivalent to*

(1) *the nerve $\Delta(\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1))$ of the poset category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$, and*
(2) *the complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$.*

For the proof, we refer to Raussen, [**Rau12b**, Theorem 4.11].

7.3.2. Trace spaces for cubical complexes with directed loops. We outline how previous metods can be adapted to trace spaces in a general cubical complex X with directed loops using suitable *coverings* of the complex X:

We exploit the d-map $s : X \to \vec{S}^1 \cong \mathbf{R}_{/\mathbf{Z}}$ introduced in Raussen [**Rau09b**] and described here in Section 6.2.1: just glue the maps $s(x_1, \dots, x_n) = \sum x_i \mod 1$ on individual cubes. Consider the pullback $\tilde{X}$ in the pullback diagram

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{S} & X \times \mathbf{R} \\ \pi \downarrow & & \downarrow id \times \exp \\ X & \xrightarrow[id \times s]{} & X \times S^1 \end{array} .$$

The map π is a covering map with unique path lifting. Since exp can be interpreted as a semi-cubical map, $\tilde{X}$ can be conceived as a cubical complex: Every cube e in X is replaced by infinitely many cubes $(e, n), n \in \mathbf{Z}$; the boundary maps are given by $\partial_-(e, n) = (\partial_- e, n), \partial_+(e, n) = (\partial_+ e, n + 1)$.

The directed paths on $\tilde{X}$ are those that project to directed paths in X under the projection map π. Remark that the maps exp and s – and hence π and $\pi_2 \circ S$ – preserve the signed L_1-arc length $l_1^{\pm}$ from Section 6.2.1. Moreover, the L_1-length $l_1^{\pm}(p)$ of a path p in X with lift $\tilde{p}$ can be expressed via the d-map $S : \tilde{X} \to X \times \mathbf{R}$ in the pullback diagram as follows:

LEMMA 7.3.3. (1) $l_1^{\pm}(p) = \pi_2(S(\tilde{p}(1))) - \pi_2(S(\tilde{p}(0)))$.
(2) *The map* $\pi_2 : \tilde{X} \to \mathbf{R}$ *is a d-map; hence:*
(3) $\tilde{X}$ *has* only trivial *directed loops.*

Another method to construct this covering is to consider the *homotopical* length map $\pi_1(X) \stackrel{l_1^{\pm}}{\to} \mathbf{Z} \to 0$ (cf Proposition 6.2.2) from the non-directed classical fundamental group of the cubical complex X. Consider the cover $\tilde{X} \downarrow X$ with fundamental group $\pi_1(\tilde{X}) = K \trianglelefteq \pi_1(X)$ the kernel of the homotopical length map $l_1^{\pm}$. It can be given the structure of a cubical complex, and every element $\mathbf{x}$ in X is covered by elements $\mathbf{x}^n \in \tilde{X}$; one for every $n \in \mathbf{Z}$. The projection map $\pi : \tilde{X} \downarrow X$ preserves the signed L_1-arc length. A path in $\tilde{X}$ is directed if and only if its projection to X is directed. There are no non-trivial directed loops in $\tilde{X}$ – these need to have L_1-length 0!

For a general cubical complex X with length cover $\tilde{X}$ we obtain the following decomposition result:

PROPOSITION 7.3.4. *For every pair of points* $\mathbf{x}_0, \mathbf{x}_1 \in X$, *trace space* $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ *is homeomorphic to the disjoint union* $\bigsqcup_{n \in \mathbf{Z}} \vec{T}(\tilde{X})(\mathbf{x}_0^0, \mathbf{x}_1^n)$.

Since the covering $\tilde{X}$ has *only trivial loops*, Proposition 7.3.4 allows us to apply the methods from Section 7.3.1 to describe the homotopy type of trace spaces $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ in an arbitrary cubical complex X.

In practice, we have so far investigated simple semaphore models with loops of the form $X = T^n \setminus F$ with $T^n = (S^1)^n$ an n-torus and F a collection of forbidden hyperrectangles. For such a space, one may consider the covering

$$\begin{array}{ccc} \tilde{X} & \hookrightarrow & \mathbf{R}^n \\ \downarrow & & \downarrow \exp \\ X & \hookrightarrow & T^n \end{array} \tag{7.3}$$

that arises as pullpack from the universal cover of the torus T^n – a far bigger gadget. The universality property ensures that (d-)paths, that are not homotopic in the torus T^n, lift to (d-)paths with different end points. The methods from Raussen [**Rau10**] can be applied to $\tilde{X}$ immediately. It is easier to get hold on periodicity properties in this setting; cf Fajstrup etal [**FGH$^+$12**], Raussen and Ziemiański [**RZ13**] and the following Section 7.4.

7.4. Explicit homology calculations for specific path spaces – with loops

It seems to be difficult to implement the considerations from Section 7.3 above in a working programme for spaces of directed paths with non-trivial directed loops. This is why we have attempted to investigate simple examples in order to find clues for calculations. One of these examples concerns path spaces in a Euclidean torus $T^n = (\vec{S}^1)^n$ from which just one rectangular hole I^n has been removed: $X = T^n \setminus I^n$.

The covering space arising from the construction (7.3) has the description $\tilde{X} = \mathbf{R}^n \setminus \bigsqcup_{\mathbf{c}-\frac{1}{2}\mathbf{1}\in\mathbf{Z}^n} I^n_{\mathbf{c}}$ with $I_{\mathbf{c}}$ a homothetic hyperrectangle centered at $\mathbf{c}$ – and edges of length less than 1. The space $\tilde{X}$ is homotopy equivalent to the $(n-1)$-skeleton of $\mathbf{R}^n$ seen as a cubical complex – with vertices in the integral points. Moreover, inclusion and retraction establishing such a homotopy equivalence can be chosen to preserve directed paths.

Also in this case it is easy to see, cf Raussen and Ziemiański [**RZ13**, Section 1.5] that the space of directed loops $\vec{P}(X)(\mathbf{x}_0, \mathbf{x}_0)$ based at $\mathbf{x}_0 \in X$ is homotopy equivalent to the disjoint union of spaces $\vec{P}(\tilde{X})(\mathbf{0}, \mathbf{k})$, $\mathbf{k} \geq \mathbf{0}, \mathbf{k} \in \mathbf{Z}^n$. For $n > 2$, this decomposition is also a decomposition into path components; for $n = 2$, $\vec{P}(X)$ is homotopy discrete cf Raussen and Ziemiański [**RZ13**, Section 1.7] and Corollary 7.4.2.

For $n = 3$, an attempt to calculate the homology of $\vec{P}(\tilde{X})^{(k,l,m)}_{\mathbf{0}}$ by "brute force" using the poset description for the cell complex of the prod-simplicial complex homotopy equivalent to that path space as in Section 7.1 – even using sophisticated homology software – failed already for $k = l = m = 3$. The prod-simplicial complex in this case has dimension $klm(n-2)$; its homological dimension is only $\min\{k, l, m\}(n-2)$. This gap was one of the motivations for looking for better descriptions of these path spaces.

In this specific case, we could show that at least the homology and the cohomology of the relevant path spaces $\vec{P}(\tilde{X})(\mathbf{0}, \mathbf{k})$ are, loosely speaking, algebraically generated by the cubical holes in the complex. Every such hole is characterized by the smallest integral vertex $\mathbf{l} \in \mathbf{Z}^n$ above it; this integral vector satisfies $\mathbf{0} \ll \mathbf{l} \leq \mathbf{k}$ (with $\mathbf{a} \ll \mathbf{b} \Leftrightarrow a_i < b_i$ for *all* i).

Let $Z^*(\tilde{X})(\mathbf{0},\mathbf{k})$ denote the free graded exterior $\mathbf{Z}$-algebra with generators the holes $\mathbf{0} \ll \mathbf{l} \leq \mathbf{k}$; every generator has grade $n-2$. Let $I(\tilde{X})(\mathbf{0},\mathbf{k})$ denote the ideal generated by products $\mathbf{l}_1\mathbf{l}_2$ with $\mathbf{l}_1 \not\ll \mathbf{l}_2$ and $\mathbf{l}_2 \not\ll \mathbf{l}_1$. Let $F^*(\tilde{X})(\mathbf{0},\mathbf{k}))$ denote the quotient algebra $F^* = Z^*/I^*$; a (graded) free abelian group with a basis consising of ordered *cube sequences* $[\mathbf{a}^*] = [\mathbf{0} \ll \mathbf{a}^1 \ll \mathbf{a}^2 \ll \cdots \ll \mathbf{a}^r \leq \mathbf{k}]$ (in dimensions $r(n-2)$) in $\tilde{X}$.

THEOREM 7.4.1. *Let $n > 2$ and $\mathbf{k} \geq \mathbf{0}$.*

(1) *Homology $H_*(\vec{P}(\tilde{X})(\mathbf{0},\mathbf{k}))$ is isomorphic to $F^*(\tilde{X})(\mathbf{0},\mathbf{k})$ as a graded abelian group.*
(2) *Cohomology $H^*(\vec{P}(\tilde{X})(\mathbf{0},\mathbf{k}))$ is isomorphic to $F^*(\tilde{X})(\mathbf{0},\mathbf{k})$ as a graded ring.*

The proof of Theorem 7.4.1 – first for homology, then refined to cohomology – relies on the fact that the path spaces $\vec{P}(X)(\mathbf{0},\mathbf{k})$ can be shown to be homotopy equivalent to the *homotopy colimit* of the system of spaces $\vec{P}(X)(\mathbf{0},\mathbf{k}-\mathbf{j})$ over the poset category $\mathcal{J}_n$ with objects the non-identical binary vectors $\mathbf{j} \in \{0,1\}^n,\ \mathbf{0} < \mathbf{j} < \mathbf{1}$. Similarly, the graded abelian group $F^*(\tilde{X})(\mathbf{0},\mathbf{k})$ is a colimit of graded abelian groups $F^*(\tilde{X})(\mathbf{0},\mathbf{k}-\mathbf{j})$ over the same category. This allows to construct a homomorphism $\Phi : F^*(\tilde{X})(\mathbf{0},\mathbf{k}) \to H_*(\vec{P}(\tilde{X})(\mathbf{0},\mathbf{k}))$.

The final steps of the proof in Raussen and Ziemiański in [**RZ13**, Section 3] establishing that Φ is indeed a graded isomorphism apply induction using a Bousfield-Kan [**BK72**] spectral sequence argument. Using Theorem 7.4.1, it is not difficult to calculate the Betti numbers of the relevant path spaces:

COROLLARY 7.4.2. *For $n > 2$ and $\mathbf{k} = (k_1, k_2, \ldots, k_n) \geq \mathbf{0}$, Betti numbers are given as*

$$\dim H_{r(n-2)}(\vec{P}(\tilde{X})(\mathbf{0},\mathbf{k})) = \binom{k_1}{r}\binom{k_2}{r}\cdots\binom{k_n}{r}.$$

Homology is trivial in all other dimensions.
For $n = 2$, $\vec{P}(\tilde{X})(\mathbf{0},(k_1,k_2))$ consists of $\binom{k_1+k_2}{k_1}$ contractible components.

The calculation of the (co-)homology of trace spaces in Theorem 7.4.1 can easily be generalized to a cubical subcomplex of $\mathbf{R}^n$ containing the $(n-1)$-skeleton (possibly with fewer holes), cf Raussen and Ziemiański [**RZ13**]. Further generalizations and other applications of the homotopy colimit constructions are still under consideration.

7.5. Outlook and discussion

Several lines of research are currently under consideration:

Section 7.1: We would like to construct a smaller poset category as a replacement for $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ such that the classifying space (nerve) has the same homotopy type. The background is that, in many cases, spaces $X_M, X_{M'}$ may be equal for different matrices $M, M' \in M_{l,n}$.

Section 7.2: So far, we have interpreted the forbidden region F_a corresponding to a semaphore of (general) arity a as a union of hyperrectangles. Instead, one may view the forbidden region $I^n \setminus F_a$ as homotopy equivalent (preserving d-paths and dihomotopies) to the a-skeleton of I^n.

Recent discussions with Fajstrup, Ottosen and Zieminaski show that the path space $\vec{P}(X)(I^n \setminus F_a)(\mathbf{0}, \mathbf{1})$ is homotopy equivalent to a *configuation space*, the so-called "a-equal manifold" with a topology (in particular its homology) has been studied by Björner and Welker [**BW95**].

Section 7.4: K. Ziemiański observed recently that it is possible to realize *every!* finite simplicial complex as the trace space of a suitable semaphore model up to homotopy. This result – quite easy to obtain – shows:

- There are no limits to the (combinatorial) "expressiveness" of linear semaphore models.
- It is undecidable whether two HDAs have the same expressiveness.

CHAPTER 8

Related Work. Outlook

8.1. Related Work

Directed Algebraic Topology has been an active research field pursued by a small community for about twenty years. As mentioned in the introduction, the original motivation was an attempt to model and study problems in concurrency theory in Computer Science. The oldest source studying connections between topology and order notions seems to be L. Nachbin's *Topology and Order* [**Nac65**] that served as background for our first attempt regarding lpo-spaces. V. Pratt's idea [**Pra91**] to use geometric and homotopy notions, later refined by van Glabbeek [**vG91**], inspired Goubault and Jensen [**GJ92**] to make use of homology tools. Gunawardena [**Gun94**] had the first clear-cut application of homotopy methods: a proof that the 2-phase locking method in database engineering is safe; cf Section 2.1.3.

From there on, developments diversified quickly. The following short guide to the literature mentions very briefly the work of several authors on mathematical perspectives motivated by concurrency and not covered in the previous sections. It is certainly biased and non-comprehensive. We start with work published by coauthors:

8.1.1. Approaches by other authors.

Lisbeth Fajstrup: Lisbeth Fajstrup (partially in collaboration with Sokołowski) investigated deadlocks and associated unsafe regions for semaphore models *with* directed loops [**Faj00,FS00**] and showed that several "deloopings" may be necessary before the unsafe region is detected correctly. She developed and investigagted the directed version of a covering space in [**Faj03, Faj10**] and gave first results for directed cubical approximation in [**Faj05**]. The paper [**FR08**] investigates directed coverings from a categorical perspective. In [**Faj13**], the author has a close look at the trace space of a torus with holes and relates it to automata theoretic methods.

Éric Goubault, Emmanuel Haucourt, Samuel Mimram and Sanjeevi Krishnan: Modeling concurrency via Higher Dimensional Automata is originally an idea of Vaughan Pratt's [**Pra91**]. It was translated into the mathematical language of (labelled) cubical complexes by Éric Goubault ([**GJ92, Gou93, Gou95, Gou01**]. We have already taken account of Goubault's survey article [**Gou00**] from 2000 followed up by another survey [**Gou03**] in 2003. Several joint papers with Emmanuel Haucourt [**GH05,GH07**] and with Sanjeevi Krishnan [**GHK09,GHK10**] and also by Haucourt [**Hau06**] are devoted to definitions and properties of *components* in varying contexts and also of categories describing relations them.

Krishnan invented streams in [**Kri09**] (cf Section 3.1.3) and proved directed simplicial and cubical approximation theorems in [**Kri13**] (cf Section 3.2.4). The implementation of many concepts and ideas in the software tool ALCOOL is due to Éric Goubault, Emmanuel Haucourt and Samuel Mimram.

Krzysztof Ziemiański: Krzysztof Ziemiański defines (directed) d-*simplicial* complexes and constructs and investigates a pre-cubical model for spaces of d-paths or traces in such a complex [**Zie12b**] – in the same spirit, but more intricate than our construction from Section 7.1. In [**Zie12a**], he shows that suitable categories of "good" d-spaces and of streams are equivalent and enjoy important properties (complete, cocomplete, Cartesian closed). A similar investigation can be found in the article [**HHH13**] authored by the Hirschowitz family.

Peter Bubenik: Peter Bubenik and Krzysztof Worytkiewicz attempt in [**BW06**] to reconcile the category of lpo-spaces (cf Section 3.1.2) with the tools and techniques of (topological) model categories. In [**Bub09**], Bubenik studies full subcategories of the fundamental category of a d-space with respect to sets of extremal points. In [**Bub12**], he produces alternative ways of giving trace spaces a combinatorial structure.

Philippe Gaucher: Philippe Gaucher has authored a long series of papers [**Gau00**, **Gau01**,**Gau02**,**Gau03d**,**GG03**,**Gau03a**,**Gau03c**,**Gau03b**,**Gau05c**,**Gau05a**,**Gau05b**, **Gau06a**,**Gau06b**,**Gau07**,**Gau08a**,**Gau08b**,**Gau09**,**Gau10a**,**Gau10b**,**Gau11**] investigating Higher Dimensional Automata and categories of such, mainly from a model category perspective. In particular, he defines categories of *flows* and two types of directed homotopy equivalences, the so-called S- and T- homotopy equivalences between flows. In recent papers, labelling and higher dimensional transition systems have been taken into consideration, as well.

Marco Grandis: The many research contributions by Marco Grandis have been focused on category theory through many years; most of them are motivated by topological and geometrical considerations in a wide range of set-ups. For several years, he has worked in directed algebraic topology; he is the author of the only book [**Gra09**] on the subject – written very systematically and with great care. This book builds on many previous articles [**Gra03a**,**Gra02**,**Gra03b**, **Gra03c**, **Gra04**, **Gra05**, **Gra06**, **Gra06b**, **Gra06a**, **Gra06c**, **Gra06**, **Gra07**] without exhausting them.

John F. Jardine: Rick Jardine is a prolific homotopy theorist, amongst others well-known for his book [**GJ99**] (with P. Goerss) on simplicial homotopy theory. He is mentioned here for his pre-print [**Jar02**] on *cubical* homotopy theory and for his investigation [**Jar10**] on path categories leading to algorithmic determinations of fundamental categories.

Thomas Kahl: Thomas Kahl is noted for two contributions [**Kah06**,**Kah12**] to directed algebraic topology. In the first, he investigates certain fibration and cofibration category structures on the category of po-spaces *under* a given pospace.

In the second, he investigates collapsing operations that may reduce the size of a d-space without affecting the homotopy type of associated trace spaces.

8.1.2. Distributed Computing and Combinatorial Algebraic Topology. Distributed Computing is a different area of theoretical Computer Science that has profited from relations with combinatorial algebraic topology. Roughly speaking, a distributed system is a software system in which components located on networked computers communicate and coordinate their actions by passing messages. A theoretical analysis becomes particularly challenging and interesting when coordination is low; in particular, when participating processors have only private memory and when they can crash without the others being able to observe that.

A typical problem that could be analyzed with tools from combinatorial algebraic topology is the algorithmic unsolvability of the so-called consensus problem (where processors have to agree on one of their inputs). Important contributions to this area earned their authors Herlihy and Shavit [**HS99**], resp. Saks and Zaharoglou [**SZ00**] the Gödel prize in 2004, awarded jointly by the European Association for Theoretical Computer Science and the Association for Computing Machinery.

A first source of inspiration for us was the survey book by Herlihy and Rajsbaum [**HR95**] followed by a long list of articles [**HRT98, CHT99, HR99, HS99, HR00, HRT00, CHLT00, GHR06, GHP09, Her10**] reasoning with tools from combinatorial algebraic topology on the (non)-solvability of (simplicial) tasks using simplicial protocol complexes extending a given simplicial input complex.

We are looking forward to the comprehensive textbook by Herlihy, Kozlov and Rajsbaum [**HKR13**] on the subject. It is a challenge to compare and to attempt finding formal relations between the use of tools from combinatorial algebraic topology in Distributed Computing and in Concurrency Theory.

8.1.3. Morse Theory. Relativity Theory. Another sort of inspiration came from the application of ideas from Morse theory applied to relativity theory – with causal or time-like curves playing the role of d-paths. This theme had been covered quite extensively in Penrose's classical monograph [**Pen72**] that includes important notions like the domain of dependence.

Rafal Wisniewski's Ph.d.-dissertation [**Wis05**] is not in relativity theory, but it applies Morse theory ideas to the analysis of (almost) flow lines of gradient- and non-gradient vector fields. He allows limited perturbations from the flow (within a cone) in the definition of a d-path. A follow up-paper by Raussen and Wisniewski appeared as [**WR07**]. The borderline between dynamical systems and directed topology deserves certainly more attention.

8.2. Outlook: Applied and Computational Algebraic Topology

The research field directed algebraic topology (with an eye to several fields of applications) is one of the areas composing the European research network ACAT: Applied and Computational Algebraic Topology. This network allows to facilitate many of the conferences in the area and to sponsor visits between researchers in the area. The network cooperates with researchers all over the world.

In the United States, applications of algebraic topology within engineering and science disciplines cover already a broad and established research spectrum. This is for instance witnessed by the entire academic year 2013 – 2014 devoted to Scientific and Engineering Applications of Algebraic Topology at the Institute of Mathematics and its Applications in Minneapolis, MN, USA.

Important research activities (apart from directed topology) in the ACAT network include

- Computational Algebraic Topology – including in particular research on persistent homology and its many applications; moreover topological aspects of visualization and shape analysis.
- Topological Robotics
- Stochastic Topology
- Applied Combinatorial Algebraic Topology.

It is impossible to cover these research areas within a few sentences; a list of survey books and articles must suffice here:

- Zomorodian [**Zom05**] and Edelsbrunner and Harer [**EH10**] on computational algebraic topology and, in particular, persistent topology.
- Farber [**Far08**] on topological robotics
- Costa, Farber and Kappeler [**CFK12**] on stochastic topology
- Kozlov [**Koz08**] on combinatorial algebraic topology

Cooperation and cross-fertilization between various subcommunities in applied and computational algebraic topology seem to be of utmost importance for future developments.

References

[Bau89] H.J. Baues, *Algebraic homotopy*, Cambridge Studies in Advanced Mathematics, no. 15, Cambridge University Press, 1989.

[BBP99] M. Bednarczyk, A. Borzyskowski, and W. Pawlowski, *Generalized congruences - epimorphisms in Cat*, Theory Appl. Categ. **5** (1999), no. 11, 266–280.

[BH81a] R. Brown and P.J. Higgins, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Algebra **22** (1981), 11–41.

[BH81b] ———, *On the algebra of cubes*, J. Pure Appl. Algebra **21** (1981), 233–260.

[BK72] A.K. Bousfield and D.M. Kan, *Homotopy limits, completions and localizations*, Lecture Notes in Math., vol. 304, Springer-Verlag, 1972.

[Bor94] T. Borceux, *Handbook of Categorial Algebra I: Basic Category Theory*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1994.

[Bub09] P. Bubenik, *Models and van Kampen theorems for directed homotopy theory*, Homology, Homotopy Appl. **11** (2009), no. 1, 185–202.

[Bub12] ———, *Simplicial models for concurrency*, Proceedings of the workshop om Geometric and Topological Methods in Computer Science (GETCO) 2010 (L. Fajstrup, É. Goubault, and M. Raussen, eds.), Electr. Notes in Theor. Comput. Sci., vol. 283, 2012, pp. 3–12.

[BW95] A. Björner and V. Welker, *The homology of "k-equal" manifolds and related partition lattices*, Adv. Math. **110** (1995), no. 2, 277–313.

[BW06] P. Bubenik and K. Worytkiewicz, *A model category for local po-spaces*, Homology, Homotopy Appl. **8** (2006), no. 1, 263–292.

[CES71] E.G. Coffman, M.J. Elphick, and A. Shoshani, *System deadlocks*, Comput. Surveys **3** (1971), no. 2, 67–78.

[CFK12] A.E. Costa, M. Farber, and Th. Kappeler, *Topics of stochastic algebraic topology*, Proceedings of the workshop om Geometric and Topological Methods in Computer Science (GETCO) 2010 (L. Fajstrup, É. Goubault, and M. Raussen, eds.), vol. 283, 2012, pp. 53–70.

[CHLT00] S. Chaudhuri, M. Herlihy, N.A. Lynch, and M.R. Tuttle, *Tight bounds for k-set agreement*, J. ACM **47** (2000), 912–943.

[CHT99] S. Chaudhuri, M. Herlihy, and M.R. Tuttle, *Wait-free implementations in message-passing systems*, Theoret. Comput. Sci. **220** (1999), 211–245.

[CR87] S.D. Carson and P.F. Reynolds, *The geometry of semaphore programs*, ACM TOPLAS **9** (1987), no. 1, 25–53.

[CZ09] G. Carlsson and A. Zomorodian, *The Theory of Multidimensional Persistence*, Discrete Comput. Geom. **42** (2009), 71–93.

[Dij68] E.W. Dijkstra, *Co-operating sequential processes*, Programming Languages (F. Genuys, ed.), Academic Press, New York, 1968, pp. 43–110.

[Dij71] ———, *Hierarchical ordering of sequential processes*, Acta Informatica **1** (1971), no. 2, 115–138.

[Dug66] J. Dugundji, *Topology*, Allyn & Bacon, 1966.

[EH10] H. Edelsbrunner and J. Harer, *Computational Topology. An Introduction*, American Mathematical Society, Providence, RI, 2010.

[Faj00] L. Fajstrup, *Loops, ditopology, and deadlocks*, Math. Structures Comput. Sci. **10** (2000), no. 4, 459–480.

[Faj03] ———, *Dicovering spaces*, Homology, Homotopy, Appl. **5** (2003), no. 2, 1–17.

[Faj05] ———, *Dipaths and dihomotopies in a cubical complex*, Adv. in Appl. Math **35** (2005), no. 2, 188–206.

[Faj06] ———, *Cubical local partial orders on cubically subdivided spaces - existence and construction*, Theor. Comput. Sci. **365** (2006), no. 3, 199–205.

[Faj10] ———, *Classification of dicoverings*, Topol. Appl. **157** (2010), 2402–2412.

[Faj13] ———, *Trace spaces of directed tori with rectangular holes*, Math. Struct. Comput. Sci. (2013), First View.

[Far02] M. Farber, *Zeros of closed 1-forms, homoclinic orbits, and Lusternik-Schnirelman theory*, Topol. Methods Nonlinear Anal. **10** (2002), 123–152.

[Far04] ———, *Topology of closed one-forms*, Mathematical Surveys and Monographs, vol. 108, American Mathematical Society, 2004.

[Far08] ———, *Invitation to Topological Robotics*, Zurich Lectures in Advanced Mathematics, EMS Publishing House, 2008.

[FGH+12] L. Fajstrup, É. Goubault, E. Haucourt, S. Mimram, and M. Raussen, *Trace Spaces: An Efficient New Technique for State-Space Reduction*, Programming Languages and Systems (H. Seidl, ed.), Lect. Notes in Comput. Sci., vol. 7211, ESOP 2012, Springer-Verlag, 2012, pp. 274–294.

[FGHR04] L. Fajstrup, É. Goubault, E. Haucourt, and M. Raussen, *Components of the fundamental category*, Appl. Categ. Structures **12** (2004), no. 1, 81–108.

[FGR98a] L. Fajstrup, É. Goubault, and M. Raussen, *Detecting Deadlocks in Concurrent Systems*, CONCUR '98; Concurrency Theory (Nice, France) (D. Sangiorgi and R. de Simone, eds.), Lect. Notes Comput. Sci., vol. 1466, Springer-Verlag, 1998, 9th Int. Conf., Proceedings, pp. 332–347.

[FGR98b] ———, *Detecting Deadlocks in Concurrent Systems*, DTA/LETI/DEIN/SLA 98-61, LETI (CEA - Technologies Avancées), Saclay, France, 1998, 25 pp.

[FGR06] ———, *Algebraic Topology and Concurrency*, Theor. Comput. Sci **357** (2006), 241–278, Revised version of Aalborg University preprint, 1999.

[FR96] L. Fajstrup and M. Raussen, *Detecting deadlocks in concurrent systems*, Technical report 1996-10, BRICS, Aarhus University, May 1996.

[FR07] U. Fahrenberg and M. Raussen, *Reparametrizations of continuous paths*, J. Homotopy Relat. Struct. **2** (2007), no. 2, 93–117.

[FR08] L. Fajstrup and J. Rosický, *A convenient category for directed homotopy*, Theory Appl. Categ. **21** (2008), 7–20.

[FS00] L. Fajstrup and S. Sokołowski, *Infinitely running concurrents processes with loops from a geometric viewpoint*, Electr. Notes Theor. Comput. Sci. **39** (2000), no. 2, 19 pp.

[Gau00] P. Gaucher, *Homotopy invariants of higher dimensional categories and concurrency in computer science*, Math. Structures Comput. Sci. **10** (2000), no. 4, 481–524.

[Gau01] ———, *Combinatorics of branchings in higher dimensional automata*, Theory Appl. Categ. **8** (2001), 324–376.

[Gau02] ———, *About the globular homology of higher dimensional automata*, Cah. Topol. Géom. Différ. Catég. **43** (2002), 107–156.

[Gau03a] ———, *Automate parallèle à homotopie près. I*, C. R. Math. Acad. Sci. Paris **336** (2003), 593–596.

[Gau03b] ———, *Automate parallèle à homotopie près. II*, C. R. Math. Acad. Sci. Paris **336** (2003), 647–650.

[Gau03c] ———, *The branching nerve of HDA and the Kan condition*, Theory Appl. Categ. **11** (2003), 75–106.

[Gau03d] ———, *A model category for the homotopy theory of concurrency*, Homology, Homotopy Appl. **5** (2003), no. 2, 549–599.

[Gau05a] ———, *Comparing globular complex and flow*, New York J. Math **11** (2005), 97–150.

[Gau05b] ———, *Flow does not model flows up to weak dihomotopy*, Appl. Categ. Structures **13** (2005), 371–388.

[Gau05c] ———, *Homological properties of non-deterministic branchings of mergings in higher dimensional automata*, Homology, Homotopy Appl. **7** (2005), 51–76.

[Gau06a] ———, *Inverting weak dihomotopy equivalence using homotopy continuous flow*, Theory Appl. Categ. **16** (2006), no. 3, 59–83.

[Gau06b] ———, *T-homotopy and refinement of observation. III. Invariance of the branching and merging homologies*, New York J. Math. **12** (2006), 319–348.

[Gau07] ———, *T-Homotopy and refinement of observation. II. Adding new T-homotopy equivalences*, Int. J. Math. Math. Sci. (2007), 20 pp.

[Gau08a] ———, *Globular realization and cubical underlying homotopy type of time flow of process algebra*, New York J. Math. **14** (2008), 101–137.

[Gau08b] ———, *Towards a homotopy theory of process algebra*, Homology, Homotopy Appl. **10** (2008), 353–388.

[Gau09] ———, *Homotopical interpretation of globular complex by multipointed d-space*, Theory Appl. Categ. **22** (2009), 588–621.

[Gau10a] ———, *Combinatorics of labelling in higher-dimensional automata*, Theoret. Comput. Sci. **411** (2010), 1452–1483.

[Gau10b] ———, *Directed algebraic topology and higher dimensional transition systems*, New York J. Math. **16** (2010), 409–461.

[Gau11] ———, *Towards a homotopy theory of higher dimensional transition systems*, Theory Appl. Categ. **25** (2011), 295–341.

[GG03] P. Gaucher and É. Goubault, *Topological Deformation of Higher Dimensional Automata*, Homology, Homotopy, Appl. **5** (2003), no. 2, 39–82.

[GH05] É. Goubault and E. Haucourt, *A Practical Application of Geometric Semantics to Static Analysis of Concurrent Programs*, CONCUR (M. Abadi and L. de Alfaro, eds.), Lect. Notes in Comput. Sci., vol. 3653, Springer, 2005, pp. 503–517.

[GH07] ———, *Components of the Fundamental Category II*, Appl. Categ. Structures **15** (2007), 387–414.

[GHK09] É. Goubault, E. Haucourt, and S. Krishnan, *Covering space theory for directed topology*, Theory Appl. Categ. **22** (2009), 252–268.

[GHK10] É. Goubault, E. Haucourt, and S. Krishnan, *Future path-components in directed topology*, Proceedings of the 26th Conference on the Mathematical Foundations of Programming Semantics (MFPS 2010), Electr. Notes Theor. Comput. Sci., vol. 265, 2010, pp. 325–335.

[GHP09] R. Guerraoui, M. Herlihy, and B. Pochon, *A topological treatment of early-deciding set-agreement*, Theoret. Comput. Sci. **410** (2009), 570–580.

[GHR06] E. Gafni, M. Herlihy, and S. Rajsbaum, *Subconsensus tasks: renaming is weaker than set agreement*, Distributed Computing, Lect. Notes in Comput. Sci., vol. 4167, 2006, pp. 329–338.

[GJ92] É. Goubault and T.P. Jensen, *Homology of higher-dimensional automata*, CONCUR '92 (Stony Brook, NY, 1992), Lect. Notes in Comput. Sci., vol. 630, 1992, pp. 254–268.

[GJ99] P.G. Goerss and J.F. Jardine, *Simplicial Homotopy Theory*, Basel: Birkhäuser, 1999.

[GM03] M. Grandis and L. Mauri, *Cubical sets and their site*, Theory Appl. Categ. **11** (2003), no. 8, 185–211.

[Gou93] É. Goubault, *Domains of higher-dimensional automata*, CONCUR '93 (Hildesheim, 1993), Lect. Notes in Comput. Sci., vol. 715, 1993, pp. 293–307.

[Gou95] ———, *The Geometry of Concurrency*, Ph.D. thesis, Ecole Normale Superieure, Paris, 1995.

[Gou00] ———, *Geometry and concurrency: A user's guide*, Math. Structures Comput. Sci. **10** (2000), no. 4, 411–425.

[Gou01] ———, *Géométrie du parallelisme, analyse statique et autres applications*, Mémoire d' habilitation à diriger les recherches, Dein/SLA, CEA/Saclay, F-91191 Gif-sur-Yvette, France, 2001.

[Gou03] ———, *Some Geometric Perspectives in Concurrency Theory*, Homology, Homotopy, Appl. **5** (2003), no. 2, 95–136.

[GR02] É. Goubault and M. Raussen, *Dihomotopy as a tool in state space analysis*, LATIN 2002: Theoretical Informatics (Cancun, Mexico) (S. Rajsbaum, ed.), Lect. Notes Comput. Sci., vol. 2286, Springer-Verlag, 2002, pp. 16–37.

[Gra06] M. Grandis, *The shape of a category up to directed homotopy*, Theory Appl. Categ. **15** (2005/06), 95–146.

[Gra02] ______, *Directed Homotopy Theory II. Homotopy Constructs*, Theory Appl. Categ. **10** (2002), no. 14, 369–391.

[Gra03a] ______, *Directed Homotopy Theory I. The Fundamental Category*, Cah. Topol. Géom. Differ. Catég. **44** (2003), 281–316.

[Gra03b] ______, *Higher fundamental groupoids for spaces*, Topology Appl. **129** (2003), 281–299.

[Gra03c] ______, *Ordinary and directed combinatorial homotopy, applied to image analysis and concurrency*, Homology, Homotopy Appl. **5** (2003), 211–231.

[Gra04] ______, *Inequilogical spaces, directed homology and noncommutative geometry*, Homology Homotopy Appl. **6** (2004), 413–437.

[Gra05] ______, *Directed combinatorial homology and noncommutative tori (the breaking of symmetries in algebraic topology*, Math. Proc. Cambridge Philos. Soc. **138** (2005), 233–262.

[Gra06a] ______, *Lax 2-categories and directed homotopy*, Cah. Topol. Géom. Differ. Catég. **47** (2006), 107–128.

[Gra06b] ______, *Modelling fundamental 2-categories for directed homotopy*, Homology, Homotopy Appl. **8** (2006), 31–70.

[Gra06c] ______, *Quotient models of a category up to directed homotopy*, Theory Appl. Categ. **16** (2006), 709–735.

[Gra07] ______, *Directed algebraic topology, categories and higher categories*, Appl. Categ. Structures **15** (2007), 341–353.

[Gra09] ______, *Directed Algebraic Topology*, New Mathematical Monographs, Cambridge University Press, Cambridge, 2009.

[Gun94] J. Gunawardena, *Homotopy and concurrency*, Bulletin of the EATCS **54** (1994), 184–193.

[Gun96] ______, *New connections between mathematics and computer science:*, Reports, Abstracts and Bibliography of a Workshop, HPL-BRIMS-96-02, Basic Research Institute in the Mathematical Sciences, HP Laboratories Bristol, 1996.

[GZ67] P. Gabriel and M. Zisman, *Calculus of fractions and homotopy theory*, Springer-Verlag, New York, 1967, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 35.

[Hat02] A. Hatcher, *Algebraic Topology*, Cambridge Univ. Press, 2002.

[Hau06] E. Haucourt, *Categories of components and loop-free categories*, Theory Appl. Categ. **16** (2006), 736–770.

[Her10] M. Herlihy, *Applications of shellable complexes to distributed computing*, CONCUR 2010 – Concurrency Theory, Lect. Notes in Comput. Sci., vol. 6269, 2010, pp. 19–20.

[HHH13] A. Hirschowitz, M. Hirschowitz, and T. Hirschowitz, *Saturating directed spaces*, J. Homotopy Relat. Struct. **8** (2013), to appear.

[HKR13] M. Herlihy, D. Kozlov, and S. Rajsbaum, *Distributed Computing through Combinatorial Topology*, Morgan Kaufmann, 2013, to appear.

[HR95] M. Herlihy and S. Rajsbaum, *Algebraic topology and distributed computing – a primer*, Computer science today, Lect. Notes in Comput. Sci., vol. 1000, Springer, 1995, pp. 203–217.

[HR99] ———, *New perspectives in distributed computing*, Mathematical Foundations of Computer Science 1999 (Szklarska Poreba) (170–186), Lect. Notes in Comput. Sci., vol. 1672, 1999.

[HR00] ———, *Algebraic spans*, Math. Structures Comput. Sci. **10** (2000), 549–573.

[HRT98] M. Herlihy, S. Rajsbaum, and M. Tuttle, *Unifying synchronous and asynchronous message-passing models*, Principles of Distributed Computing, 1998.

[HRT00] ———, *An Overview of synchronous Message-Passing and Topology*, Electronic Notes Theor. Comput. Sci. **39** (2000), no. 2, 1–17.

[HS99] M. Herlihy and N. Shavit, *The topological structure of asynchronous computability*, J. ACM **46** (1999), 858–923.

[Jar02] J.F. Jardine, *Cubical homotopy theory: a beginning*, Tech. report, Newton Institute, Cambridge, UK, 2002, NI02030-NST.

[Jar10] ———, *Path categories and resolutions*, Homology, Homotopy, Appl. **12** (2010), no. 2, 231–244.

[Kah06] Th. Kahl, *Relative directed homotopy theory of partially ordered spaces*, J. Homotopy Relat. Struct. **1** (2006), 79–100.

[Kah12] ———, *Some collapsing operations for 2-dimensional precubical sets*, J. Homotopy Relat. Struct. **7** (2012), no. 2, 281–298.

[KMM04] T. Kaczynski, K. Mischaikow, and M. Mrozek, *Computational homology*, Applied Mathematical Sciences, vol. 157, Springer-Verlag, New York, 2004.

[Koz08] D. Kozlov, *Combinatorial Algebraic Topology*, Algorithms and Computation in Mathematics, vol. 21, Springer, 2008.

[Kri09] S. Krishnan, *A convenient category of locally preordered spaces*, Appl. Categ. Structures **17** (2009), no. 5, 445–466.

[Kri13] ———, *Cubical Approximation for Directed Topology I*, Appl. Categ. Structures (2013), to appear.

[LP81] W. Lipski and C.H. Papadimitriou, *A fast algorithm for testing for safety and detecting deadlocks in locked transaction systems*, J. Algorithms **2** (1981), 211–226.

[Mil59] J. Milnor, *On spaces having the homotopy type of a CW-complex*, Trans. Amer. Math. Soc. **90** (1959), no. 2, 272–280.

[Mun75] J.R. Munkres, *Topology: a first course*, Prentice-Hall Inc., 1975.

[Nac65] L. Nachbin, *Topology and Order*, Van Norstrand, 1965.

[Pen72] R. Penrose, *Techniques of Differential Topology in Relativity*, Conference Board of the Mathematical Sciences, Regional Conference Series in Applied Mathematics, vol. 7, SIAM, Philadelphia, USA, 1972.

[Pra91] V. Pratt, *Modelling concurrency with geometry*, Proc. of the 18th ACM Symposium on Principles of Programming Languages. (1991), 311– 322.

[Rau00] M. Raussen, *On the classification of dipaths in geometric models for concurrency*, Math. Structures Comput. Sci. **10** (2000), no. 4, 427–457.

[Rau03] ______, *State spaces and dipaths up to dihomotopy*, Homology, Homotopy Appl. **5** (2003), no. 2, 257–280.

[Rau06] ______, *Deadlocks and dihomotopy in mutual exclusion models*, Theoret. Comput. Sci. **365** (2006), no. 3, 247–257.

[Rau07] ______, *Invariants of directed spaces*, Appl. Categ. Structures **15** (2007), 355–386.

[Rau09a] ______, *Reparametrizations with given stop data*, J. Homotopy Relat. Struct. **4** (2009), no. 1, 1–5.

[Rau09b] ______, *Trace spaces in a pre-cubical complex*, Topology Appl. **156** (2009), no. 9, 1718–1728.

[Rau10] ______, *Simplicial models for trace spaces*, Algebr. Geom. Topol. **10** (2010), no. 3, 1683–1714.

[Rau12a] ______, *Execution spaces for simple higher dimensional automata*, Appl. Algebra Engrg. Comm. Comput. **23** (2012), 59–84.

[Rau12b] ______, *Simplicial models for trace spaces II: General higher-dimensional automata*, Algebr. Geom. Topol. **12** (2012), no. 3, 1745–1765.

[RS71] C.P. Rourke and B.J. Sanderson, *Δ-sets I. Homotopy theory*, Quart. J. Math. Oxford Ser. (2) **22** (1971), 321–338.

[RZ13] M. Raussen and K. Ziemiański, *Homology of spaces of directed paths on Euclidean cubical complexes*, J. Homotopy Relat. Struct. (2013), 18 pages, DOI 10.1007/s40062-013-0045-4.

[SC70] A. Shoshani and E.G. Coffman, *Sequencing tasks in multiprocess systems to avoid deadlocks*, Eleventh Annual Symposium on Switching and Automata Theory (Santa Monica, CA, USA), no. 225–235, IEEE, 1970.

[Seg68] G. Segal, *Classifying spaces and spectral sequences*, Inst. Hautes Études Sci. Publ. Math. **34** (1968), 105–112.

[Ser51] J.P. Serre, *Homologie singulière des espaces fibrés*, Ann. of Math. **(2) 54** (1951), 425–505.

[Sma57] S. Smale, *A Vietoris Mapping Theorem for Homotopy*, Proc. Amer. Math. Soc. **8** (1957), no. 3, 604–610.

[Srb01] J. Srba, *On the Power of Labels in Transition Systems*, CONCUR 2001 (Aalborg, Denmark) (K.G. Larsen and M. Nielsen, eds.), Lect. Notes Comp. Science, vol. 2154, Springer-Verlag, 2001, pp. 277–291.

[Sul77] D. Sullivan, *Infinitesimal computations in topology*, Publ. Math. Inst. Hautes Études Sci. **47** (1977), 269–331.

[SZ00] M. Saks and F. Zaharoglou, *Wait-free k-set agreement is impossible: The topology of public knowledge*, SIAM J. Comput. **29** (2000), no. 5, 1449–1483.

[vG91] R. van Glabbeek, *Bisimulation semantics for higher dimensional automata*, Tech. report, Stanford University, 1991.

[vG06] ______, *On the Expressiveness of Higher Dimensional Automata*, Theor. Comput. Sci. **368** (2006), no. 1-2, 168–194.

[Wis05] R. Wisniewski, *Flow lines under perturbations within section cones*, Ph.D. thesis, Aalborg University, 2005.

[WN95] G. Winskel and M. Nielsen, *Models for concurrency*, Handbook of Logic in Computer Science, Oxford University Press, Oxford, 1995, pp. 1–148.

[WR07] R. Wisniewski and M. Raussen, *Geometric analysis of nondeterminacy in dynamical systems*, Acta Inform. **43** (2007), no. 7, 501–519.

[ZC05] A. Zomorodian and G. Carlsson, *Computing persistent homology*, Discrete Comput. Geom. **33** (2005), 249–274.

[Zie12a] K. Ziemiański, *Categories of directed spaces*, Fund. Math. **217** (2012), no. 1, 55–71.

[Zie12b] ———, *A cubical model for path spaces in d-simplicial complexes*, Topology Appl. **159** (2012), no. 8, 2127–2145.

[Zom05] A. Zomorodian, *Topology for Computing*, Cambridge monographs on applied and computational mathematics, vol. 16, Cambridge Univ. Press, 2005.

Dansk Resumé

Denne afhandling omhandler bidrag til et nyt forskningsfelt: *Algebraisk Topologi med retning*. Den giver et overblik over sytten forskningsartikler som blev publiceret inden for området i årene 1998 – 2013.

Algebraisk topologi med retning er et forholdsvis nyt forskningsområde. I "klassisk" algebraisk topologi (med sine mange forbindelser til geometri, analyse, algebra, fysik mv.) ræsonnerer man abstrakt i termer af kombinatorik, algebra og kategorier om emner af en geometrisk natur. Når man tager retninger med i betragtning, tillader man *ikke alle* geometriske stier; specielt kan man som regel ikke invertere en sti ("løbe baglæns"). Dette ødelægger muligheden for at oversætte umiddelbart til de fleste vante algebraiske strukturer.

Indtil videre kommer den vigtigste motivation for rum og stier med retning fra teorien om parallelitet ("concurrency") i den teoretiske datalogi: Når flere processer kan arbejde sig igennem hver sit program uden at der på forhånd er fastlagt en rækkefølge ligger det ikke fjernt at finde på geometriske modeller hvor stierne har en retning: Tiden kan ikke gå baglæns! Et vigtigt spørgsmål som blev behandlet i starten af arbejdet var hvordan man algoritmisk hurtigt finder "deadlocks" og tilhørende usikre regioner hvorfra man ikke kan nå i mål.

Man kan både undersøge specifikke modeller og også mere generelle rum hvis egenskaber det gælder om at kaste lys på. Det giver sig selv at man ofte *ikke umiddelbart* kan bruge metoderne fra den "klassiske" algebraiske topologi til ret meget. Man kan ikke engang definere en fundamental*gruppe* for stier med retning. Men nogle af fagets metoder kan alligevel bringes i anvendelse – når først man har stillet spørgsmålene rigtigt.

En stor del af undersøgelserne i afhandlingen handler om at beskrive og analysere *rum* af d-stier – stier har en retning, rum af stier har ikke – med udgangspunkt i viden om det såkaldte tilstandsrum som stierne bevæger sig i. Desuden ønskes information om stirummenes indbyrdes sammenhæng givet information om endepunkterne.

Man kan sige en del "kategorielt" om dette emne (se især publikationen [**Rau07**] for ret generelle d-rum). Men når man skal foretage egentlige udregninger af topologiske invarianter for stirum, så lykkes det indtil videre kun for konkrete tilstandsrum; dem som er modeller for parallele beregninger. Her er det til gengæld lykkedes at beskrive en direkte vej fra en model af tilstandsrummet til beskrivelse af stirummet (udførelser af et parallelt program) som et kombinatorisk beskrevet *simplicial kompleks* (se [**Rau10**, **Rau12a**, **Rau12b**, **FGH+12**]) med relevante topologiske invarianter. Den sidstnævnte artikel beskriver implementeringen af metoden i en praktisk anvendelig algoritme!

Appendix A: Thesis papers

Detecting Deadlocks in Concurrent Systems

Lisbeth Fajstrup[1], Eric Goubault[2*] and Martin Raußen[1]

[1] Dept of Mathematics, Aalborg University
DK-9220 Aalborg
{fajstrup,raussen} at math.auc.dk
[2] LETI (CEA - Technologies Avancées)
DEIN-SLA, CEA F91191 Gif-sur-Yvette Cedex
goubault at aigle.saclay.cea.fr

Abstract. We study deadlocks using geometric methods based on generalized process graphs [Dij68], i.e., cubical complexes or Higher-Dimensional Automata (HDA) [Pra91,vG91,GJ92,Gun94], describing the semantics of the concurrent system of interest. A new algorithm is described and fully assessed, both theoretically and practically and compared with more well-known traversing techniques. An implementation is available, applied to a toy language. This algorithm not only computes the deadlocking states of a concurrent system but also the so-called "unsafe region" which consists of the states which will eventually lead to a deadlocking state. Its basis is a characterization of deadlocks using dual geometric properties of the "forbidden region".

1 Introduction and related work

This paper deals with the detection of deadlocks motivated by applications in data engineering, e.g., scheduling in concurrent systems. Many fairly different techniques have been studied in the numerous literature on deadlock detection. Unfortunately, they very often depend on a particular (syntactic) setting, and this makes it difficult to compare them. Some authors have tried to classify them and test the existing software, like [Cor96,CCA96], but for this, one needs to translate the syntax used by each of these systems into one another, and different translation choices can make the picture entirely different. Nevertheless, we will follow their classification to put our methods in context. Notice that in this article, we go one step beyond and also derive the "unsafe region" i.e. the set of states that are bound to run into a deadlocking state after some time. This analysis is done in order to be applied to finding schedulers that help circumvent these deadlocking behaviours (and not just for proving deadlock freedom as most other techniques have been used for).

The first basic technique is a *reachability search*, i.e., the traversing of some semantic representation of a concurrent program, in general in terms of transition systems, but also sometimes using other models, like Petri nets [MR97].

* Work done partly while at Ecole Normale Supérieure and while visiting Aalborg University.

Due to the classical problem of *state-space explosion* in the verification of concurrent software, such algorithms are accompanied with state-space reduction techniques, such as *virtual coarsening* (which coalesce internal actions into adjacent external actions) [Val89], *partial-order techniques* (which alleviate the effects of representation with interleaving by pruning "equivalent" branches of search) such as *sleep sets* and *permanent (or stubborn) sets* techniques [Val91], [GPS96], [GHP95], and *symmetry techniques* (that reduce the state-space by consideration of symmetry). These techniques only reduce the state-space up to three or four times except for very particular applications.

The second most well-known technique is based on *symbolic model-checking* as in [BCM+90,GJM+97,BG96]. Deadlocking behaviors are described as a logical formula, that the model-checker tries to verify. In fact, the way a model-checker verifies such formulae is very often based on clever traversing techniques as well. In this case, the states of the system are coded in a symbolic manner (BDDs etc.) which enables a fast search.

Then many of the remaining techniques are a blend of one of these two with some abstractions, or are *compositional techniques* [YY91], or based on *dataflow analysis* [DC94], or on *integer programming techniques* [ABC+91] (but this in general only relies on necessary conditions for deadlocking behaviors).

Based on some old ideas [Dij68] and some new semantic grounds [Pra91], [vG91], [Gun94], [GJ92], [Gou95a] (see §2), we have developped an enhanced sort of reachability search (§2.3). This should mostly be compared to ordinary reachability analysis and not to virtual coarsening and symmetry techniques because these can also be used on top of ours. A first approach in the direction of virtual coarsening has actually been made in [Cri95]. Some assessments about its practical use, based on a first implementation applied to simple semaphore programs are made in §4.5. Due to the page limit, we have not fully described this algorithm. We chose to focus on the really new aspect of deadlock detection using a geometric semantics.

In §3, we propose a new algorithm based on an *abstraction* (in the sense of *abstract interpretation* [CC77,CC92]) of the natural semantics, which takes advantage of the real *geometry of the executions*. This one is an entirely different method from those in the literature.

As a matter of fact, in recent years, a number of people have used ideas from geometry and topology to study concurrency: First of all, using geometric models allows one to use spatial intuition; furthermore, the well-developed machinery from geometric and algebraic topology can serve as a tool to prove properties of concurrent systems. A more detailed description of this point of view can be found in J. Gunawardena's paper [Gun94] – including many more references – which contains a first geometrical description of *safety* issues. In another direction, techniques from algebraic topology have been applied by M. Herlihy, S. Rajsbaum, N. Shavit [HS95,HS96] and others to find new *lower bounds* and *impossibility results* for distributed and concurrent computation.

We believe that this technique, which is assessed in §4.4 and §4.5 both on theoretical grounds and on the view of benchmarks, can be applied in the static analysis of "real" concurrent programs (and not only at the PV language of §2.3) by suitable compositions and reduced products with other abstract interpretations.

The authors participated in the workshop "New Connections between Mathematics and Computer Science" at the Newton Institute at Cambridge in November 1995. We thank the organizers for the opportunity to get new inspiration. This paper is the first in a series of papers resulting from the collaboration of two mathematicians (L. Fajstrup & M. Raussen) and a computer scientist (E. Goubault).

2 Models of concurrent computation

2.1 From Discrete to Continuous

A description of deadlocks in terms of the geometry of the so-called progress graph (cf. Ex. 1) has been given earlier by S. D. Carson and P. F. Reynolds [CR87], and we stick to their terminology. The main idea in [CR87] is to model a *discrete* concurrency problem in a *continuous geometric* set-up: A system of n concurrent processes will be represented as a subset of Euclidean space $\mathbb{R}^n$ with the usual partial order. Each coordinate axis corresponds to one of the processes. The state of the system corresponds to a point in $\mathbb{R}^n$, whose i'th coordinate describes the state (or "local time") of the i'th processor. An execution is then a *continuous increasing path* within the subset from an initial state to a final state.

Example 1. Consider a centralized database, which is being acted upon by a finite number of transactions. Following Dijkstra [Dij68], we think of a transaction as a sequence of P and V actions known in advance – locking and releasing various records. We assume that each transaction starts at (local time) 0 and finishes at (time) 1; the P and V actions correspond to sequences of real numbers between 0 and 1, which reflect the order of the P's and V's. The initial state is $(0,\dots,0)$ *and the final state is* $(1,\dots,1)$. *An example consisting of the two transactions* $T_1 = P_aP_bV_bV_a$ *and* $T_2 = P_bP_aV_aV_b$ *gives rise to the two dimensional* progress graph *of Figure 1.*

The shaded area represents states, which are not allowed in any execution path, since they correspond to mutual exclusion. Such states constitute the *forbidden area*. An *execution path* is a path from the initial state $(0,0)$ to a final state $(1,1)$ avoiding the forbidden area and increasing in each coordinate - time cannot run backwards.

In Ex. 1, the dashed square marked "Unsafe" represents an *unsafe area*: There is no execution path from any state in that area to the final state $(1,1)$. Moreover, its extent (upper corner) with coordinates (Pb, Pa) represents a *deadlock*. Likewise, there are no execution paths starting at the initial state $(0,0)$ entering the *unreachable area* marked "Unreachable". Concise definitions of these concepts will be given in §2.2.

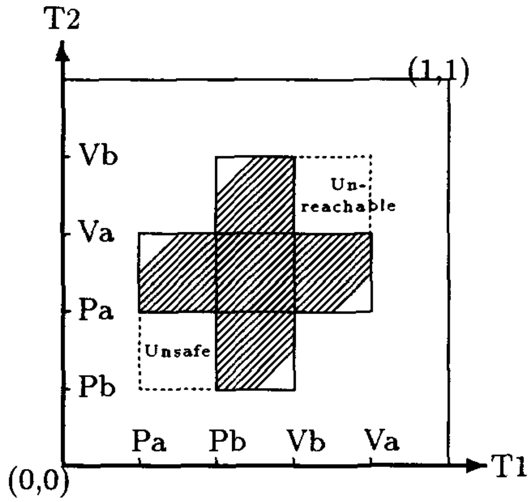


Fig. 1. Example of a progress graph

Finding deadlocks and unsafe areas is hence the geometric problem of finding n-dimensional "corners" as the one in Ex. 1. Back in 1981, W. Lipski and C. H. Papadimitriou [LP81] attempted to exploit geometric properties of forbidden regions to find deadlocks in database-transaction systems. But the algorithm in [LP81] does not generalize to systems composed of more than two processes. S. D. Carson and P. F. Reynolds indicated in [CR87] an iterative procedure identifying both deadlocks and unsafe regions for systems with an arbitrary finite number of processes.

In this section, we present a streamlined path to their results in a more general situation: Basic properties of the geometry of the state space are captured in properties of a *directed graph* – back in a discrete setting. In particular, *deadlocks* correspond to *local maxima* in the associated partial order.

This set-up does not only work for semaphore programs: In general, the forbidden area may represent more complicated relationships between the processes like for instance general k-semaphores, where a shared object may be accessed by k, but not $k+1$ processes. This is reflected in the geometry of the forbidden area F, that has to be a *union of higher dimensional rectangles* or "boxes".

Furthermore, similar partially ordered sets can be defined and investigated in more general situations than those given by Cartesian progress graphs. By the same recipe, deadlocks can then be found in concurrent systems with a variable number of processes involved or with branching (tests) and looping (recursion) abilities. In that case, one has to consider partial orders on sets of "boxes" of variable dimensions. This allows the description and detection of deadlocks in the *Higher Dimensional Automata* of V. Pratt [Pra91] and R. van Glabbeek [vG91] (cf. E. Goubault [Gou95a] for an exhaustive treatment).

In the mathematical parts below, i.e., §2.2 and §2.3, the explanations have been voluntarily simplified. The full treatment of the deadlock detection method can be found on the Web (http://www.dmi.ens.fr/ goubault/analyse.html).

2.2 The continuous setup

Let I denote the unit interval, and $I^n = I_1 \times \cdots \times I_n$ the unit cube in n-space. This is going to represent the space of all local times taken by n processes. We call a subset $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ an *n-rectangle*[1], and we consider a set $F = \bigcup_1^r R^i$ that is a finite union of n-rectangles $R^i = [a_1^i, b_1^i] \times \cdots \times [a_n^i, b_n^i]$. The interior $\overset{\circ}{F}$ of F is the "forbidden region" of I^n; its complement is $X = I^n \setminus \overset{\circ}{F}$. Furthermore, we assume that $\mathbf{0} = (0, \ldots, 0) \notin F$, and $\mathbf{1} = (1, \ldots, 1) \notin F$.

Definition 1. • 1. A continuous path $\alpha : I \to I^n$ is called a *dipath* (directed path) if *all* compositions $\alpha_i = pr_i \circ \alpha : I \to I$, $1 \le i \le n$, (pr_i being the projection on the ith coordinate of I^n) are increasing: $t_1 \le t_2 \Rightarrow \alpha_i(t_1) \le \alpha_i(t_2)$, $1 \le i \le n$.

• 2. A point $\mathbf{y} \in X = I^n \setminus \overset{\circ}{F}$ is in the *future* $J^+(\mathbf{x})$ of a point $\mathbf{x} \in X$ if there is a dipath $\alpha : I \to X$ with $\alpha(0) = \mathbf{x}$ and $\alpha(1) = \mathbf{y}$. The past $J^-(\mathbf{x})$ is defined similarly.

• 3. A point $\mathbf{x} \in I^n \setminus \overset{\circ}{F}$ is called *admissible*, if $\mathbf{1} \in J^+(\mathbf{x})$; and *unsafe* else.

• 4. Let $\mathcal{A}(F) \subset I^n$ denote the *admissible region* containing all admissible points in X, and $\mathcal{U}(F) \subset I^n$ the *unsafe region* containing all unsafe points in X.

• 5. A point $\mathbf{x} \in X$ is a *deadlock* if $J^+(\mathbf{x}) = \{\mathbf{x}\}$.

In semaphore programs, the n-rectangles R^i characterize states where two transactions have accessed the same record, a situation which is *not* allowed in such programs. Such "mutual exclusion"-rectangles have the property that only two of the defining intervals are proper subintervals of the I_j. Furthermore, serial execution should always be possible, and hence F should not intersect the 1-skeleton of I^n consisting of all edges in the unit cube. These special features will *not* be used in the present paper.

A dipath represents the continuous counterparts of the traces of the concurrent system, which must not enter the forbidden regions.

2.3 Continuous to discrete - a graph theory approach

We use geometrical ideas to construct a digraph (i.e. a directed graph) where deadlocks are the leaves (i.e. the nodes of the digraph, if any, that have no successors) and the unsafe region is found by an iterative process. The setup is as in §2.2. For $1 \le j \le n$, the set $\{a_j^i, b_j^i | 1 \le i \le r\} \subset I_j$ gives rise to a partition of I_j into at most $(2r+1)$ subintervals: $I_j = \bigcup I_{jk}$, with an obvious ordering $\le$ on the subintervals I_{jk}. The partition of intervals gives rise to a partition $\mathcal{R}$ of I^n into n-rectangles $I_{1k_1} \times \cdots \times I_{nk_n}$ with a partial ordering given by

$$I_{1k_1} \times \cdots \times I_{nk_n} \le I_{1k_1'} \times \cdots \times I_{nk_n'} \Leftrightarrow I_{jk_j} \le I_{jk_j'},\ 1 \le j \le n.$$

The partially ordered set $(\mathcal{R}, \le)$ can be interpreted as a *directed, acyclic graph*, denoted $(\mathcal{R}, \to)$: Two n-rectangles $R, R' \in \mathcal{R}$ are connected by an edge from R

[1] which has the property that all its faces are parallel to the coordinate axes. In dimension 2 this is called isothetic rectangles [Pre93]

to R' – denoted $R \rightarrow R'$ – if $R \leq R'$ and if R and R' share a face. R' is then called an *upper neighbor* of R, and R a *lower neighbor* of R'. A path in the graph respecting the directions will be denoted a *directed path*.

For any subset $\mathcal{R}' \subset \mathcal{R}$ we consider the *full* directed subgraph $(\mathcal{R}', \rightarrow)$. Particularly important is the subgraph $\mathcal{R}_{\bar{F}}$ consisting of all rectangles $R \subset X = I^n \setminus \mathring{F}$.

Definition 2. *Let $\mathcal{R}' \subset \mathcal{R}$ be a subgraph. An element $R \in \mathcal{R}'$ is a* local maximum *if it has no upper neighbors in $\mathcal{R}'$.* Local minima *have no lower neighbors. An n-rectangle $R \in \mathcal{R}_{\bar{F}}$ is called a* deadlock *rectangle if $R \neq R_1$, and if R is a local maximum with respect to $\mathcal{R}_{\bar{F}}$. An* unsafe *n-rectangle $R \in \mathcal{R}_{\bar{F}}$ is characterized by the fact, that* any *directed path α starting at R hits a deadlock rectangle sooner or later [CR87].*

In order to find the set $\mathcal{U}$ of all unsafe points – which is the union of *all* unsafe n-rectangles – apply the following. (1) Remove F from I^n giving rise to the directed graph $(\mathcal{R}_{\bar{F}}, \rightarrow)$. (2) Find the set S_1 of all deadlock n-rectangles (local maxima) with respect to $\mathcal{R}_{\bar{F}}$. Let $F_1 = F \cup S_1$. (3) Let $\mathcal{R}_{\overline{F_1}}$ denote the full directed subgraph on the set of rectangles in $I^n \setminus F_1$, i.e., after removing S_1. (4) Find the set S_2 of all deadlock n-rectangles with respect to $\mathcal{R}_{F_1}$. Let $F_2 = F_1 \cup S_2$. Carry on the same completion mechanism etc.

Notice that it is enough to search among the lower neighbors of elements in F in step 2, and that the only candidates for deadlocks in step 4 are the lower neighbors of elements of S_1. Since there are only *finitely many* rectangles, this process stops after a finite number of steps, ending with S_r and yielding the following result:

Theorem 1. • *1. The unsafe region is determined by $\mathcal{U}(F) = \bigcup_1^r S_i$.*

• *2. The set of admissible points is $\mathcal{A}(F) = I^n \setminus (\mathring{F} \cup \mathcal{U}(F))$. Moreover, any directed path in $\mathcal{A}(F)$ will eventually reach R_1.*

Implementation A prototype analyser has been programmed on the base of an HDA semantics of PV programs with the following syntax: Given a set of objects $\mathcal{O}$ (like shared memory locations, synchronization barriers, semaphores, control units, printers etc.) and a function $s : \mathcal{O} \rightarrow \mathbb{N}^+$ associating to each object a, the maximum number of processes $s(a) > 0$ which can access it at the same time, any process $Proc$ can try to access an object a by action Pa or release it by action Va, any finite number of times. In fact, processes are defined by means of a finite number of recursive equations involving process variables X in a set $\mathcal{V}$: they are of the form $X = Proc_d$ where $Proc_d$ is the process definition formally defined as,

$$\begin{aligned} Proc_d = \epsilon \mid Pa.Proc_d \quad & \mid Va.Proc_d \\ Proc_d + Proc_d & \mid Y \end{aligned}$$

(ϵ being the empty string, a being any object of $\mathcal{O}$, Y being any process variable in $\mathcal{V}$). A PV program is any parallel combination of these PV processes, $Prog =$

Proc | (*Proc* | *Proc*). The typical example in shared memory concurrent programs is $\mathcal{O}$ being the set of shared variables and for all $a \in \mathcal{O}$, $s(a) = 1$. The P action is putting a lock and the V action is relinquishing it. We will suppose in the sequel that any given process can only access once an object before releasing it. We also suppose that the recursive equations are "guarded" in the sense that for all process variables X, $Proc_X$ does not contain a summand of the form $X.T$, T being any non-empty term.

We deliver here only the theoretical and practical assessment of this "reference" algorithm (with which we are going to compare our new algorithm).

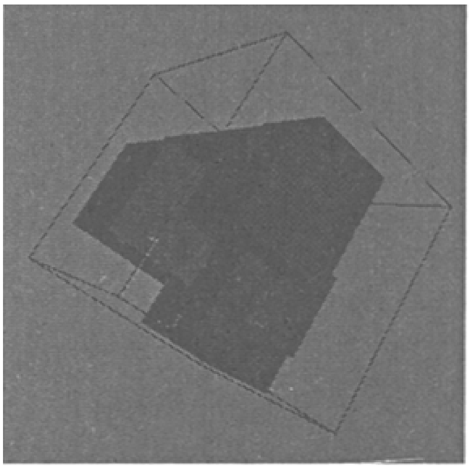


Fig. 2. The forbidden regions for 3phil

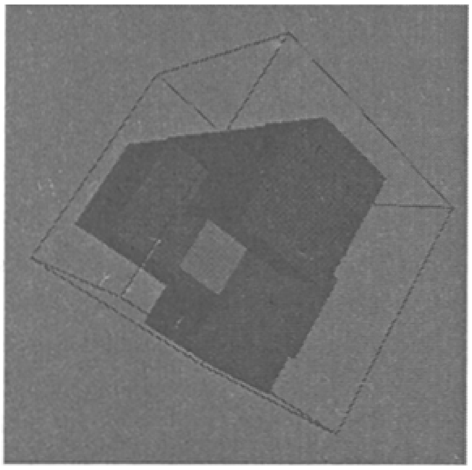


Fig. 3. Unsafe (center) region for 3phil

Algorithmic issues We let the *volume* $Vol(S)$ of a set S of nodes (n-rectangles) in $\mathcal{R}$ be the number of its elements. The dominant part of the algorithm is the removal of F and finding the deadlocks. To remove F and find S_1 one has to check for each $R \in R^i$ whether it is already marked in F. Only if the answer is no, the $2n$ operations of disconnecting R from its n sons and n parents and possibly, a single addition to, resp. removal (of R) from, the list of potential deadlocks, has to be performed. This implies:

Proposition 1. *For a pure term (i.e. no + nor any recursion) consisting of n transactions with a forbidden region* $F = \bigcup_1^r R^i$, *the worst case complexity of the algorithm of is of order* $nVol(F) + \Sigma_1^r Vol(R^i)$.

Examples reaching the worst case have a high amount of global synchronization, which in general should be avoided for good programming practice. Hence one would expect a much better behaviour in the average situation. In fact, if $nVol(F)$ is the dominating part, the complexity is at most nN (where N is the number of states).

3 Continuous to discrete - invoking the geometry

Using the combinatorial geometry of the *boundary* ∂F of the forbidden region, we are now going to describe the deadlocks in X and the unsafe regions associated to them in an efficient way.

Let again $\overset{\circ}{F} \subset I^n$ denote the forbidden region and let $X = I^n \setminus \overset{\circ}{F}$. In the sequel, we need the following *genericity* property of the n-rectangles in F:

If $i_1 \neq i_2$ and $\overset{\circ}{R^{i_1}} \cap \overset{\circ}{R^{i_2}} \neq \emptyset$, then $(a_j^{i_1} = a_j^{i_2} \Rightarrow a_j^{i_1} = 0$ and $b_j^{i_1} = b_j^{i_2} \Rightarrow b_j^{i_1} = 1,\ 1 \leq j \leq n)$.

This property ("no interior faces at the same level") is obviously satisfied for forbidden regions for "mutually exclusion" models, in particular for PV-models.

We want to include deadlocks on the boundary ∂I^n into our description: In a mutual exclusion model, points on ∂I^n stand for situations where not all processors have started their execution or where some of them already have terminated. To circumvent lengthy case studies – and with an eye to implementation – we slightly change our model in order to include the upper boundary $\partial_+(I^n) = \{\mathbf{x} \in I^n | \exists j : x_j = 1\}$ of I^n into the forbidden region. To this end, let $\tilde{I} = [0,2]$ and $I^n \subset \tilde{I}^n$.

Slightly changing the notation, let $\widetilde{R^i} = [0,2]^{i-1} \times [1,2] \times [0,2]^{n-i},\ 1 \leq i \leq n$, and shifting indices by n, $\widetilde{R^{n+1}}, \ldots, \widetilde{R^{n+r}}$ will denote the n-rectangles used in the previous model F of the forbidden region – modified to maintain genericity: If $b_j^i = 1$, then let $b_j^{i+n} = 2$. Then $\bigcup_1^n \widetilde{R^i} = \tilde{I}^n \setminus \overset{\circ}{I^n}$, and $\tilde{F} = F \cup \bigcup_1^n \widetilde{R^i} = \bigcup_{i=1}^{n+r} \widetilde{R^i}$. By an abuse of notation, we will from now on write R^i for $\widetilde{R^i}$ and F for $\tilde{F}$.

For any nonempty index set $J = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n+r\}$ define

$$R^J = R^{i_1} \cap \cdots \cap R^{i_k} = [a_1^J, b_1^J] \times \cdots \times [a_n^J, b_n^J]$$

with $a_j^J = \max\{a_j^i | i \in J\}$ and $b_j^J = \min\{b_j^i | i \in J\}$. This set is again an $n-$rectangle unless it is empty (if $a_j^k > b_j^l$ for some $1 \leq j \leq n$ and $k, l \in J$). Let $\mathbf{a}^J = [a_1^J, \ldots, a_n^J] = \min R^J$ denote the minimal point in that n-rectangle.

For every $1 \leq j \leq n$, we choose $\widetilde{a_j^J}$ as the "second largest" of the $a_j^{i_l}$, i.e., $\widetilde{a_j^J} = a_j^{i_s}$ with $a_j^{i_l} \leq a_j^{i_s} < a_j^J$ for $a_j^{i_l} \neq a_j^J$, and consider the "half-open" n-rectangle $U^J =]\widetilde{a_1^J}, a_1^J] \times \cdots \times]\widetilde{a_n^J}, a_n^J]$ "below" R^J.

Theorem 2. *1. A point* $\mathbf{x} \in X$ *is a deadlock if and only if* $\mathbf{x} \neq \mathbf{1}$ *and there is an* n-element *index set* $J = \{i_1, \ldots, i_n\}$, *with* $R^J \neq \emptyset$ *and* $\mathbf{x} = \mathbf{a}^J = \min R^J$.

2. If $\mathbf{x} = \min R^J$ *is a deadlock, then the "half-open"* $n-$*rectangle* U^J *is* unsafe, *i.e., every dipath in* I^n *from a point* $\mathbf{y} \in U^J$ *will eventually enter* $\overset{\circ}{F}$.

Proof.

1. Let $\mathbf{x} = \mathbf{a}^J = \min R^J$. Every element $\mathbf{y} = [a_1^J + \varepsilon_1, \ldots, a_n^J + \varepsilon_n]$, $\varepsilon_j \geq 0$ and $0 < \sum_1^n \varepsilon_i$ small, is contained in at least one of the sets $\overset{\circ}{R^{j_s}}$ and thus in $\overset{\circ}{F}$.

On the other hand, let $\mathbf{x} = [x_1, \dots, x_n] \in X$ be a deadlock. Then, for small values $\varepsilon > 0$, the element $\mathbf{x}^i = [x_1, \dots, x_i + \varepsilon, \dots, x_n]$ is contained in one of the sets $\overset{\circ}{R^{j_i}}$. Hence, $\mathbf{x} \in R^J$ with $J = \{j_1, \dots, j_n\}$. This set contains n different elements: If, e.g., $R^{j_1} = R^{j_2}$, then $\mathbf{x}^1 \notin \overset{\circ}{R^{j_1}}$!

Moreover, $\mathbf{x}$ is an element of the set $R^J \setminus \bigcup \overset{\circ}{R^{j_i}}$ consisting of the 2^n points with all coordinates either a_i^J or b_i^J. Obviously, the only possible deadlock point in this set is $\mathbf{x} = \mathbf{a}^J = \min R^J$.

2. Let $\alpha : I \to X$ be a dipath with $\alpha(t_0) \in U^J$ and $\alpha(t_2) \notin U^J$ for some $t_0 < t_2$. There has to be a maximal value $t_0 \le t_1 < t_2$ such that $\alpha(t_1) \in U^J$. Moreover, $\alpha(t_1) \in \partial_+ U^J = \{\mathbf{y} \in U^J | \exists k : \; y_k = a_k^J\}$, and thus $\alpha(t_1 + \varepsilon)$ is contained in one of the sets $\overset{\circ}{R^{j_i}}$ and thus in $\overset{\circ}{F}$. Contradiction!

□

As an immediate consequence, we get a criterion for deadlockfreeness that is easy to check:

Corollary 1. *A forbidden region $F = \bigcup_1^{n+r} R^i \subset I^n$ has a deadlockfree complement $X = I^n \setminus F$ if and only if for any index set $J = \{i_1, \dots, i_n\}$ with $|J| = n$*

$$R^J = R^{i_1} \cap \dots \cap R^{i_n} = \emptyset \text{ or } R^J = \{1\} \text{ or } \min R^J \in \overset{\circ}{F}.$$

In general, the n-rectangle $U_{\mathbf{a}}$ will be considerably larger than the n-rectangles from the graph algorithm; it will contain several of the n-rectangles in the partition $\mathcal{R}$. This is where we gain in efficiency: look at Figures 4, 5, 6 and 7. They describe the 3 iterations needed in the following streamlined algorithm, whereas the first algorithm needed 26 iterations (two for each thirteen unsafe 2-rectangles).

In analogy with the graph algorithm we can now describe an algorithm finding the *complete unsafe region* $U \subset I^n$ as follows: Find the set $\mathcal{D}$ of deadlocks in X and, for every deadlock $\mathbf{a} \in \mathcal{D}$, the unsafe n-rectangle $U_{\mathbf{a}}$. Let $F_1 = F \cup \bigcup_{\mathbf{a} \in \mathcal{D}} U_{\mathbf{a}}$. Find the set $\mathcal{D}_1$ of deadlocks in $X_1 = X \setminus F_1 \subset X$, and, for every deadlock $\mathbf{a} \in \mathcal{D}_1$, the unsafe n-rectangle $U_{\mathbf{a}}$. Let $F_2 = F_1 \cup \bigcup_{\mathbf{a} \in \mathcal{D}_1} U_{\mathbf{a}}$ etc.

This algorithm stops after a finite number n of loops ending with a set $U = F_n$ and such that $X_n = X \setminus U$ does no longer contain any deadlocks. The set $U \setminus \partial_-(U)$ consists precisely of the forbidden and of the unsafe points.

The example of Figure 4 demonstrates that there may be arbitrarily many loops in this second algorithm – even in the case of a 2-dimensional forbidden region associated to a simple PV-program: Obviously, the "staircase" in Figure 4 (corresponding to the PV term `example`, see Appendix A) producing more and more unsafe n-rectangles can be extended ad libitum by introducing extra rectangles R^i to F along the "diagonal".

We now show the applicability of the method by exemplifying it on our toy PV language.

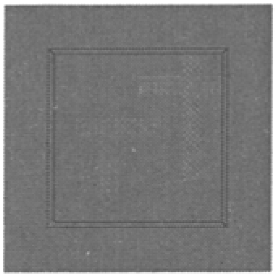


Fig. 4. The forbidden region

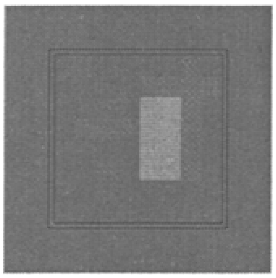


Fig. 5. First step of the algorithm

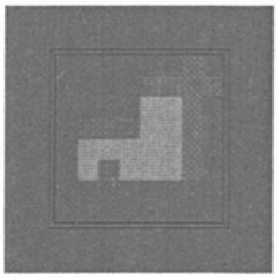


Fig. 6. Second step of the algorithm

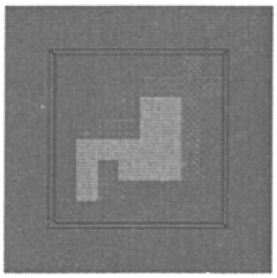


Fig. 7. Last step of the algorithm

4 Implementation of the geometric algorithm

4.1 The semantics

Now we have a dual view on PV terms. Instead of representing the allowed n-rectangles, we represent the forbidden n-rectangles only. Notice that up to now, we have only implemented the algorithm on pure terms (i.e. no recursion nor plus operator). The full treatment of the PV language and of real concurrent languages will be postponed in a forthcoming paper. Let $T = X_1 \mid \cdots \mid X_n$ (for some $n \geq 1$) be a pure term (i.e. no recursion nor plus operator) of our language such that all its subterms are pure as well. We consider here the X_i $(1 \leq i \leq n)$ to be strings made out of letters of the form Pa or Vb, $(a, b \in \mathcal{O})$. $X_i(j)$ will denote the jth letter of the string X_i. Supposing that the length of the strings X_i $(1 \leq i \leq n)$ are integers l_i, the semantics of $Prog$ is included in $[0, l_1] \times \cdots \times [0, l_n]$. A description of $[\![Prog]\!]$ from above can be given by describing inductively what should be digged into this n-rectangle. The semantics of our language can be described by the simple rule, $[k_1, r_1] \times \cdots \times [k_n, r_n] \in [\![X_1 \mid \cdots \mid X_n]\!]_2$ if there is a partition of $\{1, \cdots, n\}$ into $U \cup V$ with $card(U) = s(a) + 1$ for some object a with, $X_i(k_i) = Pa$, $X_i(r_i) = Va$ for $i \in U$ and $k_j = 0$, $r_j = l_j$ for $j \in V$.

4.2 The implementation

A general purpose library for manipulating finite unions of n-rectangles (for any n) has been implemented in C. A n-rectangle is represented as a list of n closed intervals. Regions (like the forbidden region) are represented as lists of n-rectangles. We also label some n-rectangles by associating to them a region. Labeled regions are then lists of such labeled n-rectangles. Notice that all this is quite naively implemented up to now. Much better algorithms can be devised (inspired by algorithms on isothetic rectangles [Pre93]) that reduce the complexity of intersection calculation a lot. This will be discussed in a forthcoming article.

Three arrays are constructed from the syntax in the course of computation of the forbidden region. For a process named `i` and an object (semaphore) named `j`, `tP[i][j]` is updated during the traversing of the syntactic tree to be equal to the ordered list of times at which process `i` locks semaphore `j`. Similarly tV[i][j] is updated to be equal to the ordered list of times at which process i unlocks semaphore j. Finally, an array t[i] gives the maximal (local) time that process i runs.

For all objects a, we build recursively all partitions as in §4.1 of $\{1, \cdots, n\}$ into a set U of $s(a)+1$ processes that lock a and V such that $U \cup V = \{1, \cdots, n\}$ and $U \cap V = \emptyset$. For each such partition (U, V) we list all corresponding pairs (Pa, Va) in each process X_i, $i \in U$. As we have supposed that in our programs, all processes must lock exactly once an item before releasing it, these pairs correspond to pairs (tP[i][a]$_j$,tV[i][a]$_j$) for j ranging over the elements of the lists tP[i][a] and tV[i][a]. Then we deduce the n-rectangle in the forbidden region for each partition and each such pair.

4.3 Implementation of the second deadlock algorithm

The implementation uses a global array of labeled regions called pile: pile[0], ..., pile[n-1] (n being the dimension we are interested in). The idea is that pile[0] contains at first the initial forbidden region, pile[1] contains the intersection of exactly two distinct regions of pile[0], etc., pile[n-1] contains the intersection of exactly n distinct regions of pile[0].

The algorithm is incremental. In order to compute the effect of adding a new forbidden n-rectangle S the program calls the procedure complete(S,∅). This calls an auxiliary function derive also described in pseudo-code below, in charge of computing the unsafe region generated by a possible deadlock created by adding S to the set of existing forbidden regions. The resulting forbidden and unsafe region is contained in pile[0].

```
complete(S,l)
   if S is included into an X in pile[0] return
   for i=n-2 to 0 by -1 do pile[i+1]=intersection(pile[i]\l,S)
   pile[0]=union(pile[0],S)
   for all X in pile[n-1] do pile[n-1]=pile[n-1]\X
                             derive(X)
```

The intersection of a labeled region R (such as pile[i] above) with a n-rectangle S gives the union of all intersections of n-rectangles X in R (which are also n-rectangles) labeled with the concatenation of the label of X with S (which is a region). Therefore labels of elements of regions in pile are the regions whose intersection is exactly these elements.

Now, derive(X) takes care of deriving an unsafe region from an intersection X of n forbidden or unsafe distinct n-rectangles. Therefore X is a labeled n-rectangle, whose labels is X1,...,Xn (the set of the n n-rectangles which it is the intersection of). We call X(i) the projection of X on coordinate i.

```
derive(X)
   for all i do yi=max({Xj(i) / j=1,...,n}\{X(i)})
   Y=[y1,X(1)]x...x[yn,X(n)]
   if Y is not included in one of the Xj complete(Y,(X1,...,Xn))
```

This last check is done when computing all `yi`. We use for each `i` a list `ri` of indexes `j` such that `yi=Xj(i)` (there might be several). If the intersection of all `ri` is not empty then `Y` is included into one of the `Xj`. It is to be noticed that this algorithm considers cycles (recursive calls) as representing (unbounded) finite computations.

4.4 Complexity issues

The entire algorithm consists of 3 parts: The first establishes the initial list `pile[0]` of forbidden n-rectangles, the second works out the complete array `pile` – including the deadlocks encoded in `pile[n-1]` –, and the third adds pieces of the unsafe regions, recursively.

Let again n denote the number of processes (the dimension of the state space), and r the number of n-rectangles. From a complexity viewpoint, the first step is negligeable; finding the n-rectangles involves $C^n_{s(a)+1}$ searches in the syntactic tree for every shared object a – in each of the n coordinates.

The array `pile` involves the calculation of $S(r,n) = \sum_{i=1}^{n} C^r_i$ intersections, each of them needing comparisons in n coordinates. Note that these comparisons show which of the intersections are empty, as well. To find the deadlocks, one has to compare (n coordinates of) the at most C^r_n non-empty elements in `pile[n-1]` with the r elements in `pile[0]`. Adding pieces of unsafe regions in the third step involves the same procedures with an increased number r of n-rectangles. The worst-case figure $S(r,n)$ above can be crudely estimated as follows: $S(r,n) \leq 2^r$ for all n, and $S(r,n) \leq nC^r_n$ for $r > 2n$ – which is a better estimate only for $r >> 2n$.

Remark that the algorithm above has a total complexity roughly proportional to the *geometric complexity* of the forbidden region. The latter may be expressed in terms of the *number of non-empty intersections* of elementary n-rectangles in the forbidden region. This figure reflects the degree of synchronization of the processes, and will be much lower that $S(n,r)$ for a well-written program. We conjecture, that the number of steps in *every* algorithm detecting deadlocks and unsafe regions is bounded below by this geometric complexity. On the other hand, for the analysis of big concurrent programs, this geometric complexity will be tiny compared to the number of states to be searched through by a traversing strategy.

4.5 Benchmarks

The program has been written in C and compiled using `gcc -O2` on an Ultra Sparc 170E with 496 Mbytes of RAM, 924 Mbytes of cache.

In the following table, dim represents the dimension of the program checked, #forbid2 is the number of forbidden n-rectangles found in the semantics of the program (to be compared with #forbid1, the number of unit cubes forbidden in the first semantics), t sem2 is the time it took to find these forbidden n-rectangles (respectively t sem1 is the time taken for the first semantics, looking at the enabled transitions), t unsafe2 is the time it took to find the unsafe region in the second algorithm (respectively in the first algorithm) and #unsafe is the number of n-rectangles found to be unsafe (they now encapsulate many of the "unit" n-rectangles found by the first deadlock detection algorithm). These measures have been taken on a first implementation which does not include yet the branching and looping constructs.

program	dim	#forbid2	#forbid1	t sem2	t sem1	t unsafe2	t unsafe1	#unsafe
example	2	4	14	0.020	0	0	0	3
stair2	2	6	16	0.020	0.01	0	0	15
stair3	3	18	290	0.010	0.180	0	.010	4
stair3'	3	6	80	0.030	0.640	0	0.020	0
lipsky	3	6	158	0.020	0.080	0	0	0
3phil	3	3	32	0.020	0	0	0	1
4phil	4	4	190	0.030	0.09	0	0	1
5phil	5	5	1048	0.030	0.820	0	0.020	1
6phil	6	6	5482	0.030	5.82	0	0.13	1
7phil	7	7	27668	0.040	42.35	0	0.86	1
16phil	16	16	NA	0.030	NA	0.030	NA	1
32phil	32	32	NA	0.030	NA	0.420	NA	1
64phil	64	64	NA	0.040	NA	1.520	NA	1
128phil	128	128	NA	0.100	NA	26.490	NA	1

5 Conclusion and future work

We have presented two algorithms for deadlock detection, including the computation of the set of states (the unsafe region) that will eventually lead to a deadlock. These algorithms were based on geometric intuition and techniques. They have been implemented, and the first one shows good comparison with ordinary reachability search with some state-space reduction techniques. But due to its complexity, this does not seem to be easily usable for very big programs (except if combined with clever abstract interpretations) or for a big number of processes (6 or 7 seems to be a maximum in general for practical use). The second algorithm has shown much better promise. Its complexity depends on the complexity of the synchronization of the processes, and not on a fake number of global states, as in most techniques used. In this regard it is much more practical. Dealing with 128 processes is not a problem if they are not synchronizing too much (as in the dining philosophers problem), but this is certainly intractable for reachability search with no clever partial order techniques (there are more than 10^{85} global states in that case). It should be noted also that these two algorithms could be enhanced by the use of some other well-known technique, like symmetry and (for the first one) some state-space reduction techniques. As the second algorithm is based on an abstract interpretation of the semantics, it should be developed for the use on real concurrent languages in conjunction with other well-known abstract interpretations. This is for future work. Also this should be linked with a full description of "schedules" and verification of safety properties of concurrent programs as hinted in [Gun94,Gou95b,FR96] using the geometric notions developed in this article.

Acknowledgments We used Geomview (see the Web page http://freeabel.geom.-umn.edu/software/ download/geomview.html/) to make the 3D pictures of this article (in a fully automated way).

A The examples detailed

You can check the implementations and the examples at http://www.dmi.ens.fr/-˜goubault/analyse.html.

- The dining philosophers' problem. The source below is for three philosophers, the next one is for five. The way others of these examples are generated should be obvious from these examples.

```
/* 3 philosophers ''3phil'' */
A=Pa.Pb.Va.Vb
B=Pb.Pc.Vb.Vc
C=Pc.Pa.Vc.Va
```

- This is example of Figure 4.

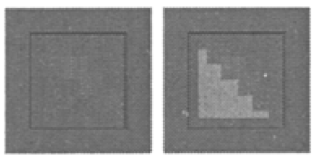
```
/* ''example'' */
A=Pa.Pb.Vb.Pc.Va.Pd.Vd.Vc
B=Pb.Pd.Vb.Pa.Va.Pc.Vc.Vd
```

- This is the classical Lipsky/Papadimitriou example (see [Gun94]) which produces no deadlock.

```
/* ''lipsky'' */
A=Px.Py.Pz.Vx.Pw.Vz.Vy.Vw
B=Pu.Pv.Px.Vu.Pz.Vv.Vx.Vz
C=Py.Pw.Vy.Pu.Vw.Pv.Vu.Vv
```

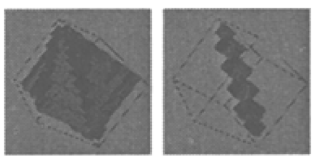
- This is a staircase (worst complexity case for the second algorithm).

```
/* ''stair2'' */
A=Pa.Pb.Va.Pc.Vb.Pd.Vc.Pe.Vd.Pf.Ve.Vf
B=Pf.Pe.Vf.Pd.Ve.Pc.Vd.Pb.Vc.Pa.Vb.Va
```



- This is a 3-dimensional staircase. Notice that if you declare all semaphores used (a, b, c, d, e and f) to be initialized to 2 (example "stair3"'), there is no 3-deadlock.

```
/* ''stair3'' */
A=Pa.Pb.Va.Pc.Vb.Pd.Vc.Pe.Vd.Pf.Ve.Vf
B=Pf.Pe.Vf.Pd.Ve.Pc.Vd.Pb.Vc.Pa.Vb.Va
C=Pf.Pe.Vf.Pd.Ve.Pc.Vd.Pb.Vc.Pa.Vb.Va
```



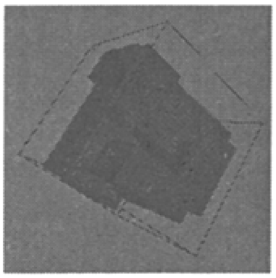


Fig. 8. The Lipsky/Papadimitriou example

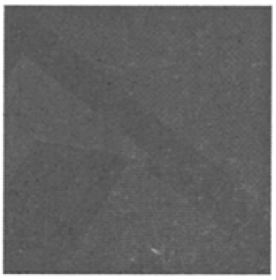


Fig. 9. A close-up to a hole in the forbidden region

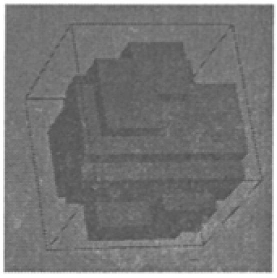


Fig. 10. Turning around

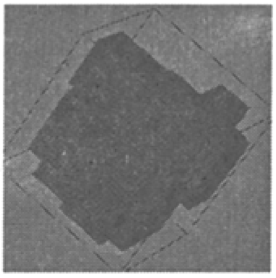


Fig. 11. Behind, notice the exit in the hole

References

[ABC+91] G. S. Avrunin, U. A. Buy, J. C. Corbett, L. K. Dillon, and J. C. Wileden. Automated analysis of concurrent systems with the constrained expression toolset. *IEEE Trans. Soft. Eng.*, 17(11):1204–1222, November 1991.

[BCM+90] J. R. Burch, E. M. Clarke, K. L. McMillan, D. L. Dill, and L. J. Hwang. Symbolic model checking: 10^{20} states and beyond. In *Proc. of the Fifth Annual IEEE Symposium on Logic and Computer Science*, pages 428–439. IEEE Press, 1990.

[BG96] B. Boigelot and P. Godefroid. Model checking in practice: An analysis of the access.bus protocol using spin. In *Proceedings of Formal Methods Europe'96*, volume 1051, pages 465–478. Springer-Verlag, Lecture Notes in Computer Science, March 1996.

[CC77] P. Cousot and R. Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction of approximations of fixed points. *Principles of Programming Languages 4*, pages 238–252, 1977.

[CC92] P. Cousot and R. Cousot. Abstract interpretation frameworks. *Journal of Logic and Computation*, 2(4):511–547, August 1992.

[CCA96] A. T. Chamillard, L. A. Clarke, and G. S. Avrunin. An empirical comparison of static concurrency analysis techniques. Technical Report 96-84, Department of Computer Science, University of Massachusetts, August 1996.

[Cor96] J. C. Corbett. Evaluating deadlock detection methods for concurrent software. *IEEE Transactions on Software Engineering*, 22(3), March 1996.

[CR87] S.D. Carson and P.F. Reynolds. The geometry of semaphore programs. *ACM TOPLAS*, 9(1):25–53, 1987.

[Cri95] R. Cridlig. Semantic analysis of shared-memory concurrent languages using abstract model-checking. In *Proc. of PEPM'95*, La Jolla, June 1995. ACM Press.

[DC94] M. B. Dwyer and L. A. Clarke. Data flow analysis for verifying properties of concurrent programs. In *Proc. of the Second Symposium on Foundations of Software Engineering*, pages 62–75, December 1994.

[Dij68] E.W. Dijkstra. Co-operating sequential processes. In F. Genuys, editor, *Programming Languages*, pages 43–110. Academic Press, New York, 1968.

[FR96] L. Fajstrup and M. Raußen. Some remarks concerning monotopy of increasing paths. unpublished manuscript, Aalborg University, 1996.

[GHP95] P. Godefroid, G. J. Holzmann, and D. Pirottin. State-space caching revisited. In *Formal Methods and System Design*, volume 7, pages 1–15. Kluwer Academic Publishers, November 1995.

[GJ92] E. Goubault and T. P. Jensen. Homology of higher-dimensional automata. In *Proc. of CONCUR'92*, Stonybrook, New York, August 1992. Springer-Verlag.

[GJM+97] H. Garavel, M. Jorgensen, R. Mateescu, Ch. Pecheur, M. Sighireanu, and B. Vivien. Cadp'97 - status, applications and perspectives. Technical report, Inria Alpes, 1997.

[Gou95a] E. Goubault. *The Geometry of Concurrency*. PhD thesis, Ecole Normale Supérieure, 1995. to be published, 1998, also available at http://www.dmi.ens.fr/~goubault.

[Gou95b] E. Goubault. Schedulers as abstract interpretations of HDA. In *Proc. of PEPM'95*, La Jolla, June 1995. ACM Press, also available at http://www.dmi.ens.fr/~goubault.

[GPS96] P. Godefroid, D. Peled, and M. Staskauskas. Using partial-order methods in the formal validation of industrial concurrent programs. *IEEE Transactions on Software Engineering*, 22(7):496–507, July 1996.

[Gun94] J. Gunawardena. Homotopy and concurrency. *Bulletin of the EATCS*, 54:184–193, 1994.

[HS95] M. Herlihy and S.Rajsbaum. Algebraic Topology and Distributed Computing. A Primer. volume 1000 of *Lecture Notes in Computer Science*. Springer-Verlag, 1995.

[HS96] M. Herlihy and N. Shavit. The topological structure of asynchronous computability. Technical report, Brown University, Providence, RI, January 1996.

[LP81] W. Lipski and C.H. Papadimitriou. A fast algorithm for testing for safety and detecting deadlocks in locked transaction systems. *Journal of Algorithms*, 2:211–226, 1981.

[MR97] S. Melzer and S. Roemer. Deadlock checking using net unfoldings. In *Proc. of Computer Aided Verification*. Springer-Verlag, 1997.

[Pra91] V. Pratt. Modeling concurrency with geometry. In *Proc. of the 18th ACM Symposium on Principles of Programming Languages*. ACM Press, 1991.

[Pre93] F. P. Preparata. *Computational Geometry, an Introduction*. Springer-Verlag, 1993.

[Val89] A. Valmari. Eliminating redundant interleavings during concurrent program verification. In *Proc. of PARLE*, volume 366, pages 89–103. Springer-Verlag, Lecture Notes in Computer Science, 1989.

[Val91] A. Valmari. A stubborn attack on state explosion. In *Proc. of Computer Aided Verification*, number 3, pages 25–41. AMS DIMACS series in Discrete Mathematics and Theoretical Computer Science, 1991.

[vG91] R. van Glabbeek. Bisimulation semantics for higher dimensional automata. Technical report, Stanford University, Manuscript available on the web as http://theory.stanford.edu/~rvg/hda, 1991.

[YY91] W. J. Yeh and M. Young. Compositional reachability analysis using process algebras. In *Proc. of the symposium on Testing, Analysis and Verification*, pages 178–187. ACM Press, October 1991.

Math. Struct. in Comp. Science (2000), *vol.* 10, *pp.* 427–457. Printed in the United Kingdom

On the classification of dipaths in geometric models for concurrency

MARTIN RAUSSEN

Department of Mathematical Sciences, Aalborg University,
Fredrik Bajers Vej 7E, DK - 9220 Aalborg Ø.
E-mail: raussen@math.auc.dk

Received 7 July 1999; revised 11 January 2000

This paper focusses on the determination of the dihomotopy classes of dipaths in cubical complexes, representing the essentially different computations in a given concurrent program due to different schedules. Several new notions have to be developed, for example, the domains of dependence (Definition 2.4), which are borrowed from relativity theory (Penrose 1972). However, it turns out that the algorithm determining deadlocks and unsafe regions described in Fajstrup *et al.* (1998a; 1998b) can be modified and applied to do the essential calculational work.

1. Introduction

1.1. *Background and history*

The use of geometric models in the description of the behaviour of concurrent systems can be traced back at least to the work of E. W. Dijkstra (Dijkstra 1968), where concurrent processes are modelled by so-called 'progress graphs' (some examples of which will be given in Section 1.2). Some properties about the order of actions (schedule) that a concurrent system can fire have been investigated in the work of S. D. Carson and P. F. Reynolds (Carson and Reynolds 1987) and W. Lipski and C. H. Papadimitriou (Lipski and Papadimitriou 1981) using these progress graphs for semaphore programs. These include the algorithmic determination of deadlock states and unreachable states. A systematic framework for studying schedules of actions of distributed computations by means of geometric properties has been proposed by V. Pratt (Pratt 1991) and subsequently R. van Glabbeek (van Glabbeek 1991). In his thesis (Goubault 1995), É. Goubault initiates a systematic study of Higher Dimensional Automata (HDA) built on cubical complexes (Serre 1951; Brown and Higgins 1981b; Brown and Higgins 1981a) employing methods from algebraic topology, in particular, homological methods. The idea is that a schedule of actions (including deadlocks and unreachables, but also serializability conditions *etc.*) is essentially the same under 'continuous deformation', that is, homotopy. Homology seemed back then to be a suitable computable invariant. Homotopical notions and arguments have been applied in the framework of concurrency by J. Gunawardena (see Gunawardena (1994)) and in Fajstrup *et al.* (1999). The introduction in the latter paper motivates the approach and its virtues, and includes several examples for applications. A new algorithm (which we will briefly recapitulate in Section 2.1) for computing deadlocks and unreachables using these geometric models of concurrency has been published in Fajstrup *et al.* (1998a; 1998b).

To put it briefly, an execution path in a concurrent system corresponds to a *dipath* (directed path), see Definition 1.5 below. Under what conditions will two execution paths (whatever the actual contents of data bases, buffers, *etc.*) give the same result; when are they essentially different, that is, risk giving different results? We argue in Fajstrup *et al.* (1999) that the corresponding dipaths have to be *dihomotopic*, see Definition 1.7 below, in the associated HDA-model (cubical complex) to guarantee equal results of the combined processes.

This paper focusses on the determination of the dihomotopy classes of dipaths in cubical complexes. In some simple cases, we can determine the number of essentially different computations in a given concurrent program due to different schedules. Several new notions have to be developed, for example, the domains of dependence (Definition 2.4), which are borrowed from relativity theory (Penrose 1972). But then it turns out that the algorithm determining deadlocks and unsafe regions described in Fajstrup *et al.* (1998a; 1998b) can be modified and applied to do the essential calculational work.

The author would like to thank the editor of this volume and the referees for many helpful suggestions during the preparation of this article.

1.2. *Ditopology*

This section contains the necessary general 'ditopological' (*di* for directed) notions assembled in Fajstrup *et al.* (1999). The essential framework for our study concerns topological spaces *with a partial order*.

1.2.1. *Basic definitions*

Definition 1.1.

1 A *partial order* $\leqslant$ on a set U is a reflexive, transitive and antisymmetric relation.

2 A partial order $\leqslant$ on a topological space X is said to be *closed* if $\leqslant$ is a closed subset of $X \times X$ in the product topology. In that case, $(X, \leqslant)$ is called a *po-space*.

Remark 1.2. Let $(X, \leqslant)$ denote a po-space.

1 For every $x \in X$, the regions $\downarrow x = \{y \in X | y \leqslant x\}$ and $\uparrow x = \{y \in X | y \geqslant x\}$ are closed.

2 For every pair of points $y_1, y_2 \in X$, the region $[y_1, y_2] = \{x \in X | y_1 \leqslant x \leqslant y_2\} = \downarrow y_2 \cap \uparrow y_1$ is closed.

3 A po-space is Hausdorff (Gierz *et al.* 1980).

Remark 1.3. To include loops in the study, one has to extend the framework to spaces with a *local* partial order, see Fajstrup *et al.* (1999, Section 3) and Fajstrup (1999). In this paper, we stick to *global* partial orders for simplicity and hope to come back to an extension of the present work to locally partially ordered spaces in future work.

Essential examples of (global or local) po-spaces are provided by cubical complexes and their geometric realizations. For precise definitions and a description of natural partial orders on those, we refer the reader to Fajstrup *et al.* (1999, Section 6). Intuitively, a cubical complex arises as the result of pasting together cubes of various dimensions along their boundaries. The (local) partial order is then the result of carefully pasting the natural partial orders on the individual cubes (or cells) together. This sometimes involves subdivisions of the original cells. In the concrete applications in the final parts of this paper, the cubical complex to be studied arises as a subcomplex X of the standard cube $I^n = [0,1]^n \subset \mathbf{R}^n$ with certain subcubes, constituting a forbidden region, deleted.

1.2.2. *Progress graphs give rise to po-spaces.* Progress graphs naturally give rise to po-spaces. Consider, for instance, a system of n concurrent processes. It will be represented as a subset of the Euclidean space $\mathbf{R}^n$ with the usual partial order. Each coordinate axis corresponds to one of the processes. The state of the system corresponds to a point in $\mathbf{R}^n$, whose ith coordinate describes the state (or 'local time') of the ith processor. An execution is then a *continuous increasing path* within the subset from an initial state to a final state.

Example 1.4. Consider a centralized database, which is being acted upon by a finite number of transactions. Following Dijkstra (Dijkstra 1968), we think of a transaction as a sequence of P and V actions known in advance – locking and releasing various records. We assume that each transaction starts at (local time) 0 and finishes at (time) 1; the P and V actions correspond to sequences of real numbers between 0 and 1, which reflect the order of the P's and V's. The initial state is $(0,\ldots,0)$ and the final state is $(1,\ldots,1)$. An example consisting of the two transactions $T_1 = P_aP_bV_bV_a$ and $T_2 = P_bP_aV_aV_b$ gives rise to the two-dimensional *progress graph* of Figure 1.4.

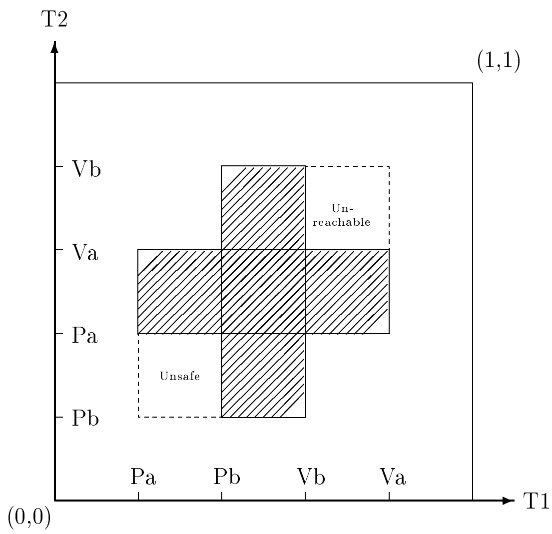


Fig. 1. Example of a progress graph

The shaded area represents states, which are not allowed in any execution path, since they correspond to mutual exclusion. Such states constitute the *forbidden region*. An *execution path* is a path from the initial state $(0,0)$ to the final state $(1,1)$, avoiding the forbidden region and increasing in each coordinate – time cannot run backwards.

In Example 1.4, the dashed square marked 'Unsafe' represents an *unsafe region*: there is no execution path from any state in that area to the final state $(1,1)$. Moreover, its extent (upper corner) with coordinates (Pb, Pa) represents a *deadlock*. Likewise, there are no execution paths starting at the initial state $(0,0)$ entering the *unreachable region* marked 'Unreachable'. Concise definitions of these concepts will be given later (Definitions 2.1 and 2.2).

More generally, the semantics of PV programs (abstraction of n concurrent processes sharing common resources) is given geometrically by some subshape of an n-dimensional cube. More precisely, let I denote the unit interval, and $I^n = \prod_{i=1}^n I$ denote the unit (hyper)cube in an n-dimensional Euclidean space. This is going to represent the space of all local times taken by n processes. We call a subset $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ a *hyperrectangle*. It actually has the property that all its faces are parallel to the coordinate axes. In dimension 2, this is called an *isothetic* rectangle in Preparata and Shamos (1988). The semantics of a PV program is given by a set $F = \bigcup_1^r R^i$, which is a finite union of hyperrectangles $R^i = [a_1^i, b_1^i] \times \cdots \times [a_n^i, b_n^i]$. The interior $int(F)$ of F is the 'forbidden region' within I^n.

Let us be a bit more formal: given a set of objects $\mathcal{O}$ (like shared memory locations, synchronization barriers, semaphores, control units, printers, *etc.*) and a function $s : \mathcal{O} \to \mathbb{N}^+$ associating to each object a, the maximum number of processes $s(a) > 0$ that can access it at the same time, any process $Proc$ can try to access an object a by action Pa or release it by action Va, any finite number of times. The programs we are considering here are processes running in parallel and defined as follows:

$$Proc \quad = \quad \epsilon \quad | \quad Pa.Proc \quad | \quad Va.Proc.$$

Its semantics is given as follows (see Fajstrup *et al.* (1998a)): let $T = X_1 \mid \cdots \mid X_n$ (for some $n \geqslant 1$) be a program. All the X_i $(1 \leqslant i \leqslant n)$ are strings made out of letters of the form Pa or Vb, $(a, b \in \mathcal{O})$. $X_i(j)$ will denote the jth letter of the string X_i.

Supposing that the length of the strings X_i $(1 \leqslant i \leqslant n)$ are integers l_i, the semantics of $Prog$ is included in the state space $[0, l_1] \times \cdots \times [0, l_n]$ (and after rescaling, it may be considered as a subspace of the unit cube I^n as above). A description of $[\![Prog]\!]$ can be given by describing inductively what should be removed from the state space. The semantics of our language can be described by the simple rule:

$[k_1, r_1] \times \cdots \times [k_n, r_n] \in [\![X_1 \mid \cdots \mid X_n]\!]$ if and only if there is a partition of $\{1, \cdots, n\}$ into $U \cup U'$ with $card(U) = s(a) + 1$ for some object a with:

for $i \in U$: $X_i(k_i) = Pa$, $X_i(r_i) = Va$ and $X_i(j) \notin \{Pa, Va\}$ for $k_i < j < l_i$;
for $j \in U'$: $k_j = 0$, $r_j = l_j$.

The space $[0, k_1] \times \cdots \times [0, k_n] \backslash [\![Prog]\!]$ together with the componentwise ordering in $\mathbf{R}^n$ is the po-space to consider.

1.2.3. *Dihomotopy and homotopy history*

Definition 1.5.

1 Let $(X, \leqslant)$ and $(Y, \leqslant')$ be *po-spaces*. A continuous map $f : X \to Y$ is called a *dimap* (directed map) if and only if it preserves partial orders, that is,

$$x_1 \leqslant x_2 \Rightarrow f(x_1) \leqslant f(x_2), \quad x_1, x_2 \in X.$$

2 A *dipath* in X is a dimap $\alpha : I \to X$ from the unit interval I with the natural (global) order $\leqslant$.
(Sometimes, we allow α to be parameterised by a different closed interval I as well).

Dipaths correspond to execution paths in HDA-models. We insist on directedness to model the irreversibility of time for every process. *Reachability* can be modelled with a *new* partial order using dipaths.

Definition 1.6. (Compare Penrose (1972) and Fajstrup *et al.* (1999)) A new local partial order $\prec$ on X is defined by $x \prec y$ if and only if there is a dipath from x to y in X.

The *equivalence* of execution paths can be modelled geometrically by an appropriate *homotopy* relation on dipaths. Again let $(X, \leqslant)$ be a partially ordered topological space, and let $X_0, X_1 \subset X$.

Definition 1.7.

1 A dipath in X from X_0 to X_1 is a dipath α in X with $\alpha(0) \in X_0$ and $\alpha(1) \in X_1$. The set of all dipaths from X_0 to X_1 will be denoted by $\vec{P}_1(X; X_0, X_1)$.
2 A *dihomotopy* from X_0 to X_1 is a continuous map $H : I \times I \to X$ such that *every* partial map $H_s : I \to X$, $H_s(t) = H(s,t)$, $s \in I$, is a dipath from X_0 to X_1.
3 Two dipaths α, β in X from X_0 to X_1 are *dihomotopic from* X_0 *to* X_1 if and only if there is a dihomotopy $H : I \times I \to X$ from X_0 to X_1 with $H_0 = \alpha$ and $H_1 = \beta$. We abbreviate this by $\alpha \vec{\simeq} \beta$ (from X_0 to X_1).
4 Dihomotopy from X_0 to X_1 is obviously an equivalence relation. The equivalence classes (*dihomotopy classes*) constitute the *dihomotopy set* $\vec{\pi}_1(X; X_0, X_1)$.

Remark 1.8.

1 In many relevant cases, X_0, respectively, X_1 will consist of a single point or a finite number of 'initial', respectively, 'final' points.
2 Concatenation of dipaths factors to yield concatenation on the level of dihomotopy classes $\vec{\pi}_1(X; X_0, X_1) \times \vec{\pi}_1(X; X_1, X_2) \to \vec{\pi}_1(X; X_0, X_2)$.
3 Dihomotopy corresponds to Pratt's *monoidal homotopy* (Pratt 1991).
4 Dihomotopy of paths also corresponds to the commutation of independent actions which plays an important role in Mazurkiewicz trace theory (Mazurkiewicz 1988).

The following notions allow one to reason about *reductions* of large *state spaces*. Two states are equivalent if every execution through one of them is equivalent to one through the other. This can be modelled by *history equivalence*, as follows.

Definition 1.9.

1 The homotopy history of a dipath $\alpha \in \vec{P}_1(X; X_0, X_1)$ is defined as

$$h\alpha := \{y \in X | \exists \text{ a dipath } \beta \in \vec{P}_1(X; X_0, X_1) \text{ through } y \text{ and } \alpha \overset{\sim}{\to} \beta\}.$$

2 Two elements x and y in X are *homotopy history equivalent* if and only if

$$x \in h\alpha \Leftrightarrow y \in h\alpha \text{ for all } \alpha \in \vec{P}_1(X; X_0, X_1).$$

3 The *diconnected components* of X from X_0 to X_1 consist of the connected components (in the classical sense) of the homotopy history equivalence classes of X.

Remark 1.10.

1 Paths in the same dihomotopy class have the same homotopy history.
2 Two points $x, y \in X$ are history equivalent if and only if *every* dipath through x is dihomotopic to one through y.
3 A homotopy history equivalence class C is an atom in the Boolean algebra generated by the homotopy histories, that is, it contains the points that are contained in the homotopy histories of certain dihomotopy classes but not in the others. More precisely, there is a subset $K \subset \vec{\pi}_1(X; X_0, X_1)$ such that

$$C = \bigcup_{\alpha \in K} h\alpha \setminus \bigcup_{\beta \notin K} h\beta.$$

Several examples of homotopy histories and homotopy equivalence classes are worked out in Fajstrup *et al.* (1999). For a po-space arising from a PV program, points correspond to the global states of all the running processes. Two global states are then homotopy history equivalent if *every* execution passing through the first global state can be replaced by an execution through the other that necessarily leads to the same output. In dimension 2 (two concurrent pocesses), two global states are homotopy history equivalent if we cannot distinguish them by just looking at all the sequences of lockings and unlockings that can *lead to* this state or that can be *done after* this state (and their effects).

For example (taken from Fajstrup *et al.* (1999)), the complement of the 'Swiss flag' in I^2 (see Figure 2) has ten homotopy history components that agree with the diconnected components of that po-space. The Swiss flag describes the semantics of the program having process $T_1 = Pa.Pb.Vb.Va$ in parallel with $T_2 = Pb.Pa.Va.Vb$ (where a and b are 1-semaphores). In region 1, we still have all possible futures (all possible access histories to a and b). In region 2, we can only go to 4 or to 6, meaning that we are going to deadlock in the future or T_2 will get a and b before T_1. In region 6, we can only come from 2 and go to 9: T_2 has got a and b before T_1. In region 9, we can 'come' from the unreachable region 7 or from 6. In region 10, we might have come from any history in the past.

2. General definitions and results

Let $(X, \leqslant)$ denote a po-space (Gierz *et al.* 1980; Fajstrup *et al.* 1999) with subspaces X_0 and X_1. Our aim is to contribute to the classification of dipaths up to dihomotopy in $\vec{\pi}_1(X; X_0, X_1)$ and, finally, to calculate this set in a special case in Section 4.

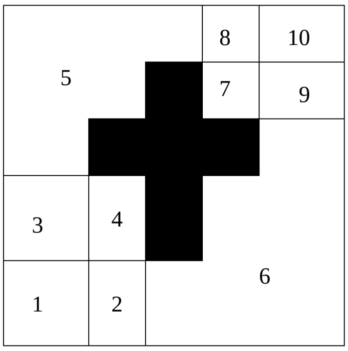


Fig. 2. The 'Swiss flag' example.

2.1. *Reachable, unsafe and unreachable regions, domains of dependence*

2.1.1. *Definitions.* We (re)define several classical notions (see Dijkstra (1968), Lipski and Papadimitriou (1981), Carson and Reynolds (1987) and Penrose (1972)) in the framework of po-spaces.

Definition 2.1. Let $(X, \leqslant)$ denote a po-space with subspaces $X_0, X_1, A_1, \dots A_k$, $k \geqslant 0$.

1 The *reachable* subspace $R(X; X_0, X_1; A_1, \dots, A_k)$ consists of all $x \in X$ such that there is a dipath $\alpha \in \vec{P}_1(X; X_0, X_1)$ and values $t_0, t_1, \dots, t_k \in I$ with $\alpha(t_0) = x$ and $\alpha(t_i) \in A_i$ for every i, $1 \leqslant i \leqslant k$.
2 The *singular* subspace $U(X; X_0, X_1; A_1, \dots, A_k)$ is the complement of the reachable subspace $R(X; X_0, X_1; A_1, \dots, A_k)$ within X.

Note that, for $k = 0$, we obtain, in particular, the definition of the subspaces $R(X; X_0, X_1)$ and $U(X; X_0, X_1)$ of reachable, respectively, unreachable points between X_0 and X_1. Particularly important are the following cases (see Figure 2 for an example).

Definition 2.2.

1 An element $x \in X$ with $\uparrow x = \{x\}$ is called a deadlock. The set of all deadlocks in X is denoted by $\mathscr{D}(X)$. (Sometimes, a particular final state is exempted from $\mathscr{D}(X)$).
2 The *unsafe region* $Uns(X; X_1) = U(X; X, X_1)$ consists of all $x \in X$ that *cannot be connected* to any point in X_1 by a dipath, that is,

$$Uns(X; X_1) = \{x \in X | \vec{P}_1(X; x, X_1) = \varnothing\} = X \setminus (\downarrow X_1) = \{x \in X | (\uparrow x) \cap X_1 = \varnothing\}.$$

3 The *unreachable region* $Unr(X; X_0) = U(X; X_0, X)$ consists of all $x \in X$ that *cannot be reached* from any point in X_0 by a dipath, that is,

$$Unr(X; X_0) = \{x \in X | \vec{P}_1(X; X_0, x) = \varnothing\} = X \setminus (\uparrow X_0) = \{x \in X | (\downarrow x) \cap X_0 = \varnothing\}.$$

4 The *singular region* $U(X; X_0, X_1)$ consists of all points that are *never touched* by a dipath from X_0 to X_1, that is,

$$U(X; X_0, X_1) = Uns(X; X_1) \cup Unr(X; X_0).$$

The singular region is the complement of the reachable region $R(X;X_0,X_1) = {=}\uparrow X_0 \cap \downarrow X_1$ within X.

5 The po-space X is called *reachable* with respect to X_0 and X_1 if $U(X;X_0,X_1) = \emptyset$, or, equivalently, if $R(X;X_0,X_1) = X$.

Remark 2.3.

1 The symbols $\uparrow$ and $\downarrow$ above have to be interpreted with respect to the partial order $\prec$ from Definition 1.6.

2 In a non-compact setting, it is sometimes more relevant to consider inextendible (future-endless, past-endless in Penrose (1972)) paths instead of paths from a given initial space to a given final space.

3 It follows from the definitions that

$$\vec{P}_1(X;X_0,X_1) = \vec{P}_1(R(X;X_0,X_1);X_0,X_1).$$

As a consequence,

$$\vec{\pi}_1(X;X_0,X_1) = \vec{\pi}_1(R(X;X_0,X_1);X_0,X_1),$$

that is, for classification purposes it is advisable first to 'delete' the singular region.

4 Consider the po-space associated to a PV-program discussed in Section 1.2.2 with final state $\mathbf{1}$. Then $Uns(X;\mathbf{1})$ corresponds exactly to the unsafe region of those states that can only reach a deadlock (different from $\mathbf{1}$).

2.1.2. *An algorithm detecting unsafe regions.* Unsafe and unreachable regions can be algorithmically determined in the PV model (Fajstrup *et al.* 1998a), and we recap here the basic idea of the algorithm.

Suppose the semantics of a PV program is given in terms of a forbidden region F containing forbidden hyperrectangles $R^i = \prod_{j=1}^n [a_j^i, b_j^i]$ (with $n \geqslant 2$). We assume, moreover, that the coordinates a_j^i are pairwise different for every $1 \leqslant j \leqslant n$ (geometrically, this is a genericity assumption). The relevant state space is $X = I^n \setminus int(F)$.

For any non-empty index set $J = \{i_1, \ldots, i_k\}$, define

$$R^J = R^{i_1} \cap \cdots \cap R^{i_k} = [a_1^J, b_1^J] \times \cdots \times [a_n^J, b_n^J]$$

with $a_j^J = \max\{a_j^i | i \in J\}$ and $b_j^J = \min\{b_j^i | i \in J\}$. This set is again an n—rectangle unless it is empty (if $a_j^k > b_j^l$ for some $1 \leqslant j \leqslant n$ and $k, l \in J$). Let $\mathbf{a}^J = [a_1^J, \ldots, a_n^J] = \min R^J$ denote the minimal point in that hyperrectangle.

For every $1 \leqslant j \leqslant n$, we choose $\widetilde{a_j^J}$ as the 'second largest' of the $a_j^{i_l}$, that is, $\widetilde{a_j^J} = a_j^{i_s}$ with $a_j^{i_l} \leqslant a_j^{i_s} < a_j^J$ for all $a_j^{i_l} \neq a_j^J$, and consider the 'half-open' hyperrectangle $U^J =]\widetilde{a_1^J}, a_1^J] \times \cdots \times]\widetilde{a_n^J}, a_n^J]$ 'below' R^J.

Then, deadlock points in the interior of I^n are exactly the points $\min R^J$ for some J of cardinality n (the number of processes, that is, the dimension of the geometric shape we are studying) such that $R^J \neq \emptyset$ and $\min R^J$ not contained in any R^i with $i \notin J$. Deadlock points on the boundary ∂I^n can be found using the same recipe after modification of the hyperrectangles used in the description (see Fajstrup *et al.* (1998a; 1998b)).

This description allows us to find the set $\mathscr{D}$ of deadlocks in X, and, for every deadlock $\mathbf{a} \in \mathscr{D}$ corresponding to a set of indices $J_{\mathbf{a}}$, the unsafe hyperrectangle $U^{J_{\mathbf{a}}}$. To detect the entire unsafe region, let $F_1 = F \cup \bigcup_{\mathbf{a} \in \mathscr{D}} U^{J_{\mathbf{a}}}$. Find the set $\mathscr{D}_1$ of deadlocks in $X_1 = X \setminus int(F_1) \subset X$, and, for every deadlock $\mathbf{a} \in \mathscr{D}_1$, the unsafe corresponding hyperrectangle $U^{J_{\mathbf{a}}}$. Let $F_2 = F_1 \cup \bigcup_{\mathbf{a} \in \mathscr{D}_1} U^{J_{\mathbf{a}}}$, and so on (see Figure 6 for an example).

This algorithm stops after a finite number l of loops ending with a set $U = F_l$ and such that $X_l = X \setminus int(U)$ no longer contain any deadlocks. The set U consists precisely of the forbidden and unsafe points.


Fig. 3. The forbidden region

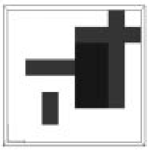


Fig. 4. First step of the algorithm

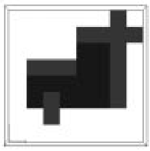


Fig. 5. Second step of the algorithm

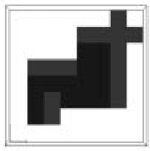


Fig. 6. Last step of the algorithm

2.1.3. *Domains of Dependence.* The following definition is adapted from Penrose (1972).

Definition 2.4. Let $S \subset X$ with $S \cap X_0 = S \cap X_1 = \emptyset$.
The *past*, the *future*, and the *total domain of dependence* of S are defined as follows:

1 $D^-(X; X_0, X_1; S) = \{x \in R(X; X_0, X_1) |$ every dipath from x to X_1 meets $S\}$.
2 $D^+(X; X_0, X_1; S) = \{x \in R(X; X_0, X_1) |$ every dipath from X_0 to x meets $S\}$.
3 $D(X; X_0, X_1; S) = D^-(X; X_0, X_1; S) \cup D^+(X; X_0, X_1; S)$.

The following lemma shows how to construct domains of dependence as *unsafe*, respectively, *unreachable* regions in *related* situations.

Lemma 2.5.

1 $D(X; X_0, X_1; S) = \{x \in R(X; X_0, X_1) |$ every dipath from X_0 to X_1 through x meets $S\}$.
2 $D^-(X; X_0, X_1; S) = R(X; X_0, X_1) \cap (S \cup Uns(X \setminus S; X_1)\}$.
3 $D^+(X; X_0, X_1; S) = R(X; X_0, X_1) \cap (S \cup Unr(X \setminus S; X_0)\}$.

Proof.

1 Clearly, both $D^-(X; X_0, X_1; S)$ and $D^+(X; X_0, X_1; S)$ are contained in the right-hand set. On the other hand, we have that the latter decomposes into its intersections with $\downarrow S$, which agrees with $D^-(X; X_0, X_1; S)$, and its intersection with $\uparrow S$ agrees with $D^+(X; X_0, X_1; S)$.
2 $\subseteq$: An element $x \in R(X; X_0, X_1)$ such that every dipath from x to X_1 meets S is either contained in S or it cannot be connected to X_1 by a dipath within $X \setminus S$.

$\supseteq$: An element $y \in X \setminus S$ that can be connected to X_1 by a dipath in X, but not in $X \setminus S$, can only be connected to X_1 by a dipath meeting S.
3 This is analogous. □

Corollary 2.6. If X is reachable with respect to X_0 and X_1 (see Definition 2.2.5), there is a simple characterization of the domains of dependence:

1 $D^-(X;X_0,X_1;S) = S \cup Uns(X \setminus S;X_1)$.
2 $D^+(X;X_0,X_1;S) = S \cup Unr(X \setminus S;X_0)$.

2.2. *Unavoidable regions and homotopy histories*

Definition 2.7.

1 Two (disjoint) subspaces $A_1, A_2 \subset X$ are called *unrelated* in X if there is no dipath connecting them, that is,

$$\vec{P}_1(X;A_1,A_2) = \vec{P}_1(X;A_2,A_1) = \emptyset.$$

2 A subset $S \subset X$ is called *unavoidable* with respect to X_0 and X_1 if every dipath α from X_0 to X_1 must hit S, that is, if $\alpha^{-1}(S) \neq \emptyset$ for every $\alpha \in \vec{P}_1(X;X_0,X_1)$. By Lemma 2.5.1, this is equivalent to the following condition: $D(X;X_0,X_1;S) = R(X;X_0,X_1)$.
3 A subspace S *filters* X with respect to X_0 and X_1 if every dipath from X_0 to X_1 hits *one and only one connected component* of S that, moreover, is invariant under dihomotopy. Equivalently, for every $\alpha \in \vec{P}_1(X;X_0,X_1)$, the homotopy history $h\alpha$ (see Definition 1.9) and every dipath in $h\alpha$ have non-empty intersection with precisely one connected component.

The following two lemmas describe the main *use* of filtering subspaces and a sufficient *condition* for an unavoidable subspace to be filtering.

Lemma 2.8. A subspace S filtering X with respect to X_0 and X_1 gives rise to a natural map $c_S : \vec{\pi}_1(X;X_0,X_1) \to \pi_0(S)$ into the set of connected components of S, associating to a dihomotopy class *the* component it hits.

Lemma 2.9. Let $S \subset X$ denote a subspace that is unavoidable with respect to X_0 and X_1. Assume, moreover, that all of the connected components S_j of S are unrelated to each other and either all open or all closed in X. Then, S filters X with respect to X_0 and X_1.

Proof. Every dipath hits *at least* one component of S, since S is unavoidable; it hits *only* one, since the components are unrelated to each other. Moreover, this component is invariant under the dihomotopy relation.

Let $H : I^2 \to X$ denote a dihomotopy, and let $I_j = \{s \in I | H_s^{-1}(S_j) \neq \emptyset\} \subset I$ for every connected component S_j. We conclude from the argument above that every $s \in I$ is contained in a uniquely determined set I_j, that is, the (homotopy) interval I decomposes into *disjoint* sets $I = \bigcup_j I_j$. With $p : I^2 \to I$ denoting the projection to the (homotopy) interval I, we obtain: $I_j = p(H^{-1}(S_j))$. Since H is continuous and p is both closed and open, the subsets $I_j \subset I$ are either all closed or all open. Since I is connected, all but one of them has to be empty. □

It is natural to ask how much of the dihomotopy information is uncovered by the map(s) c_S for one or several suitably chosen subspaces S. First of all, the component that a dipath hits gives information on the *homotopy history* (see Definition 1.9) of that dipath.

This allows us to reason about reachable regions (see Definition 2.1) in a more general context.

Let $S_1, \ldots, S_k \subset X$ denote a finite collection of mutually disjoint subspaces each filtering X with respect to X_0 and X_1. Let $A_i \subset S_i$, $1 \leqslant i \leqslant k$, denote a union of connected components of S_i, and let $B_i = S_i \setminus A_i$. Each of the regions $R(X; X_0, X_1; A_1, \ldots, A_k)$ is then a *union of homotopy histories* of dipaths, see Definitions 1.9 and 2.7.3. (Particularly interesting is the case with each A_i being just *one* connected component of S_i.) In Section 4, we show in a particular case how $A_1 \subset S_1, \ldots, A_K \subset S_k$ can be chosen such that $R(X; X_0, X_1; A_1, \ldots, A_k)$ is *a single* homotopy history.

Proposition 2.10. With the notation above, the reachable region $R(X; X_0, X_1; A_1, \ldots, A_k)$ and the domain of dependence $D(X; X_0, X_1; \bigcup_i B_i))$ are complements to each other within $R(X; X_0, X_1)$.

Proof.

1 $R(X; X_0, X_1; A_1, \ldots, A_k) \cap D(X; X_0, X_1; \bigcup_i B_i)) = \emptyset$:
Assume $x \in D(X; X_0, X_1; \bigcup_i B_i))$. By Lemma 2.5.1, every dipath α from X_0 to X_1 through x must meet at least one of the sets B_j. Since $S_j = A_j \cup B_j$ filters X, it cannot meet A_j; in particular, $x \notin R(X; X_0, X_1; A_1, \ldots, A_k)$.

2 $R(X; X_0, X_1; A_1, \ldots, A_k) \cup D(X; X_0, X_1; \bigcup_i B_i)) = R(X; X_0, X_1)$:
Assume $x \in R(X; X_0, X_1) \setminus R(X; X_0, X_1; A_1, \ldots, A_k)$. Any dipath β from X_0 to X_1 through x has to miss at least one of the A_j. Since $S_j = A_j \cup B_j$ filters X, the dipath β has to meet B_j, that is, $x \in D(X; X_0, X_1; \bigcup_i B_i))$. □

For components of filtering subspaces, this result yields the following description of domains of dependence by reachable and singular regions in related situations.

Proposition 2.11. Let S denote a subspace of X filtering X with respect to X_0 and X_1. Furthermore, let S be the disjoint union of two subsets A and B such that $A = \bigcup A_i$ is a finite union of connected components A_i of S. Then

$$D(X; X_0, X_1; B) = R(X; X_0, X_1) \cap \bigcap_i U(X; X_0, X_1; A_i).$$

Proof. By Proposition 2.10, $D(X; X_0, X_1; B) = R(X; X_0, X_1) \setminus R(X; X_0, X_1; A) = R(X; X_0, X_1) \setminus [\bigcup_i R(X; X_0, X_1; A_i)]$. By de Morgan reasoning, this latter set equals $R(X; X_0, X_1) \cap \bigcap_i \{x \in X | x \notin R(X; X_0, X_1; A_i)\} = R(X; X_0, X_1) \cap \bigcap_i U(X; X_0, X_1; A_i)$. □

2.3. (*Local*) *di-1-connectedness*

In Section 4, we will need a criterion to decide whether, in certain (sub)spaces, all dipaths between two given points are dihomotopic. In this section, we develop the necessary tools in a framework that is narrower than po-spaces, but more general than needed in these applications. Let $(X, \leqslant)$ denote a po-space.

2.3.1. *Parameterized po-spaces*

Definition 2.12.

1 A subspace $V \subset X$ is called *achronal* if for all $x, y \in V : x \leqslant y \Rightarrow x = y$; compare Penrose (1972) and Fajstrup *et al.* (1999).
2 Let $f : X \to \mathbf{R}$ denote a continous map. We write

$$X_t := f^{-1}(t), \quad X_{\geqslant t} := f^{-1}([t, \infty[), \quad X_{\leqslant t} := f^{-1}(]-\infty, t]), \quad X_I := f^{-1}(I)$$

for an interval $I \subset \mathbf{R}$.
3 A dimap $f : X \to \mathbf{R}$ is called a *parametrization* if X_t is achronal for every $t \in \mathbf{R}$. The triple $(X, \leqslant, f)$ is called a parametrized po-space.

Example 2.13. A subspace $X \subset \mathbf{R}^n$ with the natural partial order $\leqslant$ has the parametrization $f : X \to \mathbf{R}$, $f(x_1, \dots, x_n) = \sum_1^n x_i$.

Definition 2.14. Let $(X, \leqslant, f)$ denote a parametrized po-space.

1 A dipath $\alpha : J \to X$ is called *well-parameterised* if $f(\alpha(s)) = s$ for every $s \in J$, with J a closed interval.
2 A dihomotopy $H : I \times J \to X$ is called well-parameterised if every dipath $H_s = H(s, -) : J \to X$, $s \in I$ is well-parameterised.
3 A dipath $\beta : J' \to X$ is called a *reparametrisation* of another dipath α if there is a *strictly monotonic* continuous map $h : J' \to J$ such that $\beta = \alpha \circ h$.

Note that, for a well-parameterised dipath α and a subinterval $J' \subset J$, the restriction $\alpha_{J'} = \alpha|J' : J' \to X$ to J' is again well-parametrised.

Almost as in any course on elementary differential geometry, we easily get the following proposition (see Fajstrup *et al.* (1999)).

Proposition 2.15.

1 To any dipath $\alpha : J \to X$, there is exactly one well-parameterised reparametrisation $\beta : J' \to X$.
2 To any dihomotopy $H : I \times J \to X$ there is (exactly one) well-parameterised reparametrisation $\bar{H} : I \times J' \to X$.

2.3.2. *Local to global on parameterized po-spaces*

Definition 2.16. Let $(X, \leqslant, f)$ denote a parameterized po-space (see Definition 2.12). It is said to have the following properties:

uniform local di-1-connectedness property (ULD1P) if there is an $\varepsilon > 0$ such that for every $x \in X$ there is a neighbourhood $x \in U(x) \subset X$ such that for every two well-parameterized dipaths α_1, α_2 in $U(x)$ from x to $X_{f(x)+\delta}$ with $0 < \delta \leqslant \varepsilon$ there is a well-parametrized dihomotopy (from $X_{f(x)}$ to $X_{f(x)+\delta}$) connecting them.

uniform local dihomotopy extension property (ULDEP) if there is an $\varepsilon > 0$ such that for every $x \in X$ there is a neighbourhood $x \in U(x) \subset X$ with: given a *path* γ in $U(x) \cap X_{f(x)}$ and a well-parameterized *di*path α from $X_{f(x)}$ to $X_{f(x)+\delta}$ with $0 <$

$\delta \leqslant \varepsilon$ starting at the same point in $U(x) \cap X_{f(x)}$. Then there is a well-parameterized dihomotopy $H : I \times [f(x), f(x) + \delta] \to X$ with $H(0, t) = \alpha(t)$ and $H(s, 0) = \gamma(s)$.

di-1-connectedness property (D1P) between levels l **and** d if $|\pi_1(X; x_0, X_d)| \leqslant 1$ for every point $x_0 \in X_l$, that is, if there are any dipaths between x_0 and X_d, they are dihomotopic.

globally (shorter: X is di-1-connected) if X satisfies D1P between any levels $l, d \in \mathbf{R}$.

dihomotopy extension property (DEP) between levels l **and** d if, for every point $x_0 \in X_l$, every *path* γ in X_l and every well-parameterized *di*path $\alpha : [l, d] \to X$ both starting at x_0, there is a well-parameterized dihomotopy $H : [l, d] \times I \to X$ with $H(0, t) = \alpha(t)$ and $H(s, 0) = \gamma(s)$.

globally if X satisfies DEP between any levels $l, d \in \mathbf{R}$.

The dihomotopy extension property is illustrated in Figure 7.

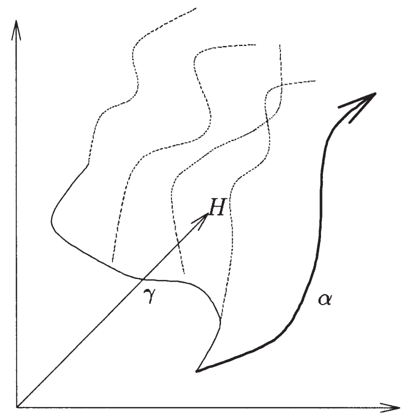


Fig. 7. The dihomotopy extension property

Remark 2.17.

1 For an interpretation of (a condition ensuring) the properties above in computer scientific terms look at Section 2.3.3, and, in particular, Remark 2.21.
2 If the po-space X is *compact*, it is enough to require that the first two properties are satisfied with an (individual) $\varepsilon(x) > 0$ for every $x \in X$. A finite refinement of the covering $U(x)$, $x \in X$, allows you to choose a uniform $\varepsilon > 0$ as the minimum of these finitely many ones.

Proposition 2.18. Let $(X, \leqslant, f)$ be a parameterized po-space satisfying ULD1P and ULDEP. Then, X is di-1-connected and satisfies DEP.

Proof. We have to prove that X satisfies D1P and DEP between arbitrary levels $l, d \in \mathbf{R}$. Fix $d \in \mathbf{R}$. Nothing has to be shown for $l > d$, and the properties are trivially true for $l = d$. For $l < d$, we are done after finitely many applications of the following inductive step starting at $l = d$:

If X satisfies D1P and DEP between levels l and d, then it satisfies D1P and DEP between levels $l-\delta$ and d (with $0<\delta\leqslant\varepsilon$ as in ULD1P and ULDEP).

This inductive step can be proved as follows:

D1P Let $x_0\in X_{l-\delta}$, and let α^1,α^2 denote two dipaths from x_0 to X_d. Using ULD1P, there is a well-parameterized dihomotopy $H^1: I\times[l-\delta,l]\to X_{\geqslant l-\delta}$ connecting $\alpha^1_{[l-\delta,l]}$ and $\alpha^2_{[l-\delta,l]}$. Let $\gamma(s)=H^1(s,l)$ denote the 'end' path in $X_l(!)$ between $\alpha^1(l)$ and $\alpha^2(l)$. Since X satisfies DEP between l and d, there is a well-parameterized dihomotopy $H^2: I\times[l,d]\to X_{\geqslant l}$ from $\alpha^1_{[l,d]}$ along γ to a dipath $\bar\alpha^1:[l,d]\to X_{\geqslant l}$ with $\bar\alpha^1(l)=\alpha^2(l)$. Since X satisfies D1P between l and d, we conclude that there is a dihomotopy $H^3: I\times[l,d]$ between $\bar\alpha^1$ and $\alpha^2_{[l,d]}$. This allows us to construct a well-parameterized dihomotopy $H: I\times[l-\delta,d]\to X_{\geqslant l-\delta}$ by the pasting scheme in Figure 8:

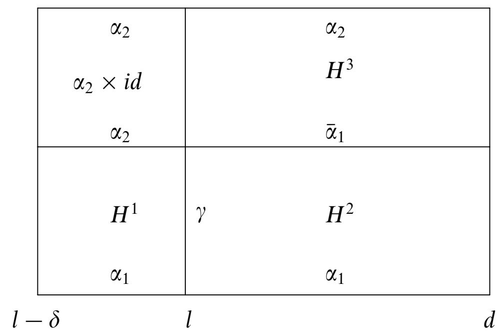


Fig. 8. Pasting scheme 1

DEP Let $\alpha:[l-\delta,d]\to X$ denote a well-parameterized dipath and $\gamma: I\to X_{l-\delta}$ denote a path, both of them starting at the same point $x_0\in X_{l-\delta}$. Using the sets $U_{\gamma(s)},\ s\in I$ from ULDEP, we obtain a covering of the interval I by sets $U_s=\gamma^{-1}(U_{\gamma(s)}),\ s\in I$, because $s\in U_s$ for every $s\in I$. Since I is compact, finitely many of them (including U_0) suffice. Choose them (leaving out unneccesary ones) in such a way that $0=s_0<s_1<\cdots<s_n$ and $U_{s_{i-1}}\cap U_{s_i}\neq\emptyset$ for $1\leqslant i\leqslant n$. By ULDEP, construct inductively well-parameterised dihomotopies $H^i:[s_{i-1},s_i]\times[l-\delta,l]\to X_{[l-\delta,l]}$ with $H^1(0,t)=\alpha(t)$, $H^i(s,0)=\gamma(s)$ and $H^i(s_i,t)=H^{i+1}(s_i,t)$. The result of a pasting is a well-parameterized dihomotopy $H^{\leqslant}:[0,1]\times[l-\delta,l]\to X$ with $H^{\leqslant}_0(t)=\alpha(t)$ and $H^{\leqslant}(s,l)$ parameterising a path γ' in X_l, see Figure 9. Since X satisfies DEP between levels l and d, there is a well-parameterised dihomotopy $H^{\geqslant}: I\times[l,d]$ with $H^{\geqslant}(0,t)=\alpha(t)$ and $H^{\geqslant}(s,l)=\gamma'(s)$. Pasting $H^{\leqslant}$ and $H^{\geqslant}$ along the trace of γ' as in Figure 9, we obtain the desired well-parameterized dihomotopy. □

The following consequence helps us to detect dihomotopy classes in Section 4.

Corollary 2.19. Let $(X,\leqslant,f)$ denote a parameterized po-space satisfying ULD1P and ULDEP.

1 Two dipaths α,β from X_l to X_d are dihomotopic if and only if they start in the same path-component of X_l.

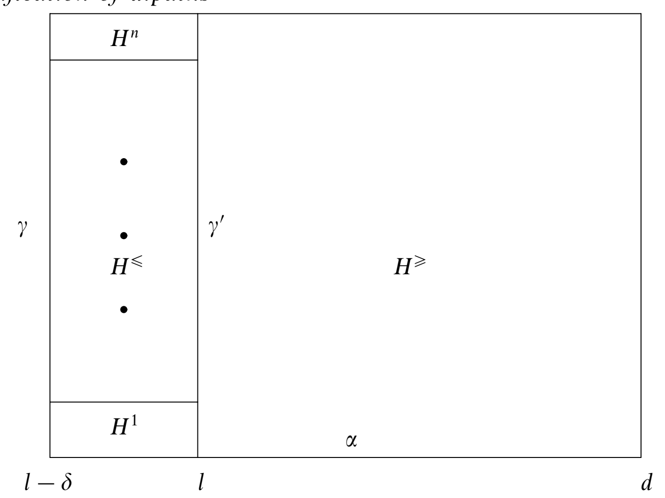


Fig. 9. Pasting scheme 2

2 Let $x_0 \in X$, and assume x_1 to be an *isolated* point in $R(X; x_0, X) \cap X_d = \uparrow x_0 \cap X_d$ ($\uparrow$ with respect to $\prec$ from Definition 1.6). Then
 — $\uparrow x_0 \cap X_d = \{x_1\}$.
 — $|\pi_1(X; x_0, x_1)| = 1$.

Proof.

1 Let γ denote a path between the start points of α and β in X_l. Since X satisfies DEP, there is a dipath α' dihomotopic ('along' γ) to α such that α' and β start at the *same* point. Since X satisfies D1P, α' and β are dihomotopic. By transitivity, α and β are dihomotopic, as well.
The other implication is trivially true.

2 Assume $x_1 \neq x_2 \in \uparrow x_0 \cap X_d$. If there were a dipath from x_0 to x_2, it would be dihomotopic to a dipath from x_0 to x_1 by Proposition 2.18. A (well-parameterized) dihomotopy connecting a (well-parameterized) dipath from x_0 to x_2 to a (well-parameterized) dipath from x_0 to x_1 would imply a path from x_1 to x_2 in $\uparrow x_0 \cap X_d$. Hence, there cannot exist any dipath from x_0 to x_2. As a consequence, $|\pi_1(X; x_0, x_1)| = |\pi_1(X; x_0, X_d)| = 1$.

□

2.3.3. *Translation to parametrized cubical complexes.* To relate these geometrical notions to the cubical complex/HDA presentation of concurrent processes mentioned in the introduction, we finish this section with a characterization of local di-1-connectedness in certain cubical complexes. We consider cubical complexes M (Serre 1951; Brown and Higgins 1981b; Brown and Higgins 1981a; Fajstrup *et al.* 1999) to be given by a family of sets $\{M_n | n \geqslant 0\}$ – the n-cubes – with face maps $\partial_i^k : M_n \to M_{n-1}$ $(1 \leqslant i \leqslant n,\ k = 0, 1)$ satisfying the semi-cubical relations

$$\partial_i^k \partial_j^l \quad = \quad \partial_{j-1}^l \partial_i^k \quad (i < j).$$

Moreover, we use the following notation: for an n-cube $y \in M_n$, let $\partial^i y = \{\partial^i_j(x) | 1 \leqslant j \leqslant n\}$, $0 \leqslant i \leqslant 1$, denote the set of *all* its lower, respectively, upper faces.

Cubical complexes model in a combinatorial manner shapes made up of n-dimensional cubes (modelling the asynchronous execution of n actions) glued together along their boundaries (glueing models combination of elementary computations).

In Fajstrup *et al.* (1999), we studied (non-self linked) cubical complexes with a natural local partial order and their geometric realization as topological spaces with a local partial order. We use $|M|$ to denote the geometric realization of the cubical complex M. It is shown in Fajstrup *et al.* (1999, Chapter 7) that cubical dipaths ('chains' in M_1), respectively, cubical dihomotopies correspond to topological dipaths, (respectively, topological dihomotopies) in the geometric realization $|M|$. On the other hand, every topological dipath in the geometric realization $|M|$ is dihomotopic to the realization of a cubical dipath (in an appropriate subdivision) of M.

Which properties of a cubical complex M make sure that its geometric realization $|M|$ satisfies ULD1P and ULDEP from Definition 2.16 – so that one can use the conclusions of Proposition 2.18 and Corollary 2.19?

Proposition 2.20. Let M denote a finite cubical complex with a parametrization extending the natural parametrization on the individual cubes, see Example 2.13. Its geometric realization $|M|$ satisfies ULD1P and ULDEP if:
For every pair of 1-cubes $y, y' \in M_1$ with $\partial^0_1 y = \partial^0_1 y' = x$ there are sequences $y_i \in M_1$ of 1-cubes and z_i of 2-cubes such that

$$\partial^0_1 y_i = \partial^0_1 y = x,\ 1 \leqslant i \leqslant k;$$

$$\partial^0 z_i = \{y_i, y_{i+1}\}, 1 \leqslant i \leqslant k-1; \quad \partial^0 z_0 = \{y, y_1\}; \quad \partial^0 z_k = \{y_k, y'\}.$$

Remark 2.21. The condition above has the following computer scientific interpretation (at least in a PV program): look at the actions that can be fired from a fixed state x. Two actions y, y' need not necessarily commute, but there have to be other actions y_i, $1 \leqslant i \leqslant k$, such that y commutes with y_1, y' commutes with y_k, and y_i commutes with y_{i+1}, $1 \leqslant i < k$, .

Proof. We only need to check the conditions for dipaths on the 1-skeleton of $|M|$ that connect two vertices: since M is finite, the set of differences $f(x_1) - f(x_0)$ with $\partial^i_1 y = x_i,\ y \in M_1$, has a positive infimum (=minimum) ε.

To prove ULD1P, observe that (the 'lower half of') a 2-cube z_i yields a *di*homotopy between the dipaths corresponding to its edges y_i and y_{i+1}, see the left-hand picture in Figure 10. By transitivity (along the 2-cubes z_i), one obtains a dihomotopy between the dipaths corresponding to y and y'.

To prove ULDEP, consider a (sufficiently short) path γ at level l from the vertex x within a k-cube $|z|$ and a 1-cube y with $\partial^0_1 y = x$. First, we construct a dihomotopy along γ from a dipath corresponding to a 1-cube y' *in the boundary of* z with $\partial y' = x$. We get it by concatenation with the 'flow' $H : I \times |z|_l \to |z|_{\geqslant l}$ given by

$$H(s; x_1, \ldots, x_k) = (x_1, \ldots, x_{j-1}, x_j + \ldots x_k + s - (k - j), 1, \ldots, 1),$$

where $1 \leqslant j \leqslant k$ is chosen such that $x_{j+1}+\ldots+x_k+s > k-j$ and $x_j+\ldots+x_k+s \leqslant k-j+1$. (See the right-hand picture in Figure 10 for an illustration in dimension 2.) The last coordinate x_k is chosen as the one rising from 0 to 1 along $|y'|$. This dihomotopy is then pasted to the dihomotopy between y and y' in the proof of ULD1P. (Literally speaking, you obtain a dihomotopy along another parametrization of γ that is constant on a part of the homotopy interval. The proof of Proposition 2.18 can easily be adapted to this situation.) □

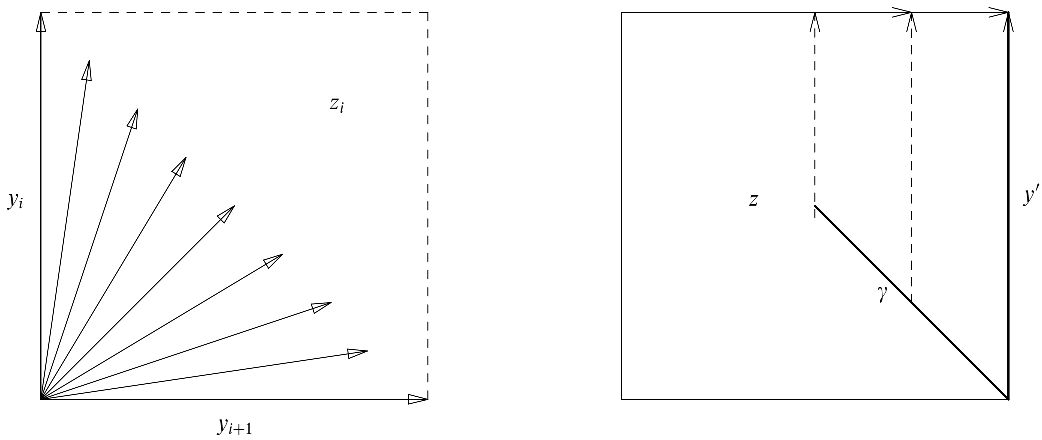


Fig. 10. Dihomotopies in Proposition 2.20

3. Dipaths in mutual exclusion (PV) models

Let $F = \bigcup_1^r R^i \subset I^n$ denote a (forbidden) region consisting of finitely many hyperrectangles $R^i \subset I^n$. In a mutual exclusion model (like the PV programs we were discussing earlier on), the state space is given as $X = I^n \setminus int(F)$. (The interior has to be understood with respect to the topology of I^n; this is important in points on ∂I^n!) We investigate dipaths from $\mathbf{0}$ to $\mathbf{1}$ in X, see Fajstrup *et al.* (1998a; 1998b) for a detailed description.

3.1. *Determination of safe regions*

First, we describe a nice application of Proposition 2.11 that determines the '*safe*' points/states in a mutual exclusion situation. For this application, the state space is allowed to contain both unsafe and unreachable points. A reachable point x is called *safe* if *no* dipath through x can end in a deadlock (see Definition 2.2), that is, *every* dipath *must end* in the final state $\mathbf{1}$.

Let $\mathscr{D} \subset I^n$ denote the set of deadlocks ($\neq \mathbf{1}$). We write $\mathbf{y} \leqslant \mathbf{x}$ if $y_i \leqslant x_i$ for every $1 \leqslant i \leqslant n$, and $[\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in I^n | a_i \leqslant x_i \leqslant b_i,\ 1 \leqslant i \leqslant n\}$ is the hyperrectangle between $\mathbf{a}$ and $\mathbf{b}$.

Corollary 3.1. The region of safe points coincides with

$$R(X; \mathbf{0}, \mathbf{1}) \cap \bigcap_{\mathbf{x} \in \mathscr{D}} [\{\mathbf{y} \in X | \mathbf{y} \nleqslant \mathbf{x}\} \cup Uns([\mathbf{0}, \mathbf{x}] \setminus F; \mathbf{x})].$$

Remark 3.2. To determine the region of safe points algorithmically, one has thus to run several rounds of the algorithm detecting deadlocks and unsafe regions from Fajstrup *et al.* (1998a; 1998b) (as recapitulated in Section 2.1). First, one determines deadlocks, unsafe and unreachable regions with respect to *all* of X. Then, for every deadlock $\mathbf{x} \in \mathcal{D}$, one determines unsafe regions for the restricted state space $[\mathbf{0}, \mathbf{x}] \setminus F$, where additional deadlocks can occur (but only on the boundary). It is algorithmically easy to calculate the intersections and unions of the unsafe regions indicated in Corollary 3.1.

The diagonally shaded rectangle in Figure 11 below illustrates that an additional component can enlarge the safe region. It corresponds to $Uns([\mathbf{0}, \mathbf{x}_2] \setminus F; \mathbf{x}_2)$, that is, that part of the safe region that is *not* contained in the set of points $\{\mathbf{y} \in X | \mathbf{y} \not\leqslant \mathbf{x}_1 \text{ and } \mathbf{y} \not\leqslant \mathbf{x}_2\}$ for the two deadlocks $\mathbf{x}_1$ and $\mathbf{x}_2$, which are safe for trivial reasons.

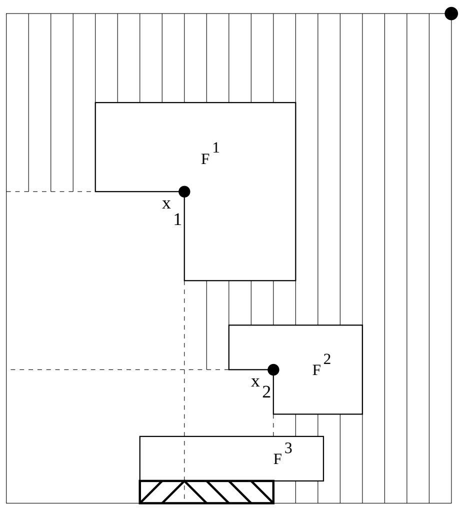


Fig. 11. 'Additional' safe points

Proof. We apply Proposition 2.11 to the case: $A = \mathcal{D}$, $B = \{\mathbf{1}\}$, and $S = A \cup B = \mathcal{D} \cup \{\mathbf{1}\}$ consists of all 'no future' points $\mathbf{x} \in X$ (with $\uparrow x = \{x\}$). The subspace S is clearly unavoidable with respect to $\mathbf{0}$ and S itself; moreover, it consists of (unrelated) one-point components, that are all closed in X. By Lemma 2.9, S filters X with respect to $\mathbf{0}$ and S. The domain of dependence $D(X; \mathbf{0}, S; \mathbf{1})$ agrees with the region of safe points, and by Proposition 2.11,

$$D(X; \mathbf{0}, S; \mathbf{1}) = R(X; \mathbf{0}, S) \cap \bigcap_{\mathbf{x} \in \mathcal{D}} U(X; \mathbf{0}, S; \mathbf{x}). \tag{1}$$

Since $D(X; \mathbf{0}, S; \mathbf{1}) \subseteq R(X; \mathbf{0}, \mathbf{1})$, we may replace $R(X; \mathbf{0}, S)$ by $R(X; \mathbf{0}, \mathbf{1})$ in (1). The singular region $U(X; \mathbf{0}, S; \mathbf{x})$ within $R(X; \mathbf{0}, S)$ deomposes into two parts consisting of all reachable points $\mathbf{y} \not\leqslant \mathbf{x}$ and of those reachable points $\mathbf{y} \leqslant \mathbf{x}$ from which $\mathbf{x}$ cannot be reached. □

3.2. *Determination of domains of dependence and of reachable regions*

The next result describes how domains of dependence and reachable regions can be determined using the algorithm detecting deadlocks und unsafe regions from Fajstrup *et al.* (1998a; 1998b). In this application, we first complete F by the associated unsafe and unreachable regions to obtain $F \subset \bar{F} = F \cup U(X;\mathbf{0},\mathbf{1})$; the resulting state space

$$\bar{X} = X \setminus U(X;\mathbf{0},\mathbf{1}) = I^n \setminus int(\bar{F}) \subset I^n \setminus int(F) = X$$

has neither unsafe nor unreachable points. By definition (see Remark 2.3), inclusion induces identities

$$\vec{P}_1(X;\mathbf{0},\mathbf{1}) = \vec{P}_1(\bar{X};\mathbf{0},\mathbf{1}); \quad \vec{\pi}_1(X;\mathbf{0},\mathbf{1}) = \vec{\pi}_1(\bar{X};\mathbf{0},\mathbf{1}).$$

In the following, we assume that F *is already completed* as described above. In particular, the state space $X = I^n \setminus int(F)$ contains neither deadlocks nor unreachable points. Applying Lemma 2.5 in this case, we obtain the following lemma.

Lemma 3.3. Let $S \subset I^n$ denote a region consisting of finitely many hyperrectangles, let $S_X = S \setminus int(F) \subset X$. Then, the domains of dependence of S_X in X satisfy:

1 $D^-(X;\mathbf{0},\mathbf{1};S_X) = S_X \cup Uns(X \setminus S_X;\mathbf{1}) = S_X \cup Uns(I^n \setminus (int(F \cup S));\mathbf{1})$, that is, one has to add to S_X the unsafe region associated to the 'new forbidden region' $F \cup S$.

2 $D^+(X;\mathbf{0},\mathbf{1};S_X) = S_X \cup Unr(X \setminus S_X;\mathbf{0}) = S_X \cup Unr(I^n \setminus (int(F \cup S));\mathbf{0})$, that is, one has to add to S_X the unreachable region associated to the 'new forbidden region' $F \cup S$.

Hence, in order to detect the domains of dependence of a region S, just apply the algorithm of Fajstrup *et al.* (1998a; 1998b) to the 'new forbidden region' $F_S = F \cup S$. The following case is of particular importance.

Corollary 3.4. Let $S_1, \ldots, S_k \subset I^n$ denote a collection of subspaces such that $S_i \cap X$ filters X with respect to $\mathbf{0}$ and $\mathbf{1}$ in X for every $1 \leqslant i \leqslant k$. Let $A_i \subset S_i$, $1 \leqslant i \leqslant k$ denote a union of connected components of S_i; its complement $B_i \subset S_i$ is also a union of connected components. Then, the reachable space $R(X;\mathbf{0},\mathbf{1},A_1,\ldots,A_k)$ is the complement of the union of unsafe and unreachable regions associated to $F \cup \bigcup_i B_i$ in I^n.

Proof. By Proposition 2.10, we have that $R(X;\mathbf{0},\mathbf{1},A_1,\ldots,A_k)$ is the complement of $D(X;\mathbf{0},\mathbf{1};\bigcup_i B_i)$, which by Lemma 3.3 identifies with the indicated unsafe and unreachable regions. □

Remark 3.5. To summarize, both domains of dependence and reachable spaces can be determined as easy extensions of the algorithm determining deadlocks and unsafe regions from Fajstrup *et al.* (1998a; 1998b). Éric Goubault and collaborators are working on implementations using dynamic segment trees as in van Kreveld and Overmars (1993), Edelsbrunner (1980), Six and Wood (1982), Preparata and Shamos (1988) and Aparici and Morezuelas (1999).

3.3. *Di-1-connectedness*

Finally, we investigate (local) di-1-connectedness of the state space X that is the complement in I^n of a forbidden region F, which in turn is a finite union of n-dimensional hyperrectangles. A hyperrectangle $R = [c_1, c_1'] \times \ldots \times [c_n, c_n']$ is determined by its corners $\mathbf{c} = (c_1, \ldots, c_n)$ and $\mathbf{c}' = (c_1', \ldots, c_n')$. A short notation is $R = [\mathbf{c}, \mathbf{c}']$. In the same spirit we define

— the lower, respectively, upper *core* of R as

$$R_- = [\mathbf{c}, \mathbf{c}'[; \quad R_+ =]\mathbf{c}, \mathbf{c}'];$$

— the lower, respectively, upper *boundaries* of R as

$$\partial^0 R = R \setminus R_+ = [\mathbf{c}, \mathbf{c}'] \setminus]\mathbf{c}, \mathbf{c}']; \quad \partial^1 R = R \setminus R_- = [\mathbf{c}, \mathbf{c}'] \setminus [\mathbf{c}, \mathbf{c}'[.$$

Now, let $R^i = [a_1^i, b_1^i] \times \cdots \times [a_n^i, b_n^i]$, and let $F = \bigcup_1^r R^i$ denote the forbidden region, let $X = I^n \setminus int(F)$ – the interior with respect to the relartive topology of I^n. We assume generically that, for a given j, the a_j^i are pairwise different. Moreover, in order to take care of special 'effects' on the boundary ∂I^n, the state space X can be described as follows (*cf.* Fajstrup *et al.* (1998a; 1998b)) taking care of the relative topology. We embed $I^n = [0,1]^n$ into the larger hypercube $[-1,2]^n$ and remove the additional (lower and upper) region again by adding the $2n$ hyperrectangles $[-1,2] \times \cdots \times [-1,2] \times [0,-1] \times [-1,2] \times \cdots \times [-1,2]$ and $[-1,2] \times \cdots \times [-1,2] \times [1,2] \times [-1,2] \times \cdots \times [-1,2]$ to the forbidden region F to yield $\tilde{F}$. Moreover, all hyperrectangles touching ∂I^n will be enlarged; the new $\tilde{r} = r + 2n$ hyperrectangles are called $\tilde{R}^i$; see Fajstrup *et al.* (1998a; 1998b) for details. The new forbidden region is $\tilde{F} = \bigcup \tilde{R}^i$, and the state space X has the descriptions $X = I^n \setminus int(F) = [-1,2]^n \setminus int(\tilde{F})$.

This state space X inherits a partial order from $[-1,2]^n$, and, moreover, we have the parametrization $f : X \subset [-1,2]^n \to \mathbf{R}$, $f(x_1, \ldots x_n) = \sum_i x_i$. The (hyperrectangular) lattice with vertices at $\mathbf{c} = (c_1, \ldots, c_n)$ with c_j one of the a_j^i, b_j^i can be interpreted as a cubical complex.

We want to decide which of the cells of that complex belong to $\tilde{F}$, respectively, to X. Let $\mathbf{c} = (c_1, \ldots, c_n)$ denote a vertex, and let $c_j' > c_j$ denote the following j-th vertex coordinate. Then, $\mathbf{c}' = (c_1', \ldots, c_n')$ denotes the upper corner of the $n - cell$ $[\mathbf{c}, \mathbf{c}']$. Let $J = \{0, 1, \ldots, \tilde{r}\}$ where $\tilde{r}$ is the number of hyperrectangles constituting $\tilde{F}$. At every vertex $\mathbf{c}$, we introduce the function

$$s_\mathbf{c} : \{1, \ldots, n\} \to J, \; s_\mathbf{c}(j) = \begin{cases} l & \Leftrightarrow c_j = a_j^l \\ 0 & \Leftrightarrow c_j = b_j^l, \; 1 \leqslant l \leqslant \tilde{r}. \end{cases}$$

The pre-images $s_\mathbf{c}^{-1}(i)$ give rise to a partition $\mathscr{P}_\mathbf{c}$ of the set $\{1, \ldots, n\}$. Let $\mathbf{c}_k \in \mathbf{R}^{|s^{-1}(k)|}$ denote the (sub)vector with coordinates a_j^k for $k \leqslant \tilde{r}$ and with coordinates $b_j^?$ for $k = \tilde{r} + 1$. Using $s_\mathbf{c}$ to reorder with respect to the partition $\mathscr{P}_\mathbf{c}$, we obtain presentations

$$[\mathbf{c}, \mathbf{c}'] = \prod_0^{\tilde{r}+1} [\mathbf{c}_j, \mathbf{c}_j'], \quad R^i = \prod_0^{\tilde{r}+1} [\mathbf{a}_k^i, \mathbf{b}_k^i].$$

Remark 3.6. The point $\mathbf{c} \in R_-^i$ if and only if $\mathbf{a}_k^i < \mathbf{c}_k < \mathbf{b}_k^i \; \forall k \neq i$. In particular, $\mathbf{c}_k = \max\{\mathbf{a}_k^i | \mathbf{c} \in R_-^i\}$.

The following lemmas then follow immediately from the definitions.

Lemma 3.7.

1 $int(R^i) \cap [\mathbf{c}, \mathbf{c}'[= \begin{cases}]\mathbf{c}_i, \mathbf{c}'_i[\times \prod_{k \neq i}[\mathbf{c}_k, \mathbf{c}'_k[& \mathbf{c} \in R^i_- \\ \varnothing & \mathbf{c} \notin R^i_- \end{cases}.$

2
$$int(\tilde{F}) \cap [\mathbf{c}, \mathbf{c}'[= \bigcup_{\mathbf{c} \in R^i_-} (]\mathbf{c}_i, \mathbf{c}'_i[\times \prod_{k \neq i}[\mathbf{c}_k, \mathbf{c}'_k[)$$

3
$$X \cap [\mathbf{c}, \mathbf{c}'[= \prod_{\mathbf{c} \in R^i_-} \partial^0([\mathbf{c}_i, \mathbf{c}'_i]) \times \prod_{\mathbf{c} \notin R^i_-} [\mathbf{c}_i, \mathbf{c}'_i[.$$

Next, we investigate 1- and 2-cubes with $\mathbf{c} \in X$ as their lower corner.

Lemma 3.8.

1 The 1-cube $\mathbf{c}^l$ with upper corner $(c_1, \ldots, c'_l, \ldots, c_n)$ is contained in X if and only if $\mathbf{c} \notin R^{s(l)}$ or if $|s_{\mathbf{c}}^{-1}(s_{\mathbf{c}}(l))| \geqslant 2$.

2 The 2-cube $\mathbf{c}^{lm}$ with upper corner $(c_1, \ldots c'_l, \ldots, c'_m, \ldots, c_n)$ is contained in X if and only if
$\mathbf{c} \notin R^{s_{\mathbf{c}}(l)} \cap R^{s_{\mathbf{c}}(m)}$ or
$\mathbf{c} \notin R^{s_{\mathbf{c}}(l)}$ and $|s_{\mathbf{c}}^{-1}(s_{\mathbf{c}}(m))| \geqslant 2$ or
$\mathbf{c} \notin R^{s_{\mathbf{c}}(m)}$ and $|s_{\mathbf{c}}^{-1}({}_{\mathbf{c}}s(l))| \geqslant 2$ or
$s_{\mathbf{c}}(l) \neq s_{\mathbf{c}}(m)$, $\mathbf{c} \in R^{s_{\mathbf{c}}(l)} \cap R^{s_{\mathbf{c}}(m)}$ and $|s_{\mathbf{c}}^{-1}(s_{\mathbf{c}}(l))|, |s_{\mathbf{c}}^{-1}(s_{\mathbf{c}}(m))| \geqslant 2$ or
$s_{\mathbf{c}}(l) = s_{\mathbf{c}}(m), \mathbf{c} \in R^{s_{\mathbf{c}}(l)}$ and $|s_{\mathbf{c}}^{-1}(s_{\mathbf{c}}(l))| \geqslant 3$.

3 The 2-cube $\mathbf{c}^{lm}$ is *not* contained in X, while its lower boundary 1-cubes $\mathbf{c}^l$ and $\mathbf{c}^m$ *are* contained in X if and only if $\mathbf{c} \in R^{s_{\mathbf{c}}(l)}$ and $s_{\mathbf{c}}^{-1}(s({}_{\mathbf{c}}l)) = \{l, m\}$.

The last result, combined with Proposition 2.18 and Proposition 2.20, allows us to give conditions for the state space X to be di-1-connected, that is, di-1-connected from level 0 corresponding to $\mathbf{0}$ to level n corresponding to $\mathbf{1}$, see Definition 2.16.

Proposition 3.9. Assume the forbidden region F is such that the minimum of every non-empty intersection $\tilde{R}^{i_1} \cap \ldots \cap \tilde{R}^{i_{n-1}}$ of $n-1$ hyperrectangles, is contained in another hyperrectangle $\tilde{R}^{i_n}$. Then X is di-1-connected.

Remark 3.10.

1 It is important to use the *extended* hyperrectangles $\tilde{R}^i$ in Proposition 3.9. Look at Example 4.6.2 and Figure 16 to see that a state space in 2 dimensions may be (even locally) di-1-connected although the forbidden region is not empty.

2 Note that this result is somehow parallel to the (necessary and sufficient) criterion for the absence of deadlocks in Fajstrup *et al.* (1998a; 1998b), where we looked at intersections of n hyperrectangles.

Proof.

Case 1: There exists $1 \leqslant i \leqslant n$ with $\mathbf{c} \notin R^{s(i)}$. By Lemma 3.8.2, the 2-cubes $\mathbf{c}^{li}$ and $\mathbf{c}^{im}$ are both contained in X and form a chain of 2-cubes from $\mathbf{c}^l$ to $\mathbf{c}^m$ as in Proposition 2.20.

Case 2: $\mathbf{c} \in \cap_1^n R^{s(i)}$, and, in particular, $\mathbf{c} = \min(\cap_1^n R^{s(i)})$ by Remark 3.6. We distinguish cases according to the size of $k = s(\{1, \ldots, n\})$.

$k = n$**:** In this case, $s_{\mathbf{c}}$ is injective. By Lemma 3.8.1, *none* of the 1-cubes $\mathbf{c}^l$ is contained in X, that is, $\mathbf{c}$ is a deadlock point; *cf.* Fajstrup *et al.* (1998a; 1998b).

$k \leqslant n-2$**:** Assume that $\mathbf{c}$ and the two 1-cubes $\mathbf{c}^l, \mathbf{c}^m, l \neq m$ are contained in X. If $s_{\mathbf{c}}(l) \neq s(m)$ or if $s_{\mathbf{c}}(l) = s(m)$ and $|s^{-1}(s(l))| \geqslant 3$, the two-cube $\mathbf{c}^{lm}$ is contained in X according to Lemma 3.8.3. and the condition in Proposition 2.20 is satisfied by virtue of $\mathbf{c}^{lm}$ alone.
It remains to consider the case $s_{\mathbf{c}}(l) = s(m)$, and $s_{\mathbf{c}}^{-1}(s(l)) = \{l, m\}$. Since $k \leqslant n-2$, there is another pair $l' \neq m'$ with $s_{\mathbf{c}}(l') = s(m') \neq s(l)$. Again from Lemma 3.8.2, we may conclude that the 2-cubes $\mathbf{c}^{ll'}$ and $\mathbf{c}^{l'm}$ are contained in X and form a chain of 2-cubes from $\mathbf{c}^l$ to $\mathbf{c}^m$ as in Proposition 2.20. □

The condition in Proposition 3.9 is certainly not necessary. For an illustration in dimension 3, see Figure 12 below: X is the complement of the interior of a forbidden region consisting of 2 hyperrectangles in I^3; dipaths leave from the right upper front $\mathbf{0}$ to the left lower back $\mathbf{1}$. Note that all dipaths in X from $\mathbf{0}$ to $\mathbf{1}$ are dihomotopic. But X does not satisfy neither of the properties ULD1P or ULDEP at the minimum of the intersection of the two forbidden hyperrectangles.

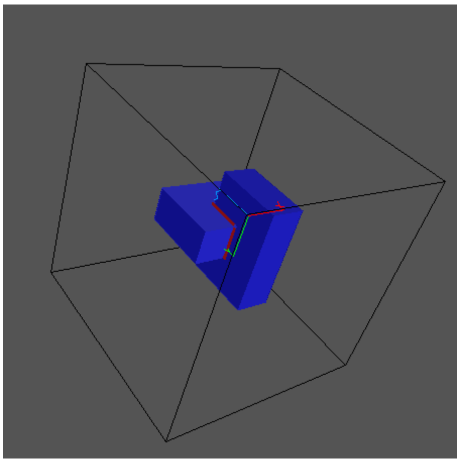


Fig. 12. Local di-1-connectedness fails at the intersection of the two cubes (Geomview drawing)

4. Classification of dipaths in 2-dimensional mutual exclusion models

Throughout this section, we assume that $n = 2$. Again let $F \subset I^2$ denote a (forbidden) region consisting of finitely many rectangles with complement $X = I^2 \backslash int(F)$. Investigating dipaths from $\mathbf{0}$ to $\mathbf{1}$ in X up to dihomotopy corresponds to finding schedules with potentially different results for PV-programs with *two* interacting transactions. In the following, we may and will assume that $U(X;\mathbf{0},\mathbf{1}) = \emptyset$, that is, that X contains neither unsafe nor unreachable points; this can always be achieved by 'completion'; see the discussion in Section 3.2.

4.1. *Significance of the connected components of the forbidden region*

The overall idea is to use the *connected components* of F to describe finitely many filtering subspaces $S \subset X$ such that the product of the component maps $c_S : \vec{\pi}_1(X;\mathbf{0},\mathbf{1}) \to \pi_0(S)$ from Lemma 2.8 yields complete information about $\vec{\pi}_1(X;\mathbf{0},\mathbf{1})$.

By assumption, the state space X contains neither deadlocks nor unreachable points. This has the following consequence.

Lemma 4.1. The connected components $F_i \subseteq \bar{F}$ possess both a unique global minimum x_0 (with $x_0 \leqslant x$ for all $x \in F_i$) and, likewise, a unique global maximum.

Proof. First, the boundary ∂F_i of every component F_i is path-connected after completion; an interior boundary would give rise to unreachable points. Moreover, ∂F_i consists of line segments (edges) parallel to the axes with orientations reflecting the partial order, which is a global order on each of these line segments. *Local* minima on F_i are contained in ∂F_i and correspond to vertices with two outgoing edges.

Assume, there are two different local minima x_1 and x_2. Since ∂F_i is path-connected, there is a chain of edges connecting them. Since, this chain starts and ends with edges leaving the end points, there has to be an intermediate vertex x_3 with two ingoing edges, the 'upper edges' of the rectangle below, which is not contained in F_i. This is why x_3 is a deadlock point within $X = I^2 \setminus int(F)$, see Figure 13. This is a contradiction! □

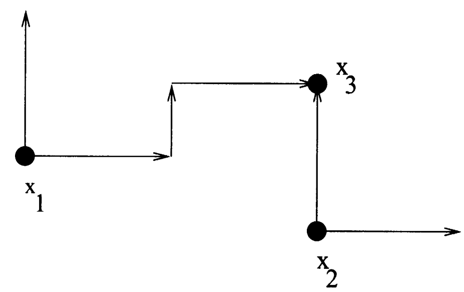


Fig. 13. Local minima separated by a deadlock

Let $\mathbf{z}_i = (x_i, y_i)$ denote the unique minimum of F_i. Choose a sufficiently small $\varepsilon > 0$, such that *none* of the rectangles that constitute F has a bounding edge within $\{(x,y)|x_i < x < x_i + 3\varepsilon \text{ or } y_i < y < y_i + 3\varepsilon\}$. This condition ensures that the subset of all points in X with $x_i < x < x_i + 3\varepsilon$ is a product with the interval $(x_i, x_i + 3\varepsilon)$, and likewise for a strip parallel to the x-axis.

We define two *unrelated* subspaces $S_i^x, S_i^y \subset X$ as follows:

$$S_i^x = \{\mathbf{z} = (x, y) \in X |\ x \leqslant x_i, y_i + \varepsilon \leqslant y \leqslant y_i + 2\varepsilon\},\ \text{and}$$

$$S_i^y = \{\mathbf{z} = (x, y) \in X |\ x_i + \varepsilon \leqslant x \leqslant x_i + 2\varepsilon, y \leqslant y_i\}.$$

These subspaces consist of two or more 'thin' rectangles; let $T_i^x \subset S_i^x$, respectively, $T_i^y \subset S_i^y$ denote those of these rectangles that touch F_i. Look at Figure 14 for an example.

The set of *all* maps ψ from $\pi_0(F)$ to the 2-element set $\{x, y\}$ will be denoted $M(\pi_0(F), \{x, y\})$. A *partial map* φ from $\pi_0(F)$ to the 2-element set $\{x, y\}$ takes values $\varphi(F_i) = x$, $\varphi(F_i) = y$, or $\varphi(F_i) = \perp$ (undefined). The set of all these partial maps will be denoted by $PM(\pi_0(F), \{x, y\})$.

All sets $S_i = S_i^x \cup S_i^y$ filter X with respect to $\mathbf{0}$ and $\mathbf{1}$ (see Definition 2.7): they are unavoidable, since $\mathbf{0}$ and $\mathbf{1}$ are contained in different connected components of $X \setminus S_i = I^2 \setminus int(F_i \cup S_i)$. Moreover, S_i^x and S_i^y are unrelated, and we can combine all the component maps c_{S_i} from Lemma 2.8 with forgetful maps in two ways to obtain maps

$$C : \vec{\pi}_1(X; \mathbf{0}, \mathbf{1}) \to M(\pi_0(F), \{x, y\})$$

$$C(\alpha)(F_i) = \begin{cases} x, & \text{if } \alpha \text{ hits } S_i^x; \\ y, & \text{if } \alpha \text{ hits } S_i^y; \end{cases} \qquad \text{and}$$

$$T : \vec{\pi}_1(X; \mathbf{0}, \mathbf{1}) \to PM(\pi_0(F), \{x, y\})$$

$$T(\alpha)(F_i) = \begin{cases} x, & \text{if } \alpha \text{ hits } T_i^x; \\ y, & \text{if } \alpha \text{ hits } T_i^y; \\ \perp & \text{if } \alpha \text{ hits another component of } S_i. \end{cases}$$

The 'local product stucture' of X mentioned above guarantees that the maps C and T are defined independently of the choice of a small ε. Note, moreover, that T factors through C. The factorization

$$f : M(\pi_0(F), \{x, y\}) \to PM(\pi_0(F), \{x, y\})$$

with $T = f \circ C$ is defined on a map $\psi : \pi_0(F) \to \{x, y\}$ with $\psi(F_i) = x$ by

$$(f(\psi))(F_i) = \begin{cases} \perp, & \exists j \text{ with } F_j \cap \{(x, y) | x < x_{i,i} + \varepsilon \leqslant y \leqslant y_i + 2\varepsilon\} \neq \emptyset \text{ and } \psi(F_j) = x; \\ x, & \text{otherwise.} \end{cases}$$

Use a similar recipe for $\psi(F_i) = y$, and check that $T = f \circ C$. Consult Figure 14 for an illustration.

Our next aim is to show that the map T – and thus also C – is injective, and to identify the image of T. To this end, we associate to a partial map $\varphi \in PM(\pi_0(F), \{x, y\})$ its *reachable subspace* $R(\varphi) \subset X$ which is (see Definition 2.1)

$$R(\varphi) = R(X; \mathbf{0}, \mathbf{1}; T_1^{\varphi(F_1)}, \ldots, T_n^{\varphi(F_n)}). \tag{2}$$

(The reachability conditions apply only to those $T_i^{\varphi(F_i)}$ with $\varphi(F_i) \neq \perp$.) In words, $R(\varphi)$ consists of the points on the traces of any dipath from $\mathbf{0}$ to $\mathbf{1}$ that hits the prescribed components $T_i^{\varphi(F_i)}$. By Proposition 2.10, the reachable region $R(\varphi)$ is the complement (within X) of the domain of dependence $D(X; \mathbf{0}, \mathbf{1}; \bigcup_i [S_i \setminus T_i^{\varphi(F_i)}])$.

A partial map $\varphi \in PM(\pi_0(F), \{x, y\})$ will be called *admissible* if $R(\varphi) \neq \emptyset$, that is, if there is a dipath from $\mathbf{0}$ to $\mathbf{1}$ in $R(\varphi)$. An admissible partial map φ will be called *complete* if

$$\varphi(F_i) = \perp \Leftrightarrow (T_i^x \cup T_i^y) \cap R(\varphi) = \emptyset, \tag{3}$$

that is, if no dipath can pass through any further component T_i^x, respectively, T_i^y. The set of all *complete partial maps* will be denoted $CPM(\pi_0(F), \{x, y\})$.

Proposition 4.2. The map $T : \vec{\pi}_1(X; \mathbf{0}, \mathbf{1}) \to PM(\pi_0(F), \{x, y\})$ yields a bijection

$$T : \vec{\pi}_1(X; \mathbf{0}, \mathbf{1}) \to CPM(\pi_0(F), \{x, y\}).$$

Proof.

1 Given $\alpha \in \vec{P}_1(\bar{X}; \mathbf{0}, \mathbf{1})$. Its trace is contained in $R(T(\alpha))$; in particular, $T(\alpha)$ is admissible. Let us show that $T(\alpha)$ is also complete. Consider a component F_i with $\mathbf{z}_i = \min F_i$ such that $T(\alpha)(F_i) = \perp$. We have to show that $R(T(\alpha)) \cap (T_i^x \cup T_i^y) = \emptyset$, or (equivalently!) that $\mathbf{z}_i \notin R(T(\alpha))$. Assume without restriction that α meets S_i in $S_i^x \setminus T_i^x$ at $\alpha(t_1)$.
Since X is closed, there is a *maximal* $t_0 < t_1$ such that $\mathbf{z}_i$ is reachable from $\alpha(t_0) = (x_0, y_0)$. The horizontal line segment $\{(x_0 + h, y_0) | h \geqslant 0\}$ must intersect the boundary of a component F_j – otherwise, $\mathbf{z}_i$ would also be reachable from $\alpha(t_0 + \varepsilon)$ for sufficiently small $\varepsilon > 0$. Choose the *first* F_j that this line segment intersects. Then $\mathbf{z}_i$ is *not* reachable from T_j^x. Since $T(\alpha)(F_j) = x$, we see that $\mathbf{z}_i \notin R(T(\alpha))$.

2 T **is onto:** Let $\varphi \in CPM(\pi_0(F), \{x, y\})$. Since φ is admissible, there is a dipath α in $R(\varphi)$ from $\mathbf{0}$ to $\mathbf{1}$ through all the $T_i^{\varphi(F_i)}$ with $\varphi(F_i) \neq \perp$. Since φ is complete, α cannot meet any $T_x^i \cup T_y^i$ with $\varphi(F_i) = \perp$. In summary, we obtain $T(\alpha) = \varphi$.

3 T **is into:** Let $\varphi \in CPM(\pi_0(F), \{x, y\})$. We show that $R(\varphi)$ is di-1-connected from $\mathbf{0}$ to $\mathbf{1}$ using Proposition 2.18. To verify that $R(\varphi)$ satisfies ULD1P and ULDEP, we apply Proposition 3.9, but first we have to represent $R(\varphi)$ as the complement of a forbidden region:
By Proposition 2.10, $R(\varphi)$ is the complement of the domain of dependence

$$D(X; \mathbf{0}, \mathbf{1}; \bigcup_i [S_i \setminus T_i^{\varphi(F_i)})].$$

By Lemma 3.3, this subspace is the union of $\bigcup_i (S_i \setminus T_i^{\varphi(F_i)})$ with the unsafe and the unreachable region corresponding to the enlarged forbidden region $F \cup \bigcup_i [S_i \setminus T_i^{\varphi(F_i)}]$. The latter is a union of subspaces $F_i \cup S_i \setminus T_i^{\varphi(F_i)}$ the local minima of which are all contained in the *boundary* ∂I^2. The same property holds for the (larger) unsafe regions. The condition in Proposition 3.9 considers precisely the minima of every such rectangle, since $n - 1 = 2 - 1 = 1$ in our case. Hence, this condition is satisfied. □

4.2. *Examples and applications*

Example 4.3. Let X denote the complement of the 4-component subspace $F = \bigcup_1^4 F_i$ from Figure 14 below. Then $\vec{\pi}_1(X; \mathbf{0}, \mathbf{1})$ can be represented by six dipaths $\alpha_1, \ldots, \alpha_6$.

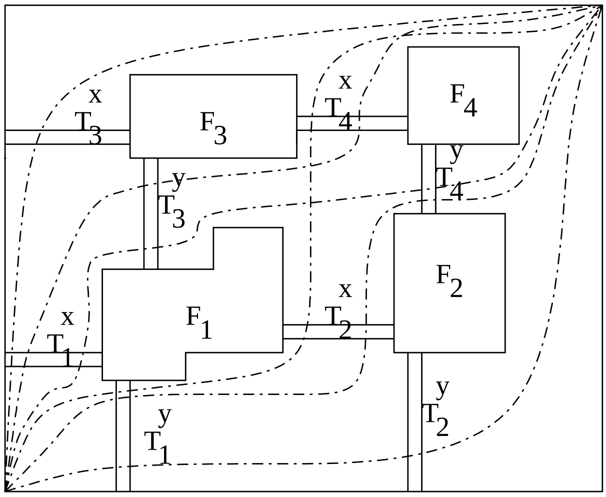


Fig. 14. Classifying dipaths in a mutual exclusion model

In this case, the maps C and T are given by the following tables:

	F_1	F_2	F_3	F_4
$C(\alpha_1)$	y	y	y	y
$C(\alpha_2)$	y	x	y	y
$C(\alpha_3)$	y	x	y	x
$C(\alpha_4)$	x	x	y	y
$C(\alpha_5)$	x	x	y	x
$C(\alpha_6)$	x	x	x	x

	F_1	F_2	F_3	F_4
$T(\alpha_1)$	y	y	$\perp$	$\perp$
$T(\alpha_2)$	y	x	$\perp$	y
$T(\alpha_3)$	y	x	$\perp$	x
$T(\alpha_4)$	x	$\perp$	y	y
$T(\alpha_5)$	x	$\perp$	y	x
$T(\alpha_6)$	x	$\perp$	x	$\perp$

Corollary 4.4.

1 The map $C : \vec{\pi}_1(X;\mathbf{0},\mathbf{1}) \to M(\pi_0(F), \{x, y\})$ is *injective*.
2 The number of dipaths in X from $\mathbf{0}$ to $\mathbf{1}$ satisfies

$$|\pi_0(F)| + 1 \leqslant |\vec{\pi}_1(X;\mathbf{0},\mathbf{1})| \leqslant 2^{|\pi_0(F)|}.$$

3 Two dipaths in X are dihomotopic if and only if they are homotopic (for $n = 2$).
4 For a dipath $\alpha \in \vec{P}_1(X;\mathbf{0},\mathbf{1})$, the homotopy history $h(\alpha)$, see Definition 1.9, coincides with the reachable region $R(T(\alpha))$. In particular, $h(\alpha)$ is di-1-connected from $\mathbf{0}$ to $\mathbf{1}$.

Proof.

1 This follows from the injectivity of T and the factorization $T = f \circ C$.

2 The second inequality follows from (1) above and from $|M(\pi_0(F), \{x, y\})| = 2^{|\pi_0(F)|}$. The first can be shown by induction over the number of components: it is obviously true for $|\pi_0(F)| = 0$, that is, $F = \emptyset$. Now, assume that a (completed) forbidden region $\bar{F}$ can be decomposed as $\bar{F} = F \cup F_{n+1}$ with F consisting of n connected components and F_{n+1} an additional connected component (intersecting neither F nor ∂I^2).

— Every dipath in $I^2 \setminus int(F)$ is dihomotopic to a dipath that avoids F_{n+1}. Indeed, a dipath through F_{n+1} can be deformed by a dihomotopy to a dipath with trace only touching the boundary.

— Dipaths that are dihomotopic in $I^2 \setminus int(F \cup F_{n+1})$ are certainly dihomotopic in $I^2 \setminus int(F)$. In particular, $|\vec{\pi}_1(I^2 \setminus int(F))| \leqslant |\vec{\pi}_1(I^2 \setminus int(F \cup F_{n+1}))|$.

— Up to reparametrization, there are precisely *two* dipaths on the boundary ∂F_{n+1} from $\min F_{n+1}$ to $\max F_{n+1}$. Those are *not dihomotopic* as paths in ∂F_{n+1} from $\min F_{n+1}$ to $\max F_{n+1}$; in fact, they even fail to be *homotopic* modulo end points since the concatenation of the first with the inverse of the second yields a generator in the first homotopy group $\pi_1(\partial F_{n+1}; \min F_{n+1}) \simeq \pi_1(S^1) \cong \mathbf{Z}$. By our general assumption at the beginning of Section 4, there are neither unsafe nor unreachable points in $I^2 \setminus int(F \cup F_{n+1})$; hence, there is a dipath from $\mathbf{0}$ to $\min F_{n+1}$ and another dipath from $\max F_{n+1}$ to $\mathbf{1}$ in that space. By concatenation with those, we obtain two dipaths in $I^2 \setminus int(F \cup F_{n+1})$ from $\mathbf{0}$ to $\mathbf{1}$: one of them intersects T^x_{n+1}; the other intersects T^y_{n+1}. Thus, they are *not* dihomotopic in $I^2 \setminus int(F \cup F_{n+1})$, whereas they are dihomotopic in $I^2 \setminus int(F)$. As a result, $|\vec{\pi}_1(I^2 \setminus int(F))| < |\vec{\pi}_1(I^2 \setminus int(F \cup F_{n+1}))|$.

3 If two dipaths α_1 and α_2 in X are homotopic, they are also homotopic in each of the larger spaces $X_i = I^2 \setminus int(F_i)$ with only one forbidden component F_i; their (non-directed) composition $\alpha_2 * (\alpha_1)^{-1}$ is thus nulhomotopic in X_i and has thus intersection number 0 with every line segment connecting F_i with ∂I^2. (This intersection number counts the intersection points with a sign ± 1 according to orientation). Hence, the traces of α_1 and α_2 have both to intersect either a line segment parallel to the x-axis through S^i_x or a line parallel to the y-axis through S^i_y; otherwise the intersection numbers would be 1 and -1, respectively. As a result, the map $C : \vec{\pi}_1(X; \mathbf{0}, \mathbf{1}) \to M(\pi_0(F), \{x, y\})$ agrees on α_1 and α_2. Use the injectivity of C (from 1 above) to conclude that the two dipaths have to be dihomotopic.

4 The homotopy history consists of the traces $\beta(I)$ of dipaths β dihomotopic to α. But $\beta(I) \subset R(T(\beta)) = R(T(\alpha))$, and hence $h(\alpha) \subseteq R(T(\alpha))$. On the other hand, $R(T(\alpha))$ is di-1-connected between $\mathbf{0}$ and $\mathbf{1}$; this follows as in the proof for the injectivity of the map T (case (3) in the proof of Proposition 4.2). By definition, every point $\mathbf{x} \in R(T(\alpha))$ is on the trace of a dipath β in $R(T(\alpha))$ from $\mathbf{0}$ to $\mathbf{1}$ within $R(T(\alpha))$. Since $\beta \overset{\sim}{\to} \alpha$, we conclude that $\mathbf{x} \in \beta(I) \subset h(\alpha)$, and thus $R(T(\alpha)) \subseteq h(\alpha)$. □

Remark 4.5.

1 The bounds in Corollary 4.4.1 are actually realised; see Figure 15 in Example 4.6.1.

2 The number $|\vec{\pi}_1(X;\mathbf{0},\mathbf{1})|$ of dipaths up to dihomotopy is one, if and only if F is empty or if every component of F intersects the boundary of I^2, see Example 4.6.2 and Figure 16. It is two, if exactly one of the components of F is contained in the interior of I^2, and greater than two, if at least two components of F are contained in the interior of I^2. Note, as an application, that a PV-schedule is serializable (every dipath dihomotopic to one of the two dipaths on ∂I^2) if and only if F is connected (or empty). This fact has already been noted in Lipski and Papadimitriou (1981).
3 Already in dimension 3, there exist dipaths in a space modelling mutual exclusion that are homotopic, but *not* dihomotopic (Fajstrup *et al.* 1999, Example 4.4).
4 It is *not true* for general cubical complexes X that all homotopy histories are di-1-connected (between a given start and end point).
5 One might think that the relative position of the minima and maxima of the components F_i is enough to get a classification of dipaths. Example 4.6.2 (due to L. Fajstrup) shows that this is *not* the case.
6 Since homotopy histories of dipaths can be determined as reachable regions, the history equivalence classes of Definition 1.9 or Fajstrup *et al.* (1999) can also be determined by taking Boolean expressions in those in an algorithmic way.

Example 4.6.

1 Figure 15 shows examples of state spaces for mutual exclusion programs with n components in the forbidden region for which $|\vec{\pi}_1(X;\mathbf{0},\mathbf{1})| = n+1$, respectively, $= 2^n$:

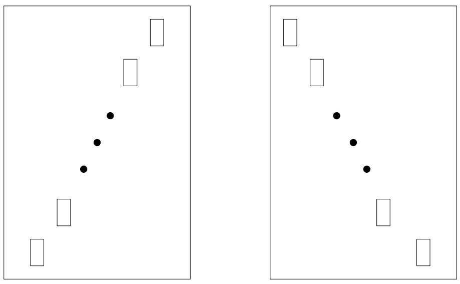


Fig. 15. Components and dipaths up to dihomotopy: two extreme cases

2 Figure 16 gives an example of a di-1-connected state space with non-empty forbidden region.
3 In both drawings in Figure 17, the forbidden region consists of three components F_1, F_2, F_3 with identical minima and maxima. But there are four dihomotopy classes in the space corresponding to the first figure, whereas there are five classes in the second.

Remark 4.7. An implementation of the method determining the dipaths $\vec{\pi}_1(X;\mathbf{0},\mathbf{1})$ up to dihomotopy for the state space $X = I^2 \setminus int(F)$ requires:

— the determination of the *unsafe* and *unreachable* regions with respect to X – Use the algorithm described in Fajstrup *et al.* (1998a; 1998b).
— the determination of the *connected components* of the enlarged forbidden region

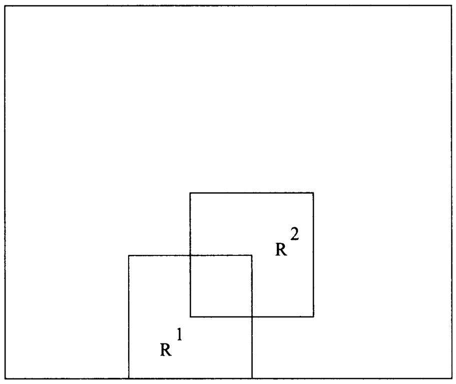


Fig. 16. A di-1-connected state space

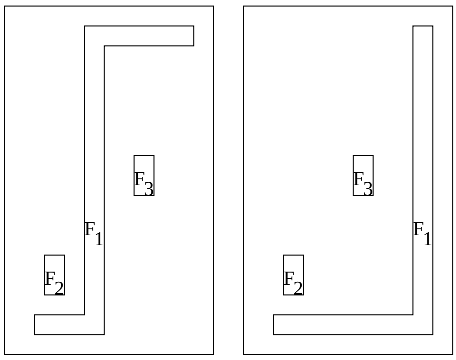


Fig. 17. Two spaces with inequivalent $\vec{\pi}_1$

(including unsafe and unreachable regions, but still consisting of unions of rectangles) and of their minima – A method is outlined in Preparata and Shamos (1988).

— the construction of the subspaces S^i and a determination of their components – This can be done in the same way.

— the determination of the *reachable* subspaces $R(\varphi)$ associated to partial maps φ – See (2) and Remark 3.5.

— to decide whether φ is *admissible* – It is enough to check whether **0** is unsafe in $R(\varphi)$.

— to decide whether φ is *complete* – One has to check whether all intersections $(T_i^x \cap T_y^i) \cap R(\varphi)$ are empty, see (3); this is easy interval arithmetic (with a better organization through dynamic segment trees as in van Kreveld and Overmars (1993), Edelsbrunner (1980), Six and Wood (1982), Aparici and Morezuelas (1999) and Preparata and Shamos (1988)).

Of course, one can avoid checking partial maps that have non-admissible restrictions or non-complete extensions. However, in general the algorithm is exponential in the number of connected components above. On the other hand, the algorithm determining deadlocks, unsafe regions, and so on, is very fast in dimension 2.

5. Outlook on higher dimensions

In dimensions greater than 2, the connected components of the (enlarged) forbidden region cannot give sufficient information about the classification of dipaths. Instead one can hope that the necessary information is related to $(n-2)$- and $(n-3)$-dimensional data of the forbidden region.

To argue for this, we compare with the non-directed case. For two dipaths α and β in X from $\mathbf{0}$ to $\mathbf{1}$, the concatenation $\alpha * \beta^{-1}$ of two dipaths yields a loop in X and thus an element in the fundamental group $\pi_1(X)$ of the state space X. If X is path-connected, the abelianization of the fundamental group is the first homology group $H_1(X)$. This homology group can be calculated in terms of cohomology groups of the forbidden region F using duality: we assume that F is approximated by a smooth manifold with boundary $\partial F = \partial_i F \cup \partial_o F$ such that the outer boundary is $\partial_o F = \partial F \cap \partial I^n$; the inner boundary $\partial_i F$ is the closure of the complement of $\partial_o F$ in ∂F; their intersection $\partial_{io} F = \partial_i F \cap \partial_o F$ is their common boundary and a manifold of dimension $n-2$.

Let U denote a one-sided collar of F inside X; it is homeomorphic to $\partial_i F \times [0,1]$. Using excision (see, for instance, Greenberg and Harper (1981) and Dold (1980)), we obtain the following sequence of isomorphisms:

$$H_1(I^n \setminus int(F)) \cong H_2(I^n, I^n \setminus int(F)) \cong H_2(F \cup U, U) \cong H_2(F, \partial_i F).$$

To analyse this last group, we use the exact relative homology sequence of the triad $(F, \partial F, \partial_i F)$, excision, and Lefschetz duality (Greenberg and Harper 1981) to get a commutative diagram

$$\begin{array}{ccccccccc}
\to & H_3(F,\partial F) & \to & H_2(\partial F,\partial_i F) & \to & H_2(F,\partial_i F) & \to & H_2(F,\partial F) & \to & H_1(\partial F,\partial_i F) \\
 & & & \cong & & & & & & \cong \\
 & \cong & & H_2(\partial_o F,\partial_{io} F) & & & & \cong & & H_1(\partial_o F,\partial_{io} F) \\
 & & & \cong & & & & & & \cong \\
\to & H^{n-3}(F) & \to & H^{n-3}(\partial_o(F)) & \to & H^{n-2}(F,\partial_o F) & \to & H^{n-2}(F) & \to & H^{n-2}(\partial_o(F))
\end{array}$$

For $n=2$, we get that $H_1(I^2 \setminus int(F)) \cong H^0(F, \partial_o(F)) = \ker(H^0(F) \to H^0(\partial_o(F))$ is the free abelian group with base the *connected components* of F not touching the boundary ∂I^2. One may speculate to what extent (reachable) representatives of $(n-2)$ dimensional (co)-homology groups (and $(n-3)$-dimensional ones on the boundary) can play a similar role in higher dimensions. Anyhow, it is possible to calculate the (co)-homology groups above from the hyperrectangles R^i that F is built from: knowing which of them have non-empty intersections allows us to construct the nerve of the covering of F that they constitute (as a simplicial complex). The nerve's homology is then the homology of F.

References

Aparici, I. and Morezuelas, S. (1999) Détection rapide de points morts. Technical Report 99-056, CEA/LETI.

Brown, R. and Higgins, P. (1981a) Colimit theorems for relative homotopy groups. *J. Pure Appl. Algebra* **22** 11–41.

Brown, R. and Higgins, P. (1981b) On the algebra of cubes. *J. Pure Appl. Algebra* **21** 233–260.

Carson, S. and Reynolds, P. (1987) The geometry of semaphore programs. *ACM TOPLAS* **9** (1) 25–53.

Dijkstra, E. (1968) Co-operating sequential processes. In: Genuys, F. (ed.) *Programming Languages*, Academic Press 43–110.

Dold, A. (1980) Lectures on Algebraic Topology. *Grundlehren der matematischen Wissenschaften* **200**, Springer-Verlag (2nd edition).

Edelsbrunner, H. (1980) Dynamic data structures for orthogonal intersectin queries. Technical Report F59, Tech. Univ. Graz, Institut für Informationsverarbeitung.

Fajstrup, L. (1999) Loops, ditopology and deadlocks. *Mathematical Structures in Computing Science*, this issue.

Fajstrup, L., Goubault, E. and Raussen, M. (1998a) Detecting Deadlocks in Concurrent Systems. DTA/LETI/DEIN/SLA 98-61, LETI (CEA - Technologies Avancées), Saclay, France.

Fajstrup, L., Goubault, E. and Raussen, M. (1998b) Detecting Deadlocks in Concurrent Systems. In: Sangiorgi, D. and de Simone, R. (eds.) CONCUR '98; Concurrency Theory. Proceedings 9th Int. Conf., Nice, France. *Springer-Verlag Lecture Notes in Computer Science* **1466** 332–347.

Fajstrup, L., Goubault, E. and Raussen, M. (1999) Algebraic topology and concurrency. Technical Report R-99-2008, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst.

Gierz, G., Hofmann, K., Keimel, K., Lawson, J., Mislove, M. and Scott, D. (1980) *A Compendium of Continuous Lattices*, Springer-Verlag.

Goubault, E. (1995) *The Geometry of Concurrency*, Ph. D. thesis, Ecole Normale Superieure, Paris.

Greenberg, M. J. and Harper, J. R. (1981) Algebraic Topology. A first Course. *Mathematics Lecture Notes Series* **58**, Benjamin/Cummings, Reading, Mass.

Gunawardena, J. (1994) Homotopy and concurrency. *Bulletin of the EATCS* **54** 184–193.

Lipski, W. and Papadimitriou, C. (1981) A fast algorithm for testing for safety and detecting deadlocks in locked transaction systems. *Journal of Algorithms* **2** 211–226.

Mazurkiewicz, A. (1988) Basic notions of trace theory. In: *Lecture notes for the REX summer school in temporal logic*, Springer-Verlag.

Penrose, R. (1972) Techniques of Differential Topology in Relativity. *Conference Board of the Mathematical Sciences, Regional Conference Series in Applied Mathematics* **7**, SIAM.

Pratt, V. (1991) Modeling concurrency with geometry. *Proc. of the 18th ACM Symposium on Principles of Programming Languages.*

Preparata, F. and Shamos, M. (1988) *Computational geometry, an introduction*, Springer-Verlag.

Serre, J. P. (1951) Homologie singulière des espaces fibrés. *Ann. of Math.* **54** (2) 425–505.

Six, W. and Wood, D. (1982) Counting and reporting intersections of d-ranges. *IEEE Transactions on computers* **31** (3) 181–187.

van Glabbeek, R. (1991) Bisimulation semantics for higher dimensional automata. Technical report, Stanford University.

van Kreveld, M. and Overmars, M. (1993) Union-copy structures and dynamic segment trees. *Journal of the Association for Computing Machinary* **40** (3).

Dihomotopy as a Tool in State Space Analysis
Tutorial

Éric Goubault[1] and Martin Raussen[2]

[1] LIST (CEA Saclay), DTSI-SLA-LSL, 91191 Gif-sur-Yvette, France,
Eric.Goubault@cea.fr
[2] Institute of Mathematical Sciences, Aalborg University,
9220 Aalborg Øst, Denmark,
raussen@math.auc.dk

Abstract. Recent geometric methods have been used in concurrency theory for quickly finding deadlocks and unreachable states, see [14] for instance. The reason why these methods are fast is that they contain in germ ingredients for tackling the state-space explosion problem. In this paper we show how this can be made formal. We also give some hints about the underlying algorithmics. Finally, we compare with other well-known methods for coping with the state-space explosion problem.

1 Introduction

In model-checking techniques, temporal formulas, expressing important properties on traces of a concurrent system one has to verify, are checked by traversing the interleaving semantics of the program. This runs unfortunately into the "*state-space explosion problem*": the number of paths to be traversed might be exponential in the number of processes involved. It has been very tempting for a number of authors to try to use the information about the *independence* of actions to decrease this number by a possibly exponential ratio. For instance, if all actions considered are completely independent, meaning that any interleaving of actions taken in this set of actions computes the same thing, as a function from (parallel or distributed) store to store, then there is no need to consider all the interleavings to check any kind of "interesting" properties, such as safety or deadlock properties.

But this is not always as simple as we show with the transition system of Figure 1. Here we suppose that a and b are independent or "commuting" actions. The problem in Figure 1 is that we might choose to traverse only path $a.b$ since it is equivalent to $b.a$ and we will have missed the branching after b, which would have lead us into transition system C, which might contain any deadlock we want for instance. In fact, there are correct ways to infer state-space reduction methods from the independence relation. A classical one explained in Section 7.1 has been originally introduced by Valmari [46] under the name of "stubborn sets", based on a notion of independence on Petri nets. These have been ameliorated later under the name of "persistent sets" by Godefroid [21], based on the notion of independence of asynchronous transition systems. We develop in this

S. Rajsbaum (Ed.): LATIN 2002, LNCS 2286, pp. 16–37, 2002.

paper new methods for finding better state-space reduction techniques, based on *global* semantical information. This is done using geometric ideas, which have recently regained impetus after the seminal work [12] and [36]. We formalize this methodology, the "diconnected components" of the geometric semantics using a category of fractions of the fundamental category of the semantics, giving all information about all possible schedules of execution.

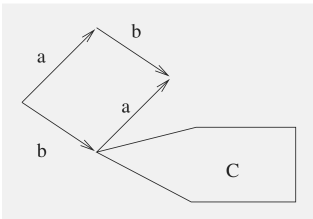


Fig. 1.

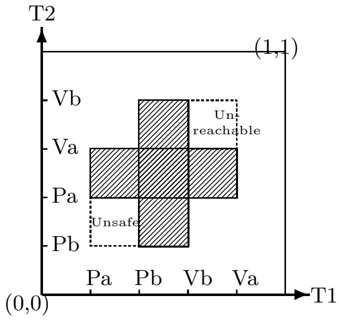


Fig. 2. Example of a progress graph

History: Towards Higher Dimensional Automata. The first "algebraic topological" model in the litterature is called *progress graph* and has appeared in operating systems theory, in particular for describing the problem of "deadly embrace" in "multiprogramming systems". *Progress graphs* are introduced in [10], but attributed to E. W. Dijkstra. The basic idea is to give a description of what can happen when several processes are modifying shared ressources. Given a shared resource a, we see it as its associated *semaphore* that rules its behaviour with respect to processes. For instance, if a is an ordinary shared variable, it is customary to use its semaphore to ensure that only one process at a time can write on it (this is mutual exclusion). Then, given n deterministic sequential processes $Q_1, \ldots, Q_n$, abstracted as a sequence of locks and unlocks on shared objects, $Q_i = R^1 a_i^1 . R^2 a_i^2 \cdots R^{n_i} a_i^{n_i}$ (R^k being P or V[1]), there is a natural way to understand the possible behaviours of their concurrent execution, by associating to each process a coordinate line in $\mathbf{R}^n$. The state of the system corresponds to a point in $\mathbf{R}^n$, whose ith coordinate describes the state (or "local time") of the ith processor.

Consider a system with finitely many processes running altogether. We assume that each process starts at (local time) 0 and finishes at (local time) 1; the P and V actions correspond to sequences of real numbers between 0 and 1, which reflect the order of the P's and V's. The initial state is $(0, \ldots, 0)$ and the final

[1] Using E. W. Dijkstra's notation P and V [12] for respectively acquiring and releasing a lock on a semaphore.

state is $(1, \ldots, 1)$. An example consisting of the two processes $T_1 = Pa.Pb.Vb.Va$ and $T_2 = Pb.Pa.Va.Vb$ gives rise to the two dimensional *progress graph* of Fig. 2. The shaded area represents states which are not allowed in any execution path, since they correspond to mutual exclusion. Such states constitute the *forbidden region*. An *execution path* is a path from the initial state $(0, \ldots, 0)$ to the final state $(1, \ldots, 1)$ avoiding the forbidden region and increasing in each coordinate - time cannot run backwards. We call these paths *directed paths* or *dipaths*. This entails that paths reaching the states in the dashed square underneath the forbidden region, marked "unsafe" are deemed to deadlock, i.e. they cannot possibly reach the allowed terminal state $(1, 1)$. Similarly, by reversing the direction of time, the states in the square above the forbidden region, marked "unreachable", cannot be reached from the initial state $(0, 0)$. Notice that all terminating paths above the forbidden region are "equivalent" in some sense: they are all characterized by the fact that T_2 gets a and b before T_1 (as far as resources are concerned, we call this a *schedule*). Similarly, all paths below the forbidden region are characterized by the fact that T_1 gets a and b before T_2 does.

In this picture, one can already recognize many ingredients that are at the center of algebraic topology, namely the classification of shapes modulo "elastic deformation". As a matter of fact, the actual coordinates that are chosen to represent the times at which Ps and Vs occur are unimportant, and these can be "stretched" in any manner, so the properties (deadlocks, schedules etc.) are invariant under some notion of deformation, or *homotopy*. This has to be a particular kind of homotopy though causing many difficulties in later work. We call it (in subsequent sections) a *directed homotopy* or *dihomotopy* in the sense that it should preserve the direction of time.

The semantics community came back to these geometric considerations with the development of "truly-concurrent" semantics, as opposed to "interleaving" semantics. The base of the argument was that interleaving semantics, i.e. the representation of parallelism by non-determinism ignores real asynchronous behaviours: $a \mid b$ where a and b are atomic is represented by the same transition system as the non-deterministic choice a then b or b then a. This fact creates problems in static analysis of (asynchronous) concurrent systems: Interleaving builds a lot of uninteresting states in the modelisation, hence induces a high cost in verification. This is called the *state-space explosion problem*. Quite a few models for true-concurrency have appeared (see in particular the account of [50]) but it is only in 1991 that geometry is proposed to solve the problem, in [36]. The diagnosis is that interleaving is only the boundary of the real picture. $a \mid b$ is really the filled-in square whose boundary is the non-deterministic choice a then b or b then a (the hollow square). The natural combinatorial notion, extension of transition systems, is that of a *cubical set*, which is a collection of points (states), edges (transitions), squares, cubes and hypercubes (higher-dimensional transitions representing the truly-concurrent execution of some number of actions). This is introduced in [36] as well as possible formalizations using n-categories, and a notion of homotopy. This is actually a combinatorial view of some kind of progress graph. Look back to Figure 2. Consider all interleavings of actions Pa,

Pb, Va and Vb: they form a subgrid of the progress graph. Take as 2-transitions (i.e. squares in the cubical set we are building) the filled-in squares. Only the forbidden region is really interleaved. Cubical sets generalize progress graphs, in that they allow any amount of non-deterministic choices as well as dynamic creation of processes. These cubical sets are called *Higher-Dimensional Automata* (HDA) following [36] because it really makes sense to consider a hypercube as some form of transition. Actually at about the same time, a bisimulation semantics was given in [47]. Notice that 2-transitions or squares are nothing but a *local commutation relation* as in Mazurkiewicz trace theory [34], *independence relation* as in asynchronous transition systems, see [2], as in trace automata, as in transition systems with independence [40], or (indirectly) as with the "confluence" relation of concurrent transition systems [45]. There are two more ingredients with HDA: the elegance and the power of the tools of geometric formalisations, and the natural generalisation to *higher* dimensions (i.e. "higher-order independence relation" or n-ary independence relations).

Example: Semaphores and progress graphs. In the rest of the paper, we will stick to one particular model which is sufficiently simple to explain, and gives sufficiently many nasty example: the shared memory model, in which asynchronous processes read and write atomically onto variables which are all in a common (shared) memory. To protect writing onto shared variables, we use mutual exclusion locks, which we put explicitely before writing a variable x, by Px, and that we release explicitely after, by Vx. It is then easy to see that writing on two distinct variables are two independant actions, as well as reading two variables (even the same one) by two processes. This model can also easily include [16] counting semaphores which are weakly synchronising objects that can be shared by n but not $n+1$ processes at the same time (for some $n > 1$). Notice that asynchronous message-passing with bounded buffers can be translated into that framework. It is therefore not only a useful example, but a quite general application indeed.

The key idea is to regard a progress graph as a topological space in which points are ordered globally through time, i.e., equipped with a (closed) partial order $\leq$. Traces of executions are continuous and increasing maps from the totally ordered unit segment to $(X, \leq)$. These are called *dipaths* for "directed paths". A dihomotopy between two dipaths f and g on X is a deformation via dipaths interpolating continuously between f and g and fixing the endpoints. The technical definitions will be given in Sect. 3. Now we can give semantics to a very simple language in which a finite number of processes can only do a deterministic sequence of lockings Px and unlockings Vx of shared resources x. So processes are just strings of P's and V's. Suppose that each semaphore x (binary or counting) is equipped with a number $s(x)$, the maximal number of processes that can share it at any time. Supposing that the length of the strings X_i $(1 \leq i \leq n)$ are integers l_i, the semantics of *Prog* is included in $[0, l_1] \times \cdots \times [0, l_n]$. A description of the progress graph $[\![Prog]\!]$ associated with *Prog* can be given by describing inductively what should be digged into this hyperrectangle. The semantics of our language can be described by the simple rule, $[k_1, r_1] \times \cdots \times [k_n, r_n] \in [\![X_1 \mid \cdots \mid X_n]\!]_2$ if there is a partition of $\{1, \cdots, n\}$ into $U \cup V$ with $card(U) = s(a) + 1$ for some object a with, $X_i(k_i) = Pa$, $X_i(r_i) = Va$ for $i \in U$ and $k_j = 0$, $r_j = l_j$ for $j \in V$. This language is somehow disappointing. To be able to consider looping and branching constructs, we are lead to the notion of local po-spaces in Sect. 3.1.

Goals of the present paper. After having explained the geometric semantics, the idea of deformation of paths of executions, and introduced the diconnected components approach to the state-space explosion problem, we compare (favorably) our technique with classical techniques such as persistent sets. We also review in Sect. 7.3 some orthogonal techniques which could still be used on top of our geometric technique.

2 The Fundamental Group of a Topological Space

In this section, we give a brief review of the fundamental group of a topological space, a very important concept from algebraic topology. See e.g. [1,5,27,35] for details. Hereafter, we develop a variation of this notion and apply it to state space analysis.

Topological spaces are abstractions of metric spaces. For a metric space X, nearness is expressed by a metric d measuring the distance between pairs of points. For a topological space Y, nearness is expressed with the aid of a collection of *open* subsets of Y. The usual definition for a continuous map between two metric spaces has the following generalisation for topological spaces: A map $f : Y \to Y'$ between topological spaces is *continuous* if and only if $f^{-1}(U) \subset Y$ is open for every open subset $U \subset Y'$.

In this paper, we will mainly be concerned with (different types of) *paths*, i.e., continuous maps $f : I \to X$ from an interval I into a topological space X. For the moment, we let $I = [0, 1]$ denote the unit interval with standard metric and topology. In general, one cannot compose paths in X. But if the endpoint $f_1(1)$ of f_1 coincides with the start point $f_2(0)$ of f_2, their *concatenation*

$$f_2 * f_1 : I \to X \text{ is defined by } (f_2 * f_1)(s) = \begin{cases} f_1(2s), & t \leq \frac{1}{2} \\ f_2(2s-1), & t \geq \frac{1}{2}. \end{cases}$$

Both paths are thus pursued with "double speed". Concatenation defines a (non-commutative, non-associative) operation on the space $\mathcal{P}(X)$ of all paths on X.

Two points $x, y \in X$ are called *path-connected*, if there exists a path f with $f(0) = x$ and $f(1) = y$. The equivalence classes of this equivalence relation are called the *path components* of X. The image $f(X_0) \subset Y$ of a path component $X_0 \subset X$ under a continuous map $f : X \to Y$ is path-connected. As a consequence, path components are completely independent of each other, and one can investigate them "one at a time". A *loop* in a topological space X is a path $f : I \to X$ such that $f(0) = f(1)$. Loops with the same start/end-point can be concatenated. A *homotopy* of paths (loops) is a continuous map $H : I \times I \to X$ with $H(t, 0) = H(0, 0)$ and $H(t, 1) = H(0, 1)$ for all $t \in I$. It should be regarded as a one-parameter family of paths $H_t : I \to X,\ H_t(s) = H(t, s)$ (with fixed end points) connecting H_0 and H_1. Two paths $f_0, f_1 : I \to X$ with the same end-points are called *homotopic* if there is a fixed end point homotopy $H : I \times I \to X$ with $H_0 = f_0$ and $H_1 = f_1$. Homotopy is an equivalence relation.

A continuous and strictly increasing map $\varphi : I \to I$ with $\varphi(0) = 0$ and $\varphi(1) = 1$ can be used to reparametrise a path, i.e., to pass from a path f in X to the (reparameterised) path $f \circ \varphi$ with the same image. Remark that φ is homotopic to the identity map on I; a homotopy is given by $H(t, s) = (1-t)\varphi(s) + ts$. As a consequence, the paths f and its reparametrisation $f \circ \varphi$ are homotopic via the homotopy $\bar{H}(t, s) = f(H(t, s))$.

A basic invariant of a topological space X is its *fundamental group*: Fix a base point $x_0 \in X$. The elements of the fundamental group $\pi_1(X; x_0)$ are the homotopy classes of *loops* $f : I \to X$ which start and end at $f(0) = f(1) = x_0$. Concatenation of loops at x_0 factorizes over homotopy to yield a 2-adic operation on $\pi_1(X; x_0)$. The homotopy class of the constant map $c : I \to X,\ c(s) = x_0,\ s \in I$, serves as the neutral element – since $f, f * c$ and $c * f$ are homotopic to each other. The inverse to the class of the loop f is given by the the class of the loop $f^- : I \to X, f^-(t) = f(1-t)$: $f^- * f$ and $f * f^-$ are both homotopic to c.

The size of the fundamental group has an interesting interpretation: A loop f can be regarded as a map from the unit circle $\bar{f} : S^1 \to X,\ \bar{f}(\exp(2\pi i s)) = f(s)$. The loop f represents the trivial element in $\pi_1(X; x_0)$ if it is homotopic to the constant loop c. A homotopy H with $H_0 = c$ and $H_1 = f$ can be transformed into an extension $\bar{H} : D^2 \to X$ of $\bar{f}$, viz. $\bar{H}(t \exp(2\pi i s)) = H(t, s)$. Conversely, an extension of $\bar{f}$ to a continuous map $\bar{H} : D^2 \to X$ gives rise to a homotopy between f and c. A homotopically trivial loop can thus be "filled in". Hence, the the fundamental group of a space "counts the numbers of holes" in it.

The fundamental group of a space does only depend on the *path component* of the base point: Let g denote an arbitrary path with $g(0) = x_0$ and $g(1) = x_1$. Then the map "conjugation with g": $\pi_1(X; x_0) \to \pi_1(X; x_1);\ [f] \mapsto [g^- * f * g]$ is a *group isomorphism* .

Examples: Proofs of the following statements can be found in almost any text-book on algebraic topology:

- The fundamental group of Euclidean space $\mathbf{R}^n$ is trivial for all n.

- The fundamental group of the unit circle S^1 is isomorphic to the integers. An explicit isomorphism $\pi_1(S^1) \to \mathbf{Z}$ associates to a loop its "winding number", i.e., it counts (with a sign) the number of times a particular value is attained. The fundamental group of an n-sphere $S^n = \{x \in \mathbf{R}^n \mid ||x|| = 1\}$ is trivial for every $n > 1$.

- The fundamental group of "the figure 8" (two circles with only a single base point in common) is the free group on two letters representing the two directed loops.

- For every group G, there is a path-connected topological space BG with $\pi_1(BG) \simeq G$.

3 The Fundamental Category of an Lpo-space

3.1 Lpo-spaces and Dipaths

There are many models for state spaces for concurrent processes and the executions on them, cf. Sect. 7. In this paper, we follow the basic idea from [16]: A po-space consists of a *topological space* X with a *partial order* $\leq \subset X \times X$. The partial order is assumed to be closed (as a subset of $X \times X$) to ensure coherence between topology and order: this makes it possible to take limits "under the $\leq$ sign". For an example of such a po-space (in fact, a progress graph) see Fig. 3; the left figure represents the state space for two processes that acquire and relinquish a lock to a single shared resource; the right figure pictures the situation where locks to two shared resources have to be acquired in reverse order by the two processes. The black areas are the "forbidden regions" of the progress graph which are *not* part of the state space.

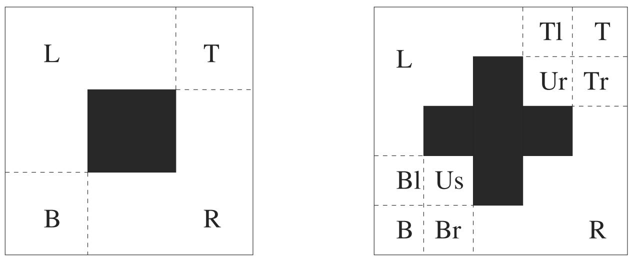


Fig. 3. Square with a hole and complement of a "Swiss flag"

If one or several of the processes contain loops, the resulting abstraction will no longer have a *global* partial order. Instead one requires for a local po-space (*lpo-space*) a relation $\leq$ on X that restricts to a partial order on sufficiently small subsets of X that form a basis for the topology. Two such relations are equivalent (and define the same local partial order) if they agree on sufficiently small open sets forming a basis for the topology. For an example, consider the relation on the unit circle $S^1 \subset \mathbf{R}^2$ given by $x \leq y \Leftrightarrow$ the angle from x to y is less than α. This relation is certainly not transitive, but it defines a local partial order if and only if $\alpha \leq \pi$; for $\alpha \leq \pi$ these are all equivalent.[2]

Traces of a concurrent system (executions) are modelled by so-called *dipaths*– di is an abbreviation for *directed*. A *short*, resp. *long* dipath in an lpo-space X is defined as an *order preserving continuous* map from the interval $\vec{I} = [0, 1]$, resp. from the non-negative reals $\mathbf{R}_{\geq 0} = \{x \in \mathbf{R} | \ x \geq 0\}$ (with the natural order) into X. A short dipath models a concurrent process from a start point to

[2] This version of the definition is due to Ulrich Fahrenberg; it is in fact equivalent to the one given in [16,17].

an end point, while a long dipath runs indefinitely (e.g., in loops) but avoiding *zeno* executions. Technically, one requires that a long dipath does not admit a limit for $t \to \infty$.

3.2 Dihomotopy

When can you be sure that two execution traces in a concurrent program provide the same result? This is the case if the corresponding dipaths $f, g : I \to X$ are *dihomotopic*. This means, that there exists a continuous order-preserving *dihomotopy* $H : I \times \vec{I} \to X$ with $H_0 = f$ and $H_1 = g$. Remark that the parameter interval is equipped with the trivial order, i.e., $(t, s) \leq (t', s') \Leftrightarrow t = t' \wedge s \leq s'$. In particular, every "intermediate" path H_t has to be a dipath. Moreover, we require a fixed start point ($H(t, 0) = H(0, 0)$) and, for short dipaths, a fixed end point ($H(t, 1) = H(0, 1)$); for long dipaths all the paths H_t have to be non-zeno.

3.3 The Fundamental Category

For an lpo-space, one should no longer watch out for a fundamental *group*. The reverse of a dipath is no longer a dipath. On a global po-space, there are no (non-trivial) directed loops at all. Instead, one has to work with the fundamental *category* of a local po-space X, or rather with two versions of it, depending on whether short or long dipaths are considered:

The *objects* of the fundamental category $\vec{\pi}_1(X)$ are the points of X. The *morphisms* between elements x and y are given as the dihomotopy classes in $\vec{\pi}_1(X; x, y)$. Composition of morphisms

$$\vec{\pi}_1(X; x, y) \times \vec{\pi}_1(X; y, z) \to \vec{\pi}_1(X; x, z)$$

is given by concatenation of dipaths – up to dihomotopy.

The category $\vec{\pi}_1^\infty(X)$ contains $\vec{\pi}_1(X)$. It has an additional maximal element ∞ with $Mor(x, \infty)$ consisting of the dihomotopy classes of long dipaths starting at x and $Mor(\infty, y) = \emptyset$ for $y \in X \cup \{\infty\}$. Concatenation of a (short) dipath from x to y with a (long) dipath from y yields a (long) dipath – up to dihomotopy via long dipaths.

Compared to the fundamental group, a fundamental category is an enormous gadget and it has a much less nice algebraic structure. On the other hand, in easy examples one has the impression, that the cardinality of the set of morphisms between two points is quite robust when these points are perturbed a little.

Example 1. For the square with one hole (Fig. 3), there is no morphism between the regions marked L, resp. R, and no morphism from T to any other region, neither a morphism from any other region to B. There are two morphisms from any point of B to any point of T. Moreover, from any point of B, certain points of B, L, R can be reached by (exactly one) morphism. Likewise, any point of T can be reached from (certain of the points in) L, R and T in one way.

For the complement of a "Swiss flag" (Fig. 3), the situation is a bit more complicated: There is no dipath leaving the unsafe rectangle *Us* and there is no dipath entering the unreachable rectangle *Ur* from the outside. It is possible to reach Us by a dipath from $B \cup Bl \cup Br$; from Ur, one can reach $Tl \cup Tr \cup T$. The only possibility for *two* classes of dipaths between points occurs when the first is in B and the second in T. Moreover, these classes can be represented by dipaths along the boundary, representing the two sequential executions.

The lesson to learn is that the complete "dynamics" of these state space can be described from the decomposition into the blocks studied above. It is the aim of this paper to define and describe these "dicomponents" in the general case and thus, in a realistically large model, to provide a "collapse" of the exponentially large state space into pieces that show the same behaviour with respect to execution paths between each other. It is then enough to study the "flow" between these "components" in order to capture the dynamics of the whole system.

4 Categories of Fractions and Components

4.1 Categories of Fractions

Next, we have to invest in a construction from category theory: We invert in a systematic way all those partial dipaths that never contribute to a decision along any dipath. The resulting category will then have many "zig-zag" isomorphisms giving rise to the components. Here is a general method [18,4]:

Let $\mathcal{C}$ denote a category. To keep things simple, assume $\mathcal{C}$ small, i.e., objects and morphisms are sets. Let $\Sigma \subset Mor(\mathcal{C})$ denote a *system of morphisms*, i.e., Σ includes all unit morphisms and is closed under composition. For any such system, one may construct the category of fractions $\mathcal{C}[\Sigma^{-1}]$ and the localization functor $q_\Sigma : \mathcal{C} \to \mathcal{C}[\Sigma^{-1}]$ [18,4] having the following universal property:

- For every $s \in \Sigma$, the morphism $q_\Sigma(s)$ is an *iso*morphism. - For any functor $F : \mathcal{C} \to \mathcal{D}$ such that $F(s)$ is an isomorphism for every $s \in \Sigma$, there is a unique functor $\theta : \mathcal{C}[\Sigma^{-1}] \to \mathcal{D}$ with $\theta \circ q_\Sigma = F$.
The objects of $\mathcal{C}[\Sigma^{-1}]$ are just the objects of $\mathcal{C}$. To define the morphisms of $\mathcal{C}[\Sigma^{-1}]$, one introduces a (formal) inverse s^{-1} to every morphism $s \in \Sigma(x, y)$. These inverses are collected in $\Sigma^{-1}(y, x), x, y \in Ob(\mathcal{C})$ and then in Σ^{-1}. Consider the closure of $Mor(\mathcal{C}) \cup \Sigma^{-1}$ under composition and the smallest equivalence relation containing $s^{-1} \circ s = 1_x$ and $s \circ s^{-1} = 1_y$ for $s \in \Sigma(x, y)$ that is compatible with composition. The equivalence classes correspond then to the morphisms of $\mathcal{C}[\Sigma^{-1}]$. In particular, if $t \circ \alpha = \beta \circ s$ for $s, t \in \Sigma$, then $\alpha \circ s^{-1} = t^{-1} \circ \beta$. A morphism in $\mathcal{C}[\Sigma^{-1}]$ can always be represented in the form

$$s_k^{-1} \circ f_k \circ \cdots \circ s_1^{-1} \circ f_1,\ s_j \in \Sigma, f_j \in Mor, k \in \mathbf{N}.$$

Let $Iso(\mathcal{C})$ denote the isomorphisms of the category $\mathcal{C}$, and let $\Sigma * Iso(\mathcal{C})$ denote the system of morphisms generated by Σ and by $Iso(\mathcal{C})$. The isomorphisms in $\mathcal{C}[\Sigma^{-1}]$ are the *zig-zag* morphisms, i.e.,

$$Iso(\mathcal{C}[\Sigma^{-1}]) = \{s_1^{-1} \circ s_2 \circ \cdots \circ s_{2k-1}^{-1} \circ s_{2k}, s_j \in \Sigma * Iso(\mathcal{C}), k \in \mathbf{N}\}.$$

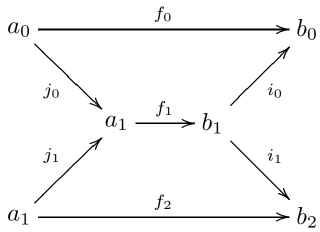
The subcategory of $\mathcal{C}[\Sigma^{-1}]$ with all objects, the morphisms of which are given by the zig-zag morphisms in $Iso(\mathcal{C}[\Sigma^{-1}])$, forms in fact a *groupoid* [30].

4.2 Components

A “compression” of the category $\mathcal{C}[\Sigma^{-1}]$ is achieved by dividing out all isomorphisms: Two objects $x, y \in Ob(\mathcal{C})$ are Σ-isomorphic or Σ-connected – $x \simeq_\Sigma y$ – if there exists a zig-zag-morphism from x to y. This definition corresponds to usual path connectedness *with respect to paths representing isomorphisms only – but regardless orientation.* Σ-connectivity is an equivalence relation; the equivalence classes are called the *Σ-connected components* – the path components with respect to Σ-zig-zag paths, viz. the components of the groupoid above.

Next, consider the equivalence relation on the morphisms of $\mathcal{C}[\Sigma^{-1}]$ generated (under composition) by $\alpha \simeq \alpha \circ s, \quad \alpha \simeq t \circ \alpha$ with $\alpha \in Mor(x, y), s \in Inv(\mathcal{C}[\Sigma^{-1}])(x', x), t \in Inv(\mathcal{C}[\Sigma^{-1}])(y, y')$. Remark that equivalent morphisms no longer need to have the same source or target. Remark moreover, that any two zig-zag morphisms from x to y are equivalent to each other; in particular, they are equivalent to the unit-morphisms in both x and y.

Example 2. If $i_0, i_1, j_0, j_1 \in Inv(\mathcal{C}[\Sigma^{-1}])$, then $f_0, f_1, f_2 \in Mor(\mathcal{C})$ in the following diagram are equivalent to each other in $\mathcal{C}[\Sigma]^{-1}$:



The objects of the *component category* $\pi_0(\mathcal{C}; \Sigma)$ are by definition the Σ-connected components of $\mathcal{C}$; the morphisms from $[x]$ to $[y]$, $x, y \in Ob(\mathcal{C})$, are the equivalence classes of morphisms in $\bigcup_{x' \simeq_\Sigma x, y' \simeq_\Sigma y} Mor_{\mathcal{C}[\Sigma^{-1}]}(x', y')$. The composition of $[\beta] \circ [\alpha]$ for $\alpha \in Mor_{\mathcal{C}[\Sigma^{-1}]}(x, y)$ and $\beta \in Mor_{\mathcal{C}[\Sigma^{-1}]}(y', z)$ is given by $[\beta \circ s \circ \alpha]$ with s any zig-zag morphism from y to y'. The equivalence class of that composition is independent of the choices taken.

Remark 1. These constructions decompose the study of the morphisms of $\mathcal{C}$ into two pieces: Firstly, the study of the groupoid $Inv(\mathcal{C}[\Sigma^{-1}])$ which can be performed separately on each of the Σ-connected components. For the fundamental category, all these morphisms represent executions that can be performed and/or backtracked without global effects. Secondly, certainly more important for applications, the study of the component category, which encompasses the global effects of irreversibility. In the case of the component category of the fundamental category, representatives of all non-unit dipath classes may have (different) global effects – backtracking along such a dipath class may therefore change the result of a computation.

5 Applications to State Space Analysis

In this section, we apply the preceeding constructions to our models of the state space and the space of executions (from a given initial state) of a concurrent program. The key task is to single out the relevant system of morphism Σ that is to be inverted. It should consist of morphisms that, *from a global point of view*, do not contribute with any *decision* to the outcome of the concurrent program. Here, we give the key definitions (in a general categorial framework), their motivation, and a few elementary examples. For algorithms in low dimensions, cf. Sect. 6.

5.1 Extensions of Morphisms

For a small category, let $X_0, X_1 \subset Ob(\mathcal{C})$. Let $Mor_{0,1} = \bigcup_{x_0 \in X_0, x_1 \in X_1} Mor(x_0, x_1)$ denote the set of all morphisms between X_0 and X_1. We associate to a morphism $f \in Mor(x, y)$ with $x, y \in Ob(\mathcal{C})$, the set of all its *extensions*

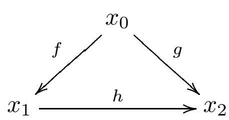
$$\mathcal{E}(f) = \{g \circ f \circ h \mid h \in Mor(X_0, x), g \in Mor(x, X_1)\} \subset Mor_{01}$$

from X_0 to X_1. This set consists of all morphisms from X_0 to X_1 that *factor through* f. It is empty if $Mor(X_0, x) = \emptyset$ or if $Mor(y, X_1) = \emptyset$. In the particular case $f = 1_x$, the unit at $x \in Ob(\mathcal{C})$, the set $\mathcal{E}(x) = \mathcal{E}(1_x)$ consists of all morphisms from X_0 to X_1 factoring through x.

For concatenable morphisms f_1, f_2 it is obvious that $\mathcal{E}(f_2 \circ f_1) \subseteq \mathcal{E}(f_1) \cap \mathcal{E}(f_2)$. The geometric example Ex. 2.1 in [38] shows that the left hand side may be a *proper* subset of the right hand side.

5.2 Components on the Space of Executions

The space of partial executions of a concurrent program is modelled as the set of morphisms from the initial point x_0 in the fundamental category $\vec{\pi}_1(X)$. More generally, one may associate to any category $\mathcal{C}$ and any object $x_0 \in Ob(\mathcal{C})$ the comma category $(x_0 \downarrow \mathcal{C})$ of *objects under* x_0 [33]: Its objects are the morphisms in $Mor(x_0, x)$, $x \in Ob(\mathcal{C})$, and its morphisms are the *commutative* triangles



with x_0 in the top and $h \in Mor(x_1, x_2)$.

If $\mathcal{C}$ is the fundamental category $\vec{\pi}_1(X)$ and x_0 an initial element, the comma categories $(x_0 \downarrow \vec{\pi}_1(X))$ and $(x_0 \downarrow \vec{\pi}_1^{\infty}(X))$ have as objects the dihomotopy classes of dipaths starting at x_0 : a partial dipath $h \in \vec{\pi}_1(x_1, x_2)$ with $x_1 \in X$ and $x_2 \in X \cup \{\infty\}$ induces a map $\vec{\pi}_1(x_0, x_1) \to \vec{\pi}_1(x_0, x_2)$ by concatenation.

Assume given a (minimal) object x_0 such that $X_0 = \{x_0\}$ and a set X_1 of maximal objects in a category $\mathcal{C}$. For the fundamental category $\vec{\pi}_1(X)$, this set X_1 should be chosen as a discrete set of final accepting states[3], for the fundamental category $\vec{\pi}_1^{\infty}(X)$, *the* maximal object should be chosen as ∞.

[3] which could include deadlocking points

Definition 1. *A morphism s from $f \in Mor(x_0, x)$ to $g \in Mor(x_0, y)$ belongs to Σ_1 if and only if $\mathcal{E}(f) = \mathcal{E}(g) \subseteq Mor_{01}$.*

It is clear that either every or no morphism from f to g is contained in Σ_1. Obviously, Σ_1 contains the units and is closed under composition. For $\mathcal{C} = \vec{\pi}_1(X)$, a dipath s extending f to g is contained in Σ_1 if no "decision" has been made in between – all "careers" in $\vec{\pi}_1(X; x_0, X_1)$ open to f are still open to g. No (globally detectable) branching occurs between f and g.

A detection of the component category wrt. Σ_1 entails the following benefit:

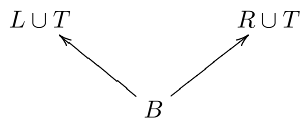
Proposition 1. *Two dipaths f and g from x_0 to x_1 that proceed through the same Σ_1-components are dihomotopic.*

We illustrate the resulting component categories by two elementary examples:

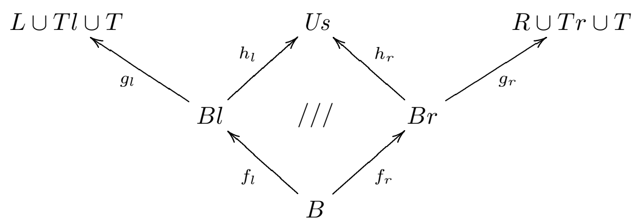
Example 3. Let x_0, resp. x_1 denote the minimal, resp. the maximal element in the po-space X, the square with one hole from Fig. 3. Then Mor_{01} has two elemens represented by dipaths f_L, f_R touching R, resp. L. Any dipath within B and any dipath within $L \cup R \cup T$ is in Σ_1. No dipath starting within B and ending in L or R is in Σ_1.

The category $(x_0 \downarrow \vec{\pi}_1(X))$ consists of three Σ_1-connected components: the dipaths ending in B; those touching L and those touching R. Observe: There is *no* zig-zag path $t^{-1} \circ s$ from a dipath to R via T to a dipath to L since there are no dipaths f from x_0 to R and g from x_0 to L with $t * g$ dihomotopic to $s * f$.

The component category $\pi_0(x_0 \downarrow \vec{\pi}_1(X)), \Sigma_1)$ can – apart from the units – be represented by (end points in)



Example 4. Let Y denote the "Swiss flag" po-space from Fig. 3. Let x_0 and x_1 denote the minimal, resp. the maximal elements, and let y denote the deadlock point (maximal within the unsafe region *Us*). The set of accepting states is $X_1 = \{x_1, y\}$, and Mor_{01} consists of three elements – there is also a dihomotopy class with end point in y. The component category $\pi_0(x_0 \downarrow \vec{\pi}_1(Y)), \Sigma_1)$ is represented by the diagram (with obvious morphisms between the given regions/components)



with $h_l \circ f_l = h_r \circ f_r$.

Remark 2. In [43], S. Sokołowski introduced a somehow similar approach resulting in the *fundamental poset* $\Omega_1(X)$ of a po-space X. Using our terminology, a preorder on $(x_0 \downarrow \vec{\pi}_1(X))$ is defined by:

$$f \in \vec{\pi}_1(X; x_0, x) \sqsubseteq g \in \vec{\pi}_1(X; x_0, y) \Leftrightarrow \forall h \in \vec{\pi}_1(X; y, z)$$

$$\exists\, a, j_1 \in \vec{\pi}_1(X; x, a), j_2 \in \vec{\pi}_1(X; z, a) \text{ with } j_1 * f = j_2 * h * g \in \vec{\pi}_1(X; x_0, a).$$

The equivalence classes given by "$\sqsubseteq$ and $\sqsupseteq$" are the elements of the poset $\Omega_1(X)$, equipped with the partial order induced by $\sqsubseteq$.

If one considers morphisms Mor_{01} corresponding to a set X_1 of maximal elements, it is easy to see that $f \sqsubseteq g \Leftrightarrow \mathcal{E}(f) \supseteq \mathcal{E}(g)$, and hence the Σ_1-connected components agree with the elements of $\Omega_1(X)$. The partial order between equivalence classes in $\Omega_1(X)$ corresponds to the *existence* of morphisms in $\mathcal{C}[\Sigma^{-1}]$ between elements of these classes. The component category contains more information. It allows to compare factorisations of two given morphisms and to discuss in which parts of the po-space they agree and in which they differ.

P. Gaucher [19] has a quite different categorical approach to branching and merging, not only for dipaths, but also for their higher-dimensional analoga.

5.3 Components of the State Space

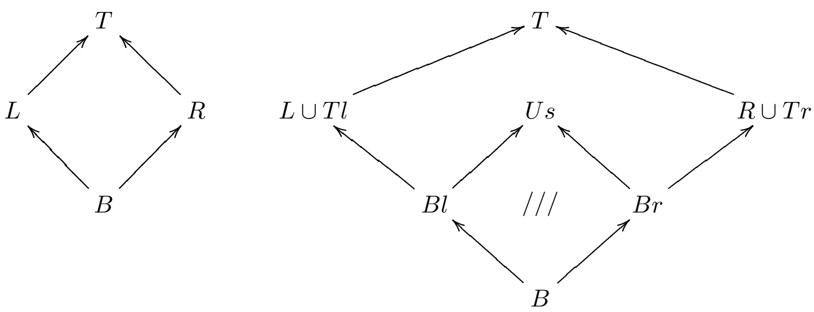
Next, we shift attention to the entire state space of a concurrent program, modelled by an lpo-space X with a minimal element x_0, the only element of X_0, and a (discrete) subset X_1 of maximal elements. For an element $x \in X$, we ask: Which essentially different execution paths pass through x? How does this information develop throughout the state space?

Definition 2. *The system $\Sigma_2 \subset Mor(\mathcal{C})$ consists of all morphisms $s \in Mor(x, y)$ with $x, y \in Ob(\mathcal{C})$ satisfying*

$$\mathcal{E}(f) = \mathcal{E}(s \circ f), \mathcal{E}(g) = \mathcal{E}(g \circ s) \text{ for all } f \in Mor(-, x) \text{ and } g \in Mor(y, -). \quad (1)$$

Obviously, Σ_2 contains the units, and it is closed under composition. It is easy to see that it suffices to require (1) for all $f \in Mor(x_0, x)$ and $g \in Mor(x, X_1)$. Along a morphism $s \in \Sigma_2(x, y)$, no decisions with global effects are taken: concatenation with s does not alter any of the extension sets of morphisms with source in y or target in x. States in the *same* Σ_2-component (or *dicomponent*) cannot be distinguished by the results of executions passing through them.

Example 5. Let again X denote the square with one hole and Y the "Swiss flag" from Fig. 3 with minimal and maximal elements x_0 (and y), resp. x_1. The component categories $\pi_0(\vec{\pi}_1(X); \Sigma_2)$ and $\pi_0(\vec{\pi}_1(Y); \Sigma_2)$ are then of the form



5.4 Relation to History Equivalence

The Σ_2-components refine the notion of a *dicomponent* of an lpo-space defined earlier in [16,37]. Those were only defined as sets and lacked the dynamical perspective given by the component category:

Definition 3. *The* history hf *of a morphism* $f \in Mor(X_0, X_1)$ *is defined as*

$$hf = \{x \in Ob(\mathcal{C}) | \ \exists f_0 \in Mor(X_0, x), f_1 \in Mor(x, X_1) \text{ with } f = f_1 \circ f_0\}.$$

Two objects $x, y \in Ob(\mathcal{C})$ *are* history equivalent *if and only if* $x \in hf \Leftrightarrow y \in hf$ *for all* $f \in Mor(X_0, X_1)$.

A history equivalence class $C \subset Ob(\mathcal{C})$ is thus a primitive element of the Boolean algebra generated by the histories, i.e., an intersection of histories and their complements such that for all $f \in Mor(X_0, X_1)$ either $C \subseteq hf$ or $C \cap hf = \emptyset$.

The following argument shows that a morphism $s \in \Sigma_2(x, y)$ has history equivalent source and target:

$$x \in hf \Leftrightarrow f \in \mathcal{E}(x) = \mathcal{E}(i_x) = \mathcal{E}(s) = \mathcal{E}(i_y) = \mathcal{E}(y) \Leftrightarrow y \in hf.$$

Hence, a Σ_2-component is contained in a path-component of a history equivalence class.

6 Algorithms for 2-dimensional Mutual Exclusion Models

In this section, we confine ourselves to the progress graphs described in the introduction. Classifying dipaths up to dihomotopy in these mutual exclusion models corresponds to finding out which (and how many) schedules for a given concurrent program can potentially yield different results. An algorithm arriving at such a classification in dimension two, i.e, for semaphore programs with just *two* interacting transactions, was described in [37]; the results in this section rely on the methods described there.

In this case, the state space has $X = I^2 \setminus int(F)$ as a model, i.e., a unit square from which a forbidden region F (e.g., the region in black in Fig. 3) is deleted. This region is a union of rectangles that are parallel to the axes. Since we are interested in dipaths connecting the minimal point to the maximal point, we may assume that X does not contain neither unsafe nor unreachable points; this can always be achieved by a completion process, cf. [37]. As a consequence ([37], Lemma 4.1), every path-component $F_i \subset F$ has a global minimum $\mathbf{z}_i = (x_i, y_i)$ and a global maximum $\mathbf{z}^i = (x^i, y^i)$. We define line segment subspaces $S_i^x, S_i^y, S_x^i, S_y^i \subseteq X$ emerging horizontally and vertically from these minima and maxima as follows:

$$S_i^x = \{\mathbf{z} = (x,y) \in X |\ x \leq x_i, y = y_i\} \quad S_i^y = \{\mathbf{z} = (x,y) \in X |\ x = x_i, y \leq y_i\}$$
$$S_x^i = \{\mathbf{z} = (x,y) \in X |\ x \geq x^i, y = y^i\} \quad S_y^i = \{\mathbf{z} = (x,y) \in X |\ x = x^i, y \geq y^i\}.$$

All these subspaces consist of one or several line segments, that may be broken up into pieces by other components of the forbidden region. Let $T_i^x \subseteq S_i^x, T_i^y \subseteq S_i^y, T_x^i \subseteq S_x^i$ and $T_y^i \subseteq S_y^i$ denote the segment touching F_i, cf. Fig. 4. The unions of these separating line segments will be called $T_- = \bigcup_i (T_i^x \cup T_i^y), T^- = \bigcup_i (T_x^i \cup T_y^i)$ and $T = T^- \cup T_-$. A dipath $f : I \to X$ from x to y is said *to cross* T_- if there exists an i such that $\emptyset \neq f^{-1}(T_-)$ is contained in the interior of I, i.e., if its image contains points on both sides of one of the segments. Similarly, one defines crossing wrt. T^- and to T. We can now detect which of the dipath classes in X are inverted in the two categories of fractions of Sect. 5:

Proposition 2. *Let $s : I \to X$ denote a (partial) dipath with $f(0) = x$ and $f(1) = y$. Its dihomotopy class $[s] \in \vec{\pi}_1(X; x, y)$ is contained in Σ_1 if and only if f does not cross T_-; in Σ_2 if and only if f does not cross T.*

Example 6. In the example of Fig. 4 with a forbidden region consisting of four components F_i, there are six dihomotopy classes of dipaths between $\mathbf{x}_0$ and $\mathbf{x}_1$, cf. [37], Fig. 14. The upper figures contain the line segments T_*^* that, together with boundary segments of the forbidden region, cut out the components in the two cases discussed in Sect. 5. For instance, the component marked C in the rightmost figure is characterised by two ingoing non-unit morphisms that, each have two extensions. The lower figures show the associated component categories with morphisms going upward. In this example, there are no non-trivial relations.

A similar analysis in dimensions higher than two is certainly more demanding. Not only the components of F, but also their finer topological properties will certainly play a role, cf. the discussion in [37], Sect. 5.

7 Classical State-Space Reduction Techniques

7.1 Persistent Sets

Let $(S, i, E, Tran, I)$ be an asynchronous transition system [2]. This means that $(S, i, E, Tran)$ is a transition system and that $I \subseteq E \times E$ is a relation between labels E, the "independence relation" between two actions. We will not give a precise axiomatics for I here, and will keep on simple grounds. Basically, I should satisfy the following conditions (taken from [22]):

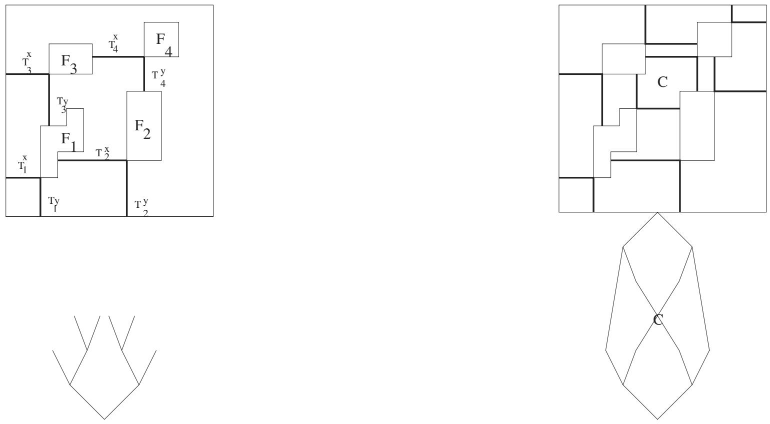


Fig. 4. Components and their categories in a 2-dimensional mutual exclusion model

- if t_1 (respectively t_2) is enabled in s and $s \to^{t_1} s'$ (respectively $s \to^{t_2} s'$) then t_2 (respectively t_1) is enabled in s if and only if t_2 (t_1) is enabled in s' (independent transitions can neither disable nor enable each other); and,

- if t_1 and t_2 are enabled in s, then there is a unique state s' such that both $s \to^{t_1} s_1 \to^{t_2} s'$ and $s \to^{t_2} s_2 \to^{t_1} s'$ (commutativity of enabled independent transitions).

In our technique, I is just a set of squares, or 2-transitions, or in the topological sense, they are elementary surfaces, enabling us to continuously deform dipaths. We extend in an intuitive manner I to sets of actions by putting AIB if for all $a \in A$, for all $b \in B$, aIb. We identify a with the singleton $\{a\}$.

Let T be a set of actions, $T \subseteq E$, and $p \in S$ be a state. We say that T is persistent in state p if, T contains only actions which are enabled at p, and, for all traces t beginning at p containing only actions q out of T, qIT. Suppose we have a set of persistent actions T_p for all states p in an asynchronous transition system. Then let us look at the following set of traces PT (identified with a series of states) in $(S, i, E, Tran, I)$ defined inductively as follows: $(i) \in PT$, and if $(p_1, \ldots, p_n) \in PT$, then $(p_1, \ldots, p_n, q) \in PT$ where $p_n \to^{t'} q \in Tran$ and $t' \notin T_{p_n}$. Deadlock detection can be performed on this subset PT of traces instead of the full set of traces of $(S, i, E, Tran, I)$. At least when $(S, i, E, Tran, I)$ is acyclic, PT is enough for checking LTL temporal formulas (and you can modify the method so that it works generally). We exemplify the method on the process $Pb.Pa.Vb.Va \mid Pa.Pb.Va.Vb$. A standard interleaving semantics would be as sketched in Figure 5, showing the presence of deadlocking state 13. One set of persistent sets is $T_1 = \{Pa\}$, $T_2 = \{Pb\}$, $T_5 = \{Pa, Pb\}$, $T_6 = \{Pb, Va\}$, $T_8 = \{Pa, Va\}$, $T_{13} = \emptyset$, $T_9 = \{Vb\}$, $T_{12} = \{Va\}$, $T_{17} = \{Pb\}$, $T_{18} = \{Va\}$, $T_{22} = \{Vb\}$, $T_{23} = \emptyset$, $T_7 = \{Pb, Vb\}$, $T_{14} = \{Vb\}$, $T_{15} = \{Pb\}$, $T_{16} = \{Pa\}$,

$T_{20} = \{Vb\}$, $T_{21} = \{Va\}$, and we show the corresponding traces PT in Figure 6. We have not indicated the persistent sets corresponding to 3, 4 etc. since in a persistent set search, they will not be reached anyway, so their actual choice is uninteresting. In Figure 5 there are 16 paths from 1 to be traversed if no selective search was used. Six of them lead to the deadlock 13, and 10 (5 above the hole, 5 below the hole) are going to the final point 23. In Figure 6, one can check that there are only 8 paths to be traversed if one uses the persistent sets selective search (3 to state 13, 1 to state 23 below the hole and 4 to state 23 above the hole).

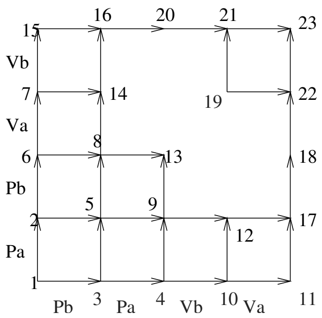


Fig. 5.

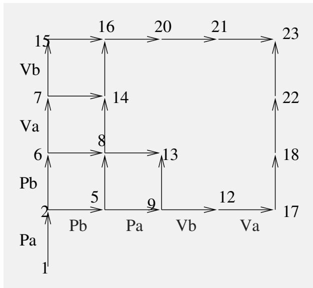


Fig. 6.

How did we find this set of persistent sets? In the PV case this can be done quite easily as follows. First the independence relation can be found out right away. Px and Py stand respectively for the query for a lock on x and y (nothing is committed yet) so they are independent actions, whatever x and y are. We should rather declare Px and Vy dependent in general: if $x = y$ this is clear, and for $x \neq y$ this can come from the fact locks on x and y are causally related (precisely as in the case of Figure 5 with $x = a$ and $x = b$). This is slightly different from the more usual case of atomic reads and writes languages in which the independance relation can be safely determined as: actions are independent if and only if they act on distinct variables. The most elaborated technique known in this framework is that of "stubborn sets" see [46], and its adaptation to the current presentation, see [22] for a precise definition. The example of persistent set we gave in Figure 6 is in fact a stubborn set. As one can see as well, the persistent set approach here reduces the 5 paths below the hole into 1, which is a representant modulo dihomotopy of these 5 dipaths. In the diconnected components approach, one finds the set of 7 diconnected components and the corresponding graph of regions pictured in Figure 7.

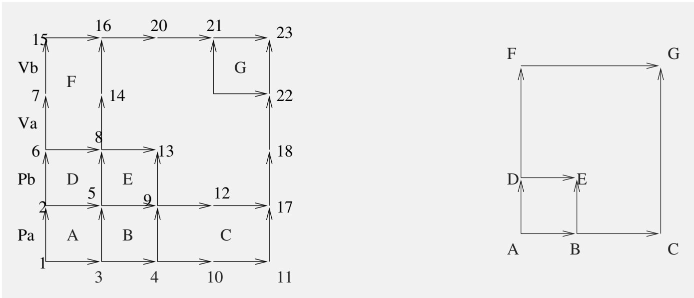


Fig. 7.

7.2 Comparison with Geometric Techniques

There are 4 dipaths to be traversed in the graph of diconnected regions to determine the behaviour of this concurrent system (two of them leading to state 13 being dihomotopic). In fact, there are two explanations why the method of diconnected components is superior to the persistent set approach. In the latter, the independence relation does not in general depend on the current state (even if this might be changed by changing the set of labels), Our notion of independence is given by a 2-transition, which depends on the current state (see for instance [25] where the link is made formal). The second and more important reason is that the diconnected graph algorithm determines regions using *global* properties, whereas the persistent sets approach uses only (in general syntactic) local criteria for reducing the state-space. Conversely, it is relatively easy to see the following: For every state p in our asynchronous transition system (or by [25], in a 2-dimensional cubical set), all traces t composed of actions outside T_p is such that all its actions are independent with T_p. So any trace from p made up of any action (those of T_p as well as those outside T_p) can be deformed (by dihomotopy, or "is equivalent to") into a trace firing first actions from T_p and then actions outside T_p. Therefore the selective search approach using only actions from T_p (for all p) is only traversing some representatives of the dihomotopy classes of paths. The persistent search approach is a particular case of dihomotopic deformation (not optimal in general).

7.3 Miscellaneous Techniques

Sleep sets. The sleep sets technique can be seen as a mere amelioration of the traversal we saw, and therefore can be combined with the method of persistent sets (as well as ours). The problem we had in Section 7.1 is that quite a few of the paths we are traversing go through the same states at some point, and have a common suffix (like paths (1,2,6,8,14,16,20,21,23) and (1,2,5,8,14,16,20,21,23) in Figure 6). It is obviously not necessary to traverse again common suffixes if we want to check future tense logical formulas. The sleep sets S_p (p a state in our asynchronous transition system) use the information about the current traversals made, and not any information of any semantic kind to cope with this problem: see [21] for the precise definition. This reduces the number of transitions but not the number of states as shown in Figure 8. We now produce only 5 paths. The same method is applied on Figure 7 to give the optimal result of Figure 9 (3 paths).

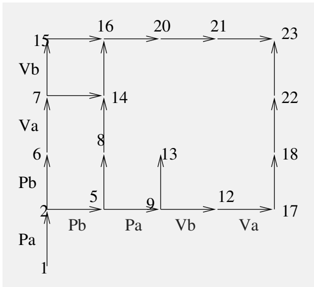


Fig. 8.

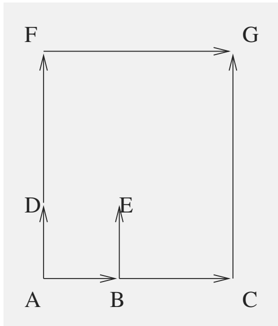


Fig. 9.

This does not entirely solve the suffix problem. A classical means to complete this method is to use "state space hashing" [20]. In this method (again orthogonal to "dihomotopy reduction"), one tries to maintain the set of states (and the set of transitions sometimes) already traversed avoiding to visit again the same sequences of states and transitions. As this database might be very big, one uses standard hashing techniques which quickly decide if two states are equal, but which might "identify" unequal states. This was used for instance in one of our simple implemented analyzers described in [14]. Again, for cyclic transition systems, this transformation does not change the deadlock(s) or LTL (future tense temporal logics) formulae that are true.

Covering steps or Virtual Coarsening. Another idea, called "covering step" in [49] or sometimes called in other situations "virtual coarsening", is to group together interleavings of independent actions by "multiset" transitions or "covering steps". For instance, when possible (basically, you should check that you are not in the situation of Figure 1), one replaces an interleaving square of two actions by the transition $s_0 \to^{\{a,b\}} s_3$. This is obviously subsumed by our notion of n-transition (formally, this can be done in the style of [39]).

Algorithmics of the representation of the state-space. A number of clever algorithmic methods have been used to reduce the representation of the state-space in memory without discarding any transition nor state, rather by compressing the representation. A very much used technique in model-checking is the representation of the transition relation with binary decision diagrams (BDDs) or QDDs as in [3] associated with symbolic representations of states [28]. Some amount of work has been devoted to "on the fly" techniques, also in model checking, see for instance [31]: Only a part of the state-space is represented during the analysis, because there is no need in general to construct first the whole state-space and then traverse it. Last but not least some techniques involving reducing the state-space using symmetry arguments have been proposed and successfully used, see [9]. All these techniques could be equally applied to our diconnected components approach, and should be exemplified in future papers. For instance, symmetry techniques are quite well studied in geometry and should apply straightforwardly to our geometric approach.

8 Concluding Remarks

Two further arguments in favor of our geometric techniques should be developped: We should be able to gain much more when the dimension of the problem (i.e. the number of processes involved) increases. The persistent sets types of methods basically use local transpositions, or in our geometric phrasing, faces of dimension 2, to equate some of the equivalent dipaths. Geometrically speaking, we can use sometime shorter deformation paths, like any hypercube, i.e. any cyclic permutation. The other argument is that geometric methods do cooperate well with abstraction mechanisms (in the sense of abstract interpretation [11]). It is in particular shown in [15] that the upper-approximation (or lower-approximation) of the forbidden regions can be carried out simply for a variety of languages, using classical abstract interpretation domains. These give lower (respectively upper) approximations of the "interesting" schedules or paths to be traversed.

References

1. M.A. Armstrong, *Basic Topology*, Springer-Verlag, 1990.
2. M. A. Bednarczyk. *Categories of asynchronous systems.* PhD thesis, University of Sussex, 1988.
3. Bernard Boigelot and Patrice Godefroid. Symbolic verification of communication protocols with infinite state spaces using QDDs. *Formal Methods in System Design: An International Journal*, 14(3):237–255, May 1999.
4. T. Borceux, *Handbook of Categorial Algebra I: Basic category theory*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1994.
5. Glen E. Bredon, *Topology and Geometry*, GTM, vol. 139, Springer-Verlag, 1993.
6. R. Brown and P.J. Higgins, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Algebra **22** (1981), 11–41.
7. ———, *On the algebra of cubes*, J. Pure Appl. Algebra **21** (1981), 233–260.
8. S.D. Carson and P.F. Reynolds, *The geometry of semaphore programs*, ACM TOPLAS **9** (1987), no. 1, 25–53.
9. E. M. Clarke, E. A. Emerson, S. Jha, and A. P. Sistla. Symmetry reductions in model checking. In *Proc. 10th International Computer Aided Verification Conference*, pages 145–458, 1998.

10. E. G. Coffman, M. J. Elphick, and A. Shoshani. System deadlocks. *Computing Surveys*, 3(2):67–78, June 1971.
11. P. Cousot and R. Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction of approximations of fixed points. *Principles of Programming Languages 4*, pages 238–252, 1977.
12. E.W. Dijkstra, *Co-operating sequential processes*, Programming Languages (F. Genuys, ed.), Academic Press, New York, 1968, pp. 43–110.
13. L. Fajstrup, *Dicovering spaces*, Tech. Report R-01-2023, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst, November 2001.
14. L. Fajstrup, É. Goubault, and M. Raussen, *Detecting Deadlocks in Concurrent Systems*, CONCUR '98; Concurrency Theory (Nice, France) (D. Sangiorgi and R. de Simone, eds.), Lect. Notes Comp. Science, vol. 1466, Springer-Verlag, September 1998, 9th Int. Conf., Proceedings, pp. 332 – 347.
15. L. Fajstrup, E. Goubault, and M. Raussen. Detecting deadlocks in concurrent systems. Technical report, CEA, 1998.
16. ———, *Algebraic topology and concurrency*, Tech. Report R-99-2008, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst, June 1999, conditionally accepted for publication in Theoret. Comput. Sci.
17. L. Fajstrup and S. Sokolowski, *Infinitely running concurrents processes with loops from a geometric viewpoint*, Electronic Notes Theor. Comput. Sci. **39** (2000), no. 2, 19 pp., URL: http://www.elsevier.nl/locate/entcs/volume39.html.
18. P. Gabriel and M. Zisman, *Calculus of fractions and homotopy theory*, Springer-Verlag New York, Inc., New York, 1967, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 35.
19. P. Gaucher, *Homotopy invariants of higher dimensional categories and concurrency in computer science*, Math. Structures Comput. Sci. **10** (2000), no. 4, 481–524.
20. P. Godefroid, G. J. Holzmann, and D. Pirottin. State space caching revisited. In *Proc. 4th International Computer Aided Verification Conference*, pages 178–191, 1992.
21. P. Godefroid and P. Wolper, *Using partial orders for the efficient verification of deadlock freedom and safety properties*, Proc. of the Third Workshop on Computer Aided Verification, vol. 575, Springer-Verlag, Lecture Notes in Computer Science, 1991, pp. 417–428.
22. P. Godefroid and P. Wolper. Partial-order methods for temporal verification. In *Proc. of CONCUR'93*. Springer-Verlag, LNCS, 1993.
23. É. Goubault, *The Geometry of Concurrency*, Ph.D. thesis, Ecole Normale Superieure, Paris, 1995.
24. ———, *Geometry and Concurrency: A User's Guide*, Electronic Notes Theor. Comput. Sci. **39** (2000), no. 2, 16 pp., URL: http://www.elsevier.nl/locate/entcs/volume39.html.
25. ———, *Cubical sets are generalized transition systems*, avail. at `http://www.di.ens.fr/~goubault`, 2001.
26. J. Gunawardena, *Homotopy and concurrency*, Bulletin of the EATCS **54** (1994), 184–193.
27. A. Hatcher, *Algebraic Topology*, electronically avialable at `http://math.cornell.edu/~hatcher/# ATI`
28. T. A. Henzinger, O. Kupferman, and S. Qadeer. From pre-historic to post-modern symbolic model checking. In *Proc. 10th International Computer Aided Verification Conference*, pages 195–206., 1998.
29. M. Herlihy, S. Rajsbaum, and M. Tuttle, *An Overview of synchronous Message-Passing and Topology*, Electronic Notes Theor. Comput. Sci. **39** (2000), no. 2.

30. P.J. Higgins, *Categories and Groupoids*, Mathematical Studies, vol. 32, van Nortrand Reinhold, London, 1971.
31. G. Holzmann. On-the-fly model checking. *ACM Computing Surveys*, 28(4es):120–120, December 1996.
32. W. Lipski and C.H. Papadimitriou, *A fast algorithm for testing for safety and detecting deadlocks in locked transaction systems*, Journal of Algorithms **2** (1981), 211–226.
33. S. Mac Lane, *Categories for the working mathematician*, Graduate Texts in Mathematics, vol. 5, Springer-Verlag, New York, Heidelberg, Berlin, 1971.
34. A. Mazurkiewicz. Trace theory. In G. Rozenberg, editor, *Petri Nets: Applications and Relationship to Other Models of Concurrency, Advances in Petri Nets 1986, PART II, PO of an Advanced Course*, volume 255 of *LNCS*, pages 279–324, Bad Honnefs, September 1986. Springer-Verlag.
35. J.R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley, 1984.
36. V. Pratt, *Modelling concurrency with geometry*, Proc. of the 18th ACM Symposium on Principles of Programming Languages. (1991).
37. M. Raussen, *On the classification of dipaths in geometric models for concurrency*, Math. Structures Comput. Sci. **10** (2000), no. 4, 427–457.
38. ———, *State spaces and dipaths up to dihomotopy*, Tech. Report R-01-2025, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst, November 2001.
39. V. Sassone and G. L. Cattani. Higher-dimensional transition systems. In *Proceedings of LICS'96*, 1996.
40. V. Sassone, M. Nielsen, and G. Winskel. Relationships between models of concurrency. In *Proceedings of the Rex'93 school and symposium*, 1994.
41. J. P. Serre, *Homologie singulière des espaces fibrés.*, Ann. of Math. **(2) 54** (1951), 425–505.
42. A. Shoshani and E. G. Coffman. Sequencing tasks in multiprocess systems to avoid deadlocks. In *Conference Record of 1970 Eleventh Annual Symposium on Switching and Automata Theory*, pages 225–235, Santa Monica, California, Oct 1970. IEEE.
43. S. Sokołowski, *Categories of dimaps and their dihomotopies in po-spaces and local po-spaces*, Preliminary Proceedings of the Workshop on Geometry and Topology in Concurrency Theory GETCO'01 (Aalborg, Denmark) (P.Cousot et al., ed.), vol. NS-01, BRICS Notes Series, no. 7, BRICS, 2001, pp. 77 – 97.
44. J. Srba, *On the Power of Labels in Transition Systems*, CONCUR 2001 (Aalborg, Denmark) (K.G. Larsen and M. Nielsen, eds.), Lect. Notes Comp. Science, vol. 2154, Springer-Verlag, 2001, pp. 277–291.
45. A. Stark. Concurrent transition systems. *Theoretical Computer Science*, 64:221–269, 1989.
46. A. Valmari, *A stubborn attack on state explosion*, Proc. of Computer Aided Verification, no. 3, AMS DIMACS series in Discrete Mathematics and Theoretical Computer Science, 1991, pp. 25–41.
47. R. van Glabbeek, *Bisimulation semantics for higher dimensional automata*, Tech. report, Stanford University, 1991.
48. R. van Glabbeek and U. Goltz. Partial order semantics for refinement of actions. *Bulletin of the EATCS*, (34), 1989.
49. F. Vernadat, P. Azema, and F. Michel. Covering step graph. *Lecture Notes in Computer Science*, 1091, 1996.
50. G. Winskel and M. Nielsen. *Models for concurrency*, volume 3 of Handbook of Logic in Computer Science. Oxford University Press, 1994.

Homology, Homotopy and Applications, vol.5(2), 2003, pp.257–280

STATE SPACES AND DIPATHS UP TO DIHOMOTOPY

MARTIN RAUSSEN

(*communicated by Gunnar Carlsson*)

Abstract

Geometric models have been used by several authors to describe the behaviour of concurrent sytems in computer science. A concurrent computation corresponds to an oriented path (*dipath*) in a (locally) partially ordered state space, and *di*homotopic dipaths correspond to equivalent computations. This paper studies several invariants of the state space in the spirit of those of algebraic topology, but taking partial orders into account as an important part of the structure. We use several categories of fractions of the fundamental category of the state space and define and investigate the related quotient categories of "components". For concurrency applications, the resulting categories can be interpreted as a dramatic reduction of the size of the state space to be considered.

1. Introduction

1.1. Background and history

The use of geometric models in the description of the behaviour of concurrent systems in computer science can be traced back at least to the work of E.W. Dijkstra [**6**], where concurrent processes are modeled by so-called *progress graphs*; cf. for instance Fig. 1. For so-called semaphore programs (explained below), these progress graphs have been exploited for an algorithmic determination of deadlocs and unreachable states [**23, 5, 9**]. A systematic framework for studying schedules of actions of distributed computations by means of geometric properties was proposed by V. Pratt [**25**] and subsequently R. van Glabbeek [**30**]. In his thesis [**16**], É. Goubault initiated a systematic study of Higher Dimensional Automata (HDA) built on cubical sets [**27, 4, 3**] employing methods from algebraic topology, in particular homological methods. The idea is that a schedule of actions (including deadlocks and unreachables, but also serializability conditions etc.) is essentially invariant under "continuous deformation", i.e. some sort of homotopy. This point of view has been exploited in a database framework in [**20**] and later in [**11**].

Acknowledgements. The author wishes to thank the referee for several corrections to the original version of this paper as for suggestions improving the presentation. Thanks are also due to Laboratoire d'Informatique, École Polytechnique, Paris, for its hospitality while the final revision of this paper was undertaken.
Received November 1, 2001, revised September 9, 2002; published on April 22, 2003.
2000 Mathematics Subject Classification: 51H15, 54E99.
Key words and phrases: abstract homotopy theory, dihomotopy theory.

Relevant models have to reflect the *irreversibility of time*, and this is why partial orders have to play an important role. A prototypical example (the "Swiss flag" in Fig. 1) models the concurrent execution of two programs (on the axes) both locking (P) and releasing (V) by a *semaphore* two shared objects a and b, but in reverse order. An execution path in this model has to be a "*dipath*", i.e., a continuous path with *monotone* projection to each axis – modelling the progress of an individual program; moreover, it has to start at the minimal point $(0, 0)$ and to end at the maximal point $(1, 1)$, and it has to *avoid* the shaded forbidden region ("Swiss flag") modelling concurrent access to a or b. A dipath entering the "*unsafe*" region cannot end up at $(1, 1)$ – likewise, no dipath from $(0, 0)$ can ever enter the "*unreachable*" region in Fig. 1. Moreover, there are two possible outcomes of a run of the concurrent program: Either T_1 locks both a and b before T_2 can access any of them, or T_2 uses b and a before T_1 does. These two runs correspond to dipaths that pass "under", resp. "over" the forbidden region, but without any further restrictions. The example suggests that (some sort of) homotopy can capture the essential difference between two dipaths or executions.

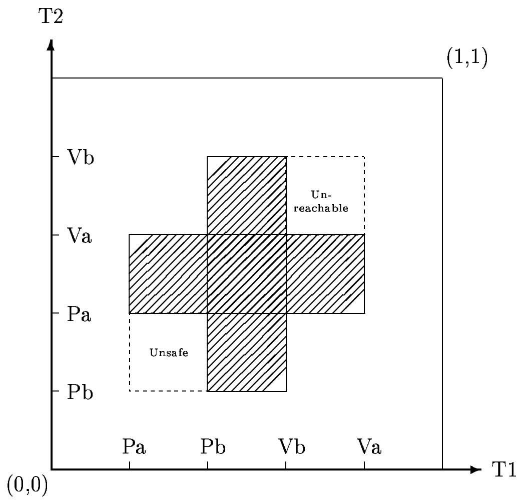


Figure 1: Example of a progress graph

1.2. Partial orders and dipaths

With the intention to employ topological methodology in this framework, we proposed [**11**] to use partially ordered topological spaces, cf. [**24**] for an early and detailed reference, or rather a local version, as a base for further analysis:

A topological space X with a partial order $\leqslant$ is called a *po-space* if and only if the relation $\leqslant \subset X \times X$ is *closed*. A po-space is automatically Hausdorff [**24**].

The definition of a locally partially ordered space (for short *lpo-space*) formally resembles that of a manifold, using covers of a Hausdorff space X by open po-subsets such that the partial orders on those agree on suitable (po-)neighbourhoods of every element. Two local partial orders are equivalent if their union is still a local partial order. See [**11**] – slightly uncorrect in the preprint version – or [**12**] for details.

The structure preserving maps between lpo-spaces are the *dimaps* [**11**], i.e., continuous maps respecting partial orders within sufficiently small neighbourhoods of every point. The most important dimaps for our purposes are the *dipaths*: Let $\vec{I} := [0,1]$ denote the unit interval and let $\mathbf{R}_{\geqslant 0} := \{t \in \mathbf{R} | t \geqslant 0\}$, both equipped with the natural order; let X denote an lpo-space, and let $x_0, x_1 \in X$. A dipath from x_0 to x_1 is a dimap $f : \vec{I} \to X$ with $f(0) = x_0$ and $f(1) = x_1$. An *infinite* dipath from x_0 is a dimap $f : \mathbf{R}_{\geqslant 0} \to X$ with $f(0) = x_0$ and such that $\lim_{t\to\infty} f(t)$ does *not* exist. Infinite dipaths model execution paths that run indefinitely without "dying slowly" (thus avoiding the so-called "zeno" executions) in *forward* semantics of concurrent programs. Analogous problems in *backwards* semantics can be handled likewise by considering infinite dipaths defined on $\mathbf{R}_{\leqslant 0} = \{t \in \mathbf{R} | t \leqslant 0\}$ or on $\mathbf{R}$.

Higher Dimensional Automata (cf. Sect. 1.1) and their dynamics can be seen as particular lpo-spaces; executions of programs on these state spaces correspond to finite or infinite dipaths on those. Just recently, alternative frameworks for handling the properties of HDAs have been proposed and discussed. In particular, the *flows* of P. Gaucher[**14**] and the *d-spaces* of M. Grandis[**18, 19**] – many of them arising from lpo-spaces – admit nicer categorical and homotopy theoretical properties.

Classical concurrency uses mainly techniques of a combinatorial or graph theoretical nature. All of the approaches mentioned above have in common an attempt to employ topological techniques to enhance our understanding; these are in particular useful to model higher dimensional connections and relations.

1.3. Dihomotopy

To capture equivalent behaviour (ensuring the same results of computations etc.) along executions, V. Pratt [**25**] suggested to use "monoidal homotopies" as equivalence relation on spaces of executions. Examples of 3-dimensional progress graphs (cf. [**11**]) showed that it is *not* enough to consider standard homotopies between dipaths; instead, one has to modify the definition in a rather obvious way:

Definition 1.1. Let X denote an lpo-space with $x_0, x_1 \in X$.

1. A *dihomotopy* from x_0 to x_1 is a continuous map $H : I \times \vec{I} \to X$ such that $H_s = H(s,-) : \vec{I} \to X$ is a dipath from x_0 to x_1 for every $s \in I$. Two dipaths $f, g : \vec{I} \to X$ from x_0 to x_1 are dihomotopic to each other if there exists a dihomotopy from x_0 to x_1 such that $H_0 = f$ and $H_1 = g$. We denote by $\vec{\pi}_1(X)(x_0, x_1)$ the set of dihomotopy (equivalence) classes of dipaths from x_0 to x_1.
2. An infinite dihomotopy from x_0 is a continuous map $H : I \times \mathbf{R}_{\geqslant 0} \to X$ such that $H_(s,-) : \mathbf{R}_{\geqslant 0} \to X$ is an infinite dipath from x_0 for every $s \in I$. We denote by $\vec{\pi}_1(X)(x_0, \infty)$ the set of dihomotopy classes of infinite dipaths from x_0. Likewise, one defines $\vec{\pi}_1(X)(-\infty, x_1)$ and $\vec{\pi}_1(X)(-\infty, \infty)$.

Remark that the paths $H(-,t),\ t \in \vec{I}$, in general, are *not* directed. Otherwise, dihomotopy would not be an equivalence relation. It is quite obvious how to generalise these definitions from the dihomotopy of paths with fixed end points to the dihomotopy of dipaths with end points moving in specified subspaces X_0 and X_1 – yielding equivalence classes $\vec{\pi}_1(X; X_0, X_1)$ – or to the dihomotopy of dimaps, cf. [**11**].

Concatenation on the level of dipaths factors over dihomotopy and induces compositions

$$\vec{\pi}_1(X)(x_0, x_1) \times \vec{\pi}_1(X)(x_1, x_2) \to \vec{\pi}_1(X)(x_0, x_2) \text{ and}$$

$$\vec{\pi}_1(X)(x_0, x_1) \times \vec{\pi}_1(X)(x_1, \infty) \to \vec{\pi}_1(X)(x_0, \infty),$$

$$(f, g) \mapsto g * f,$$

satisfying the associativity conditions. In this paper $g * f$ means: "first f, then g".

There is an alternative ("combinatorial") approach to dihomotopy: An *elementary* dihomotopy in X is a dimap $H : \vec{I}^2 \to X$ defined on the partially ordered square $\vec{I}^2$. The two dipaths $H(1,t) * H(s,0)$ and $H(s,1) * H(0,t)$ on the boundary of the square are then elementarily dihomotopic to each other. This relation is clearly reflexive and symmetric. It is not difficult to define concatenations of elementary dihomotopies with matching faces; in this context, we insist on directedness "horizontally", whereas directions may shift "vertically". The relation *combinatorial dihomotopy* is then defined as the transitive closure of the relation elementary dihomotopy.

Combinatorial dihomotopy is the relation suggested by concurrency models. The interpretation of an elementary dihomotopy is the independence of two transitions τ_0 and τ_1, i.e., first τ_0 and then τ_1 is equivalent to first τ_1 and then τ_0; moreover any interleaving of partial executions of these two transitions has to yield the same result.

Remark 1.2. It is clear, that an elementary dihomotopy is a particular dihomotopy which is directed along both parameters. As a consequence, combinatorial dihomotopy implies dihomotopy. A combinatorial dihomotopy is a dihomotopy with the special property that the paths $H(-,t),\ t \in \vec{I}$, are concatenations of actual dipaths and of dipaths "in the wrong direction" (zig-zags).

In general, dihomotopy does *not* imply combinatorial dihomotopy, as the following example shows: Let $\vec{\Sigma}X$ denote the *unreduced* suspension of a topological space X with the partial order coming *exclusively from the suspension coordinate.* This is the po-space introduced in [**15**] – for different purposes – under the term *Glob*(X). All dipaths from the minimal to the maximal point have the form $\alpha_x : I \to \vec{\Sigma}X,\ t \mapsto [(x,t)]$ for a fixed $x \in X$ – or are monotone reparametrizations of those. The dipaths α_x and $\alpha_{x'}$ from the bottom to the top cannot be connected by a combinatorial homotopy for $x \neq x'$: For $t \notin \partial I$, the only zig-zag paths connecting (x,t) and (x',t) have to pass through the minimal or through the maximal point. Since the endpoints have to be kept fix, it is not possible to construct a *continuous* combinatorial dihomotopy between α_x and $\alpha_{x'}$. On the other hand, these two dipaths are obviously dihomotopic in the sense of Def. 1.1 if just x and x' are contained in the same path component of X.

There is evidence, that the two relations agree for "nice enough" po-spaces: L. Fajstrup has recently proved [**8**] that two dihomotopic dipaths in a *cubical complex* – the geometric realisation of a cubical set[**4, 3**] – are combinatorially dihomotopic as well.

1.4. Aims and Structure

The transition from (directed) topology to algebra is more complicated than in the classical situation, since the reverse of a dipath is no longer directed. Hence, dipaths up to dihomotopy neither form a fundamental group nor a fundamental groupoid. Instead, one has to work with *fundamental categories*. These are huge gadgets, and this paper searches for representations of the essential dihomotopy information in more compressed ways. To this aim, we propose to use categories of fractions of a fundamental category with respect to suitably chosen sytems of morphisms and to investigate quotient categories of those with objects the path components with respect to these systems.

In Sect. 2, we discuss the fundamental category of an lpo-space and of a related quotient category retaining only "globally relevant information". Sect. 3 reviews the main tool, categories of fractions with respect to systems of morphisms, and proposes to investigate certain "component categories". Sect. 4 describes and investigates several relevant systems of morphisms within a fundamental category and the associated component categories. In Sect. 5, we propose a similar scheme for an investigation of "higher dihomotopy". Finally, Sect. 6 discusses the (lack of) naturality of the component categories.

The original stimulus for this study was the interesting paper [**28**] by S. Sokołowski who defined a functor Ω_1 associating to a po-space a *partial order* on the dihomotopy classes of dipaths with given start point; moreover, he defined in that paper higher dimensional functors Ω_n. I would like to thank him and also L. Fajstrup and É. Goubault for many clarifying discussions.

2. The fundamental category and its relatives

2.1. The fundamental category

Let X denote an lpo-space or a d-space, cf. [**18, 19**], i.e., a topological space X with a specified set of dipaths within the path set PX including the constant paths, which is closed under concatenation and invariant under monotone reparameterizations. A d-space may have arbitrarily small loops; in particular, the dipaths do not give rise to a locally antisymmetric relation. The dihomotopy relation investigated by Grandis corresponds to our combinatorial dihomotopy.

Definition 2.1. 1. The objects of the fundamental category $\vec{\pi}_1(X)$ are the points of X. The morphisms between elements x and y are given as the dihomotopy classes in $\vec{\pi}_1(X)(x, y)$.

2. The category $\vec{\pi}_1^{\infty}(X)$ contains $\vec{\pi}_1(X)$. It has an additional maximal element ∞ with $Mor(x, \infty) = \vec{\pi}_1(X)(x, \infty)$ for $x \in X$, $Mor(\infty, y) = \emptyset$ for $y \in X$ and $Mor(\infty, \infty) = 1_{\infty}$.

In both cases, composition of morphisms with matching target, resp. source is given by concatenation of dipaths – up to dihomotopy.

Compared to a fundamental group, a fundamental category is an enormous gadget and it has a much less nice algebraic structure. On the other hand, from simple examples one gets the impression, that the cardinality of the set of morphisms between two points is quite robust when these points are only perturbed a little bit:

Example 2.2. 1. For the square with one hole (left part of Fig. 2), there are no dipaths between the regions marked L and R, there is no dipath from T to any other region, neither is there a morphism from any other region to B. There are, up to dihomotopy, two dipaths from any point of B to any point of T. Moreover, from any point of B, certain points of B, L, R can be reached by (exactly one) dipath up to dihomotopy. Likewise, any point of T can be reached from (certain of) the points in L, R and T in essentially one way.

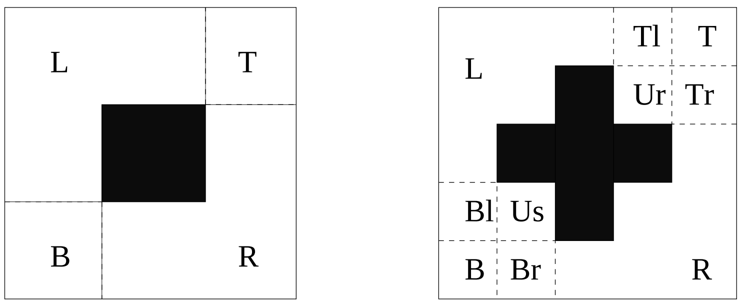


Figure 2: Square with a hole and complement of a "Swiss flag"

2. For the complement of a "Swiss flag" (right part of Fig. 2), the situation is a bit more complicated: There is no dipath leaving the unsafe rectangle *Us* and there is no dipath entering the unreachable rectangle *Ur* from the outside. It is possible to reach *Us* by essentially one dipath from $B \cup Bl \cup Br$ up to dihomotopy, and from *Ur*, one can reach points in $Tl \cup Tr \cup T$ in essentially one way. The only possibility for *two* classes of dipaths between points occurs when the first is in B and the second in T. Moreover, these classes can be represented by dipaths along the boundary – representing the two sequential executions.

In general, it is not easy to calculate fundamental categories of an lpo-space or a d-space. For the spaces arising from 2-dimensional mutual exclusion models, tools for the calculation are contained in [**26**]. In a much more general direction, M. Grandis quite recently adapted the usual proof of the Seifert-van Kampen theorem to the case of d-spaces ([**18**], Thm. 3.6) exhibiting the fundamental category of a (suitable) union of subspaces as a pushout (in **Cat**) of the fundamental categories of the subspaces. With the result of L. Fajstrup (cf. Rem. 1.2), this theorem is also valid for the fundamental categories of cubical sets or complexes with dihomotopy as defined in Def. 1.1. It is still not quite clear though how to use this pasting theorem algorithmically to calculate fundamental categories for interesting classes of lpo-spaces.

2.2. Cancellation problems

In general, cancellation is *not* possible in a fundamental category.

Example 2.3. Consider the po-space $X = \partial\vec{I}^3 \setminus int(\vec{I}^2 \times \{0\}) \subset \mathbf{R}^3$, the boundary of a standard cube from which the interior of the bottom face is removed. For $x_0 = (0,0,0)$ and $x_1 = (1,1,z)$, the dihomotopy set $\vec{\pi}_1(X)(x_0, x_1)$ consists of *two* elements for $z < 1$. They yield the unique element in $\vec{\pi}_1(X)(x_0,(1,1,1))$ after composition with a dipath class from x_1 to $(1,1,1)$.

One way to handle non-cancellation is to neglect all information that is not "visible" for dipaths from a set of initial point to a set of final points. In the applications, one is mainly interested in executions (dipaths) from a specified subset $X_0 \subset X \cup \{-\infty\}$ of initial points to a specified subset $X_1 \subset X \cup \{\infty\}$ of final points – or infinitely running executions (infinite dipaths) from a set of initial points. The reason for insisting on sources and targets being subspaces instead of just points (as in [**17**]) is that inductive calculations may require to cut dipaths and dihomotopies into pieces: "below X_0, between X_0 and X_1 and above X_1". In many applications, these subsets are *achronal*, i.e., $\vec{\pi}_1(X_i)(x,y) = \emptyset$ for $x \neq y, x, y \in X_i$, or even *discrete*.

The following example – with one point sets X_0 and X_1 – shows that the fundamental category often contains information that is not relevant for dipaths starting at X_0 and ending at X_1:

Example 2.4. Let $\vec{J} \subset \vec{I}$ denote an open subinterval, and let Y_n[1]$= \vec{I}^n \setminus \vec{J}^n$, a set with minimal point $\mathbf{0} = (0,\ldots,0)$ and maximal point $\mathbf{1} = (1,\ldots,1)$. It is easy to see that, for $n > 2$, all dipaths in Y from $\mathbf{0}$ to $\mathbf{1}$ are dihomotopic. But the fundamental category $\vec{\pi}_1(Y_n)$ is not trivial. Let $I_- = \{t \in I | t \leqslant \inf J\}$ and $I_+ = \{t \in I | t \geqslant \sup J\}$. Then, $\pi_1(Y_n)(x,y) = \emptyset$ if there is an i with $x_i > y_i$ or if there is an i with $x_i \in I_-, y_i \in I_+$ and all $x_k, y_k \in \vec{J}, k \neq i$. Otherwise, $\vec{\pi}_1(Y_n)(x,y)$ has one element unless there are precisely two coordinates $1 \leqslant i < j \leqslant n$ such that $x_i, x_j \in I_-, y_i, y_j \in I_+$[2] and all other $x_k, y_k \in \vec{J}$; in this case, there are two dihomotopy classes of dipaths from x to y.

To get rid of cancellation problems and of superfluous information, we proceed as follows: Two dihomotopy classes $\beta_1, \beta_2 \in \vec{\pi}_1(X)(x,y)$ are called equivalent if

$$\gamma * \beta_1 * \alpha = \gamma * \beta_2 * \alpha \in \vec{\pi}_1(X)(x_0,x_1) \text{ for } \textit{all } \alpha \in \vec{\pi}_1(X)(x_0,x)$$
$$\text{and } \textit{all } \gamma \in \vec{\pi}_1(X)(y,x_1),\ x_i \in X_i.$$

The equivalence class of an element $\beta \in \vec{\pi}_1(X)(x,y)$ will be denoted by $[\beta]$, the set of all such equivalence classes by $\vec{\pi}_1(X;[X_0,X_1])(x,y)$. Remark that the equivalence

[1]This space models a shared objects that can be accessed by at most $n-1$ out of n competing processes at the same time.

[2]This corresponds to $(x_i,x_j) \in B, (y_i,y_j) \in T$ in the square with a hole from Fig. 2.

relation is compatible with concatenation. We arrive at a category $\vec{\pi}_1(X;[X_0,X_1])$ whose objects are the elements $x \in X$ between X_0 and X_1, i.e., with $\vec{\pi}_1(x)(X_0,x) \neq \emptyset \neq \vec{\pi}_1(x,X_1)$ and with equivalence classes in $\vec{\pi}_1(X;[X_0,X_1])(x,y)$ as morphisms from x to y.

For the equivalence classes of dihomotopy classes, one has then a weak form of cancellation: If

$$[\gamma] * [\beta_1] * [\alpha] = [\gamma] * [\beta_2] * [\alpha] \in \vec{\pi}_1(X)(x_0,x_1)$$

for *all* $\alpha \in \vec{\pi}_1(X)(x_0,x)$ and $\gamma \in \vec{\pi}_1(X)(y,x_1)$, $x_i \in X_i$, then $[\beta_1] = [\beta_2]$.

2.3. Aims

It is the aim of this paper to relate dipaths (up to dihomotopy) contributing to the same global information although possibly having *different end points*, and hereby to define and describe – several versions of – the "components" (cf. Ex. 2.2 and Fig. 2) for general lpo-spaces or *d*-spaces. As a result, one may compress the fundamental category to one or several component categories that are much smaller – often discrete – but that still contain the essential information.

3. Categories of fractions and components

Since there is nothing special about the *fundamental* category in the following analysis, this section will be formulated for a general (small) category $\mathcal{C}$.

3.1. The category of fractions

Definition 3.1. A subset $\Sigma \subseteq Mor(\mathcal{C})$ is called a *system of morphisms* if

1. Σ is closed under composition.
2. $1_x \in \Sigma$ for every $x \in Ob(\mathcal{C})$.

with 1_x denoting the identity on x. The elements of Σ are sometimes called *weakly invertible.*

Examples for interesting systems of morphisms within a fundamental category will be given in Sect. 4.

For a system Σ of $\mathcal{C}$-morphisms, one may define the *category of fractions* $\mathcal{C}[\Sigma^{-1}]$ and the localization functor $q_\Sigma : \mathcal{C} \to \mathcal{C}[\Sigma^{-1}]$ [**13, 2**] having the following universal property:

- For every $s \in \Sigma$ the morphism $q_\Sigma(s)$ is an isomorphism.
- For any functor $F : \mathcal{C} \to \mathcal{D}$ such that $F(s)$ is an isomorphism for every $s \in \Sigma$ there is a unique functor $\theta : \mathcal{C}[\Sigma^{-1}] \to \mathcal{D}$ with $\theta \circ q_\Sigma = F$.

It is not too difficult to construct such a category of fractions, cf. [**2**] for details. Briefly, the objects of $\mathcal{C}[\Sigma^{-1}]$ are just the objects of $\mathcal{C}$. To define the morphisms of $\mathcal{C}[\Sigma^{-1}]$, one introduces an inverse s^{-1} to every morphism $s \in \Sigma(x,y) = \Sigma \cap Mor(x,y)$. These inverses are collected in $\Sigma^{-1}(y,x), x,y \in Ob(\mathcal{C})$ and then in Σ^{-1}. Consider the closure of $Mor(\mathcal{C}) \cup \Sigma^{-1}$ under composition and the smallest equivalence relation containing $s^{-1} \circ s = 1_x$ and $s \circ s^{-1} = 1_y$ for $s \in \Sigma(x,y)$ that is compatible with composition. The equivalence classes constitute the morphisms of $\mathcal{C}[\Sigma^{-1}]$. A morphism in $\mathcal{C}[\Sigma^{-1}]$ can always be represented [**13, 2**] in the form

$$s_k^{-1} \circ f_k \circ \cdots \circ s_1^{-1} \circ f_1,\ s_j \in \Sigma, f_j \in Mor, k \in \mathbf{N}.$$

In the context of homotopy theory – with topological spaces as objects, continuous maps as morphisms and the weak equivalences as the system of morphisms – categories of fractions are often called the *homotopy category* of $\mathcal{C}$, cf. e.g. [**1, 21**].

3.2. The component category

Any morphism of the form $s_1^{-1} \circ s_2 \circ \cdots \circ s_{2k-1}^{-1} \circ s_{2k}, s_j \in \Sigma, k \in \mathbf{N}$ is called a Σ-*zig-zag* morphism. The set $ZZ(\Sigma)$ of all Σ-*zig-zag* morphisms forms a system of morphisms contained in the invertibles of the category of fractions, denoted $Inv(\mathcal{C}[\Sigma^{-1}])$. Equality holds if Σ contains the invertibles $Inv(\mathcal{C})$ of the original category $\mathcal{C}$. The subcategory of $\mathcal{C}[\Sigma^{-1}]$ with all objects, the morphisms of which are given by the zig-zag morphisms $ZZ(\Sigma)$, forms in fact a *groupoid*.

Two objects $x, y \in Ob(\mathcal{C})$ are called Σ-*connected* – $x \simeq_\Sigma y$ – if there exists a zig-zag-morphism from x to y. This definition corresponds to usual path connectedness *with respect to paths in* Σ *only – but regardless of orientation.* Σ-connectivity is an equivalence relation; the equivalence classes will be called the Σ-*connected components* – the path components with respect to Σ-zig-zag paths, i.e., the components of the groupoid above.

Next, consider the smallest equivalence relation on the morphisms of $\mathcal{C}[\Sigma^{-1}]$ generated (under composition) by

$$\alpha \simeq \alpha \circ s^j, \quad \alpha \simeq t^j \circ \alpha \text{ for } \alpha \in Mor(x, y), s \in \Sigma(x', x), t \in \Sigma(y, y'), j = \pm 1. \quad (3.1)$$

Remark that equivalent morphisms no longer need to have the same source or target. In particular, every morphism in Σ is equivalent to the identities in both its source and its target; hence, all zig-zag morphisms within a component are equivalent to each other.

Dividing out the morphisms in Σ within $\mathcal{C}$, we arrive at a *component category*: The objects of the component category $\pi_0(\mathcal{C}; \Sigma)$ are by definition the Σ-connected components of $\mathcal{C}$; the morphisms from $[x]$ to $[y]$, $x, y \in Ob(\mathcal{C})$, are the equivalence classes of morphisms in $\bigcup_{x' \simeq_\Sigma x, y' \simeq_\Sigma y} Mor_{\mathcal{C}[\Sigma^{-1}]}(x', y')$. The composition of $[\beta] \circ [\alpha]$ for $\alpha \in Mor_{\mathcal{C}[\Sigma^{-1}]}(x, y)$ and $\beta \in Mor_{\mathcal{C}[\Sigma^{-1}]}(y', z)$ is given by $[\beta \circ s \circ \alpha]$ with s any zig-zag morphism from y to y'. The equivalence class of that composition is independent of the choices of representatives α and β (by definition) and of the choice of the zig-zag path s by the preceeding remark.

The overall idea is thus as follows: Having fixed a suitable system Σ of "weakly invertible" morphisms, we decompose the study of $\mathcal{C}$ into the study of

- the component category encompassing the global effects of irreversibility and
- the components with a *groupoid* structure given by the Σ-zig-zags.

The original category $\mathcal{C}$ and the component category $\pi_0(\mathcal{C}; \Sigma)$ are related by a functor $\pi_0(\Sigma) : \mathcal{C} \xrightarrow{q_\Sigma} \mathcal{C}[\Sigma^{-1}] \to \pi_0(\mathcal{C}; \Sigma)$; the last arrow is the quotient functor. Particularly interesting are systems Σ for which $\pi_0(\Sigma)$ is injective on the morphism sets and bijective on non-empty morphism sets.

3.3. Morphisms between given sources and targets

For a description of components of the quotient category $\vec{\pi}_1(X;[X_0,X_1])$ from Sect. 2.2, we need a modification: Let $X_0, X_1 \subset Ob(\mathcal{C})$ denote nonempty sets of objects such that the morphisms in $Mor(\mathcal{C})$ satisfy the following weak cancellation property for $\beta_i \in Mor(x,y)$:

$$\gamma \circ \beta_1 \circ \alpha = \gamma \circ \beta_2 \circ \alpha \text{ for all } \alpha \in Mor(x_0,x), \gamma \in Mor(y,x_1),\ x_i \in X_i \Rightarrow \beta_1 = \beta_2. \tag{3.2}$$

Let $Mor(X_0,X_1) = \{f \in Mor(x_0,x_1) \mid x_0 \in X_0, x_1 \in X_1\}$. We wish to analyse the structure of $Mor(X_0,X_1)$ up to an equivalence relation given by a system Σ of morphisms in $\mathcal{C}$. For a given such system, let $\Sigma^j := \{s \in \Sigma(x,y) |\ x,y \in X_j,\ j = 0,1\}$.

Definition 3.2.

1. An *elementary equivalence* between $f \in Mor(x_0,x_1)$ and $g \in Mor(x_0',x_1'),\ x_0,x_0' \in X_0,\ x_1,x_1' \in X_1$ consists of a pair of $s \in \Sigma^0(x_0,x_0'), t \in \Sigma^1(x_1,x_1')$ such that

$$\begin{array}{ccc} x_1 & \xrightarrow{t} & x_1' \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle g} \\ x_0 & \xrightarrow{s} & x_0' \end{array}$$

commutes.

2. The symmetric and transitive closure of this relation is called *equivalence* and compares morphisms from X_0 to X_1 under *zig-zag* morphisms:

$$\begin{array}{ccccccccc} \cdots & \longleftarrow & x_1 & \xrightarrow{t} & x_1' & \xleftarrow{t'} & x_1'' & \longrightarrow & \cdots \\ & & {\scriptstyle f}\uparrow & & {\scriptstyle f'}\uparrow & & {\scriptstyle f''}\uparrow & & \\ \cdots & \longleftarrow & x_0 & \xrightarrow{s} & x_0' & \xleftarrow{s'} & x_0'' & \longrightarrow & \cdots \end{array}$$

3. The equivalence classes form the sets $Mor_{01} = \bigcup_{x_0 \in X_0, x_1 \in X_1} Mor(x_0,x_1)_{/\sim}$.
4. $Mor(X_0,x) = Mor(X_0,\{x\})$ for $x \in X$.

As in the case of the fundamental category of an lpo-space, we want to define systems of morphisms and associated component categories that inherit the essential information in the category $\mathcal{C}$ from the perspective of Mor_{01}. In many cases of interest, Σ^j will consist only of the identity morphisms on the objects in X_i – e.g., if $\mathcal{C}$ is the fundamental category of a n lpo-space and the X_i are achronal subsets of X, cf. Sect. 2.2. In that case, $Mor_{01} = \bigcup_{x_0 \in X_0, x_1 \in X_1} Mor(x_0,x_1)$.

3.4. Induced morphisms. Representations of morphisms

Let $X_0, X_1 \subset Ob(\mathcal{C})$. By composition, a morphism $s \in Mor(x,y)$ induces maps

$$\begin{array}{rccc} s_\# : & Mor(X_0,x) & \to & Mor(X_0,y) \\ & f & \mapsto & s \circ f \end{array} \qquad \begin{array}{rccc} s^\# : & Mor(y,X_1) & \to & Mor(x,X_1) \\ & g & \mapsto & g \circ s. \end{array}$$

Since composition is associative, these induced maps are adjoints under the composition pairings c_x at x and c_y at y:

$$\begin{array}{ccccc} Mor(X_0,x) & \times & Mor(x,X_1) & \xrightarrow{c_x} & Mor_{01}, \\ \big\downarrow{\scriptstyle s_\#} & & \big\uparrow{\scriptstyle s^\#} & \nearrow{\scriptstyle c_y} & \\ Mor(X_0,y) & \times & Mor(y,X_1) & & \end{array}$$

or equivalently

$$\begin{array}{ccc} Mor(X_0,x) & \xhookrightarrow{\Lambda(c_x)} & Mor_{01}^{Mor(x,X_1)} \\ \big\downarrow{\scriptstyle s_\#} & & \big\downarrow{\scriptstyle (s^\#)^*} \\ Mor(X_0,y) & \xhookrightarrow{\Lambda(c_y)} & Mor_{01}^{Mor(y,X_1)} \end{array} \qquad \begin{array}{ccc} Mor(y,X_1) & \xhookrightarrow{\Lambda(c_y)} & Mor_{01}^{Mor(X_0,y)} \\ \big\downarrow{\scriptstyle s^\#} & & \big\downarrow{\scriptstyle (s_\#)^*} \\ Mor(x,X_1) & \xhookrightarrow{\Lambda(c_x)} & Mor_{01}^{Mor(X_0,x)} \end{array}.$$

If the category $\mathcal{C}$ satisfies weak cancellation (3.2), the maps $\Lambda(c_x)$ and $\Lambda(c_y)$ are *injections.*

We associate with a morphism $f \in Mor(x,y)$ the set of all its *extensions*

$$\mathcal{E}(f) = \{[g \circ f \circ h] \mid h \in Mor(X_0,x), g \in Mor(y,X_1)\} \subset Mor_{01}$$

from X_0 to X_1 up to equivalence. Collecting these, we obtain maps into the power set $2^{Mor_{01}}$:

$$\mathcal{E}_{xy} : Mor(x,y) \to 2^{Mor_{01}}, \quad \mathcal{E}_{xy}(f) = \mathcal{E}(f).$$

Likewise, one obtains extension maps $\mathcal{E}_{0y} : Mor(X_0,y) \to Mor_{01}$. For $f \in Mor(x,y)$ and $g \in Mor(y,z)$, one has obviously

$$\mathcal{E}_{xz}(g \circ f) \subset \mathcal{E}_{xy}(f) \cap \mathcal{E}_{yz}(g). \tag{3.3}$$

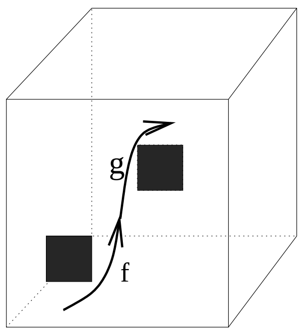


Figure 3: Dipaths on the surface of a cube with two holes

Example 3.3. Even in easy geometric examples, equality does *not* hold in (3.3). In Fig. 3, we consider the surface of a cube with two squares on the front face punched out. The dipaths f and g on the front face can both be extended to the same two (out of three) dihomotopy classes of dipaths from the left front bottom vertex to the right rear top vertex, whereas their concatenation $g * f$ only can be extended to one of them.

4. Applications

4.1. Classes of weakly invertible morphisms

First some trivial cases: If Σ consists of the identity morphisms only, then obviously $\mathcal{C}[\Sigma^{-1}]$ and the component category $\pi_0(\mathcal{C};\Sigma)$ are equivalent to $\mathcal{C}$. If $\Sigma = Mor$, all morphisms in $\mathcal{C}[\Sigma^{-1}]$ are invertible, and the Σ-connected components are the usual path components of $\mathcal{C}$ – regarded as a non-oriented graph. The component category $\pi_0(\mathcal{C};\Sigma)$ has only identity morphisms.

We will now list several more interesting classes Σ_i of weakly invertible morphisms. Comments on the respective categories of fractions and component categories, in particular for categories of the form $\mathcal{C} = \vec{\pi}_1(X;[X_0,X_1])$ will be given in Sect. 4.2. Let always $\mathcal{C}$ denote a small category. Let X_0 and X_1 denote non-empty subsets of $Ob(\mathcal{C})$ of source, resp. target objects.

1. Let $(X_0 \downarrow \mathcal{C})$ denote the associated comma category of morphisms under X_0 - if X_0 contains just one object, this is just the usual comma category [**22**]. Let $f \in Mor(X_0,x), g \in Mor(X_0,y)$ denote objects in $(X_0 \downarrow \mathcal{C})$. Then

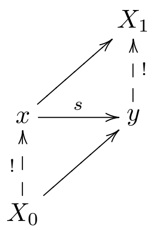
$$\Sigma_1(f,g) = \begin{cases} Mor(f,g) & \mathcal{E}(f) = \mathcal{E}(g) \\ \emptyset & \text{else} \end{cases}$$

   with $\mathcal{E}$ the extension functor from Sect. 3.4.

2. Now, we turn to the category $\mathcal{C}$ itself. For $x,y \in Ob(\mathcal{C})$, a morphism $s \in Mor(x,y)$ is contained in $\Sigma_2(x,y)$ if and only if $s^{\#} : Mor(y,X_1) \to Mor(x,X_1)$ is a *bijection.*

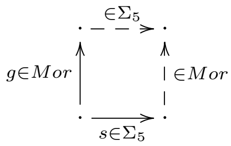
3. Dually, we let $\Sigma_3(x,y)$ consist of all morphsisms $s \in Mor(x,y)$ such that $s_{\#} : Mor(X_0,x) \to Mor(X_0,y)$ is a *bijection.*

4. $\Sigma_4 = \Sigma_2 \cap \Sigma_3 \subset Mor$.



5. Σ_5 is a system of morphisms satisfying the *extension condition* that every

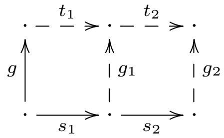
diagram



can be completed, i.e.,

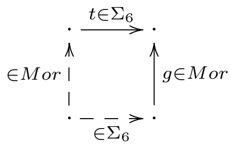
$$(\Sigma_5 \circ g) \cap (Mor \circ s) \neq \emptyset \text{ for } s \in \Sigma_5, g \in Mor.$$

The diagram



shows how to fill in the diagram for a composition of morphisms in Σ_5 by the composition of two "solutions" in Σ_5.

6. Likewise, Σ_6 is a system of morphisms satisfying the *extension condition* that every diagram



can be completed, i.e.,

$$(f \circ \Sigma_6) \cap (t \circ Mor) \neq \emptyset.$$

Same remarks as for Σ_5.

7. Σ_7 is a system satisfying both extension conditions above.
8. For $x, y \in Ob(\mathcal{C})$, let $\Sigma_8(x,y) = \emptyset$ if $Mor(x, X_1) \neq \emptyset = Mor(y, X_1)$ and $\Sigma_8(x,y) = Mor(x,y)$ else. Dually, one may compare reachability from X_0.

Particularly interesting are the *maximal* systems satisfying the requirements for $\Sigma_i, 5 \leqslant i \leqslant 7$. Maximality makes sense because the system generated by (finitely or infinitely many) such systems under composition satisfies the extension properties, as can be seen from the composition diagrams above, cf. (5).

4.2. Properties and examples

1. For the fundamental category $\mathcal{C} = \vec{\pi}_1(X)$, the comma category $(X_0 \downarrow \vec{\pi}_1(X))$ has as objects the dihomotopy classes of dipaths starting in X_0. A partial dipath s with $g = s \circ f$ is contained in Σ_1 if no "decision" has been made between f and g – all "careers" in $\vec{\pi}_1(X; [X_0, X_1])$ open to f are still open to g. Walking along a zig-zag path does not alter the extension sets of the execution paths en route. "No branching occurs between f and g" is another slogan explaining Σ_1.

The component category $\pi_0(\vec{\pi}_1(X;\{x_0\},\{\infty\}]),\Sigma_1)$ – with x_0 an initial point – induces the partially ordered set $\Omega_1(X)$ defined and investigated by S. Sokołowski, cf. [**28, 29**].

2. If $s\in\Sigma_2(x,y)$, then $\mathcal{E}(s_\#(f))=\mathcal{E}(f)$ *for all* $f\in Mor(X_0,x)$. If $t\in\Sigma_3(x,y)$, then $\mathcal{E}(t^\#(f))=\mathcal{E}(f)$ *for all* $f\in Mor(y,X_1)$.
3. The conditions for Σ_2- and Σ_3-morphisms are not independent. For a category satisfying weak cancellation (3.2), the (adjunction) diagrams at the end of Sect. 3 show:
$$s_\# \text{ onto} \Rightarrow (s_\#)^* \text{ injective} \Rightarrow s^\# \text{ injective}$$
$$s^\# \text{ onto} \Rightarrow (s^\#)^* \text{ injective} \Rightarrow s_\# \text{ injective.}$$
4. Here is how to interpret the conditions for Σ_2 if $\mathcal{C}=\vec{\pi}_1(X;[X_0,X_1])$:
   (a) For every $f\in\vec{\pi}_1(X)(x,X_1)$ there exists a "factor" $g\in\vec{\pi}_1(X)(y,X_1)$ such that $f\circ h=g\circ s\circ h$ for all $h\in\vec{\pi}_1(X)(X_0,x)$.
   (b) Factorisation is unique: Two such factors $g_1,g_2\in\vec{\pi}_1(X)(y,X_1)$ satisfying $g_1\circ s\circ h=g_2\circ s\circ h$ for all $h\in\vec{\pi}_1(X)(X_0,x)$ have the property: $g_1\circ h'=g_2'\circ h'$ for all $h'\in\vec{\pi}_1(X)(X_0,y)$.

   Analogously for Σ_3.
5. The systems $\Sigma_i, 1\leqslant i\leqslant 4$ enjoy the "2 out of 3 property": if two out of $s,t,t\circ s$ are contained in Σ_i, then so is the last.
6. In a category with weak cancellation (3.2) with respect to X_0 and X_1, one may cancel elements in Σ_2 on the left: Let $f,g\in Mor(x',x), s\in Mor(x,y)$ such that $s\circ f=s\circ g\in Mor(x',y)$. As a consequence, $k\circ s\circ f\circ h=k\circ s\circ g\circ h$ for all $h\in Mor(X_0,x'), k\in Mor(y,X_1)$. Since $s\in\Sigma_2$, any morphism $k'\in Mor(x,X_1)$ can be written in the form $k\circ s$, whence $k'\circ f\circ h=k'\circ g\circ h$ for all $h\in Mor(X_0,x'), k'\in Mor(x,X_1)$. By weak cancellation (3.2), we conclude: $f=g$. By the same argument, elements in Σ_3 may be cancelled on the right.
7. For *2-dimensional mutual exclusion models*, an algorithm for determining the Σ_i components, $i=2,3,4$ has been described in [**17**] using results of [**26**].
8. Every morphism in $\mathcal{C}[\Sigma_5^{-1}]$ can be represented in the form $s^{-1}\circ f$ with $s\in\Sigma$ and $f\in Mor$: It is easy to see (cf. e.g. [**2**]) that the composition of two morphisms of this type can be rechristened as a morphism of that same type. Similarly, every morphism in $\mathcal{C}[\Sigma_6^{-1}]$ can be represented in the form $g\circ t^{-1}$ with $t\in\Sigma$ and $g\in Mor$.
9. By successive application of the definitions, one obtains: Let $x\simeq_{\Sigma_5}x'\in Ob(\mathcal{C})$ and let $Mor_{\mathcal{C}}(x,y)\neq\emptyset$. Then there exists $y\simeq_{\Sigma_5}y'\in Ob(\mathcal{C})$ with $Mor_{\mathcal{C}}(x',y')\neq\emptyset$. Likewise, let $y\simeq_{\Sigma_6}y'\in Ob(\mathcal{C})$ and let $Mor_{\mathcal{C}}(x,y)\neq\emptyset$. Then there exists $z\simeq_{\Sigma_6}x'\in Ob(\mathcal{C})$ such that $Mor_{\mathcal{C}}(x',y')\neq\emptyset$. In particular, for Σ_7-components, the existence of morphisms between components can be investigated by examining one arbitrarily chosen object in each component.
10. The conditions for $\Sigma_i, i=5,6,7$ are stronger than one might think at first glance: Call X_0, resp. X_1 Σ_i-closed if
$$y_0\in X_0,\ \Sigma_6(x_0,y_0)\neq\emptyset\Rightarrow x_0\in X_0,\ \text{resp. } x_1\in X_1,\ \Sigma_5(x_1,y_1)\neq\emptyset\Rightarrow y_1\in X_1.$$

For a Σ_5-closed set X_1, the extension property has the consequence that $s^{\#} : Mor(y, X_1) \to Mor(x, X_1)$ is *onto* for a morphism $s \in \Sigma_5(x, y)$. Likewise, for X_0 Σ_6-closed, an element $s \in \Sigma_6(x, y)$ induces a surjection $s_{\#} : Mor(X_0, x) \to Mor(X_0, y)$. Using (3) above, we conclude: $s \in \Sigma_7 \Rightarrow s^{\#}$ and $s_{\#}$ are bijections, and thus $\Sigma_7 \subseteq \Sigma_4$.
In particular, we have for $f \in Mor, s, t \in \Sigma_7$: $\mathcal{E}(f) = \mathcal{E}(s \circ f) = \mathcal{E}(f \circ t)$, cf. (2) above. Moreover, the extension properties show that for $g, h \in Mor$, there exist $g', h' \in Mor$ such that $\mathcal{E}(g \circ f) = \mathcal{E}(g' \circ s \circ f)$, resp. $\mathcal{E}(f \circ h) = \mathcal{E}(f \circ t \circ h')$. In other words, not only is there a correspondance of the set of extensions for f and $s \circ f$, but there is a similar correspondance for all their "prolongations".

11. In a category with weak cancellation with respect to sets of initial objects X_0 and final objects X_1, a system Σ_7 of morphisms admits a left and a right calculus of fractions [**2**] generalising (8) above: Since the extension properties are the defining property for Σ_7, we need only check the following properties [**2**] for $f, g \in Mor(x, y)$:
$$\begin{array}{lcl} s \in \Sigma_7, s \circ f = s \circ g & \Rightarrow & \exists s' \in \Sigma_7 \text{ with } f \circ s' = g \circ s' \text{ and} \\ t \in \Sigma_7, f \circ t = g \circ t & \Rightarrow & \exists t' \in \Sigma_7 \text{ with } t' \circ f = t' \circ g. \end{array}$$
Since $\Sigma_7 \subseteq \Sigma_4$, we can use (6) to cancel s and t and conclude even more than necessary: $f = g$.

12. The system Σ_8 is relevant for the analysis of deadlocks and unsafe regions; the dual version for the analysis of unreachable regions, cf. [**10, 11, 26**].

Remark 4.1. A straightforward modification of the definitions of weakly invertible systems of morphisms without mentioning subsets of sources and targets (in particular for the fundamental category $\vec{\pi}_1(X)$ of an lpo-space X) does not give satisfactory results. Recent discussions with E. Haucourt and É. Goubault indicate a solution. This theme will be taken up elsewhere.

Example 4.2. 1. Several examples determining the component categories of simple po-spaces with respect to the systems Σ_i, $i \leqslant 4$, are given in [**17**].

2. The following example shows that, in general, Σ_4 does not satisfy the extension conditions for a Σ_6-system. Consider again the po-space X that is given as the surface of a cube with two holes on the front face in Fig. 4. The elements x_0 and x_2 are contained in the bottom face. It is easy to see, that all of the sets $\vec{\pi}_1(X)(x_0, x_2), \vec{\pi}_1(X)(x_1, x_2), \vec{\pi}_1(X)(\mathbf{0}, x_i)$ and $\vec{\pi}_1(X)(x_i, \mathbf{1})$ consist of a single element. In particular, the unique element $s_j \in \vec{\pi}_1(X)(x_j, x_2)$ is contained in $\Sigma_4(x_j, x_2)$, $0 \leqslant j \leqslant 1$. On the other hand, the diagram
$$\begin{array}{ccc} x_0 & \xrightarrow{s_0} & x_2 \\ & & \uparrow {\scriptstyle s_1} \\ & & x_1 \end{array}$$
cannot be completed to a square by Σ_4-morphisms: Any element $x \leqslant x_0, x_1$ is contained in the segment of the front edge and "ahead of" x_0. In particular, $\vec{\pi}_1(X)(x, \mathbf{1})$ consists of at least two elements, and hence $\Sigma_4(x, x_j) = \emptyset$, $0 \leqslant j \leqslant 1$.

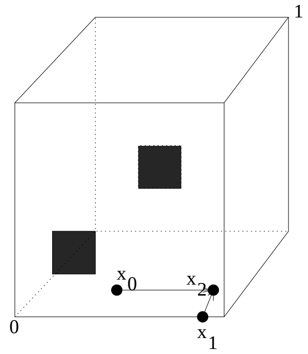


Figure 4: Invertibility on the surface of a cube with two holes

4.3. Relation to history equivalence

In [**10**], we introduced the *homotopy history* of a dipath f in X from X_0 to X_1 and the associated *history equivalence classes*. In a categorical framework, those definitions read as follows:

Definition 4.3. Let $f \in Mor(X_0, X_1)$.

1. The *history* hf of f is defined as

$$hf = \{x \in Ob(\mathcal{C})|\ \exists f_0 \in Mor(X_0, x), f_1 \in Mor(x, X_1) \text{ with } f = f_1 \circ f_0\}.$$

2. Two objects $x, y \in Ob(\mathcal{C})$ are *history equivalent* if and only if $x \in hf \Leftrightarrow y \in hf$ for all $f \in Mor(X_0, X_1)$.

A history equivalence class $C \subset Ob(\mathcal{C})$ is thus a primitive element of the Boolean algebra generated by the histories, i.e., an intersection of histories and their complements such that either $C \subseteq hf$ or $C \cap hf = \emptyset$ for all $f \in Mor(X_0, X_1)$.

Proposition 4.4. *Let $x, y \in Ob(\mathcal{C})$ and $f \in Mor(X_0, X_1)$.*

1. *$\Sigma_2(x, y) \neq \emptyset$ implies: $x \in hf \Rightarrow y \in hf$.*
2. *$\Sigma_3(x, y) \neq \emptyset$ implies: $y \in hf \Rightarrow x \in hf$.*
3. *Every Σ_4-component is contained in a path component of a history equivalence class.*

Proof. (1) Let $s \in \Sigma_2(x, y)$ and let $f = f_1 \circ f_0$ with $f_0 \in Mor(X_0, x), f_1 \in Mor(x, X_1)$. There exists $g_1 \in Mor(y, X_1)$ such that $f_1 = g_1 \circ s$. Hence $f = g_1 \circ (s \circ f_0)$, i.e., $y \in hf$.
(2) is proved similarly.
(3) For $s \in \Sigma_4(x, y)$, we have thus: $x \in hf \Leftrightarrow y \in hf$ for every $f \in Mor(X_0, X_1)$, and hence: $x \in C \Leftrightarrow y \in C$ for every history equivalence class C. The path s connects x and y. □

Prop. 4.4 suggests a method for a start of the construction of the Σ_4-components: If you know the dihomotopy classes in $\vec{\pi}_1(X; [X_0, X_1])$, find the history equivalence classes and their path components with respect to zig-zag dipaths in $Mor_{\mathcal{C}}$ (those were called the diconnected components in [**10**]); a further refinement might be necessary. In Ex. 2.2, there are two dihomotopy classes $l, r \in \vec{\pi}_1(X)(\mathbf{0}, \mathbf{1})$ of dipaths from the bottom to the top. It is easy to see, that $hl = B \cup L \cup T$ and $hr = B \cup R \cup T$. Hence, $hl \cap hr = B \cup T, hl \cap (X \setminus hr) = L, hr \cap (X \setminus hl) = R$, and the remaining intersection of complements is empty. The subspace $hl \cap hr$ consists of the two Σ_4-components B and T.

5. Higher homotopy categories

A first serious attempt to bring *higher* homotopy into the discussion of po-spaces via methods from algebraic topology was formulated by S. Sokołowski in [**28**]. In this section, I would like to give a presentation of the definitions and of first results in the categorical framework of this paper.

For a topological space Z (made into a po-space with equality as the partial order) and a local po-space (or d-space) X, let X^Z denote the mapping space with the compact-open topology. Maps in X^Z come equipped with the pointwise (local) partial order, i.e.,

$$f \leqslant g \Leftrightarrow f(z) \leqslant g(z) \text{ for all } z \in Z \tag{5.1}$$

or with an induced d-space structure. A *dicylinder*, cf. [**28**] for Z a sphere, is a dimap $F : Z \times \vec{I} \to X$; equivalently, it may be regarded as a dipath from $f = F_0$ to $g = F_1$ in X^Z with respect to the partial order (5.1).

We can now define a category $[Z : X]_1$ which has the maps in X^Z as objects. The morphisms between f and g in $[Z : X]_1$ are the *fixed end dihomotopy classes* of dicylinders; i.e., two dicylinders F and G from f to g are dihomotopic, if there is a dihomotopy $H : Z \times I \times \vec{I} \to X$ with $H(z, t, 0) = f(z), H(z, t, 1) = g(z)$ and $H(z, 0, s) = F(z, s), H(z, 1, s) = G(z, s)$ for all $z \in Z, t \in I$ and $s \in \vec{I}$. Concatenation along g allows us to compose a dicylinder from f to g with a dicylinder from g to h. This concatenation is compatible with dicylinder dihomotopy and thus gives rise to the category $[Z : X]_1$– which is equivalent to the fundamental category of the mapping space X^Z. An analogue to the higher fundamental groups is given by the special cases $Z = S^{n-1}, n > 1$. We call $[S^{n-1} : X]_1$ the n-th category of X.

Studying higher homotopy invariants of a po-space X means studying component categories of its nth category. With a source subspace $X_0 \subset X$ and a target subspace $X_1 \subset X$, one would like to structure the dihomotopy classes of dimaps

$$f : (S^{n-1} \times \vec{I}; S^{n-1} \times \{0\}, S^{n-1} \times \{1\}) \to (X; X_0, X_1).$$

Again, the results will depend on the definition of the "weakly invertible" morphisms. Details will be worked out elsewhere. We rephrase and comment some of the findings and examples of S. Sokołowski in [**28**]:

1. Even if the po-space X does not have any deadlock point x (i.e., $\vec{\pi}_1(X)(x, X_1) \neq \emptyset$ for all $x \in X$, cf. [**23, 5, 9**]), the mapping spaces very often have lots of them. If X is the po-space from the left part of Fig. 1, a map $S^1 \to X$ whose image intersects both L and R cannot be the bottom of a dicylinder with top

the constant map from S^1 into the top point.

2. The nth categories can discriminate between po-spaces with equivalent fundamental categories (with given source and target). For an example, let $X = \vec{I}^3 \setminus \vec{J}^3$ denote the po-space from Ex. 2.4, i.e., a 3-dimensional cube with an open subcube removed. All dipaths from **0** to **1** are dihomotopic to each other. Hence, the associated component category $\pi_0(\pi_1(X; [\mathbf{0}, \mathbf{1}]), \Sigma_4)$ is trivial. A dicylinder $f : S^1 \times (\vec{I}; \mathbf{0}, \mathbf{1}) \to (X; \mathbf{0}, \mathbf{1})$ from the bottom to the top induces a map $S^2 \simeq \Sigma S^1 \to X$ and is classified (up to dihomotopy) by the integral mapping degree of that latter map. The Σ_4-component category of the second category of X contains a bottom and a top element (represented by constant maps) and, for every $k \in \mathbf{Z}$, one class inbetween. There are no morphisms between components corresponding to different values $k \neq l$. Both the fundamental category and the second category of $Y = \vec{I}^3$ are trivial.
3. The nth categories come with additional structure that ought to be exploited: Evaluation at a base point $* \in S^{n-1}$ yields a functor from the nth category of a po-space X to its fundamental category. On the fibre of that functor over a chosen *dipath* in X, the dicylinders can be concatenated using a suspension coordinate in S^{n-1}.

6. Naturality questions

Let $f : X \to Y$ denote a dimap (continuous and preserving local partial orders) between lpo-spaces. It is obvious that f induces a map $f_* : \vec{\pi}_1(X) \to \vec{\pi}_1(Y)$ between the fundamental categories. If f also preserves base points or base spaces, one may ask whether there is an induced map on the component categories, as well. This is in general *not* the case:

Example 6.1. Consider the space Y (square with one hole) from Ex. 2.2.1 and the inclusion $i : X \to Y$ of the subspace $X = B \cup L \cup T$. Since there is only one dihomotopy class from the bottom point ($X_0 = Y_0 = \{\mathbf{0}\}$) to the top point ($X_1 = Y_1 = \{\mathbf{1}\}$) in X, all morphisms belong to any of the relevant systems of weakly invertible morphisms: For $\mathcal{C} = \vec{\pi}_1(X; [\mathbf{0}, \mathbf{1}])$, we get: $\Sigma_i = Mor, 1 \leqslant i \leqslant 8$ (for $i = 1$, we consider the morphisms of the comma category). In particular, X consists of a single Σ_i-component. On the other hand, X viewed as a subset of Y decomposes into two or three components – depending on the choice of $\Sigma_i, 1 \leqslant i \leqslant 7$ – with respect to $\mathcal{C} = \vec{\pi}_1(Y; [\mathbf{0}, \mathbf{1}])$.

There is a simple reason for this failure of naturality: In general, f_* does not map $\Sigma_i(X)$ into $\Sigma_i(Y)$. In particular, there is no reason to expect our systems of morphisms to be preserved unless $f_* : \vec{\pi}_1(X; [X_0, X_1]) \to \vec{\pi}_1(Y; [Y_0, Y_1])$ is surjective. For another view on this naturality problem, compare S. Sokołowski's [**29**].

Is there an intermediate level (between the fundamental category and one of the component categories) on which one can talk about naturality?

6.1. Equivalences of categories with systems of morphisms

In this section, we look at categories $\mathcal{C}$ equipped with a system of morphisms $\Sigma \subset Mor(\mathcal{C})$ and an associated equivalence relation (3.1).

Definition 6.2. A functor $\Phi : (\mathcal{C}, \Sigma_{\mathcal{C}}) \to (\mathcal{D}, \Sigma_{\mathcal{D}})$ – with $\Phi(\Sigma_{\mathcal{C}}) \subseteq \Sigma_{\mathcal{D}}$ – is called an *equivalence* if

1. For every $g \in Mor_{\mathcal{D}[\Sigma_{\mathcal{D}}^{-1}]}$ there exists $f \in Mor_{\mathcal{C}[\Sigma_{\mathcal{C}}^{-1}]}$ such that $\Phi(f) \simeq_{\Sigma_{\mathcal{D}}} g$;
2. for $f_1, f_2 \in Mor_{\mathcal{C}[\Sigma_{\mathcal{C}}^{-1}]}$, one has: $\Phi(f_1) \simeq_{\Sigma_{\mathcal{D}}} \Phi(f_2) \Rightarrow f_1 \simeq_{\Sigma_{\mathcal{C}}} f_2$.

Pairs $(\mathcal{C}, \Sigma_{\mathcal{C}}), (\mathcal{D}, \Sigma_{\mathcal{D}})$ related by an equivalence or a (zig-zag) sequence of equivalences are called equivalent.

Applying the definition to identity morphisms in $\mathcal{D}$, one requires in particular every object in $\mathcal{D}$ to be $\Sigma_{\mathcal{D}}$-connected to an object in the image of Φ. More generally, an equivalence Φ induces an isomorphism $\Phi_* : \pi_0(\mathcal{C}; \Sigma_{\mathcal{C}}) \to \pi_0(\mathcal{D}; \Sigma_{\mathcal{D}})$ between the component categories.

In particular, the quotient functor $\pi_0(\Sigma) : (\mathcal{C}, \Sigma) \to (\pi_0(\mathcal{C}, \Sigma), I)$ – with I consisting only of the identity morphisms on the components – from Sect. 3.2 is an equivalence, by definition. More generally, let $\Sigma' \subset \Sigma$ denote a (closed) subsystem of morphisms, and let Σ/Σ' denote the system of equivalence classes. Then, we get a triangle of quotient functors

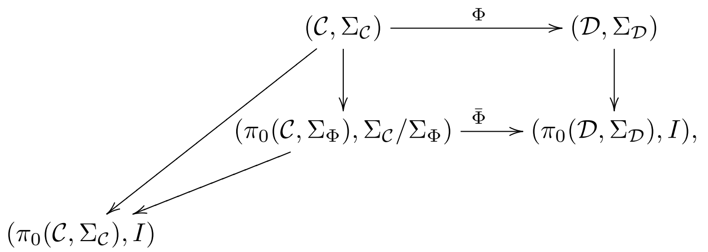
$$\begin{array}{ccc} (\mathcal{C}, \Sigma) & & \\ \downarrow & \searrow & \\ (\pi_0(\mathcal{C}; \Sigma'), \Sigma/\Sigma') & \longrightarrow & (\pi_0(\mathcal{C}, \Sigma), I). \end{array}$$

The diagonal functor is an equivalence. Hence, the vertical functor satisfies (2), and by definition, it satisfies (1), as well. As a result, the horizontal functor has to be an equivalence, as well.

6.2. Induced functors

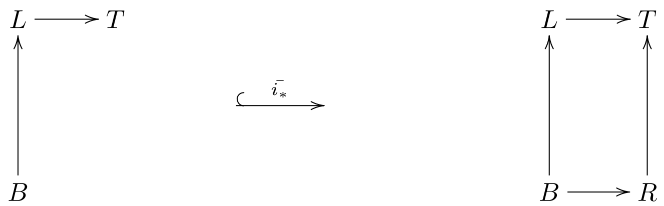
The following construction allows us to represent a functor $\Phi : \mathcal{C} \to \mathcal{D}$ that does not necessarily respect chosen systems $\Sigma_C \subset Mor_{\mathcal{C}}$ and $\Sigma_D \subset Mor_{\mathcal{D}}$ by a functor $\bar{\Phi}$ between equivalent categories of a "smaller" size inbetween the original and the component category. Here, two functors are considered equivalent if they can be "conjugated" into each other by a (zig-zag) sequence of equivalences of categories and systems on both sides.

We define the system $\Sigma(\Phi) := \Sigma_C \cap \Phi^{-1}(\Sigma_D)$ to consist of those morphisms, that are weakly invertible in $\mathcal{C}$ and whose images are weakly invertible in $\mathcal{D}$. It follows immediately from the definition that $\Phi(\Sigma_\Phi) \subseteq \Sigma_D$. We obtain a commutative diagram of functors



which "conjugates" Φ into the equivalent functor $\bar{\Phi}$.

Example 6.3. In the case of the functor i_* induced by inclusion $i : X = B \cup L \cup T \to Y$ from Ex. 6.1 on the fundamental categories, $\Sigma_4(Y)$ consists of the dipaths entirely contained in one of the domains B, L, R, resp. T. Hence, $\Sigma_4(i_*)$ consists of the dipaths entirely contained in one of the domains B, L, resp. T. Hence, $\bar{i_*}$ is the inclusion of categories



Alternatively, one might consider the lattice of systems contained in a particular system Σ and ask a functor to map "sufficiently" small systems of one lattice into systems of the other. This possibility is currently under investigation.

6.3. An application to the (non)existence of dimaps

An analysis of components and histories (cf. Sect. 4.3) can help to find restrictions to the existence of dimaps between lpo-spaces with specific properties. We use essentially the fact, that a dimap $\varphi : (X; X_0, X_1) \to (Y; Y_0, Y_1)$ preserves histories and finite intersections of these: $\varphi(hf) \subseteq h(\varphi \circ f)$ for f a dipath in X from X_0 to X_1.

Example 6.4. Let X and Y denote the two po-spaces from Fig. 5 together with their component categories $\pi_0(X; \Sigma_4)$ and $\pi_0(Y; \Sigma_4)$ with non-commuting and commuting squares (indicated by semicircular arrows). Both spaces X and Y admit exactly four dihomotopy classes of dipaths from the bottom to the top; those on X are given by $g_i * f_j$, $0 \leqslant i, j \leqslant 1$. Which abstract maps $\Phi : \vec{\pi}_1(X)(\mathbf{0}, \mathbf{1}) \to \vec{\pi}_1(Y)(\mathbf{0}, \mathbf{1})$ can be realised by a dimap $\varphi : (X; \mathbf{0}, \mathbf{1}) \to (Y; \mathbf{0}, \mathbf{1})$?

The intersection of the (homotopy) histories of all four dihomotopy classes in X from $\mathbf{0}$ to $\mathbf{1}$ consists of the union of the components $B \cup M \cup T$. Every dihomotopy class in Y is characterised by the particular "antidiagonal" component R_i, $1 \leqslant i \leqslant 4$, that it touches. In Y, the union of the six intersections of pairs of histories corresponding to the four dihomotopy classes is not pathwise connected (in the usual sense). Its two path components, denoted C_0 and C_1, consist of the six Σ_4-components below, resp. above the antidiagonal.

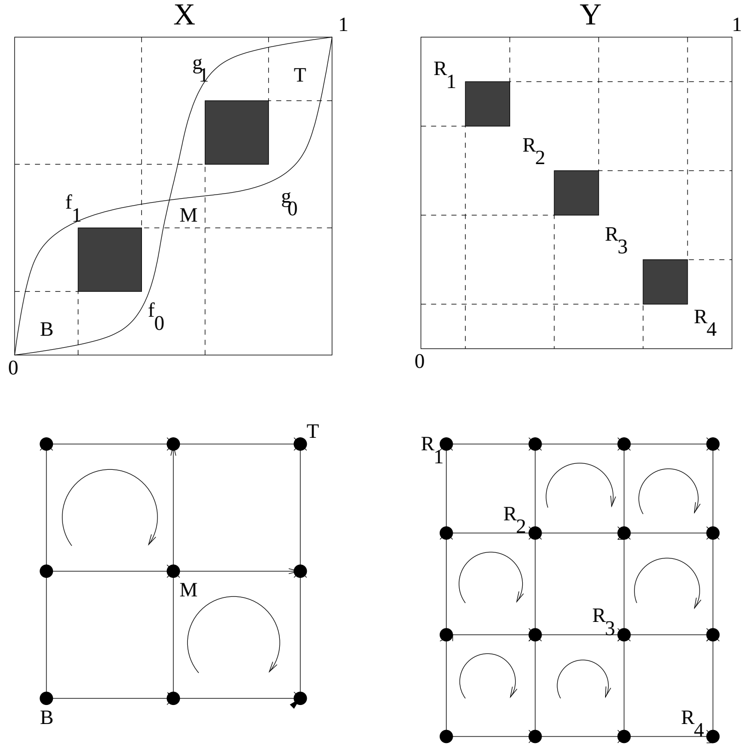


Figure 5: Two po-spaces and their component categories

If the image of $\varphi_* : \vec{\pi}_1(X)(\mathbf{0},\mathbf{1}) \to \vec{\pi}_1(Y)(\mathbf{0},\mathbf{1})$ contains at least two elements, then either M, and thus its "past" $\downarrow M$ are mapped into C_0 – or M and its "future" $\uparrow M$ are mapped into C_1. In the first case, $\varphi_*([g_i * f_0]) = \varphi_*([g_i * f_1])$, in the second $\varphi_*([g_1 * f_i]) = \varphi_*([g_0 * f_i])$, $0 \leqslant i \leqslant 1$. We conclude, that the image of φ_* has *at most two* elements. In particular, there is *no* surjective dimap $\varphi : (X; \mathbf{0}, \mathbf{1}) \to (Y; \mathbf{0}, \mathbf{1})$.

7. Concluding remarks

Lisbeth Fajstrup has worked on a translation of the covering concept to categories of lpo-spaces [**7**]. It turns out, that these "dicoverings", in general, have fibers with non-constant cardinality. It seems that cardinality is constant along the Σ_3-morphisms of the approach of this paper. It is an obvious task to work out an analogue to covering theory, i.e., to relate the combinatorics of (the component categories of) the fundamental categories to the topological investigation.

Certainly, the naturality problems touched upon in Sect. 6 deserve further investigation; a satisfactory framework seems to be crucial for several applications connected to the simulation and bisimulation concepts from concurrency theory. Combined with Grandis' version [**18**] of the Seifert-van Kampen theorem, we hope to be able to achieve algorithmic calculations of the (component categories) of fundamental categories, at least for spaces arising from the Higher Dimensional Automata mentioned in the introduction. This is the subject of ongoing work by L. Fajstrup, É. Goubault, E. Haucourt and the author.

References

[**1**] M. Aubry, *Homotopy theory and models*, Birkhäuser Verlag, Basel, 1995, Based on lectures held at a DMV seminar in Blaubeuren by H. J. Baues, S. Halperin and J.-M. Lemaire.

[**2**] T. Borceux, *Handbook of Categorial Algebra I: Basic Category Theory*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1994.

[**3**] R. Brown and P.J. Higgins, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Algebra **22** (1981), 11–41.

[**4**] ______, *On the algebra of cubes*, J. Pure Appl. Algebra **21** (1981), 233–260.

[**5**] S.D. Carson and P.F. Reynolds, *The geometry of semaphore programs*, ACM TOPLAS **9** (1987), no. 1, 25–53.

[**6**] E.W. Dijkstra, *Co-operating sequential processes*, Programming Languages (F. Genuys, ed.), Academic Press, New York, 1968, pp. 43–110.

[**7**] L. Fajstrup, *Dicovering spaces*, Tech. Report R-01-2021, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst, 2001, revised version to appear in Homology Homotopy Appl.

[**8**] L. Fajstrup, *Personal communication*, 2002.

[**9**] L. Fajstrup, É. Goubault, and M. Raussen, *Detecting Deadlocks in Concurrent Systems*, CONCUR '98; Concurrency Theory (Nice, France) (D. Sangiorgi and R. de Simone, eds.), Lect. Notes Comp. Science, vol. 1466, Springer-Verlag, September 1998, 9th Int. Conf., Proceedings, pp. 332 – 347.

[**10**] ______, *Detecting Deadlocks in Concurrent Systems*, DTA/LETI/DEIN/SLA 98-61, LETI (CEA - Technologies Avancées), Saclay, France, August 1998, 25 pp.

[**11**] ______, *Algebraic topology and concurrency*, Tech. Report R-99-2008, Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg Øst, June 1999, conditionally accepted for publication in Theoret. Comput. Sci.

[12] L. Fajstrup and S. Sokolowski, *Infinitely running concurrents processes with loops from a geometric viewpoint*, Electronic Notes Theor. Comput. Sci. **39** (2000), no. 2, 19 pp., URL: http://www.elsevier.nl/locate/entcs/volume39.html.

[13] P. Gabriel and M. Zisman, *Calculus of fractions and homotopy theory*, Springer-Verlag, New York, 1967, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 35.

[14] P. Gaucher, *Whitehead's Theorem in Homotopy Theory of Concurrency*, Tech. report, IRMA Univ. Strasbourg, 2002.

[15] P. Gaucher and É. Goubault, *Topological Deformations of Higher Dimensional Automata*, Tech. Report 01760, arXiv:math.AT, 2001, to appear in Homology Homotopy Appl.

[16] É. Goubault, *The Geometry of Concurrency*, Ph.D. thesis, Ecole Normale Superieure, Paris, 1995.

[17] É. Goubault and M. Raussen, *Dihomotopy as a tool in state space analysis*, LATIN 2002: Theoretical Informatics (Cancun, Mexico) (S. Rajsbaum, ed.), Lect. Notes Comput. Sci., vol. 2286, Springer-Verlag, April 2002, pp. 16 – 37.

[18] M. Grandis, *Directed Homotopy Theory I. The Fundamental Category*, Tech. Report 443, Dip. di Matematica dell' Univ. di Genova, 2001, to appear in Cahiers Top. Géom. Diff. Catég.

[19] ______, *Directed Homotopy Theory II. Homotopy Constructs*, Tech. Report 446, Dip. di Matematica dell' Univ. di Genova, 2001, to appear in Theory Appl. Categ.

[20] J. Gunawardena, *Homotopy and concurrency*, Bulletin of the EATCS **54** (1994), 184–193.

[21] M. Hovey, *Model categories*, Mathematical Surveys and Monographs, vol. 63, American Mathematical Society, 1999.

[22] S. Mac Lane, *Categories for the working mathematician*, Graduate Texts in Mathematics, vol. 5, Springer-Verlag, New York, Heidelberg, Berlin, 1971.

[23] W. Lipski and C.H. Papadimitriou, *A fast algorithm for testing for safety and detecting deadlocks in locked transaction systems*, Journal of Algorithms **2** (1981), 211–226.

[24] L. Nachbin, *Topology and Order*, Van Norstrand, 1965.

[25] V. Pratt, *Modelling concurrency with geometry*, Proc. of the 18th ACM Symposium on Principles of Programming Languages. (1991), 311–322.

[26] M. Raussen, *On the classification of dipaths in geometric models for concurrency*, Math. Structures Comput. Sci. **10** (2000), no. 4, 427–457.

[27] J. P. Serre, *Homologie singulière des espaces fibrés*, Ann. of Math. **(2) 54** (1951), 425–505.

[28] S. Sokołowski, *Classifying holes of arbitrary dimension in partially ordered cubes*, Manuscript. Kansas State University, February 2000.

[29] ______, *Categories of dimaps and their dihomotopies in po-spaces and local po-spaces*, Preliminary Proceedings of the Workshop on Geometry and Topology in Concurrency Theory GETCO'01 (Aalborg, Denmark) (P.Cousot et al., ed.), vol. NS-01, BRICS Notes Series, no. 7, BRICS, 2001, pp. 77 – 97.

[30] R. van Glabbeek, *Bisimulation semantics for higher dimensional automata*, Tech. report, Stanford University, 1991.

This article may be accessed via WWW at http://www.rmi.acnet.ge/hha/
or by anonymous ftp at
ftp://ftp.rmi.acnet.ge/pub/hha/volumes/2003/n2a9/v5n2a9.(dvi,ps,pdf)

Martin Raussen `raussen@math.auc.dk`

Department of Mathematical Sciences,
Aalborg University,
Fredrik Bajers Vej 7G, DK - 9220 Aalborg Ø

Applied Categorical Structures **12:** 81–108, 2004.

Components of the Fundamental Category *

L. FAJSTRUP[1], M. RAUSSEN[1], E. GOUBAULT[2] and E. HAUCOURT[2]
[1]*Department of Mathematical Sciences, Aalborg University, Denmark.*
e-mail: {fajstrup,raussen}@math.auc.dk
[2]*CEA/Saclay, France. e-mail: eric.goubault,emmanuel.haucourt@cea.fr*

Abstract. In this article we study the fundamental category (Goubault and Raussen, 2002; Goubault, 2000) of a partially ordered topological space (Nachbin, 1965; Johnstone, 1982), as arising in, e.g., concurrency theory (Fajstrup et al., 1999). The "algebra" of dipaths modulo dihomotopy (the fundamental category) of such a po-space is essentially finite in a number of situations: We define a component category of a category of fractions with respect to a suitable system, which contains all relevant information. Furthermore, some of these simpler invariants are conjectured to also satisfy some form of a van Kampen theorem, as the fundamental category does (Goubault, 2002; Grandis, 2001). We end up by giving some hints about how to carry out some computations in simple cases.

Mathematics Subject Classifications (2000): 18A32, 54F05, 55Q05, 55U40, 68N30, 68Q85.

Key words: po-space, dihomotopy, fundamental category, category of fractions, component, invertible morphism, lr-system, pure system, weakly invertible morphism.

1. Introduction

The aim of this paper is to show how to compute some algebraic topological invariants relevant to questions about concurrent and distributed systems.

A class of examples, which will be used throughout this text, generating geometrical invariants, arises from a toy language manipulating semaphores. Using Dijkstra's notation [3], we consider processes to be sequences of locking operations Pa on semaphores a and unlocking operations Va. In this introduction, we consider only binary semaphores, ensuring mutual exclusion of accesses, but in further examples, we will also model and use counting semaphores, or k-semaphores ($k > 1$) which can be accessed concurrently by up to k processes.

* A preliminary version of this work was presented and discussed at the Seminar *Structures for Computable Topology and Geometry* at Schloss Dagstuhl. Part of this work was accomplished while the first two authors were visiting Ecole Polytechnique, France.

This work was completed during an academic visit of the third author to the computer science department of Macquarie University, Sydney. Acknowledgments are due to the members of Sydney Category Theory Seminar for numerous discussions on the subject. Part of this work was done while the fourth author was visiting Aalborg University, with support from BRICS and Paris 7.

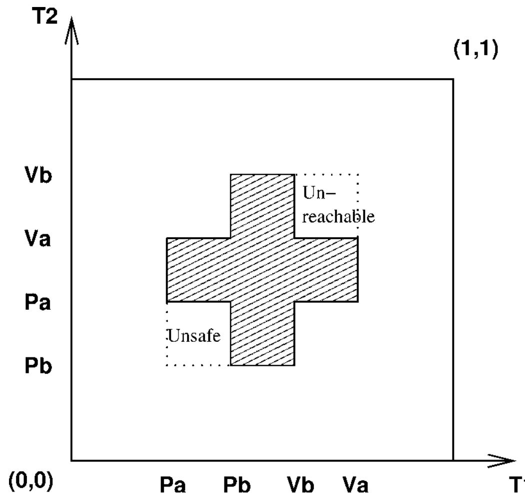


Figure 1. The Swiss Flag example – two processes sharing two resources.

In the example where two processes share two resources a and b:

$$T_1 = Pa.Pb.Vb.Va,$$
$$T_2 = Pb.Pa.Va.Vb$$

the geometric model is the "Swiss flag", Figure 1, regarded as a subset of $\mathbb{R}^2$ with the componentwise partial order $(x_1, y_1) \leq (x_2, y_2)$ if $x_1 \leq y_1$ and $x_2 \leq y_2$. The (interior of the) horizontal dashed rectangle comprises global states that are such that T_1 and T_2 both hold a lock on a: this is impossible by the very definition of a binary semaphore. Similarly, the (interior of the) vertical rectangle consists of states violating the mutual exclusion property on b. Therefore both dashed rectangles form the *forbidden region*, which is the complement of the space X of (legal) states. This space with the inherited partial order provides us with a particular po-space X [16, 13], as defined in Section 2.

Moreover, legal execution paths, called *dipaths*, are increasing maps from the po-space $\vec{I}$ (the unit segment with its natural order) to X. The partial order on X thus reflects (at least) the time ordering on all possible execution paths.

Many different execution paths have the same global effect: In the "Swiss Flag" example, for any execution path shaped like the one at the left of Figure 2, T_1 gets hold of locks a and b before T_2 does. To give a specific example, assume that T_1 does $b := b + 1$ and T_2 does $b := b * 2$, and that we start with 2 as an initial value for b. All execution paths below the hole will end up with the value $b = 6$, since T_1 will do $b = 2 + 1 = 3$ and then only after will T_2 do $b = 3 * 2 = 6$. In fact, there are only two essentially different execution paths from the initial point $(0, 0)$ to the final point $(1, 1)$, that fully determine the computer-scientific behaviour of the system: one is the type of dipaths just discussed, the other one runs to the left and above the central hole (see picture at the right-hand side of

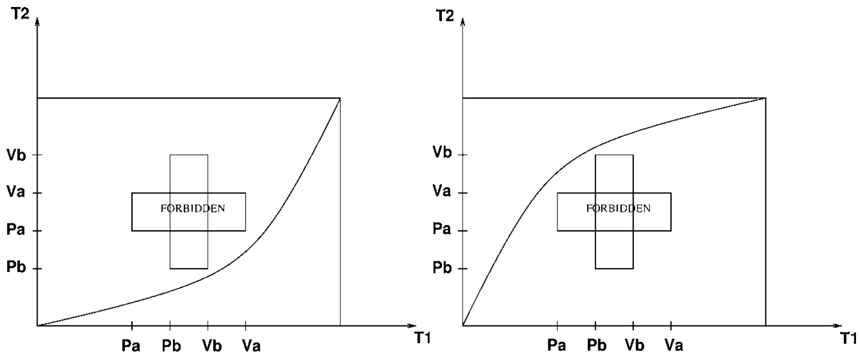


Figure 2. Essential schedules for the swiss flag.

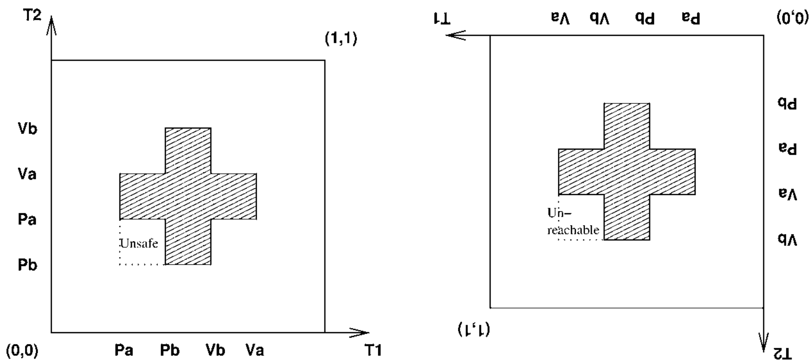


Figure 3. Deadlocks and unreachables.

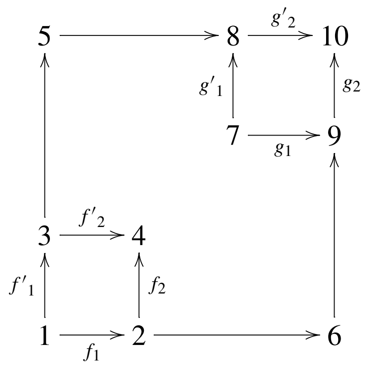
Figure 2). In terms of schedules of executions, the latter corresponds to executions in which T_2 is the first to read and write (after having got the corresponding locks) on a and b, before T_1 does (ending up with result $b = 2 * 2 + 1 = 5$). These are in fact the only two classes of dipaths from $(0, 0)$ to $(1, 1)$ modulo "continuous deformations" that do not reverse time, i.e., up to *dihomotopy* as defined in [5] and in Section 2, and these are the only paths of execution which are relevant to the computation of the possible final values that our toy concurrent program can reach. This fact is indeed general, and is not at all limited to the example. For determining the possible outcome of a concurrent program (modelled in a suitable way, as for our PV programs), only the dihomotopy classes of dipaths count, and it is thus natural, both on the mathematical side and for computer-scientific purposes, to try to characterize these classes.

We are not only interested in maximal dipaths modulo deformation. Other interesting dipaths, in our example space start in the initial point $(0, 0)$ and end in a deadlock, cf. the first picture of Figure 3, or start in an unreachable state and end in the final point $(1, 1)$, cf. the right-hand side of Figure 3. In fact, all continuous increasing paths entering the lower concavity below the holes is bound to end at the intersection of the two forbidden rectangles, which is the deadlock. This lower concavity is called the *unsafe* region. Now, formally reversing the order of time on the two coordinates results in the picture, on the right-hand side of Figure 3. Notice that the unsafe region, once time has been reversed, is in fact what we call the *unreachable* states. It is the set of points which cannot be reached from the initial point (with the initial ordering).

In general, one of the important invariants of a concurrent system is its *fundamental category* [9, 10], classifiying dipaths between any pair of points up to dihomotopy, i.e., a directed version of the fundamental groupoid [2] of a topological space. A drawback of the fundamental category is that it is less easy to compute than the fundamental groupoid or the fundamental group. There are similarities though, for instance there is a van Kampen theorem in the directed case [9, 11].

Our aim is to go further in the study of the *algebraic properties* of the fundamental category in order to manipulate and compute it for a variety of systems. In nice cases, the relevant information in the fundamental category is essentially finite. This is shown using a construction based on categories of fractions [6], which are briefly explained in Section 3. The principle is to formally "invert" systems of "inessential" morphisms in the fundamental category. Of course, we should be able to deduce from this construction, applied to the Swiss flag example, at least the regions of unsafe and unreachable states, and also that we have two classes of maximal dipaths.

In fact we want a little more than that. Our aim is to decompose the fundamental category into big chunks as the regions 1 to 10 in Figure 4. Basically, inside these regions, or components, nothing important happens: first of all, there is at most one dihomotopy class of dipaths between any two points in the same component. Moreover, composing with morphisms (= dihomotopy classes of dipaths) within these regions does not affect the "shape" of the future nor of the past. We will consider the category of fractions with these morphisms formally inverted. A certain quotient of the fundamental category with respect to this system of "inessential" morphisms forms then the category of components, which, in our example is the following finite category:



together with relations $g'_2 \circ g'_1 = g_2 \circ g_1$ and $f'_2 \circ f'_1 = f_2 \circ f_1$ (compare with, e.g., [8]). In some sense, this category of components finitely presents the fundamental

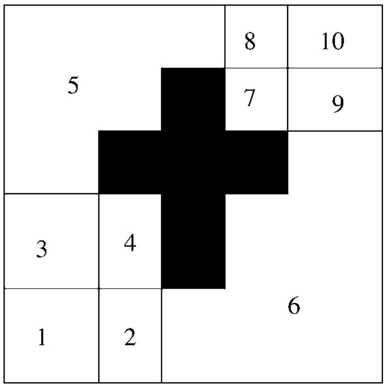


Figure 4. The components of the Swiss flag.

category. In particular, we can infer from this component category, all dihomotopy information because of a lifting property, see Propositions 3 and 7.

In general it is not obvious how to characterize the inessential morphisms, i.e., the morphisms which should be inverted formally. This leads to a more specific calculus of fractions, in particular left and right categories of fractions as defined in Section 4. Moreover, as shown in [6], finite limits and finite colimits are preserved when taking left and right categories of fractions. We can view equalizers, sums and products (when they exist, at least locally in some subcategories) as expressing particular equations between dipaths modulo dihomotopy which the category of fractions we construct has to preserve.

We then apply this "abstract nonsense" to various topological situations, arising from questions regarding dihomotopies. Last but not least, we give some hints about how to compute these invariants for simple spaces like some compact subsets of $\mathbb{R}^n$ with the componentwise ordering. This is done in Section 5.

There are two importants points that need to be outlined, particularly with respect to earlier work [18, 10]:

First, the "inessential" morphisms used to be defined with respect to sets of initial and final points. Dipaths were considered as pieces of dipaths between a set of initial points and a set of final points. The new definitions allow us to be more natural, without any reference to specific sets of points. This also tackles some of the problems with the "homotopy history equivalence" relation as defined in [5] which also needed to be bipointed by sets of initial and final points. In some sense, Proposition 6 shows that the new definition of inessential morphisms allows us to encompass all possible choices of initial and final sets.

Another modification arises from the fact that the fundamental category does not satisfy cancellation properties, in general. This is the reason for introducing the additional concept of a *pure* system (cf. Definition 4.3): For instance, the fundamental category of a cube minus an inner cube (see Section 6) is *not* trivial (as its fundamental group). Close to the inner deleted cube, there are local obstructions to directed homotopy of directed paths. But these are cancelled out under any long enough extensions, in the future as well as in the past. In general it is not clear whether composites of essential morphisms can become inessential. To avoid this, we ask a system of morphims to be inverted to satisfy the pureness property from Definition 4.3. And furthermore, this property has some very nice consequences as we show in Section 5. Unfortunately, it is not clear in general how to construct significant systems of inessential morphisms satisfying this pureness property.

2. Basic Definitions

The framework for the applications we have in mind is mostly based on the simple notion of a po-space:

DEFINITION 1. (1) A po-space is a topological space X with a (global) closed partial order $\leq$ (i.e., $\leq$ is a closed subset of $X \times X$).

(2) A dimap $f : X \to Y$ between po-spaces X and Y is a continuous map that respects the partial orders (is non-decreasing).

(3) A dipath $f : \vec{I} \to X$ is a dimap whose source is the interval $\vec{I}$ with the usual order.

Po-spaces and dimaps form a category. To a certain degree, our methods apply to the more general categories of lpo-spaces [5] (with a local partial order), of flows [7] and of d-spaces [11], but for the sake of simplicity, we stick to po-spaces in the present paper. Dihomotopies between dipaths f and g (with fixed extremities α and β in X) are dimaps $H : \vec{I} \times I \to X$ such that for all $x \in \vec{I}$, $t \in I$,

$$H(x, 0) = f(x), \qquad H(x, 1) = g(x), \qquad H(0, t) = \alpha, \qquad H(1, t) = \beta.$$

A dihomotopy is to be understood as a 1-parameter family of dimaps without order requirements in the second I-coordinate.* Now, we can define the main object of study of this paper:

DEFINITION 2. The fundamental category is the category $\vec{\pi}_1(X)$ with:

- as objects: the points of X,
- as morphisms, the dihomotopy classes of dipaths: a morphism from x to y is a dihomotopy class $[f]$ of a dipath f from x to y.

Concatenation of dipaths factors over dihomotopy and yields the composition of morphisms in the fundamental category. A dimap $f : X \to Y$ between po-spaces induces a functor $f_\# : \vec{\pi}_1(X) \to \vec{\pi}_1(Y)$, and we obtain thus a functor $\vec{\pi}_1$ from the category of po-spaces to the category of categories.

* This is slightly different for d-spaces, but coincides in important cases.

The fundamental category of a po-space generalizes the fundamental group $\pi_1(X)$ of a topological space X (a single object = base point; morphisms = homotopy classes of loops). It is often an enormous gadget (with uncountably many objects and morphisms) and possesses less structure than a group. It is the aim of this paper to "shrink" the essential information in the fundamental category to an associated component category, that in many cases is finite and possesses a comprehensible structure.

3. Categories of Fractions and Component Categories

Many of the tools we need for the study of the fundamental category can in fact be applied to at least all small categories. These are the notions of categories of fractions, of left and right calculi of fractions and of pure systems. The first two notions are well-known in the category theory literature [6, 1] and were already applied to the analysis of fundamental categories in [18, 10]. The new notion in this paper is that of *pure* systems yielding far more satisfactory applications.

3.1. CATEGORIES OF FRACTIONS

In the sequel, we will only consider *small* categories (most of the results would still hold with locally small categories [15], but we do not need these in the applications to the fundamental category).

DEFINITION AND LEMMA ([1]). Let $\mathcal{C}$ be a category.

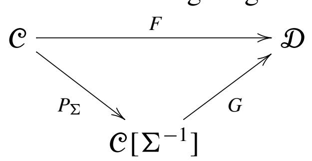
(1) A subset $\Sigma \subset Mor(\mathcal{C})$ is called a *system of morphisms* of $\mathcal{C}$ if

(i) $\forall x$ object of $\mathcal{C}$, $Id_x \in \Sigma$,
(ii) $\forall \sigma_1 : x \longrightarrow x', \sigma_2 : x' \longrightarrow x'' \in \Sigma, \sigma_2 \circ \sigma_1 \in \Sigma$.

(In other words, the objects of $\mathcal{C}$ together with Σ form a wide subcategory of $\mathcal{C}$.)

(2) Given a system Σ of morphisms* in $\mathcal{C}$, there is, up to isomorphism of categories, a unique category (denoted $\mathcal{C}[\Sigma^{-1}]$) and a functor $P_\Sigma : \mathcal{C} \longrightarrow \mathcal{C}[\Sigma^{-1}]$, such that:

- $\forall \sigma \in \Sigma$, $P_\Sigma(\sigma)$ is an isomorphism of $\mathcal{C}[\Sigma^{-1}]$.
- For any functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ such that if σ is an isomorphism of $\mathcal{C}$ then $F(\sigma)$ is an isomorphism of $\mathcal{D}$, there is a unique functor $G : \mathcal{C}[\Sigma^{-1}] \longrightarrow \mathcal{D}$ such that the following diagram commutes:



* Note that the assumptions in (1) are not necessary for the existence of a category of fractions. Considering only those Σ that are subcategories of $\mathcal{C}$ will make things simpler in the rest of the paper, and we do not lose generality by this.

In fact, each morphism of $\mathcal{C}[\Sigma^{-1}]$ can be represented in the form $\sigma_1^{-1} \circ a_1 \circ \cdots \circ \sigma_{k-1}^{-1} \circ a_k$ where each a_i is a morphism of $\mathcal{C}$ and σ_i^{-1} denotes the formal inverse of $\sigma_i \in \Sigma$, cf. [6, 1].

EXAMPLE 1. In algebraic topology, one considers the category of CW-complexes or of simplicial sets with formal inverses to the system of "weak equivalences", i.e., those maps which induce isomorphisms of all homotopy groups. This category of fractions is called the *homotopy category* or the category of "homotopy types" [6].

3.2. COMPONENT CATEGORIES

Any morphism of the form $s_1^{-1} \circ s_2 \circ \cdots \circ s_{2k-1}^{-1} \circ s_{2k}$, $s_j \in \Sigma, k \in \mathbb{N}$ is called a Σ-*zig-zag* morphism. The set $ZZ(\Sigma)$ of all Σ-*zig-zag* morphisms forms a system of morphisms contained in the invertibles of the category of fractions, denoted $Inv(\mathcal{C}[\Sigma^{-1}])$. Equality holds if Σ contains the invertibles $Inv(\mathcal{C})$ of the original category $\mathcal{C}$. In fact, $\mathcal{C}[(\Sigma \cup Inv(\mathcal{C}))^{-1}] = \mathcal{C}[\Sigma^{-1}]$. The subcategory of $\mathcal{C}[\Sigma^{-1}]$ with all objects, the morphisms of which are given by the zig-zag morphisms $ZZ(\Sigma)$, forms in fact a *groupoid*.

Two objects $x, y \in Ob(\mathcal{C})$ are called Σ-*related* – $x \simeq_\Sigma y$ – if there exists a zig-zag-morphism from x to y. This definition corresponds to usual path connectedness *with respect to paths in* Σ *only – but regardless of orientation*. Being Σ-related is an equivalence relation; the equivalence classes will be called the Σ-*connected components* – the path components with respect to Σ-zig-zag paths, i.e., the components of the groupoid above. This can be rephrased by saying that $\simeq_\Sigma$ is the equivalence relation generated by the relation $x \equiv y$ if there exists $\sigma : x \to y$ in Σ.

Next, consider the smallest equivalence relation on the morphisms of $\mathcal{C}[\Sigma^{-1}]$ generated (under composition, when they make sense) by

$$\alpha \simeq \alpha \circ s, \qquad \alpha \simeq t \circ \alpha \quad \text{for } \alpha \in Mor(x, y).$$

Remark that equivalent morphisms no longer need to have the same source or target. In particular, every morphism in Σ is equivalent to the identities in both its source and its target; hence, all zig-zag morphisms within a component are equivalent to each other.

Dividing out the morphisms in Σ within $\mathcal{C}$, we arrive at a *component category*: The objects of the component category $\pi_0(\mathcal{C}; \Sigma)$ are by definition the Σ-connected components of $\mathcal{C}$; the morphisms from $[x]$ to $[y]$, $x, y \in Ob(\mathcal{C})$, are the equivalence classes of morphisms in $\bigcup_{x' \simeq_\Sigma x, y' \simeq_\Sigma y} Mor_{\mathcal{C}[\Sigma^{-1}]}(x', y')$. The composition of $[\beta] \circ [\alpha]$ for $\alpha \in Mor_{\mathcal{C}[\Sigma^{-1}]}(x, y)$ and $\beta \in Mor_{\mathcal{C}[\Sigma^{-1}]}(y', z)$ is given by $[\beta \circ s \circ \alpha]$ with s any zig-zag morphism from y to y'. The equivalence class of that composition is independent of the choices of representatives α and β (by definition) and of the choice of the zig-zag path s by the preceeding remark.

The overall idea is thus as follows: Having fixed a suitable system Σ of "weakly invertible" morphisms, we decompose the study of $\mathcal{C}$ into the study of

- the component category encompassing the global effects of irreversibility and
- the components with a *groupoid* structure given by the Σ-zig-zags.

The original category $\mathcal{C}$ and the component category $\pi_0(\mathcal{C};\Sigma)$ are related by a functor $\pi_0(\Sigma) : \mathcal{C} \overset{P_\Sigma}{\to} \mathcal{C}[\Sigma^{-1}] \to \pi_0(\mathcal{C};\Sigma)$; the last arrow is the quotient functor.

As noticed by Marco Grandis and the anonymous referee, this extends the case of quotients of groupoids by normal groupoids (see [12]), at least in the case when the only endomorphisms in the category $\mathcal{C}$ are the identities (hence every subgroupoid is normal), as will be the case in most of what follows. Also, the component category can be seen as a pushout: let T be the functor which associates to a set S the trivial groupoid on S (one invertible arrow for each element of S). Let K be the set of connected components of Σ. Let R_Σ be the functor which to each arrow σ of Σ associates the identity on the component of the domain and range of σ in the groupoid $T(K)$. Then $\pi_0(\mathcal{C},\Sigma)$ is part of the following pushout diagram:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\subseteq} & \mathcal{C} \\ {\scriptstyle R_\Sigma}\downarrow & & \downarrow{\scriptstyle \pi_0(\Sigma)} \\ T(K) & \longrightarrow & \pi_0(\mathcal{C};\Sigma) \end{array}$$

3.3. FUNCTORS

Let Σ denote a system of morphisms in the category $\mathcal{C}$ and Υ a system of morphisms in the category $\mathcal{D}$. To ensure that a functor $F : \mathcal{C} \to \mathcal{D}$ induces a well-defined functor between the categories of fractions $\mathcal{C}[\Sigma^{-1}] \to \mathcal{D}[\Upsilon^{-1}]$ and then between the categories of components $\pi_0(\mathcal{C},\Sigma)$ and $\pi_0(\mathcal{D},\Upsilon)$, we need to assume that $F(\Sigma) \subseteq \Upsilon$. This is not at all automatically satisfied in easy geometric examples with systems of weakly invertible morphisms. But one can always refine a given system to ensure this condition:

LEMMA 1. *F induces*

- *a functor $F_{\Sigma,\Upsilon}$ from $\mathcal{C}[(\Sigma \cap F^{-1}(\Upsilon))^{-1}]$ to $\mathcal{D}[\Upsilon^{-1}]$,*
- *a functor $\pi_0 F_{\Sigma,\Upsilon}$ from $\pi_0(\mathcal{C}, \Sigma \cap F^{-1}(\Upsilon))$ to $\pi_0(\mathcal{D},\Upsilon)$.*

Proof. Obvious. □

Particularly important is the case of an inclusion $i : \Sigma_1 \hookrightarrow \Sigma_2$ of systems of morphisms within the *same* category $\mathcal{C}$. The identity on $\mathcal{C}$ leads immediately to the functor $i_{\Sigma_1\Sigma_2} : \mathcal{C}[\Sigma_1]^{-1} \to \mathcal{C}[\Sigma_2]^{-1}$ and to the functor

$$\pi_0 i_{\Sigma_1\Sigma_2} : \pi_0(\mathcal{C},\Sigma_1) \to \pi_0(\mathcal{C},\Sigma_2)$$

which reflects an inverse to a *refinement*. In general, it is useful to understand the structure of wide sub-categories of fractions of $\mathcal{C}[\Sigma^{-1}]$ where we invert fewer morphisms than the ones of Σ.

LEMMA 2. *Let L_Σ be the poset of categories of the form $\mathcal{C}[\Lambda^{-1}]$ where $\Lambda \subseteq \Sigma$, with the inclusion of morphisms as partial order. L_Σ is a complete lattice.*

Proof. Let $(\Sigma_i)_{i\in I}$ a family of systems of morphisms. It is easy to see that $\mathcal{C}[(\bigcap_{i\in I}\Sigma_i)^{-1}]$ is the greatest lower bound in L_Σ.

Let the least system of morphisms stable under composition of the underlying category, containing all Σ_i be denoted by $\biguplus_{i\in I}\Sigma_i$. Then $\mathcal{C}[(\biguplus_{i\in I}\Sigma_i)^{-1}]$ is the least upper bound of the families of categories $\mathcal{C}[\Sigma_i^{-1}]$. □

In fact, the induced functor of Lemma 1 is the largest functor agreeing with F, meaning that it is couniversal with respect to inclusion maps $\mathcal{C}[\Lambda^{-1}] \hookrightarrow \mathcal{C}[\Sigma^{-1}]$ (which are the maps induced by set-theoretic inclusion maps $\Lambda \hookrightarrow \Sigma$). If one uses the components with respect to a system Σ of morphisms as the basis for a topology on the objects, then Lemma 1 states that we can always take the greatest such topology making F continuous.

4. Calculi of Fractions

4.1. WEAKLY INVERTIBLE MORPHISMS

In the case of the fundamental category $\mathcal{C} = \vec{\pi}_1(X)$ of a po-space $(X, \leq)$, we want to define morphisms to be "weakly invertible" if no "decision" is taken. This means that composition on the left and on the right with such morphims induce (natural) bijections between sets of morphisms. This idea can be formulated for general small categories:

Let $\mathcal{C}_{\to x}$ (respectively $\mathcal{C}_{y\to}$) denote the full subcategory of $\mathcal{C}$ consisting of objects z such that $\mathcal{C}(z, x) \neq \emptyset$ (respectively $\mathcal{C}(y, z) \neq \emptyset$), and consider first the Yoneda functor: $\mathcal{Y}_\mathcal{C} : \mathcal{C} \to \hat{\mathcal{C}}$, where $\hat{\mathcal{C}}$ is the category of presheaves over $\mathcal{C}$.

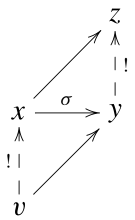
DEFINITION 3. We say that a morphism $\sigma : x \to y$ in $\mathcal{C}$ is *weakly invertible* on the left (respectively on the right) if for all objects z, $\mathcal{Y}_\mathcal{C}(\sigma)$ (respectively $\mathcal{Y}_{\mathcal{C}^{op}}(\sigma)$) is a natural isomorphism when restricted to $\mathcal{C}_{\to x}$ (respectively on $\mathcal{C}^{op}_{\to y} = \mathcal{C}_{y\to}$). We say that σ is weakly invertible if σ is weakly invertible both on the left and on the right.*

Less abstractly formulated, we ask that all maps (for all $v, z \in \mathcal{C}$):

$$\begin{array}{ccc} \mathcal{C}(y,z) \longrightarrow \mathcal{C}(x,z) & \qquad & \mathcal{C}(v,x) \longrightarrow \mathcal{C}(v,y) \\ & & \\ g \longrightarrow g\circ\sigma & & h \longrightarrow \sigma\circ h \end{array}$$

* The fact, that we look only at *restrictions* of the Yoneda functor on $\mathcal{C}_{\to x}$ and $\mathcal{C}_{y\to}$ is of primary importance: otherwise we would define the weakly invertible morphism to be the isomorphisms in the *original* category, by Yoneda's lemma.

are set-theoretic bijections:



whenever $\mathcal{C}(y, z) \neq \emptyset$ (respectively, on the right-hand side, $\mathcal{C}(v, x) \neq \emptyset$). Obviously,

LEMMA 3. *The weakly invertible morphisms in* $\mathcal{C}$ *form a system of morphisms* (*cf. Definition* 1).

As an example, consider the po-space of Figure 5, which is $\vec{I} \times \vec{I}$ minus the interior of a square.

All morphisms with end-points within the closed square region A or D are weakly invertible in the sense above. Similarly, all morphisms with both end points within the open regions B, resp. C, are weakly invertible.

4.2. CALCULI OF LEFT AND RIGHT FRACTIONS

Whether a morphism $s \in \mathcal{C}(x, y)$ is weakly invertible or not depends only on the morphisms with a target reachable from y, resp. with a source that can reach x. This condition is thus, in general too weak to compare objects with respect to all ingoing and outgoing morphisms.

EXAMPLE 2. In Figure 6, the "vertical" morphism s is weakly invertible, but "taking" this morphism represents a decision (in particular to end in a deadlock or to have the possibility of ending in the final state).

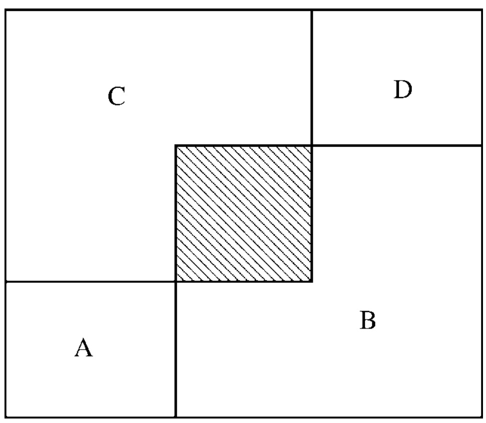
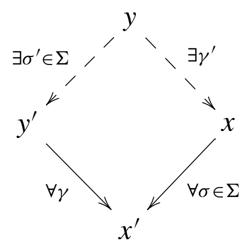


Figure 5. A simple po-space and components containing only weakly invertible morphisms.

This defect can be repaired by an extra condition to the system Σ to be chosen. This "lr" condition moreover allows us to represent morphisms in the category of fractions and in the component category in a much easier way:

DEFINITION 4 ([1]). Let $\mathcal{C}$ be a category. A system Σ of morphisms in $\mathcal{C}$ is said to *admit a right calculus of fractions* (for short: is an r-system) if it satisfies (in addition to properties (i) and (ii) from Definition 4):

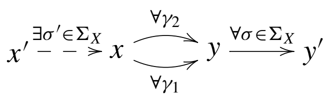
(iii) $\forall\gamma : y' \longrightarrow x', \forall\sigma : x \longrightarrow x' \in \Sigma, \exists\sigma' : y \longrightarrow y', \exists\gamma' : y \longrightarrow x$ such that $\sigma \circ \gamma' = \gamma \circ \sigma'$, i.e. the following diagram is commutative:



(iv) $\forall\gamma_1, \gamma_2 : x \longrightarrow y, \forall\sigma : y \longrightarrow y' \in \Sigma$ such that

$$\sigma \circ \gamma_1 = \sigma \circ \gamma_2,\ \exists\sigma' : x' \longrightarrow x \in \Sigma$$

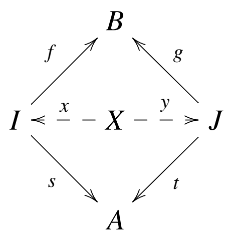
such that $\gamma_1 \circ \sigma' = \gamma_2 \circ \sigma'$



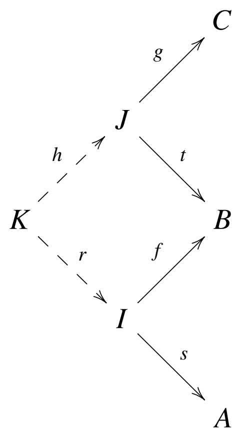
Property (iii) will be called the extension property for calculi of right fractions. A left calculus of fractions is defined similarly. We will also call such a system Σ an r-system, respectively l-system, respectively lr-system for left and right fractions.

A straightforward consequence of the extension property for calculi of right fractions – and explaining the name – is that every morphism of $\mathcal{C}[\Sigma^{-1}]$ can be written as $[(a\sigma^{-1})]$ for certain morphisms a of $\mathcal{C}$ and $\sigma \in \Sigma$.

As for ordinary fractions, fs^{-1} and gt^{-1} can represent equivalent morphisms in the category of fractions $\mathcal{C}[\Sigma^{-1}]$. In fact they do if one can find morphisms $x : X \to I$ and $y : X \to J$ in $\mathcal{C}$ such as in the following commutative diagram:



and such that sx (= ty) is in Σ. Now the composite of equivalence classes of $fs^{-1} : A \to B$ with $gt^{-1} : B \to C$ is the class of morphism $(g \circ h)(s \circ r)^{-1}$ as pictured in the following diagram:



The object K and the morphisms h and r arise from the extension property of calculi of right fractions.

For the properties of the component category with respect to an lr-system, cf. Section 5.2.

LEMMA 4. *The class of weakly invertible morphisms on the right satisfy property* (iv) *of calculi of right fractions. The class of weakly invertible morphisms on the left satisfy property* (iv) *of calculi of left fractions.*

Proof. Let $s : z \to x$ be a weakly invertible on the right and $f, g : x \to y$ such that $f \circ s = \mathcal{Y}_{\mathcal{C}^{op}}(s)(f) = g \circ s = \mathcal{Y}_{\mathcal{C}^{op}}(s)(g)$.

As s is weakly invertible on the right, $\mathcal{Y}_{\mathcal{C}^{op}}(s)$ is a bijection from $\mathcal{C}(x, y)$ to $\mathcal{C}(z, y)$ so we must have $f = g$. Just take $t = Id_y$ which is weakly invertible on the right (by (i)): this gives property (iv) of right-fractions.

The dual of (iv) is proven similarly by using $\mathcal{Y}_{\mathcal{C}}(s)$. □

It is *not* true in general that the class of weakly invertible morphisms is a calculus of left or right fractions. An example is the fundamental category of the Swiss flag again, Figure 1. Every morphism is weakly invertible in regions 1, 2, 3, 5, 6, 8, 9 and 10 of Figure 4 but there is no way to "detect" regions 4 and 7. Look at Figure 6, if we suppose s to be weakly invertible, p is the dipath shown on this figure, we cannot find a way that property (iii) is satisfied. So if we impose the lr properties, then we are bound to subdivide furthermore the regions, to find regions 4 and 7. There are examples for which the only left and right calculus of fractions included in the weak-invertibles is the set of identities, making the retract of the fundamental category no simpler than the fundamental category itself (see Figure 9).

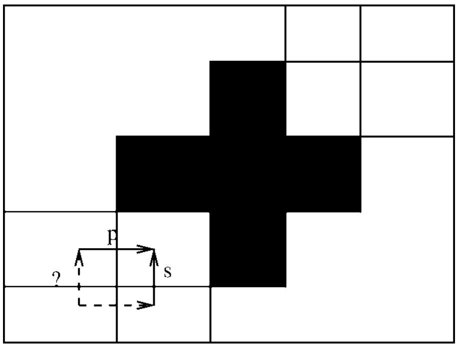


Figure 6. How to find the components with lr conditions.

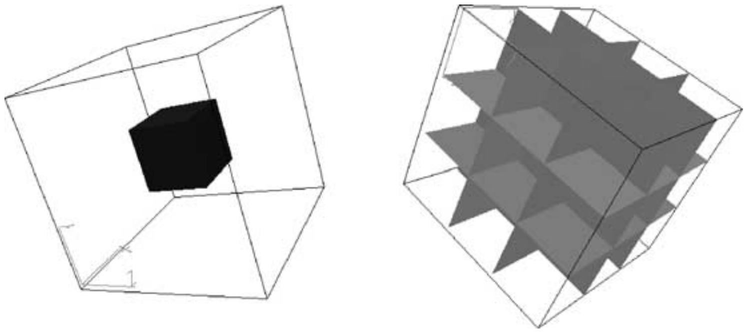


Figure 7. Weakly invertible morphisms need not be pure.

4.3. PURE SYSTEMS

Why not just stop here? If we look at simple examples, the category of components seems alright. For instance the component category of the weakly invertible morphisms (defining here a left and right category of fractions of Definition 3) for the po-space from Figure 5 is just the free category on the graph delineated in Figure 8.

But now, consider the po-space consisting of the left part of Figure 7, i.e., $\vec{I} \times \vec{I} \times \vec{I}$ minus the interior of a cube. Then all morphisms in the interior of the 26 regions delineated in the right-hand side of the same figure are weakly invertible. But any dipath from the initial to the final point is weakly invertible as well, composed of the composition of a number of non-weakly invertible morphisms going from one of the 26 regions to a neighbouring one (we will come back to the full

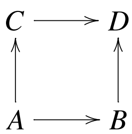


Figure 8. The category of components of a simple po-space.

calculation in Section 6). This means that the component category given by inverting the weakly invertible morphisms would actually have some endomorphisms which would not be the identity. In the concrete case here, when we impose the lr-conditions, the problem disappears, but it is not at all clear, that this is true in general.

So we try to eliminate such pathologies by imposing some extra condition at the calculus of invertibles, that we now define:

DEFINITION 5. A system of morphisms Σ within a category $\mathcal{C}$ is called *pure* if

(v) for all $\sigma \in \Sigma$, if we can decompose σ as $g \circ f$ then g and f must be in Σ too.

Another way to phrase this property is that no invertibles can be decomposed using a non-invertible, in particular $Mor(\mathcal{C}) \setminus \Sigma$ is closed under composition. In some ways, a real inessential morphism should be a morphism that does not make any decision, not only from start to end point but also on the way: some decisions cancel out but we want to have the "atomic" ones. This technical condition will also prove extremely useful in the proofs of most propositions in Sections 5.2 and 5.3.

4.4. MAXIMAL SYSTEMS

In the following, we will mainly be concerned with pure lr-systems of morphisms. A system of morphisms always contains a largest subsystem which is lr, but there is no reason that there should be a maximal pure lr subsystem.

Given a system of morphisms Σ in a category $\mathcal{C}$. The least subsystem $\Sigma' \subset \Sigma$ that has the l-, r-, lr-property, consists just of the identity morphisms Id_x. But there is also always a greatest such system:

LEMMA 5. *Let* $(\Sigma_i)_{i \in I}$ *a family of systems of morphisms so that* $\forall i \in I$, Σ_i *satisfies all the conditions of the right* (*left, respectively*) *calculus of fractions and* $\Sigma_i \subset \Sigma$. *Then* $\biguplus_{i \in I} \Sigma_i \subset \Sigma$ *satisfies all the conditions of the right* (*left, respectively*) *calculus of fractions.*

Proof. The first and second conditions are clearly satisfied. The rest is easily done by induction on the number of compositions of morphisms. □

If we are lucky enough to start with a pure subcategory Σ (meaning that factorization of morphisms are always within Σ) of weakly invertible morphisms, then it is the case that the greatest lr-subcategory is still pure. This is of course in that case the "maximal pure lr-system" in Σ:

PROPOSITION 1. *Let* Σ *be any pure subcategory of the weakly invertible morphisms in a category* $\mathcal{C}$ (*with unique identity endomorphisms*). *Then the greatest left and right calculus of fractions in* Σ *for* $\mathcal{C}$ *is a pure calculus of fractions.*

Proof. We take any sub-lr-system Σ' of Σ and we suppose it is not pure. Then there exists $\sigma = f_1 \circ f_2$ with f_1 or f_2 not in Σ'. Consider Σ'' the category generated by Σ' and f_1 (similarly for f_2). Then it can be shown that it is a lr-system, including strictly Σ', but included in Σ. This proves that the greatest lr-system (which always exists by Lemma 5) in Σ has to be pure. □

In the general situation, one may proceed as follows: We work with pairs $(\mathcal{C}, \Sigma)$ of categories and admissible systems of weakly invertible morphisms (admissible means lr, pure lr etc.) Call a functor $F : (\mathcal{C}_1, \Sigma_1) \to (\mathcal{C}_2, \Sigma_2)$ with $F(\Sigma_1) \subseteq \Sigma_2$ an elementary equivalence if

- $\mathcal{C}_1 = \mathcal{C}_2$ or
- $\pi_0 F_{\Sigma_1 \Sigma_2} : \pi_0(\mathcal{C}_1, \Sigma_1) \to \pi_0(\mathcal{C}_2, \Sigma_2)$ is an equivalence of categories

the first of which reflects refinements, cf. Section 3. Regard the equivalence relation on these pairs generated by elementary equivalences. Then the component category with respect to *any* admissible system of morphisms is equivalent to the original category with the system of identities. Of course, it is still interesting to ask for as large as possible admissible systems (as coarse as possible component categories), although these might not be unique.

5. Properties of Systems of Weakly Invertible Morphisms and Corresponding Components

5.1. WEAKLY INVERTIBLE MORPHISMS AND HISTORIES

There is a strong link between factorization properties in $\mathcal{C}$ and weak invertibility (as categories of fractions and factorization systems do have in general, see [1]), which will have a strong geometric and concurrency theoretical meaning: the homotopy histories of [5]. First, we need a definition:

DEFINITION 6. Given two objects x_0, x_1 in $\mathcal{C}$ and $f : x_0 \to x_1$ a morphism in $\mathcal{C}$, the history $h_{x_0,x_1}[f]$ of f is defined as

$$h_{x_0,x_1}[f] = \{x \in \mathcal{C} \mid \exists f_0 : x_0 \to x,\ f_1 : x \to x_1 \text{ with } f = f_1 \circ f_0\}.$$

Two objects $x, y \in \mathcal{C}$ are history equivalent (for a given x_0 and x_1, noted $x \sim_{x_0,x_1} y$, if $x \in h_{x_0,x_1}[f] \Leftrightarrow y \in h_{x_0,x_1}[f]$ for all $f : x_0 \to x_1$).

LEMMA 6. *Let $\mathcal{C}$ be a category where the only endomorphisms are identities. Then, $\sigma : x \to y$ has surjective composition on the left* (*as weakly invertible on the left, in Definition* 3, *except we only require surjectivity*) *and on the right if and only if,*

- *$x \sim_{x_0,x_1} y$ for all x_0 and x_1 such that $\mathcal{C}(x_0, x) \neq \emptyset$ and $\mathcal{C}(y, x_1) \neq \emptyset$,*
- *σ is the only morphism from x to y in $\mathcal{C}$.*

Proof. Easy diagram chasing. The second condition of the lemma is of course needed geometrically: take for instance the square minus one square in Figure 5. Regions A and D are in the same homotopy-history equivalence classes. This is implied by the surjectivity of left composition with σ by the same argument as the one of Proposition 2. □

This implies that weak invertibility refines the notion of homotopy history equivalence of [5]. In the case where $\mathcal{C}$ is the fundamental category of a sub-pospace of $\mathbb{R}^2$, the two notions are in fact equivalent (see [17]).

When applied to the fundamental category of po-spaces, this means that the essential schedules in a concurrent system are separated out by the notion of component. Notice that this is true also for partial schedules and not just the maximal ones as in, e.g., [10] and [5].

5.2. GENERAL PROPERTIES OF SYSTEMS OF WEAKLY INVERTIBLE MORPHISMS

In this section, we state and prove essential properties of the partition into components of a small category $\mathcal{C}$ that a system Σ of weakly invertible morphisms induces. The most important case we have in mind is the fundamental category $\mathcal{C} = \vec{\pi}_1(X)$ of a po- or d-space X together with a maximal pure lr-system of weakly invertible morphisms in $\mathcal{C}$ from Section 4.4. Several properties are true for more general systems, and they will thus be stated with a minimal set of conditions.

First an easy formal consequence of the definitions:

LEMMA 7. *The weakly invertible morphisms in $\mathcal{C} \times \mathcal{D}$ are products of weakly invertible morphisms in $\mathcal{C}$ with weakly invertible morphisms in $\mathcal{D}$.*

PROPOSITION 2. *Let Σ consist of weakly invertible morphisms and let $s \in \Sigma(x, y)$. Then the maps*

$$\mathcal{C}(x,x) \xrightarrow{s_\#} \mathcal{C}(x,y) \xleftarrow{s^\#} \mathcal{C}(y,y)$$

are bijections.

*If, in particular, $\mathcal{C}(x,x) = \{Id_x\}$,** *then $\mathcal{C}(x,y) = \Sigma(x,y) = \{s\}$. In other words: The components in the component category $\pi_0(\mathcal{C};\Sigma)$ have unique endomorphisms.*

Proof. Immediate from definitions. □

PROPOSITION 3. *Let Σ denote an l-system of morphisms in $\mathcal{C}$.*

(1) *For every $f \in \mathcal{C}(x, y)$ and every $x' \sim_\Sigma x$ there exists $y' \sim_\Sigma y$ and $f' \in \mathcal{C}(x', y')$ such that $f' \sim_\Sigma f$.*

* This condition is always satisfied for the fundamental category of a po-space.

(2) *Let* $[f]_\Sigma \in \pi_0(\mathcal{C};\Sigma)([x]_\Sigma,[y]_\Sigma)$ *and let* $x' \in [x]_\Sigma$. *Then there exists* $y' \in [y]_\Sigma$ *and* $f' \in \mathcal{C}(x',y')$ *such that* $[f']_\Sigma = [f]_\Sigma$.

Statement (2) should be interpreted as a lifting property, lifting morphisms from the component category $\pi_0(\mathcal{C};\Sigma)$ to the original category $\mathcal{C}$.

Proof. Immediate from the definition of an l-system. □

There is of course an analogous statement for liftings in r-families of morphisms.

PROPOSITION 4. *Let* Σ *denote a* pure l-*system of weakly invertible morphisms. Let* $\mathcal{C}(x,x)=\{Id_x\}$ *and let* $x \sim_\Sigma y$. *If*

$$x \xrightarrow{f} z \xrightarrow{g} y,$$

then $f, g \in \Sigma$ *and* $z \sim_\Sigma x$.

Proof. Since $x \sim_\Sigma y$ with Σ an l-system, there exist morphisms $\sigma \in \Sigma(x,u)$, $\tau \in \Sigma(y,u)$ in the diagram

$$\begin{array}{ccccc} x & \xrightarrow{f} & z & \xrightarrow{g} & y \\ & \searrow^{\sigma} & & \swarrow_{\tau} & \\ & & u & & \end{array}.$$

By Proposition 2, these are the only morphisms beween x and u, resp. y and u. In particular, $\tau \circ g \circ f = \sigma$. Since σ is pure, we conclude that $f, g \in \Sigma$. □

Again, there is an analogous property (with an analogous proof) for pure r-families of weakly invertible morphisms.

The result can be understood as a "diconvexity" property of the components. Here is a negative formulation of the result: If $x \not\sim_\Sigma z$ and $\mathcal{C}(z,y) \neq \emptyset$, then $x \not\sim_\Sigma y$. You never return to a component that you have left.

PROPOSITION 5. *Let* Σ *denote a* pure l-*system of morphisms in* $\mathcal{C}$, *let* $\sigma, \tau \in \Sigma(x,-)$. *There exists a solution of the extension problem*

$$\begin{array}{ccc} \cdot & \overset{\sigma'}{\dashrightarrow} & \cdot \\ {\scriptstyle\tau}\uparrow & & \uparrow{\scriptstyle\tau'} \\ x & \xrightarrow[\sigma]{} & \cdot \end{array}$$

with both morphisms $\sigma', \tau' \in \Sigma$.

Proof. From the extension condition we get $\sigma' \in \Sigma$ and $\tau' \in \mathcal{C}$ such that $\tau' \circ \sigma = \sigma' \circ \tau \in \Sigma$. Since Σ is pure, τ' has to be in Σ, as well. □

Again, there is an analogous statement for extensions with respect to pure r-families of morphisms.

COROLLARY 1. *Let* Σ *denote a* pure l-*system* (*resp. r-system*) *of morphisms in* $\mathcal{C}$. *Every morphism in the subcategory generated by* Σ *in* $\mathcal{C}[\Sigma]^{-1}$ *can be represented in the form* $\sigma_1^{-1} \circ \sigma_2$ (*resp.* $\sigma_1 \circ \sigma_2^{-1}$), $\sigma_i \in \Sigma,\ 1 \leq i \leq 2$.

Proof. The same proof as for the expression of general morphisms using the Σ-extension property from Proposition 5. □

For a pure system Σ, the lifing property from Proposition 3 can be sharpened:

PROPOSITION 7. *Let* Σ *be a* pure l-*system* (*or pure r-system*) *within* $\mathcal{C}$. *Let* $C_1, C_2 \subset Ob(\mathcal{C})$ *denote two components such that the set of morphisms* (*in* $\pi_0(\mathcal{C}; \Sigma)$) *is* finite. *Then, for every* $x_1 \in C_1$ *there exists* $x_2 \in C_2$ *such that the quotient map*

$$\mathcal{C}(x_1, x_2) \to \pi_0(\mathcal{C}; \Sigma)(C_1, C_2),\ f \mapsto [f]$$

is onto. If Σ *is a pure lr-system of weakly invertible morphisms with*

$$\mathcal{C}(x, x) = \{Id_x\}$$

for all $x \in Ob(\mathcal{C})$*, the quotient map is even a bijection.*

Proof. By repeated application of Proposition 3, all n morphisms from C_1 to C_2 can be lifted to morphisms $f_1, \ldots, f_n$ with source in x_1 and targets in $y_1, \ldots, y_n \in C_2$. By repeated application of Corollary 1, there exist morphisms $\sigma_i \in \Sigma$ from y_i to the *same* target $x_2 \in C_2$, $1 \leq i \leq n$. The quotient map is onto, since $\sigma_i \circ f_i \simeq f_i$, $1 \leq i \leq n$.

To prove injectivity, assume $f_i \in \mathcal{C}(x_1, x_2)$ with

$$[f_1] = [f_2] \in \pi_0(\mathcal{C}; \Sigma)(C_1, C_2).$$

Then, there exist $x_0 \in C_1, x_3 \in C_2$ and morphisms $\sigma_i \in \Sigma(x_0, x_1)$, $\tau_i \in \Sigma(x_2, x_3)$, $1 \leq i \leq 2$, such that $\tau_1 \circ f_1 \circ \sigma_1 = \tau_2 \circ f_2 \circ \sigma_2 \in \mathcal{C}(x_0, x_3)$. By Proposition 2, $\sigma_1 = \sigma_2$ and $\tau_1 = \tau_2$. Since σ_1 and τ_1 are weakly invertible, we conclude: $f_1 = f_2 \in \mathcal{C}(x_1, x_2)$. □

Another useful interpretation is as follows: the components are the appropriate counterparts of connected components in the classical case. As a matter of fact, dihomotopy classes can be read if we use a bipointing, as in, e.g., [11], and not just a base point as in the classical case. Ideally we would like to be able to change the pair of base points chosen, without changing the classes of dipaths between these points, as long as they stay in the same components *and* as long as they are consistent: the "end" base point should be reachable (by a dipath) from the "start" end point. Proposition 7 does not quite give this for a general pure lr-system, but we suggest that one should look for extra requirements which make it hold – for instance being in an appropriate subcategory of po-spaces. In the 2-dimensional mutual exclusion models, cf. [17] and [10], it is certainly true. The set of components should be called the set of diconnected components (a $\vec{\pi}_0$ in some sense) and the algebraic structure of dipaths between diconnected components is really the counterpart of the fundamental group (or groupoid). In fact, in the classical case of the fundamental groupoid π_1, π_0 can be read from it as its set of connected components (see, e.g., [6]).

Another application of Proposition 5 shows that any component can possess at most one maximal, resp. minimal object:

DEFINITION 7. Let $D \subseteq Ob(\mathcal{C})$. An object $m \in D$ is a *minimal* element in D, if $\mathcal{C}(x, m) \neq \emptyset \Rightarrow x = m$ or $x \notin D$. A maximal element is defined similarly.

COROLLARY 2. *Let* Σ *be a* pure l-*system* (*or pure r-system*) *within* $\mathcal{C}$. *Every component with respect to* Σ *can at most have one maximal and one minimal element.*

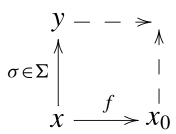
REMARK 1. From easy geometric examples as Example 5, we know that a component in general need not possess a minimal or maximal element. Question: Is there always an infimum (supremum) for every component? Are those unique? We conjecture that some of the results of [14] could be useful for proving this.

In the presence of maximal or minimal elements for the objects of the whole category, several nice properties can be proved to hold without the pureness assumption. The first lemmas are relevant for the unsafe regions in deadlock analysis (cf. [4]):

DEFINITION 8. For $x \in Ob(\mathcal{C})$, let $x^{\rightarrow}$, resp. $x^{\leftarrow}$ denote the set of maximal (minimal) elements $x_0, y_0 \in Ob(\mathcal{C})$ with $\mathcal{C}(x, x_0) \neq \emptyset$ ($\mathcal{C}(y_0, x) \neq \emptyset$), i.e., the maximal elements reachable from x (minimal elements that can reach x).

LEMMA 8. *Let* Σ *denote an r-system of morphisms in* $\mathcal{C}$, *and let* $\Sigma(x, y) \neq \emptyset$. *Then* $x^{\rightarrow} = y^{\rightarrow}$. *If* Σ *is an l-system, then* $x^{\leftarrow} = y^{\leftarrow}$.

Proof. Consider an extension problem



with x_0 a maximal element. The right vertical arrow has to be an endomorphism of x_0, and hence x_0 can be reached from y. □

LEMMA 9. *Let* Σ *denote an l-system of morphisms in* $\mathcal{C}$ *and let* x_0 *denote a maximal element of* $Ob(\mathcal{C})$. *If* $\Sigma(x, x_0) \neq \emptyset$ *and* $\mathcal{C}(x, y) \neq \emptyset$, *then* $\Sigma(y, x_0) \neq \emptyset$.

In other words: If x and x_0 are Σ-equivalent, then every y reachable from x is Σ-equivalent to x.

Proof. Consider an extension problem

$$\begin{array}{ccc} y & \dashrightarrow & \\ {\scriptstyle f}\uparrow & & \uparrow \\ x & \xrightarrow{\sigma\in\Sigma} & x_0. \end{array}$$

Again, the right vertical arrow has to be an endomorphism of x_0, and hence there is a Σ-morphism from y to x_0. □

PROPOSITION 8. *Let Σ denote an l-system of weakly invertible morphisms in $\mathcal{C}$ and let x_0 denote a maximal element of $Ob(\mathcal{C})$ with $\mathcal{C}(x_0, x_0) = \{Id_{x_0}\}$. If $\Sigma(x, x_0) \neq \emptyset$ and $\mathcal{C}(x, y)$, $\mathcal{C}(y, z)$, $\mathcal{C}(y, x_0)$ are all non-empty, then all these sets consist of a single element.*

In other words: If x and x_0 are Σ-equivalent, the category is trivial between x and x_0 (no non-trivial dihomotopy between x and x_0 in the fundamental category).

Proof. Any composition $x \stackrel{f}{\to} y \stackrel{g}{\to} z \stackrel{h}{\to} x_0$ is equal to the unique morphism $\sigma \in \Sigma(x, x_0)$ by Proposition 2. By Lemma 9, $\Sigma(y, x_0) \neq \emptyset \neq \Sigma(z, x_0)$, and by Proposition 2, $\mathcal{C}(z, x_0) = \Sigma(z, x_0) = \{h\}$, $\mathcal{C}(y, x_0) = \Sigma(y, x_0) = \{h \circ g\}$. Since $h \circ g$ is weakly invertible, f is the only element of $\mathcal{C}(x, y)$; since h is weakly invertible, g is the only element of $\mathcal{C}(y, z)$. □

5.3. TOPOLOGICAL PROPERTIES

More precise information on components relies on topological properties on top of the categorical ones. From now on we investigate a *topological category* $\mathcal{C}$, i.e., the objects $Ob(\mathcal{C})$ form a *topological space* X. Additionally, a system Σ of morphisms in $\mathcal{C}$ is given.

DEFINITION 9. (1) Let U denote an open set in X and $x, y \in U$. A morphism $f \in \mathcal{C}(x, y)$ is called a *U-morphism* if

$$f = f_2 \circ f_1,\ f_1 \in \mathcal{C}(x, z),\ f_2 \in \mathcal{C}(z, y) \Rightarrow z \in U.$$

The set of all such U-morphisms from x to y will be denoted $\mathcal{C}_U(x, y)$.

(2) An open set $U \subset X$ is called Σ-simple if

(a) $x, y \in U \Rightarrow |\mathcal{C}_U(x, y)| \leq 1$.
(b) For all $x \sim_\Sigma y \in U$ there exists $z \in U$ such that $\mathcal{C}_U(x, z \cap \Sigma(x, z) \neq \emptyset \neq \mathcal{C}_U(y, z \cap \Sigma(y, z)$.

DEFINITION 10. Two Σ-components C_1 and C_2 are called *neighbours* if there are

(1) a morphism with source in C_1 and target in C_2 and
(2) a Σ-simple open set U containing an element $p \in C_1 \cap \partial C_2$ such that every $g \in \mathcal{C}(p, x_2)$, $x_2 \in C_2$ decomposes as $g = g_2 \circ g_1$ with $g_1 \in \mathcal{C}_U(p, z)$ and $z \in C_2$ (or the symmetric condition with $p \in \partial C_1 \cap C_2$).

PROPOSITION 9. *Let* $(\mathcal{C}, \Sigma)$ *denote a category with an lr-system of weakly invertible morphisms. Let* C_1 *and* C_2 *denote two neighbouring* Σ*-components such that* $\mathcal{C}(x_i, x_i) = \{Id_{x_i}\}$ *for some* $x_i \in C_i$*. Then,* $Mor(C_1, C_2)$ *has exactly one element in the component category* $\pi_0(\mathcal{C}; \Sigma)$.

Proof. Choose an element $p \in C_1 \cap \partial C_2$ as in Definition 10(2). By Proposition 3, every morphism from C_1 to C_2 is equivalent to one with source p; by Definition 10(1), there exists such a morphism. By assumption in Definition 10(2), every such morphism decomposes as $s \circ f$ with f a U-morphism and s a Σ-morphism within C_2 (use Proposition 2) and is hence equivalent to the U-morphism f.

Consider two U-morphisms f, f' with source p. By Definition 9, there are U-morphisms $s, s' \in \Sigma$ such that $s \circ f = s' \circ f'$. Thus $f \simeq s \circ f = s' \circ f' \simeq f'$. □

The result allows to interpret the component category as a directed graph (rather than a multigraph) with relations. If, in particular, every morphism decomposes into morphisms between neighbour components (as for the fundamental category of a po-space), one may use the classes of these unique morphisms between neighbour components as *generators* for the component category.

6. Examples

In the case of Figure 9, the *only* left and right calculus of fractions included in the weakly invertible morphisms is easily shown to consist of the identities only. As a consequence, the category of components with respect to this greatest lr-system of weakly invertible morphisms in this case is isomorphic to the original category!

The po-spaces arising from 2-*dimensional mutual exclusion models*, i.e., a square, from which a number of isothetic rectangles (with edges parallel to the square) have been deleted (as the forbidden region), are handled completely in [17] and [10]: A system of morphisms is a pure lr-system of weakly invertible morphisms if no such morphism crosses a system of line segments emerging from (certain of) the minima, resp. maxima of the rectangles that constitute the forbidden region.

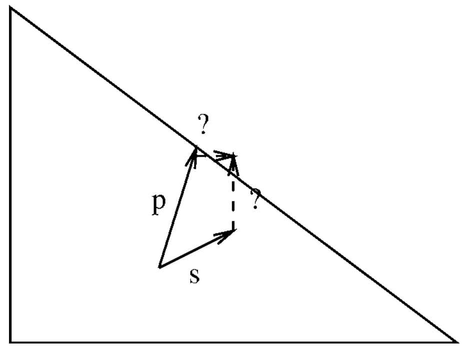


Figure 9. A po-space with no nontrivial weakly invertibles.

6.1. THE SURFACE OF A 3-CUBE

Now for a more intricate example, treated in full details. The faces of the 3-cube C, or equivalently, the 3-cube minus an interior 3-cube, has 26 components. Points on the faces of the 3-cube are $\{(x, y, z) \in \vec{I}^3 \mid \{x, y, z\} \cap \{0, 1\} \neq \emptyset\}$. Let $\mathcal{C} = \vec{\pi}_1(C)$. Observe that

- There are two elements in $\mathcal{C}((0, 0, a), (1, 1, a))$, $\mathcal{C}((0, a, 0), (1, a, 1)))$ and $\mathcal{C}((a, 0, 0), (a, 1, 1))$ when $a \notin \{0, 1\}$. For instance the composition of arrows from $(0, 0, a)$ to $(0, 1, a)$ and then to $(1, 1, a)$ is different from going from $(0, 0, a)$ to $(1, 0, a)$ and then to $(1, 1, a)$. They are different, since all dipaths from $(0, 0, a)$ to $(1, 1, a)$ has the third coordinate $a \in]0, 1[$, and since the interior of the cube is missing.
- Between other pairs of points, there are at most one morphism.
- All morphisms $\alpha : (x_1, x_2, x_3) \to (y_1, y_2, y_3)$ such that $x_i = 0 \Rightarrow y_i = 0$ and $x_i = 1 \Rightarrow y_i = 1$ are weakly invertible. This is easy to see by the geometry of the cube – the future and past of (x_1, x_2, x_3) and (y_1, y_2, y_3) have the same geometry.

By the last property, we can restrict attention to the 26 classes of points represented by $(x, y, z) \in \{0, -, 1\}^3 \setminus (-, -, -)$ where the coordinate $-$ just means an interior point of $]0, 1[$. We will see, that none of the morphisms between these points are in the system Σ. We will omit the commas and write $(0 - 1)$ for $(0, -, 1)$.

6.1.1. *The Weakly Invertible Morphisms*

We will find the arrows which are *not* weakly invertible. Since $\mathcal{C}((00-), (11-))$ has 2 elements and $\mathcal{C}((00-), (1--))$ has one element, the arrow $(1--) \to (11-)$ is not weakly invertible. Similarly, there is only one element in $\mathcal{C}((00-), (111))$, so the arrow $(11-) \to (111)$ is not weakly invertible. Hence (the lack of) weak invertibility implies that no arrow from an upper face $(1--)$, $(-1-)$ or $(--1)$ to an upper edge $(11-)$, $(1-1)$ or (-11) is in Σ, and similarly for all maps from lower edges to lower faces. Lack of weak invertibility also implies that maps from upper edges to $(1, 1, 1)$ or from $(0, 0, 0)$ to lower edges are not in Σ.

Similarly, $\Sigma((xy-), (11-)) = \emptyset$, when $xy \neq 11$ and $\Sigma((00-), (xy-)) = \emptyset$, when $xy \neq 00$ since there are no weakly invertible morphisms. Permuting the coordinates gives 24 other instances of this.

Notice that the system of weakly invertible morphisms is not pure.

6.1.2. *The Maximal lr-System in the Weakly Invertible Morphisms*

1. We study maps from any $(abc) \neq (111)$, which is not an upper edge, to (111). These are weakly invertible, but they are not in Σ: Suppose $s : (abc) \to (111)$ is in Σ and suppose $c \neq 1$. Let $f : (abc) \to (11-)$. Then the lr-property implies that we can complete the diagram

$$\begin{array}{ccc} (111) & \overset{g}{\dashrightarrow} & (xyz) \\ {\scriptstyle s}\uparrow & & \uparrow{\scriptstyle \sigma} \\ (abc) & \xrightarrow{f} & (11-) \end{array}$$

with $\sigma \in \Sigma$. Since (xyz) has to be (111), g is the identity.
But $\Sigma((11-), (111)) = \emptyset$, so the diagram cannot be completed with $\sigma \in \Sigma$. Hence $\Sigma((abc), (111)) = \emptyset$ when $(abc) \neq (111)$ and similarly for maps from (000). Notice that this is a concrete example of Proposition 8.
2. Any morphism $s : (ab0) \to (11-)$ is weakly invertible, since $(ab0)$ is not reachable from $(00-)$. But suppose $s \in \Sigma$. Then the lr-property is violated: Let $f : (00-) \to (11-)$ be one of the two morphisms. Then the lr-property says that there are maps $\sigma \in \Sigma$ and g completing the diagram:

$$\begin{array}{ccc} (00-) & \xrightarrow{f} & (11-) \\ {\scriptstyle \sigma}\uparrow & & \uparrow{\scriptstyle s} \\ (xyz) & \overset{g}{\dashrightarrow} & (ab0) \end{array}$$

but since (xyz) is below $(00-)$ and $(ab0)$, we conclude that $(xyz) = (000)$ and we know that $\Sigma((000), (00-)) = \emptyset$. Hence $s \notin \Sigma$. Together with what we found in 1), we have seen that no map to an upper edge, and symmetrically, no map from a lower edge is in Σ.
3. Now for maps from and to intermediate edges: Suppose first, we map to an upper face. The morphism $t : (1-0) \to (1--)$ is weakly invertible. If it is in Σ, then we can complete the diagram

$$\begin{array}{ccc} (1-0) & \xrightarrow{t} & (1--) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ (110) & \overset{\sigma}{\dashrightarrow} & (xyz) \end{array}$$

with $\sigma \in \Sigma$. But then $(xyz) \in \{(11-), (111)\}$ and we know there is no such σ. So $t \notin \Sigma$. The only other option, which is not already covered above, is to map to an intermediate vertex $s : (1-0) \to (110)$; use the diagram above – now assuming $f = s \in \Sigma$. Symmetrically, no map from or to an intermediate edge is in Σ.

4. Maps from and to faces: Maps *from* an upper face or *to* a lower face are covered above. Now suppose $s : (abc) \to (1--)$ is in Σ, and suppose (abc) is not an edge – these are covered above. Then suppose

   $(abc) < (10-)$ (else $(abc) < (01-)$,

   so this case is similar). Let $f : (abc) \to (10-)$ and do the diagram. There are no (nontrivial) morphisms from $(10-)$, so $s \notin \Sigma$. The other cases follow in a similar way.
5. The last case we have to check is maps between intermediate vertices, since maps to and from all other types is covered above: Let $s : (100) \to (110)$. If $s \in \Sigma$, the diagram with $f : (100) \to (1-0)$ should have a completion with $\sigma : (1-0) \to (xyz) \neq (1-0)$ and there are no such morphisms from an intermediate edge.

Hence, in this case, the biggest lr-system in the weakly invertible morphisms is in fact pure, since the morphisms between the 26 types of points are all in the complement of Σ.

7. Conclusion and Future Work

We hope to achieve an effective calculation of the component categories of the fundamental category of reasonable po-spaces by applying Marco Grandis' directed version [11] of a van Kampen theorem for directed spaces. More precisely, let $X = X_1 \cup X_2$ and let Σ_0, Σ_1 and Σ_2 denote admissible (lr, pure lr) systems of weakly invertible morphisms in the fundamental categories $\vec{\pi}_1(X_1 \cap X_2), \vec{\pi}_1(X_1)$ and $\vec{\pi}_1(X_2)$. The task is to derive an admissible system Σ_{12} of weakly invertible morphisms in $\vec{\pi}_1(X)$ – and thus derive a suitable component category for the union. For the time being, we can only state the following conjecture:

Let $i_1 : X_1 \cap X_2 \to X_1$ and $i_2 : X_1 \cap X_2 \to X_2$ be the canonical inclusion morphisms (respectively

$$i_1^* : \vec{\pi}_1(X_1 \cap X_2) \to \vec{\pi}_1(X_1) \text{ and } i_2^* : \vec{\pi}_1(X_1 \cap X_2) \to \vec{\pi}_1(X_2)$$

the induced functors between the corresponding fundamental categories).

Our claim is:

CONJECTURE 1 (van Kampen on components). *The greatest left and right calculus of fractions* (*cf.* *Lemma* 5) *in the pushout* Δ *of* Σ_1 *and* Σ_2 *above* $I_{12} := i_1^{*-1}(\Sigma_1) \cap i_2^{*-1}(\Sigma_2) \cap \Sigma_0$ *as below:*⋆

⋆ The induced functors from i_1^* and i_2^* on the invertible morphisms, still denoted the same way, are the ones of Lemma 1.

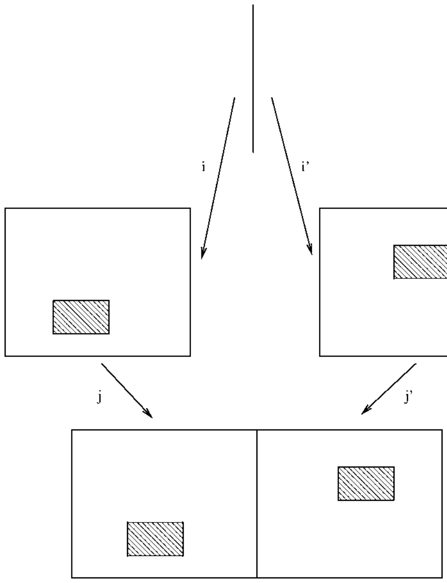


Figure 10. A pushout of two po-spaces.

$$\begin{array}{ccc} I_{12} & \xrightarrow{i_1^*} & \Sigma_1 \\ \downarrow{\scriptstyle i_2^*} & & \downarrow \\ \Sigma_2 & \longrightarrow & \Delta \end{array}$$

is denoted $\bar{\Sigma}_{12}$*. Then the system*

$$\Sigma_{12} = \{s \in \vec{\pi}_1(X) \mid P_{\bar{\Sigma}_{12}}(s) \in ZZ(\bar{\Sigma}_{12})\},$$

containing all morphisms that are identified with zig-zag morphisms in the category of fractions with respect to $\bar{\Sigma}_{12}$*, is the admissible system describing the "inessential" morphisms of the fundamental category* $\vec{\pi}_1(X)$*.*

As an example of this conjectural van Kampen theorem, consider the situation of Figure 10 with two copies of Figure 5 glued together along a common boundary. Figure 10 shows the corresponding pushout diagram of po-spaces X_1 and X_2.

The left part of Figure 11 shows the union of the components of X_1 and X_2. Extension properties imposed by the property to be left and right systems imply that some of the inessential morphisms should *no longer* be considered as inessential in the union of the two spaces. The greatest left and right system (which is pure) is shown in the right part of Figure 11.

As a second example, consider a rectangle X as the union of X_1, a rectangle without an inner square (Figure 5), and X_2 filling in that inner square (with a collar). The intersection $X_1 \cap X_2$ is dihomeomorphic to X_1. This example shows

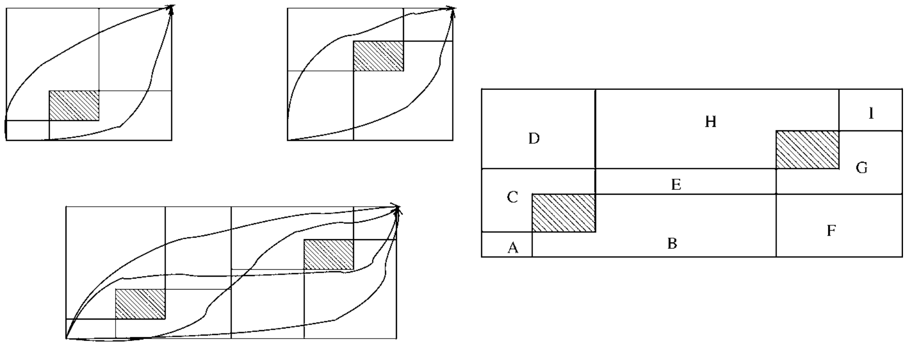


Figure 11. The inessential morphims after pushout.

that it is necessary to "complete" $\bar{\Sigma}_{12}$ in the category of fractions to arrive at the result $\Sigma_{12} = \pi_1(X)$.

The system Σ_{12} in the conjecture is an lr-system almost by definition. Probably, one needs additional assumptions (e.g., X_1 "below" X_2 or vice versa) to make sure that it consists of weakly invertible morphisms and/or satisfies pureness.

For application purposes, we would like to exploit the van Kampen conjecture to arrive at a geometrically based algorithm detecting the components in a mutual exclusion model (cf. Section 1) from a description of the forbidden region, as a generalisation of our algorithm detecting deadlocks, unsafe and unreachable regions [4].

Last but not least, we believe that our construction based on a category of fractions of the fundamental category of a po-space has close connections to some kind of universal covering of the fundamental category of a pospace (as defined in, e.g., [12]). In fact, the category of components enjoys a certain lifting property. The set of weakly-invertibles itself is defined through suitable representations of the fundamental category: they are well-known to be in one to one correspondence with coverings of that category [12].

Acknowledgements

We would like to thank Marco Grandis for his advice and encouragement on behalf of this paper, and the anonymous referee for helpful comments.

References

1. Borceux, F.: *Handbook of Categorical Algebra 1: Basic Category Theory*, Cambridge University Press, 1994.
2. Brown, R.: *Elements of Modern Topology*, McGraw-Hill, Maidenhead, 1968.
3. Dijkstra, E.: *Cooperating Sequential Processes*, Academic Press, New York, 1968.
4. Fajstrup, L., Goubault, E. and Raussen, M.: Detecting deadlocks in concurrent systems, in *Proceedings of the 9th International Conference on Concurrency Theory*, Springer-Verlag, 1998, pp. 332–447, also available at http://www.dmi.ens.fr/~goubault.

5. Fajstrup, L., Goubault, E. and Raussen, M.: Algebraic topology and concurrency, Submitted to *Theoretical Computer Science*, also Technical Report, Aalborg University, 1999.
6. Gabriel, P. and Zisman, M.: *Calculus of Fractions and Homotopy Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete 35, Springer-Verlag, 1967.
7. Gaucher, P.: A convenient category for the homotopy theory of concurrency, Preprint, 2002, available at math.AT/0201252.
8. Gaucher, P. and Goubault, E.: Topological deformation of higher dimensional automata, *Homotopy Homology Appl.* **5**(2) (2003), 39–82.
9. Goubault, E.: Some geometric perspectives in concurrency theory, *Homology Homotopy Appl.* **5**(2) (2003), 95–136.
10. Goubault, E. and Raussen, M.: Dihomotopy as a tool in state space analysis, in S. Rajsbaum (ed.), *LATIN 2002: Theoretical Informatics* (Cancun, Mexico), Lecture Notes in Computer Science 2286, Springer-Verlag, 2002, pp. 16–37.
11. Grandis, M.: Directed homotopy theory, I. The fundamental category, *Cahiers Topologie Géom. Différentielle Catég.* **44**, Preliminary version: Dip. Mat. Univ. Genova, Preprint 443, 2003.
12. Higgins, P. J.: *Categories and Groupoids*, Van Nostrand Reinhold Mathematical Series 32, 1971.
13. Johnstone, P. T.: *Stone Spaces*, Cambridge University Press, 1982.
14. Kelly, G. M., Lack, S. and Walters, R. F. C.: Coinverters and categories of fractions for categories with structure, *Appl. Categ. Structures* (1993).
15. Mac Lane, S.: *Categories for the Working Mathematician*, Springer-Verlag, 1971.
16. Nachbin, L.: *Topology and Order*, Van Nostrand, Princeton, 1965.
17. Raussen, M.: On the classification of dipaths in geometric models for concurrency, *Math. Structures Comput. Sci.* **10**(4) (2000), 427–457.
18. Raussen, M.: State spaces and dipaths up to dihomotopy, *Homology Homotopy Appl.* **5**(2) (2003), 257–280.

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Theoretical Computer Science 357 (2006) 241–278

Theoretical Computer Science

www.elsevier.com/locate/tcs

Algebraic topology and concurrency

Lisbeth Fajstrup[a,*], Martin Raußen[a], Eric Goubault[b]

[a]*Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg Ø, Denmark*
[b]*LIST (CEA-Technologies Avancées), DTSI-SLA, CEA F91191 Gif-sur-Yvette Cedex, France*

Abstract

We show in this article that some concepts from homotopy theory, in algebraic topology, are relevant for studying concurrent programs. We exhibit a natural semantics of semaphore programs, based on partially ordered topological spaces, which are studied up to "elastic deformation" or homotopy, giving information about important properties of the program, such as deadlocks, unreachables, serializability, essential schedules, etc. In fact, it is not quite ordinary homotopy that has to be used, but rather a "directed homotopy" that does not reverse the flow of time. We show some of the essential differences between ordinary and directed homotopy through examples. We also relate the topological view to a combinatorial view of concurrent programs closer to transition systems, through the notion of a cubical set. Finally we apply some of these concepts to the proof of the safeness of a two-phase protocol, well-known and used in concurrent database theory. We end up with a list of problems from both a mathematical and a computer-scientific point of view.

Keywords: Homotopy; Concurrency; Cubical set; Higher-dimensional automata; Partial order; Partially ordered space

1. Introduction

This article is intended to provide some new insights about concurrency theory using ideas from geometry, and more specifically from algebraic topology. The aim of the paper is twofold: we justify applications of geometrical methods in concurrency through some chosen examples and we give the mathematical foundations needed to understand the geometric phenomenon that we identify. In particular we show that the usual notion of homotopy has to be refined to take into account a partial ordering describing the time flow. This gives rise to new interesting mathematical problems and it provides common grounds to computer-scientific problems that have not been precisely related otherwise in the past.

The organization of the paper is as follows.

We first explain the interest of using geometric ideas for semantical reasons, in Section 2, pointing to some examples that will be given throughout the paper, and in Section 10.

In Section 3, we give the first few definitions needed for modelling the topological spaces already arising in Section 2. Basically, we need to define a topological space containing all traces of executions of the concurrent systems we want to

* Corresponding author.
E-mail addresses: fajstrup@math.aau.dk (L. Fajstrup), raussen@math.aau.dk (M. Raußen), Eric.Goubault@cea.fr (E. Goubault).

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2006.03.022

characterize plus the information about how time flows. This is the main difference from standard topological reasoning in which there is no information about relation "in time" among points.

The central notion here is that of a local po-space, which is a topological space with a local partial order of time on it. Some examples are given. In Section 6, it will be pointed out that cubical sets (or higher-dimensional automata, [34,65]) give rise to such spaces in a natural way, hence most "combinatorial" concurrency models are instances of these local po-spaces. It is worth noting that some models in general relativity [64] consider timed spaces, and the authors were inspired by these physical concepts when developing the theory.

Section 4 gives the first definitions of the new homotopy theory we need in order to define equivalence of paths in accordance with the intuitive explanations from Section 2. A central notion here is that of homotopy history components, which contains the relevant information for computer-scientific applications, as well as for classification of local po-spaces modulo "directed" homotopy. We give examples that show that this directed homotopy is finer than usual homotopy in the sense that it can distinguish homotopy equivalent (in the standard sense) topological spaces.

We study in Section 5 a particular subcategory of local po-spaces: those which are locally Euclidean, i.e. the local partial order is inherited from $\mathbb{R}^n$ (for some n). A central statement is that we can take "still pictures" of the dynamics on such spaces, i.e. look at cuts which contain points not related through time, and this can give obstructions to deformation in the directed sense.

We carry on in Section 6 by investigating cubical sets (or higher-dimensional automata, HDA) and show that they are in some sense a combinatorial counterpart of these local po-spaces (at least, of a large subcategory). We refer the reader to [34] or to the more recent [21] for actual semantics of some concurrent systems using these cubical sets. A "combinatorial" deformation theory in cubical sets is developed and related to the directed homotopy in the continuous case, using in particular the notion of subdivision, in Section 6.8.

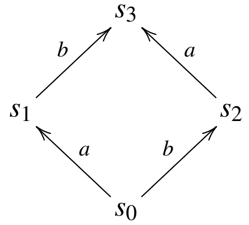
A major application, serializability in concurrent databases, is fully treated in Section 7. It is an application of the preceding theory and a refinement and extension of the result in [48]. There are also similar problems in scheduling microinstructions in chips, as will be briefly explained in Section 10.

Some mathematical directions are given in Section 8, and related computer-scientific perspectives are listed in Section 9. Finally we refer the reader to related work where algebraic topology is at the centre of computer-scientific modelling and proofs, and to application areas which should be examined in detail in future work, in Section 10.

Part of this was presented by two of the co-authors at the 14th conference on the Mathematical Foundations of Programming Semantics (London, May 1998). It has taken a long time to assemble this full version of our work, and the development has gone on since then. This is why we have to give some forward references, some of them based on earlier versions of this paper, both as footnotes and in the concluding sections 8–10.

2. Motivation and examples of applications

Without the ambition to be complete, we can trace back the use of geometrical models and properties to the beginning of theoretical computer science, in the use of graph theory, or of partial orders to describe the semantics of systems. For instance, sequential machines can be studied by examining their operational behaviours—that is by looking at their state transition graphs. One of the fundamental properties that we might want to study is confluence of the performed computation. This is obviously a property of a highly geometric nature: we must be able to complete all non-deterministic applications of conflicting reductions by some other reductions that all converge to the same result; i.e. we must have diamond shapes in the state transition graphs describing the sequences of operations of our sequential machines. For concurrent machines, the geometric properties of computation include those of sequential machines but they are even more intricate. Purely (interference free) asynchronous executions of two processes are confluent and therefore recognizable geometrically as diamonds (or squares). For example, the operational semantics of the interference free parallel composition of two actions a and b is as in the figure below:



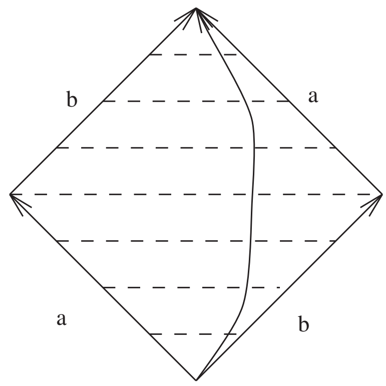


Fig. 1. New traces.

Let us take a closer look at the geometry of transition systems used in concurrency. We only have to think of a concurrent execution of two actions *a* and *b* on two processors P_1 and P_2 as a curve in $\mathbb{R}^2$ whose points have as abscissa (respectively, ordinate) the local time of P_1 taken to execute *a* (respectively, the local time of P_2 taken to execute *b*). This gives new traces, other than just the interleavings, as in Fig. 1, which are all increasing paths in the two coordinates (because we cannot invert the time flow) included in the square delineated by the interleavings of *a* and *b*.

Here we are confronted with two presentations of essentially the same phenomenon. The first one is the geometry of continuous paths (like in the study of mechanics). The second is a discretization of it, which in general comes first in the semantics applications (but not in others; see Section 9). Basically, abstract all paths inside an *n*-dimensional cube by the interior of the cube itself, then describe the geometry of executions as the amalgamation (or pasting) of all the different *k*-dimensional cubes entering into play (for describing, in a similar way as above, *k* actions executing asynchronously), as shown in Fig. 10. This is precisely what is called a cubical set in combinatorial algebraic topology (see Section 6) or a HDA [65] in computer science. But here, we are considering cubical sets with an orientation given by the time flow (by orienting the segments constituting it), whereas in ordinary algebraic topology, shapes have no preferred order. So the continuous counterpart of such a discretization is more than a topological space, it contains also (partial) order relations. This is developed in Section 3 under the name of po-spaces (and lpo-spaces) whose formalization and understanding is the main objective of this article.

The reader might wonder why we go from a typically discrete world (transition systems for instance) to a continuous world (topological subspaces of $\mathbb{R}^n$ for instance), especially when trying to tackle the state-space explosion problem! In some situations, going back and forth between discrete and continuous helps a lot: continuous situations are much easier to "deform and retract elastically" whereas it is combinatorially very demanding and less intuitive to do the same in a discrete world. On the converse, discrete situations make some arguments simpler, that otherwise have to be abstracted by hard topological properties. Even when the continuous image is questionable, since there is no immediate interpretation of infinitesimals for instance, it can help a lot, as in the case of quantum mechanics. Moreover, continuous models have a better chance of giving models to real-time (and hybrid) parallel systems, for this see for instance [36] (slightly outdated) or better, [17].

Last but not least, this way of looking at concurrent processes relies on well-established theories, which we have to transform, but still, quite a lot of the technical difficulties are handled by constructions in these theories (homotopy and homology theory). Certainly, the geometric intuition helped a lot in understanding essential phenomena of concurrency theory: for instance, there are new examples/counter-examples (Examples 4.4 (4) or 5.11) that could not have been obtained without these geometric ideas.

3. Partially ordered spaces and local po-spaces

The geometric model which is already implicitly used in the pictures in Section 2 is a topological space with a (time) direction. For codes without loops, the right model is a po-space, a topological space with a partial order Definition 3.1. When there are loops involved, we cannot get a global partial order: the notion of clockwise order on the circle is not a partial order—all points are both before and after all other points. Hence we need a *locally partially ordered space* or an *lpo-space*, cf. Definition 3.4.

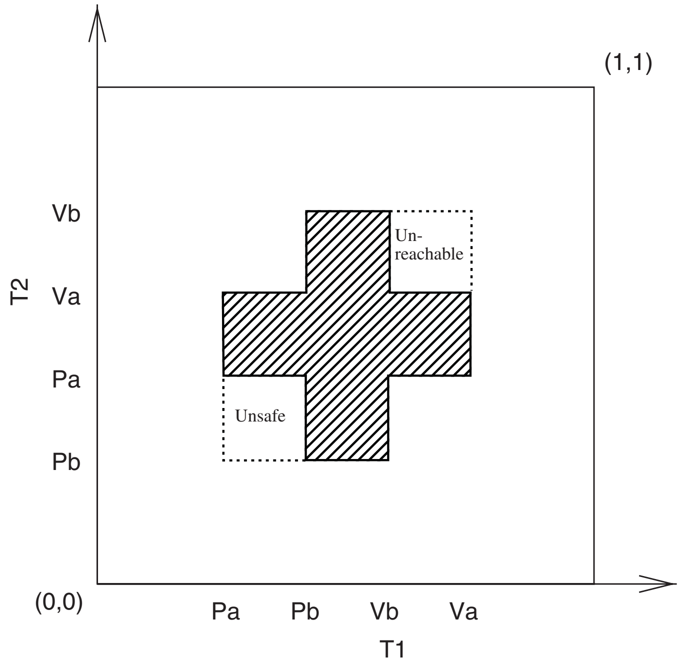


Fig. 2. The Swiss flag example—two processes sharing two resources.

In this geometric model, we can now give geometric characterizations of an *execution*, a *deadlock*, an *unreachable point* etc. As a class of examples throughout this text, we use a toy language manipulating semaphores. Using Dijkstra's notation [15]. Processes are sequences of locking operations *Pa* on semaphores *a* and unlocking operations *Va*. For the time being we consider only binary semaphores, ensuring mutual exclusion of accesses, but in further examples, we will also model and use counting semaphores, or *k*-semaphores ($k > 1$) which can be accessed concurrently by up to *k* processes.

In the example where two processes share two resources *a* and *b*:

$$T1 = Pa.Pb.Vb.Va,$$

$$T2 = Pb.Pa.Va.Vb$$

the geometric model is the "Swiss flag", Fig. 2. This is a subset of $\mathbb{R}^2$ with partial order $(x_1, y_1) \leqslant (x_2, y_2)$ if $x_1 \leqslant y_1$ and $x_2 \leqslant y_2$. As a matter of fact, the (interior of the) horizontal dashed rectangle comprises global states that are such that T_1 and T_2 would hold a lock on *a*: this is impossible by the very definition of a binary semaphore. Similarly, the (interior of the) vertical rectangle is a path from the initial point consists of states violating the mutual exclusion property on *b*. Therefore both dashed rectangles are *forbidden regions*, i.e. are not part of the space of (legal) states. This provides us with a particular po-space, as defined below. The idea is that a po-space is a topological space in which points are ordered globally through time.

Definition 3.1. (1) A *partial order* $\leqslant$ on a set *U* is a reflexive, transitive and antisymmetric relation. We write $x < y$ for ($x \leqslant y$ and $x \neq y$).

(2) A partial order $\leqslant$ on a topological space *X* is called *closed* if $\leqslant$ is a closed subset of $X \times X$ in the product topology. In that case, $(X, \leqslant)$ is called a *pospace*.

Example 3.2. The partial order on $\mathbb{R}^2$ used in the Swiss flag example above generalizes to $\mathbb{R}^n$: $(x_1, \ldots, x_n) \leqslant (y_1, \ldots, y_n)$ if $x_i \leqslant y_i$ for all $i = 1, \ldots, n$.

Remark 3.3. Let $(X, \leqslant)$ denote a pospace.
(1) For every $x \in X$, the sets $\downarrow x = \{y \in X | y \leqslant x\}$ and $\uparrow x = \{y \in X | y \geqslant x\}$ are closed.
(2) For every pair of points $y_1, y_2 \in X$, the set $[y_1, y_2] = \{x \in X | y_1 \leqslant x \leqslant y_2\} = \downarrow y_2 \cap \uparrow y_1$ is closed.

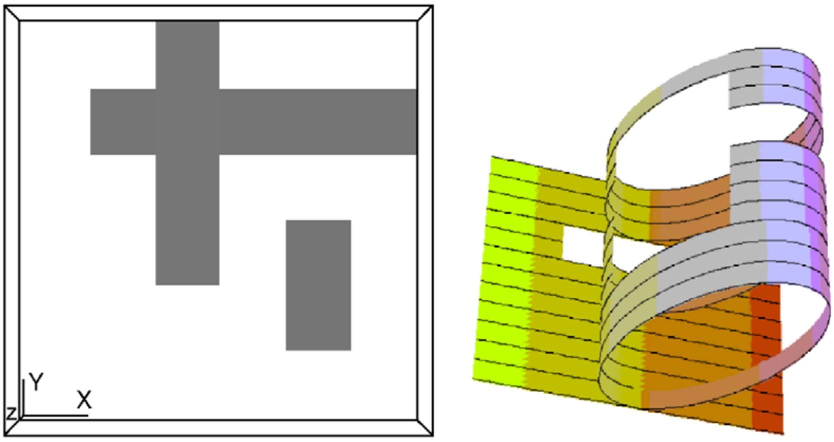


Fig. 3. Folding of a po-space to give semantics to a loop.

(3) A partially ordered topological space is a pospace if and only if whenever $a \not\leqslant b$, there exist open sets U and V with $a \in U$ and $b \in V$ such that for all $x \in U$ and $y \in V$ $x \not\leqslant y$. Hence a pospace is Hausdorff (see [31,63]).

When there are loops involved, having a partial order is too strong an assumption. This fact will lead us to local po-spaces (Definition 3.4). For instance, consider the following two processes in parallel:

```
A=Pa.(Pb.Vb.Pc.Vc)*.Va
B=Pc.Pb.Vc.Pa.Va.Vb
PROG=A|B
```

`A` can loop any number of times after `Pa`. As a topological space of states, this consists in folding the po-space on the left of Fig. 3 so that the start time of the left parenthesis is identified with the final time of execution of the right parenthesis, giving the topological space on the right of the same figure. Notice that this topological space cannot possibly be globally ordered, but is of course locally ordered as in the following definition (expressing the flow of time, locally).

Definition 3.4. Let X be a topological space.

- A collection $\mathcal{U}(X)$ of pairs $(U, \leqslant_U)$ with partially ordered open subsets U covering X is a *local partial order* on X if for every $x \in X$ there is a non-empty open neighbourhood $W(x) \subset X$ with a partial order $\leqslant_{W(x)}$ such that the restrictions of $\leqslant_U$ and $\leqslant_{W(x)}$ to $U \cap W(x)$ coincide for all $U \in \mathcal{U}(X)$ with $x \in U$, i.e.,

$$y \leqslant_U z \iff y \leqslant_{W(x)} z \quad \text{for all } U \in \mathcal{U}(X) \text{ such that } x \in U \text{ and for all } y, z \in W(x) \cap U. \tag{1}$$

- Two local partial orders on X are equivalent if their *union* is a local partial order.
- A topological space X together with an equivalence class of local partial orders is called a *locally partially ordered space* and a neighbourhood $W(x)$ as above is called a *partial order neighbourhood*. If, moreover, X is Hausdorff and there is a covering $\mathcal{U}$ such that for all $(U, \leqslant_U) \in \mathcal{U}$ the order $\leqslant_U$ on U is a closed relation ($(U, \leqslant_U)$ is a pospace), then X together with an equivalence class of coverings by po-spaces is a *local po-space*, or *lpo-space*.

When X is a local po-space, a neighbourhood $W(x)$ as in Definition 3.4, s.t. the partial order on $W(x)$ is closed, is called a *po-neighbourhood* of x.

What we gain here is the ability to consider loops and points which you can come across in a trace of execution (infinitely) many times.

Remark 3.5. (1) The partial orders $\leqslant_U$ need not coincide on their intersection, but on a sufficiently small neighbourhood, $W(x)$ of each point, all $\leqslant_U$ where $x \in U$ agree.

(2) This is a sort of *germ* or *sheaf* type definition of a local partial order and in particular of the monotone functions (below).

(3) The transitive hull of the partial orders given on subsets does not in general give rise to an interesting relation on X. If X is circle (3.6) with local partial order given by a chosen direction, then the hull of the relation is the trivial relation $x \leqslant y$ for any pair x and y. The same is true for the torus.

(4) By an abuse of notation, we will henceforth denote a locally partially ordered space $(X, \mathcal{U})$ without the $\leqslant_U$.

(5) The equivalence of local partial orders is an equivalence relation. To prove transitivity, suppose $\mathcal{U}$ and $\mathcal{V}$ are equivalent and that $\mathcal{V}$ and $\mathcal{T}$ are equivalent and let $x \in X$. Now let $(W_1(x), \leqslant_{W1})$ be a neighbourhood of x such that whenever $x, y, z \in W_1(x) \cap R$, where $R \in \mathcal{U} \bigcup \mathcal{V}$, $y \leqslant_R z \Leftrightarrow y \leqslant_{W1} z$ and let $W_2(x)$ have the analogous properties for $\mathcal{V} \bigcup \mathcal{T}$. Suppose moreover wlog, that there are $U \in \mathcal{U}$, $V \in \mathcal{V}$ and $T \in \mathcal{T}$ such that $W_1(x) \in U \cap V$ and $W_2(x) \in V \cap T$. Then the partial orders $\leqslant_{W1}$ and $\leqslant_{W2}$ agree on $W_1(x) \cap W_2(x)$, since for $y, z \in W_1(x) \cap W_2(x)$, $y \leqslant_{W1} z \Leftrightarrow y \leqslant_V z \Leftrightarrow y \leqslant_{W2} z$. The neighbourhood $W_1(x) \cap W_2(x)$ with this partial order satisfies the definition above for $\mathcal{U} \cup \mathcal{V} \cup \mathcal{T}$.

(6) There are various constructions of local po-spaces from other local po-spaces—glueing is used in the construction in Section 6. Another example is the product: given two local po-spaces, the product inherits a local po-structure from the product of the coverings—we leave the details to the reader.

(7) In an earlier version of this paper, we defined local partial orders in a slightly different way. A covering $\mathcal{U}$ as above was called a local partial order, if any x has a neighbourhood $W(x)$ (*with no ordering required*) such that whenever $U, V \in \mathcal{U}$ and $y, z \in W(x) \cap U \cap V$, $y \leqslant_U z \Leftrightarrow y \leqslant_V z$.

Since one may assume that $W(x) \subset U$ for some $U \in \mathcal{U}$, and hence the restriction of $\leqslant_U$ gives $\leqslant_{W(x)}$ which agrees with the order on all $V \cap W(x)$. Hence the fact that we require an ordering on $W(x)$ in the new definition does not make any difference. The difference is, that in the new definition, we only ask that elements from the cover which are *neighbourhoods of x* have partial orders which agree on $W(x)$. In the old definition, we required $W(x)$ to be a po-neighbourhood of all its points, not just x.

The cover $U_1 = \{e^{i\theta} \in S^1 | 0 < \theta < 2\pi\}$ and $U_2 = \{e^{i\theta} \in S^1 | \pi < \theta < 3\pi\}$ ordered by increasing θ gives a local partial order on the circle $S^1 = \{e^{i\theta} \in \mathbb{C}\}$ with the new definition, but not with the old one. There is no neighbourhood around $x = (1, 0)$ where the two orders $\leqslant_{U_1}$ and $\leqslant_{U_2}$ agree. But since only U_1 is a neighbourhood of $(1, 0)$; for the new definition, we could let $W((1, 0)) = U_1$.

In Example 3.6, we give another covering of the circle, which is a local partial order with both definitions.

Given a covering satisfying the new definition, one could ask whether there is an equivalent covering, which satisfies the old definition, (“shrinking” the elements in the cover, as in the circle example) but we did not go into this. For the local partial order which we introduce for cubical sets in Section 6 the answer is yes. After a subdivision, the local partial order given by the cover of (new and smaller) stars of vertices does in fact satisfy the old definition.

Example 3.6. (1) The circle $S^1 = \{e^{i\theta} \in \mathbb{C}\}$ has a local partial order: $U_1 = \{e^{i\theta} \in S^1 | \pi/4 < \theta < 7\pi/4\}$ has a (partial) order given by the order of the θ and $U_2 = \{e^{i\theta} \in S^1 | 5\pi/4 < \theta < 11\pi/4\}$ is (partially) ordered by the order on the θ's.

(2) The torus T^2 is $\mathbb{C}$ modulo a lattice $z \equiv z + \mathrm{i}p + q \ \forall (p, q) \in \mathbb{Z} \times \mathbb{Z}$ and hence it inherits a local partial order from the standard partial order on $\mathbb{C} \cong \mathbb{R}^2$. This is equivalent to choosing a local partial order on each of the two generators of the torus.

(3) Let X be a disjoint union of four copies of the unit square I^2. We get inequivalent global partial orders on X by considering

$$X = [0, 1] \times [0, 1] \bigcup [0, 1] \times [4, 5] \bigcup [4, 5] \times [0, 1] \bigcup [4, 5] \times [4, 5]$$

or

$$X = [0, 1] \times [0, 1] \bigcup [0, 1] \times [4, 5] \bigcup [4, 5] \times [0, 1] \bigcup [2, 3] \times [2, 3]$$

both with the partial order induced from $\mathbb{R}^2$. Considered as local partial orders, these are equivalent. A common refinement is defined by letting all four copies of I^2 have the partial order induced from $\mathbb{R}^2$ and no further relations.

The geometric counterpart of an execution is a path from $(0, 0)$ which is increasing in both coordinates. An execution of the whole program is an increasing path from $(0, 0)$ to the final point in the upper right hand corner. Just looking at the picture if Fig. 2, it is clear that:

- The point at $(2/5, 2/5)$ (which corresponds to (P_b, P_a)) is a deadlock—there are no executions from there. Everything stops.
- Once an execution path has entered the area marked "unsafe", there is no way of continuing to the final point.
- The points marked "unreachable" will never be reached by an execution path initiating in (0, 0)

We give a formal definition of these concepts here.

Definition 3.7. (1) Let $(X, \mathcal{U})$ and $(Y, \mathcal{V})$ be *locally partially ordered spaces*. A continuous map $f : X \to Y$ is called a *dimap* (directed map) if for any $x \in X$ there are partial order neighbourhoods $W(x)$ and $W(f(x))$ such that

$$x_1 \leqslant_{W(x)} x_2 \Rightarrow f(x_1) \leqslant_{W(f(x))} f(x_2)$$

whenever $x_1, x_2 \in f^{-1}(W(f(x))) \cap W(x)$.

(2) A *dipath* in *X* is a dimap $f : I \to X$ from the unit interval *I* with the natural (global) order $\leqslant$.

The notion of dipath corresponds to the intuitive idea of traces of executions as explained in Figs. 1 and 2 for instance. Of course, there is a major algebraic structure on the set of dipaths, which is concatenation:

Remark 3.8. (1) Let $f_1, f_2 : I \to X$ denote two dipaths with $f_1(1) = f_2(0)$. Their *concatenation* $f_1 * f_2$ given by

$$(f_1 * f_2)(t) = \begin{cases} f_1(t), & t \leqslant 0.5; \\ f_2(2t-1), & t \geqslant 0.5. \end{cases}$$

is again a dipath.

(2) One might look at dimaps from *arbitrary* intervals and allow equivalence classes with respect to strictly increasing homeomorphisms between intervals.

Definition 3.9 (*Compare Penrose [64]*). Let *X* be a locally partially ordered space. We define a new relation $\prec$ on *X* by $x \prec y$ if there is a dipath from *x* to *y* in *X*.

Lemma 3.10. *If X has a global partial order*, $\leqslant$ *the relation* $\prec$ *is a new partial order*.

Proof. The relation $\prec$ is coarser than the relation $\leqslant$, i.e., $x \prec y \Rightarrow x \leqslant y$. Hence, $\prec$ is antisymmetric. Concatenation of dipaths shows the transitivity of $\prec$. □

This is a geometric "reachability" relation which underlies most proofs in semantics.

Remark 3.11. If *X* is locally partially ordered, the relation $\prec$ is still transitive, but it is not necessarily antisymmetric as one can see from Example 3.6 with the oriented circle. Moreover, the relation $\prec$ may be closed while $\leqslant$ is not and vice versa. See [18] for examples of this.

Definition 3.12. Let $y \in X$, $S \subset X$.

(1) The set $J^+(y) := \{x \in X | y \prec x\}$ is called the *future* of *y*; likewise, one defines the *past* $J^-(y)$. The set $J(y) := J^-(y) \cup J^+(y)$ is called the *history*.

(2) $J^+(S) := \bigcup_{x\in S} J^+(x)$, $J^-(S) := \bigcup_{x\in S} J^-(x)$, $J(S) := J^-(S) \cup J^+(S)$ are called the *future*, *past*, *history* of *S*.

(3) *x* is called an *initial* point if $J^-(x) = \{x\}$; *x* is called a *final* point if $J^+(x) = \{x\}$.

Remark 3.13. Initial points, resp. final points are local maxima, resp. minima with respect to the partial order $\prec$. An initial point is *unreachable* from any other initial point. A final point is unreachable from any other final point. Hence a *deadlock* is a final point, which is not among the final points representing successful outcomes of the computations.

4. Dihomotopy and diconnected components

The objective of this section is twofold: first of all, we want to formalize the concept of *deformation* of a path used in the Introduction. Technically, we speak of dihomotopies between dipaths in a given lpo-space. Secondly, we describe how the dihomotopy concept can be used to split a po-space into "di-connected components" such that points (states) in the same component have the same properties with respect to dipaths (executions) visiting them.

Throughout this section, $(X, \leqslant)$ is supposed to be a locally partially ordered topological space (lpo-space).

4.1. Dihomotopy

It is important that the dipaths to be considered cannot be shrinked under a deformation. This is taken care of in

Definition 4.1. (1) A dipath $\alpha : \vec{I} \to X$ is called *inextendible*, if there is no dipath $\beta : \vec{J} \to X$ such that $\alpha(\vec{I}) \subset \beta(\vec{J})$ and $\alpha(\vec{I}) \neq \beta(\vec{J})$.

(2) The set of all inextendible dipaths in X is denoted as $\vec{P}_1(X)$.

In particular, an inextendible dipath α starts at an initial point and ends at a final point of the lpo-space.

Let $\vec{I}$ denote the ordered unit interval, and let I denote another unit interval (with equality as partial order, i.e., without using order properties).

Definition 4.2. (1) A continuous map $H : \vec{I} \times I \to X$ is called a *dihomotopy* if every partial map $H_t = H(-, t) : \vec{I} \to X,\ t \in \vec{I}$, is an inextendible dipath.

(2) Two inextendible dipaths $\alpha, \beta : \vec{I} \to X$ are called *dihomotopic* if there is a dihomotopy $H : \vec{I} \times I \to X$ with $H_0 = \alpha$ and $H_1 = \beta$. We write: $\alpha \overset{\rightarrow}{\simeq} \beta$.

(3) The set of dihomotopy classes of inextendible dipaths in X is denoted as $\vec{\pi}_1(X)$.

Remark 4.3. (1) A homotopy of paths is just a 1-parameter deformation. What is new in the definition above, is that we insist that *all intermediate* paths H_t preserve the orientation.

(2) The local partial order need *not* be preserved wrt to the variables in I.

(3) Dihomotopy is obviously an equivalence relation.

(4) Without the restriction to *inextendable* dipaths, every dipath $\alpha : \vec{I}$ would be dihomotopic to a constant dipath via the dihomotopy

$$H(s, t) = \alpha((1 - t)s), \quad s \in \vec{I},\ \ t \in I.$$

Then all dipaths in the same connected component of X would be dihomotopic to each other.

(5) Assume that both the set of initial points and the set of final points in X are *discrete*. Then, if H_0 starts in the initial point x, then (by inextendibility and continuity), all H_t have to start in the same point x. Similarly, they have to end in the same final point. Hence, in this case, $\vec{\pi}_1(X)$ splits into subsets corresponding to every particular pair of an initial and a final point.

Why is dihomotopy a relevant concept? It can be argued in fact that schedules (or dipaths) which can be deformed into each other by a continuous family of executions are in fact equivalent (yield same results) *for all* possible actual values that the semaphore (in fact, the variable they protect) can take. Since, in our geometric model, schedules correspond to dipaths, dihomotopic dipaths correspond to equivalent executions.

To see why this holds, consider for instance Fig. 2. Any dipath below the central hole (with shape of a cross), see the second picture of Fig. 4, is such that T_1 gets hold of locks a and b before T_2 does. If we suppose that all assignments on a and b are protected by the corresponding P operations, it is clear that only the order of accesses to the shared variables on an execution path counts for determining what is actually computed on this path, as shown in the second picture of Fig. 4. In fact in this system, there are only two essential behaviours (that do not go into a deadlock see first picture of Fig. 4): one is the type of dipaths just discussed, the other one is the class of dipath going above the central hole (see third picture of Fig. 4). In terms of schedule of executions, the latter corresponds to executions in which T_2

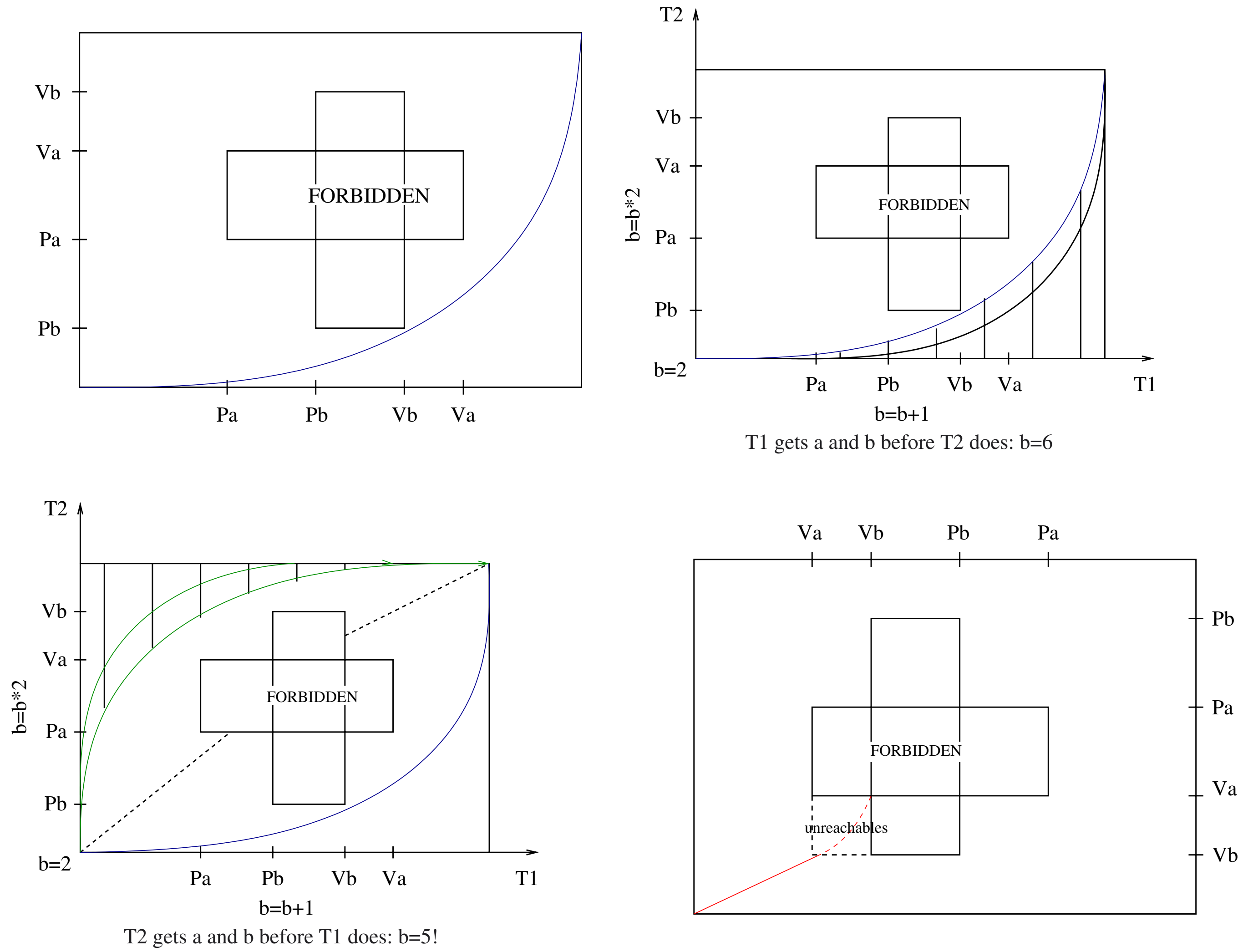


Fig. 4. Essential schedules for the Swiss flag.

is the first to read and write (after having got the corresponding locks) on a and b, before T_1 does. These are of course the only two dipaths from (0, 0) to (1, 1) modulo "continuous deformations".

There is another rather combinatorial point of view on dipaths and dihomotopies, in which local dihomotopies are modelled by rectangles (with the semantical meaning: first a then b is equivalent to first b then a, this is very much similar to Mazurkiewicz trace theory [61], and this provides another reason for dihomotopy to provide us with the "right" equivalence relation on dipaths). In a more abstract setting, see also our Section 7, this version of dihomotopy has been investigated by Grandis [46,45]. For nice enough lpo-spaces (like the cubical sets from Section 6, both dihomotopy concepts are shown to be equivalent by Fajstrup [19]).

Previously, it had been conjectured [65] in a similar context, that homotopy (instead of our dihomotopy) was the proper notion to discriminate essentially different schedules. Our Example 4.4 (4) and its interpretation in Section 4.3 shows that homotopy without directions is not sufficient.

Example 4.4. The following examples show several crucial effects of the directedness (monotonicity) requirements in the definitions:

(1) Fig. 5 represents a path from an initial to a final point that cannot be homotoped to a dipath (with fixed end-points).
(2) Fig. 6 shows two po-spaces (a square with two holes in different positions) that are homeomorphic (even fixing the end points) but that show different behaviour with respect to dihomotopy. For the po-space on the left, there are three classes of dihomotopy classes of dipaths from the bottom to the top; for the one on the right, there are four

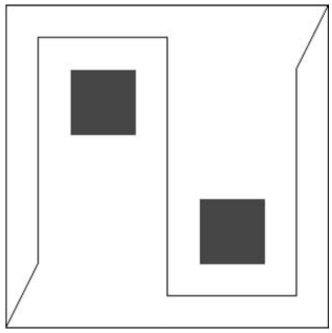


Fig. 5. A path that is not homotopic to a dipath.

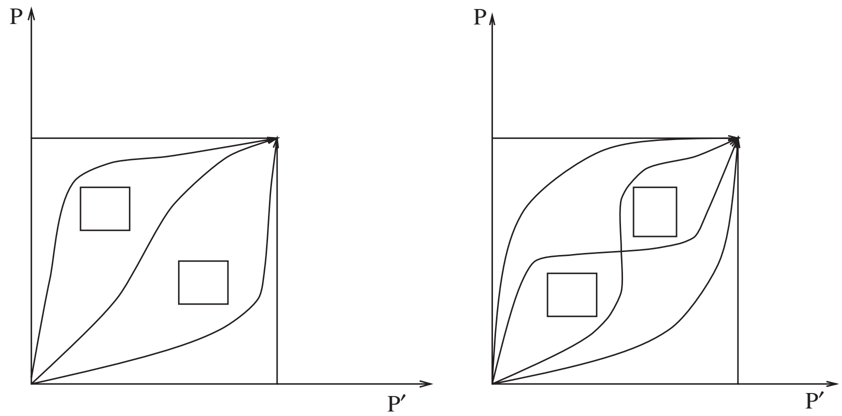


Fig. 6. The two possible relative configurations of holes.

dihomotopy classes. In particular, the two spaces cannot be dihomeomorphic to each other (relative to end points), i.e., there is no invertible *di*map from one to the other. Here is a practical interpretation for this distinction: for the program on the left side, there is no schedule allowing the vertical process to arrive first at the left obstruction as well as the horizontal process to arrive first at the right obstruction (Fig. 6).

(3) For the complement of the Swiss flag in Fig. 2, there are four classes of dihomotopy classes: two from the bottom to the top, one from the bottom to the deadlock point, and one from the lower corner of the unreachable region to the top point.

(4) Fig. 7 shows a "room with three barriers" and two dipaths from the unique initial to the unique final point. The po-space in question is a solid cube with three blocks removed. The two vertical rectangular blocks stretch from the bottom surface to the top surface. Without the third middle block removed, this corresponds to a product of the rightmost space from Fig. 2 with an interval $\vec{I}$—which would again give rise to four dihomotopy classes. But now we remove additionally a third middle box, that does not touch any boundary surface, but such that its extremities overlap with those of the two other blocks when projected to the front face.

Consider the two dipaths in Fig. 7: both of them pass the first block on the right and the last block on the left. But the first dipath passes under and the second over the middle block. When regarded as paths from the initial point to the final point, i.e., forgetting about orientations, these two paths are homotopic, even fixing the end points. One may connect the two dipaths by a homotopy (1-parameter family) of *paths* which pass by either to the left or to the right of the middle block. But no such homotopy can be directed. It has to contain an intermediate path which changes direction—which is prohibited for a dipath. This is not just a geometric artefact. A concurrent program with that geometric model with different results for the schedules modelled by the two dipaths will be given in Section 4.3.

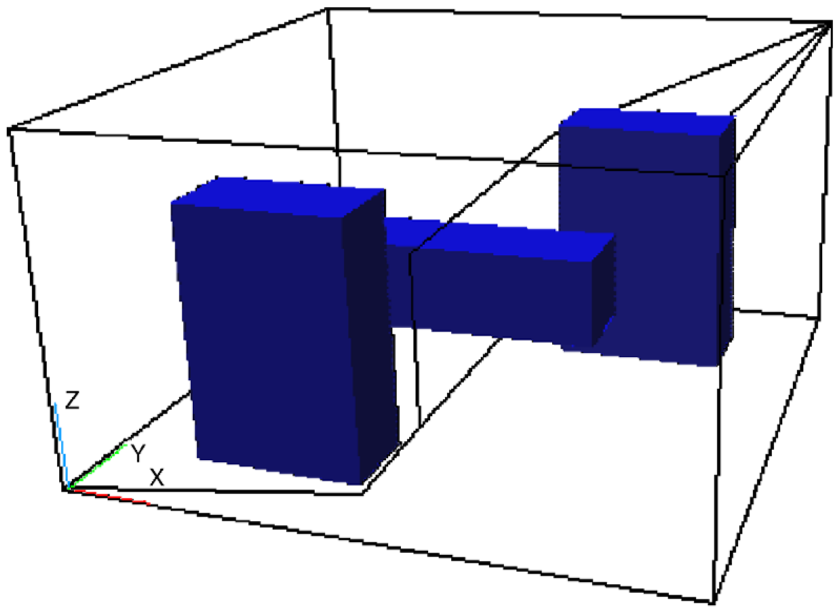


Fig. 7. Room with three barriers and two non-dihomotopic dipaths.

4.2. Diconnected components

Which points (states) can be visited by a dipath dihomotopic to a given one? To answer this question, we need to introduce a few concepts:

Definition 4.5. (1) The *homotopy history* of an inextendible dipath $\alpha : I \to X$ is defined as

$$h\alpha := \{y \in X \mid \exists \text{ a dipath } \beta \text{ through } y \text{ and } \alpha \overset{\rightsquigarrow}{\simeq} \beta\}.$$

(2) Two points $x, y \in X$ are *homotopy history equivalent* if

$$x \in h\alpha \quad \Leftrightarrow \quad y \in h\alpha \quad \text{for all } \alpha \in \vec{P}_1(X).$$

(3) The *diconnected components* of X consists of the path connected components (in the classical sense) of the homotopy history equivalence classes of X.

Remark 4.6. (1) Inextendible paths in the same dihomotopy class have the same homotopy history.

(2) Two points $x, y \in X$ are history equivalent if and only if *every* dipath through x is dihomotopic to one through y and vice versa.

(3) Two points $x, y \in X$ may be history equivalent, but there is no path inbetween them passing only through points history equivalent to them. Then, they do not belong to the same diconnected component, cf. Example 4.7.

(4) If there are finitely many dihomotopy classes of dipaths in X, the Boolean algebra generated by the homotopy histories is atomic and a homotopy history equivalence class C is an atom, i.e., it contains the points that are contained in the homotopy histories of certain dihomotopy classes but not in others, i.e., there is a subset $K \subset \vec{\pi}_1(X)$ such that

$$C = \bigcup_{\alpha \in K} h\alpha \setminus \bigcup_{\beta \notin K} h\beta.$$

(5) In subsequent work [20], the concept of a diconnected component has been refined. Using the *fundamental category* of a po-space, we define and analyse its *component category*.

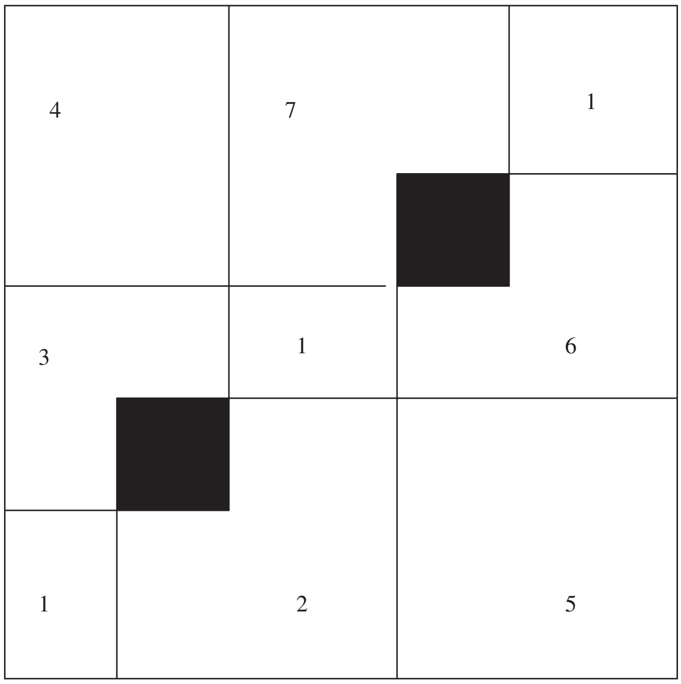


Fig. 8. “Two partially ordered holes”.

Point 4 of the remark above is of primary importance for program analysis. Each homotopy history $h\alpha$ corresponds to some property of accesses of shared resources in the PV model. The decomposition of the state space X shows that there are elementary regions which are separated by the properties that executions visiting them can have in the future (and in the past). We give examples below:

Example 4.7. (1) The complement of the “Swiss flag” in $\vec{I}^2$ (see Fig. 2) has 10 homotopy history components. All of them are pathwise connected. This gives the semantics of the program having process $T_1 = Pa.Pb.Vb.Va$ in parallel with $T_2 = Pb.Pa.Va.Vb$ (where a and b are 1-semaphores). In region 1, we still have all possible futures (all possible access histories to a and b). From region 2, we can proceed to 4 or to 6, meaning that we are either going to deadlock (this is the unsafe region of the first picture of Fig. 4) in the future or that T_2 will get a and b before T_1. Region 6 can only be reached from 2 and continue to 9: T_2 has got a and b before T_1. Region 9 can be reached from the unreachable region 7 (that we already identified as the fourth picture of Fig. 4) or from 6. In region 10, we might have come from any history in the past.

(2) The complement of “two partially ordered holes” in $\vec{I}^2$ (see Fig. 8) has seven homotopy history components. Region 1 contains both the initial point **0**, the final point **1**, and an area in the middle. This homotopy history class decomposes into three diconnected components, all the others are pathwise connected. This pictures gives the semantics of the term $Pa.Va.Pb.Vb|Pa.Va.Pb.Vb$.

(3) The “room with three barriers” in I^3 from Example 4.4 (4) has nine homotopy history components. This can be seen by comparing with Fig. 8 by a projection. In fact, the region over the middle component (labelled 1) splits into three history components according to which of the two “critical” dihomotopy classes can visit (the first, the second or both).

4.3. A weak synchronization example

The rest of this section is devoted to an example elaborating on Example 4.4 (4)—the “room with three barriers” and shows that the two homotopic but not dihomotopic dipaths may correspond to essentially different executions in a concrete situation:

For this and further examples, we have to enhance our toy programming languages slightly so that,

- shared objects can be "weakly synchronizing", i.e., they can be shared by k processes but not $k+1$ at the same time, for any $k \geqslant 1$. Examples of such objects can be redundant functional units (for instance, in microprocessors, or in workshop modelizations), or communication buffers of fixed size (in the case of asynchronous message passing), or shared FIFO queues (in shared-memory systems). We choose to think of these objects `s` in the convenient form of "k-places buffers", on which we can do actions `push(x,s)` where `x` is any integer value and `read(y,s)` where `y` is any local (to the process executing the instruction) integer array variable.
- `read(y,s)` gives an atomic snapshot of the shared buffer s in the local memory, in the local array y. Then any array operation like access at the `i`th element, "`s[i]`", can be performed locally by doing $y[i]$.
- `push(x,s)` corresponds to asking to take one of the free places of the buffer s (in FIFO order here) to put value x: if the buffer is full then it pushes the values so that the first value entered is discarded. If two or more `push` instructions are executed right at the same time the semantics is not defined (anything can happen, in practice at the hardware level if no locks are used, this corresponds to some kind of short-circuit). In order to protect the integrity of the messages, we are using instructions `Ps` and `Vs` to acquire (respectively, relinquish) one of the locks on the buffer. One place buffers are just the same as ordinary integer variables.

To explain the different meanings of the two dipaths in Example 4.4 (4) in computer-scientific terms, consider the following three programs:.

- $T_1 = Pa.\texttt{push}(1,a).Va.Pb.\texttt{read}(u,b).\texttt{read}(v,a).\texttt{push}(u[1]+v,b).Vb.\ Pc.\texttt{read}(u,b).$ $\texttt{push}(u[1]+u[2],c).Vc$,
- $T_2 = Pb.\texttt{read}(v,a).\texttt{push}(v+1,b).Vb$,
- $T_3 = Pa.\texttt{push}(3,a).Pb.\texttt{read}(v,a).\texttt{push}(0,b).Va.Pc.\texttt{read}(u,b).\texttt{push}(u[1]*v,b).\ \texttt{push}(u[1]+u[2],c).$ $Vb.Vc$.

Basically,

- T_1: sets value 1 to a (remember, this is just a shared variable), then pushes the current value of $b[1]$ by the current value of a on the 2-cell buffer b, and finally sets c (a shared variable) to the sum of the elements in b.
- T_2: just pushes the value of a plus one in b.
- T_3: first sets a to be equal to 3, then pushes 0 on b, then the value of a times the value of the first cell of b, then puts the sum of the entries of b in c.

Notice that the functions that are computed by each of the three processes are sensitive to the order in which values are written in `b`, hence, as we will see, are able to distinguish the two schedules pictured in Fig. 7. The first dipath (the one below the central hole in Fig. 7) corresponds to the following schedule (where T_3 gets a before T_1 and T_2 gets into b after (T_1, T_3)),

T_1	T_2	T_3	$Values$
–	–	Pa	
–	–	Pb	$a=3$
–	–	Va	$b=(\square_3 0, 0)$
Pa	–	–	
Va	–	–	$a=1$
Pb	–	–	
Vb	–	–	$b=(\square_1 1, \square 0)$
Pc	–	–	
Vc	–	–	$c=1$
–	–	Pc	
–	–	Vb	$b=(\square_3 3, 1)$
–	–	Vc	$\underline{c=4}$
–	Pb	–	
–	Vb	–	$b=(\square_1 2, 3)$

where "boxed" values (for b) are the places of this buffer which are holding a lock. The second dipath corresponds to (where T_3 gets a before T_1 and T_1 gets into b after (T_2, T_3)),

T_1	T_2	T_3	*Values*
–	–	Pa	
–	–	Pb	$a = 3$
–	–	Va	$b = (\Box_3 0, 0)$
Pa	–	–	
Va	–	–	$a = 1$
–	Pb	–	
–	Vb	–	$b = (\Box_2 2, \Box_3 0)$
Pb	–	–	
Vb	–	–	$b = (\Box_1 3, \Box_2 2)$
Pc	–	–	
Vc	–	–	$c = 5$
–	–	Pc	
–	–	Vb	$b = (\Box_3 9, 3)$
–	–	Vc	$\underline{c = 12}$

Let us assume the purpose of this program was to give a value for c then we see that these two homotopic (in the classical sense) but not dihomotopic dipaths give different results.

5. Parametrized and Euclidean partial orders

In this section, we look at a particular subcategory of lpo-spaces, where locally, the time ordering is the component-wise ordering in $\mathbb{R}^n$. These spaces are a special case of parametrized spaces in which it is possible to "cut transversally to time". This technique allows sometimes, but not always, to find out that two dipaths are *not* dihomotopic: this is the case, when those pass through *distinct connected components* of the cuts.

5.1. Definitions

Definition 5.1. Let U be a set with a partial order $\leqslant$. A subset $V \subset U$ is called *achronal* if for all $x, y \in V : x \leqslant y \Rightarrow x = y$; compare [64].

Example 5.2. In the po-space $(\mathbb{R}^2, \leqslant)$ with the natural partial order, the *anti-diagonal* $\bar{\Delta} = \{(t, -t)| \ t \in \mathbb{R}\}$ is achronal.

Achronal cuts are much like instantaneous snapshots that are at the base of some recent geometric proofs of non-existence of some fault-tolerant protocols for distributed systems (see for instance [50]). The idea is that, under some circumstances (for instance, studying a distributed system with a global clock), one can observe a system at some given "still moments", and that one can find interesting *classical* algebraic topological invariants that help understand the possible states that a system can reach. We refer to this work more in detail in Section 10.

Definition 5.3. Let $(X, \leqslant)$ denote a po-space.

(1) We call $(X, \leqslant)$ *parametrized* if there is a (parameter) dimap $F : X \to \mathbb{R}$ such that every *cut* $X_t := F^{-1}(t)$ is achronal for every $t \in \mathbb{R}$.
(2) We call $(X, \leqslant)$ *Euclidean* if there are finitely many dimaps $f_i : X \to \mathbb{R}$ such that

$$\forall x, y \in X : \quad x \leqslant y \ \Leftrightarrow \ \forall i : \quad f_i(x) \leqslant f_i(y); \forall i : f_i(x) = f_i(y) \ \Leftrightarrow \ x = y.$$

(3) A *local* partial order on a topological space X is called *parametrized*, resp. *Euclidean* if it (or a refinement of it) consists of parametrized, resp. Euclidean partial orders.

Remark 5.4. For a Euclidean partial order, the relation $\leqslant$ is given by comparison (via the map $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^n$) with the natural partial order on Euclidean space $\mathbb{R}^n$ from Example 3.2.

Lemma 5.5. (1) *A parametrized po-space gives rise to a new partial order* $\leqslant'$ *defined by*: $x \leqslant' y \Leftrightarrow x = y$ or $F(x) < F(y) \in \mathbb{R}$.
It satisfies: $x \leqslant y \Rightarrow x \leqslant' y$.
(2) *A Euclidean partial order is parametrized*.

Proof. (1) Obvious.
(2) Let X and $f_i : X \to \mathbb{R}$ be given as in Definition 5.3.2. The function $F = \sum f_i : X \to \mathbb{R}$ is a dimap. For every $t \in \mathbb{R}$, the preimage $X_t := F^{-1}(t) \subset X$ is achronal. □

5.2. Well-parametrized dipaths and dihomotopies

For the rest of this section, let $(X, \leqslant, F)$ denote a parametrized po-space *foliating* X into the cuts X_t, $t \in \mathbb{R}$. We arrive at a better understanding when the parameter map F is used to reparametrize dipaths and dihomotopies to yield new parametrizations matching with that foliation:

Definition 5.6. Let $\vec{J}$ denote a closed real interval.
(1) A dipath $\alpha : \vec{J} \to X$ is called *well-parametrized* if $F(\alpha(t)) = t$ for every $t \in \vec{J}$.
(2) A dihomotopy $H : \vec{J} \times I \to X$ is called well-parametrized if $F(H_t(0))$ is constant and if every dipath $H_t : \vec{J} \to X$, $t \in I$ is well-parametrized.
(3) A dipath $\beta : \vec{J} \to X$ is called a *reparametrization* of another dipath α if there is a *monotonic* map $h : \vec{I} \to \vec{J}$ such that $\alpha = \beta \circ h$.

Remark, that also $F(H_t(r))$ is constant along a well-parametrized dihomotopy H.
Almost as in any course on elementary differential geometry, we get:

Proposition 5.7. (1) *To any dipath* $\alpha : \vec{I} \to X$, *there is a unique well-parametrized reparametrization* $\beta : \vec{J} \to X$.
(2) *To any dihomotopy* $H : \vec{I} \times I \to X$ *from one cut to another* (*i.e.*, $F(H_t(1))$ *and* $F(H_t(0))$ *are both constants*), *there is a unique reparametrization* $\bar{H} : \vec{J} \times I \to X$, *i.e.*, *such that every dipath* $\bar{H}_t$ *is a well-parametrized reparametrization of* H_t *for every* $t \in I$.

Proof. The following proofs are relatively complicated, since we do not assume, that the dipaths involved are strictly monotonic. For strictly monotonic maps, one arrives at more elementary proofs via the *dihomeomorphism* $h := F \circ \alpha$.
(1) The map h given by $h(s) := (F \circ \alpha)(s)$ is monotonic and continuous; its image is the interval $\vec{J} = [F(\alpha(0)), F(\alpha(1))]$. For $s \in \vec{J}$, choose $0 \leqslant t_s \in \vec{I}$ such that $h(t_s) = h(0) + s$ using the continuity of α. The assignment $\beta(s) = \alpha(t_s)$ yields a well-defined monotonic map β. From the definition, we get immediately, that $\alpha = \beta \circ h$. The map β is continuous, as well: let $U \subset X$ be an open non-empty set. Then $\alpha^{-1}(U)$ is either empty or it contains an open interval, on which α is not constant. Hence, $\beta^{-1}(U)$ contains an open interval, as well.
(2) For every $t \in I$, let $h_t : \vec{I} \to \vec{J}$ denote the monotonic map given by $h_t(s) := (F \circ H_t)(s)$. Proceeding as in (1) for every $t \in I$, we arrive at a unique map $\bar{H} : \vec{J} \times I \to X$ such that $H_t = \bar{H}_t \circ h_t$. Each map $\bar{H}_t$ is a well-parametrized dipath; the trouble is to ensure that the map $\bar{H}$ is continuous:
Let again $U \subset X$ denote an open set and let $\bar{H}(\bar{s}_0, t_0) \in U$. Choose $s_0 \in \vec{I}$ such that $H(s_0, t_0) = \bar{H}(\bar{s}_0, t_0)$. Since H_t is a dipath, one can choose $0 < \varepsilon, \delta_1, \delta_2$ such that $H([s_0 - \delta_1, s_0 + \delta_2]) \subset U$ and
$F(H(s_0 - \delta_1, t_0)) = F(H(s_0, t_0)) - \varepsilon, \; F(H(s_0 + \delta_2, t_0)) = F(H(s_0, t_0)) + \varepsilon.$

The continuity of H and a compactness argument allows to choose $0 < \delta_i' < \delta_i,\ 1 \leqslant i \leqslant 2,\ 0 < \rho$, such that $H([s_0 - \delta_1', s_0 + \delta_2'] \times]t_0 - \rho, t_0 + \rho[) \subset U$ and

$$F(H(s_0 - \delta_1', t_0 + \tau)) < F(H(s_0, t_0)) - \frac{\varepsilon}{2},\ F(H(s_0 + \delta_2', t_0 + \tau)) > F(H(s_0, t_0)) + \frac{\varepsilon}{2}$$

for $|\tau| < \rho$. As a consequence,
$\bar{H}(]\bar{s}_0 - \frac{\varepsilon}{2}, \bar{s}_0 + \frac{\varepsilon}{2}[) \times]t_0 - \rho, t_0 + \rho[) \subseteq H(]s_0 - \delta_1', s_0 + \delta_2'[\times]t_0 - \rho, t_0 + \rho[) \subseteq U$. □

5.3. Components of cuts distinguish dihomotopy classes

It is in general not easy to find out whether two given dipaths are dihomotopic or not. This is the main motivation for investigating cuts $X_t := F^{-1}(t)$ with respect to a parametrization $F : X \to \mathbb{R}$. We show here that dipaths proceeding through *different* (classical) connected components of a particular cut can never be dihomotopic:

Proposition 5.8. *Let $\alpha_1, \alpha_2 : (\vec{I}, 0, 1) \to (X; X_a, X_b)$ denote dimaps that are dihomotopic by a dihomotopy $H : (\vec{I}, 0, 1) \times I \to (X; X_a, X_b)$, which preserves the "end cuts" X_a and X_b. Then, the uniquely determined elements $x_t^i \in \alpha_i(\vec{I}) \cap X_t,\ 1 \leqslant i \leqslant 2$, are contained in the same path component of X_t for $t \in [a, b]$.*
In other words: if α_1 and α_2 pass through different path components of X_t for some parameter $t \in \mathbb{R}$, then α_1 and α_2 cannot be dihomotopic.

This proposition is an immediate consequence of the existence of a reparametrization from Proposition 5.7 and of the following particular case:

Proposition 5.9. *Let $H : \vec{J} \times I \to X$ denote a well-parametrized dihomotopy between two well-parametrized dipaths $\alpha_1, \alpha_2 : I \to X$. Then, $\alpha_1(t)$ and $\alpha_2(t)$ are contained in the same path component of X_t for every $t \in I$.*

Proof. The map $H(-, t) : I \to X_t$ is a path from $\alpha_1(t)$ to $\alpha_2(t)$. □

Remark 5.10. The criterion in Proposition 5.8 is exploited to give a complete classification of dihomotopy classes for a 2D mutual exclusion model in [67]. Two different parameter maps, i.e., the two projections to the axes, have to be employed for this task.

Unfortunately, the criterion from Proposition 5.8 is in general *not* sufficient to distinguish different dihomotopy classes. An easy example showing that a family of *diagonal* cuts alone is not enough is:

Example 5.11. Let X be the subset $[0, 3] \times [0, 3] \times [0, 3] \setminus [1, 2] \times [1, 2] \times [0, 3]$ in $\mathbb{R}^3$ with the standard partial order, i.e., a cube with an interior box removed, as pictured in Fig. 9. This po-space models the PV program:

```
A=Pa.Va
B=Pa.Va
C=Pb.Vb
```

where `a`, `b` and `c` are all binary semaphores.

There are two dihomotopy classes of paths from (0, 0, 0) to (3, 3, 3), but the cuts induced by $F(x, y, z) = x + y + z$ are all connected; in fact, they are polygons with or without a hole.

Hence to get full information about dihomotopy classes, it does not suffice to study *just one* family of cuts. One might still ask whether the non-existence of a dihomotopy—within nice po-spaces, e.g., those arising from cubical sets (treated in next section)—can always be detected if one knows *all* families of cuts and their connected components.

6. Cubical sets as local po-spaces

Cubical sets were defined by Serre [72] and a theoretical framework is developed by Brown and Higgins [7], and by Jardine [55], see also [34] for their use as models for HDA. We show here that they are the natural combinatorial

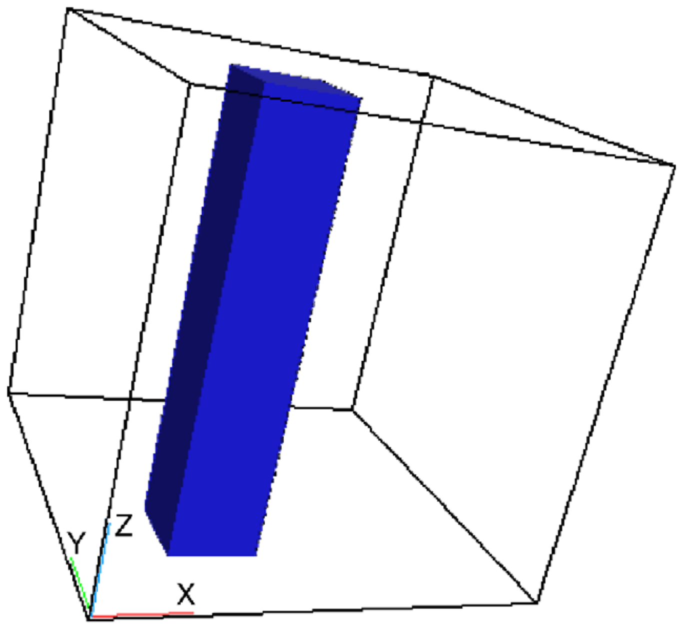
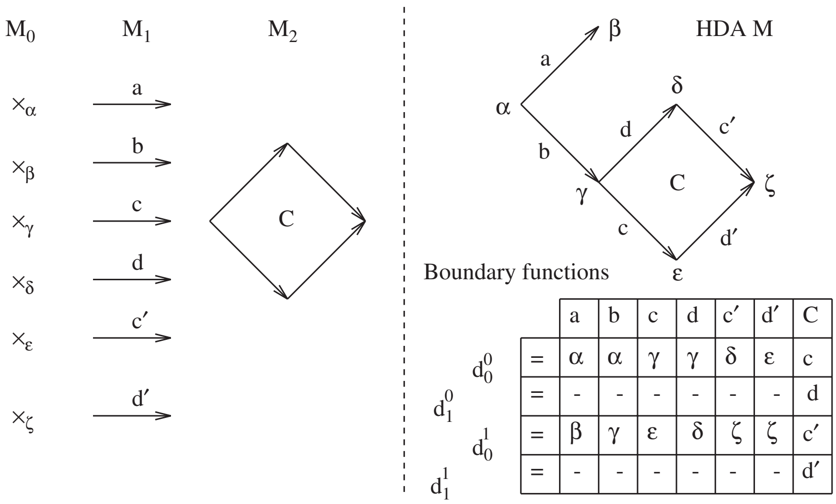


Fig. 9. A simple example where instant snapshots do not distinguish dihomotopy classes.



		a	b	c	d	c′	d′	C
d_0^0	=	α	α	γ	γ	δ	ε	c
d_1^0	=	-	-	-	-	-	-	d
d_0^1	=	β	γ	ε	δ	ζ	ζ	c′
d_1^1	=	-	-	-	-	-	-	d′

Fig. 10. The glueing of elementary cubes.

counterpart of local po-spaces (the centre of this is Theorem 6.23 and Proposition 6.38). This makes the link with more standard combinatorial techniques for reasoning about concurrent systems (interleaving ones or truly concurrent ones like HDA). For the correspondence with transition systems, see Goubault [40].

6.1. Cubical sets

In Fig. 10, the connection between the combinatorial description and the geometric description of an HDA is pictured. In the following, the reader should keep in mind, that this is what we are defining: a combinatorial way of describing the relations between a set of cubes of varying dimensions, glued along common boundaries or faces.

Definition 6.1. A pre-cubical set M is a family of sets $\{M_n | n \geqslant 0\}$ with face maps $\partial_i^k : M_n \to M_{n-1}$ $(1 \leqslant i \leqslant n, k = 0, 1)$ satisfying the pre-cubical relations:

$$\partial_i^k \partial_j^l = \partial_{j-1}^l \partial_i^k \quad (i < j).$$

The (iterated) face maps determine glueings of cubes, e.g, if $\partial^l_j z = \partial^k_i y = x$, then the cubes z and y share a face, x. If $k = 1$, then x is an upper face of y and if $k = 0$, it is a lower face. The commutator rules follow from this point of view.

Remark 6.2. In Brown et al. [8,7], a cubical set is defined—our precubical sets differ from cubical sets in that they do not have degeneracies.

Definition 6.3. Let *M* and *N* be two pre-cubical sets, and *f* a family $f_n : M_n \to N_n$ of functions. f is a morphism of pre-cubical sets if

$$f_n \circ \partial^0_i = \partial^0_i \circ f_{n+1},$$

$$f_n \circ \partial^1_i = \partial^1_i \circ f_{n+1}$$

for all $n \in \mathbb{N}$ and $1 \leqslant i \leqslant n+1$.

This defines the category Υ_{sr} of pre-cubical sets.

We write Υ^n_{sr} for the full subcategory of Υ_{sr} consisting of pre-cubical sets whose elements are cubes of dimension less than or equal to n.

6.2. The geometric realization of a pre-cubical set

Given a pre-cubical set *M*, we want to construct the corresponding geometric object $|M|$ by glueing cubes. For this purpose, let $\square_n$ be the standard cube in $\mathbb{R}^n$ $(n \geqslant 0)$, $\square_n = \{(t_1, \ldots, t_n) | \forall i, 0 \leqslant t_i \leqslant 1\}$, $\square_0 = \{0\}$.

Notice that for cubes $y = \{(t_1, \ldots, t_n) \in \square_n\}$ and $x = \{(u_1, \ldots, u_{n-1}) \in \square_{n-1}\}$, the statement x is the ith upper face of y could be expressed $\partial^1_i y = x$ or by defining an inclusion $\delta^1_i : x \to y$,

$$\delta^1_i(u_1, \ldots, u_{n-1}) = (u_1, \ldots, u_{i-1}, 1, u_i, \ldots, u_{n-1}).$$

This dual point of view is given here:

Let $\delta^k_i : \square_{n-1} \to \square_n$, $1 \leqslant i \leqslant n$, $k = 1, 2$, be the continuous functions $(n \geqslant 1)$

$$\begin{array}{ccc} \square_n & \xleftarrow{\ \delta^0_i\ } & \square_{n-1} \\ {\scriptstyle \delta^1_i}\uparrow & & \\ \square_{n-1} & & \end{array}$$

defined by

$$\delta^k_i(t_1, \ldots, t_{n-1}) = (t_1, \ldots, t_{i-1}, k, t_i, \ldots, t_{n-1}).$$

Then,

Lemma 6.4.

$$\delta^k_i \delta^l_j = \delta^l_{j+1} \delta^k_i \quad (i \leqslant j).$$

Proof. Left to the reader.

We notice that, as indicated above, the commutator relations satisfied by the δ^k are dual to the commutator relations satisfied by ∂^k in a pre-cubical set.

Definition 6.5. Let *M* be a precubical set and let $\mathbf{R}(M) = \coprod_n M_n \times \square_n$ with disjoint sum topology induced by the discrete topology in M_n and the topology from $\mathbb{R}^n$ on $\square_n$.

Let $\equiv$ be the equivalence relation induced by the identities:

$$\forall k, i, n, \forall x \in M_{n+1}, \forall t \in \square_n, n \geqslant 0, (\partial_i^k(x), t) \equiv (x, \delta_i^k(t)).$$

Let $|M| = \mathbf{R}(M)/\equiv$ have the quotient topology. The topological space $|M|$ is called the *geometric realization* of M.

Another point of view is that M is a labelling of a subdivision of $|M|$ into cubes. An element $y_m \in M_m$ is the label of an m-cube in $|M|$. Let $p \in |M|$, then there is a minimal cube in the subdivision of $|M|$ containing p, namely the unique cube $x \times \square_k$ which has p in the interior. We call x the *carrier* of p.

6.3. An open covering of $|M|$ and a subclass of precubical sets

We want to define a local partial order on $|M|$ using the partial order on each cube and taking a "local" transitive closure. For this we need an open covering.

Definition 6.6. Let M be a pre-cubical set and y an element of M. Then $x \in M$ is a *face* of y if there exists a (possibly empty) collection of indices $k_1, \ldots, k_1$ being 0 or 1 and $l_1, \ldots, l_i$ integers such that $x = \partial_{l_1}^{k_1} \cdots \partial_{l_i}^{k_i} y$.

Definition 6.7. The open star of a point $p \in |M|$ with respect to the subdivision M is

$$St(p, M) = \{q \in |M| | \text{carrier } (p) \text{ is a face of carrier } (q)\}.$$

For a cube $x \in M_n$ we define the open star combinatorially

$$St(x, M) = \{y \in M | \exists (k_1, l_1), \ldots, (k_i, l_i), \partial_{l_1}^{k_1} \ldots \partial_{l_i}^{k_i}(y) = x\}.$$

Remark 6.8. (1) If $x \in M_n$ is the carrier of $p \in |M|$, then $St(p, M)$ is the union of the interiors of cubes in $|M|$ which are labelled by an element of $St(x, M)$.

(2) By an abuse of notation, we will omit M and write $St(p)$ and $St(x)$ if there is no risk of confusion. And moreover, when x is a cube, we will write $St(x)$ for the geometric object $St(p)$ where $p \in |M|$ and x is the carrier of p.

(3) If p and q in $|M|$ have the same carrier, i.e., if they are in the interior of the same cube, then $St(p) = St(q)$

(4) For $p \in |M|$, and $x \in M_n$ its carrier, there is a unique $t \in \mathring{\square}_n$, s.t. $p = [(x, t)] \in R(M)/\equiv$. This means that there is a canonical representative of the equivalence classes in $R(M)/\equiv$.

The stars of vertices $v_i \in M_0$ define an open covering of $|M|$ and the partial order on each of these should be the transitive hull of the partial order in the (open) cubes.

There are some obvious problems with this, such as if the upper and lower vertex of an interval $\square_1$ are identified, in which case the star of this vertex is a loop and the transitive hull of the relation relates all points to all points. To avoid this and also some less obvious problems, we restrict the class of precubical sets:

Definition 6.9. Let M be a pre-cubical set. M is a non-self-linked cubical set if for all faces x of an n-cube y, x can be written uniquely as $x = \partial_{l_1}^{k_1} \cdots \partial_{l_i}^{k_i} y$ with $k_1 \cdots k_i = 0, 1$ and $l_1 < l_2 < \cdots < l_i$ ("*canonical form*").

Remark 6.10. One may still have loops in $|M|$, but they will always consist of more than one cube. Hence each $y \in M_n$ has a full subtree of iterated boundaries with $2\binom{n}{k}$ vertices in M_{n-k}, as does indeed an n-cube and its iterated boundaries.

Given $y \in M_n$ and its tree of boundaries $\partial_{l_1}^{k_1} \ldots \partial_{l_m}^{k_m} y$ we may think of this as the n-cube $\square_n$ in $|M|$ labelled y and its iterated boundaries. To be precise, $x = \partial_i^k y$ means that the $(n-1)$-cube $\square_{n-1}$ labelled x is identified with $\{(t_1, \ldots, t_n) \in \square_n | t_i = k\}$. The commutator rules for the boundaries ensure that this works for iterated boundaries:

Lemma 6.11. *Let M be a precubical set. Let $x = \partial_{l_1}^{k_1} \ldots \partial_{l_m}^{k_m} y$. Then there is a (not necessarily unique) canonical form $\partial_{l'_1}^{k'_1} \ldots \partial_{l'_m}^{k'_m} y$ with $l'_1 < l'_2 < \cdots < l'_m$, and x is identified with $\{(t_1, \ldots, t_n) \in \square_n | t_{l'_i} = k'_i$ for $i = 1, \ldots, m\}$.*

Proof. Use the commutator relations to get a canonical form. □

When the complex is non-self-linked, the canonical form is unique. And there is another unique form:

Lemma 6.12. *Let M be a non-self-linked precubical complex, and x, y be elements of M. Suppose x is a face of y. Then x can be written uniquely as*

- $x = \partial^0_{l'_1} \cdots \partial^0_{l'_j} \partial^1_{l'_{j+1}} \cdots \partial^1_{l'_i} y$ *with* $l'_1 < l'_2 < \cdots < l'_j$ *and* $l'_{j+1} < l'_{j+2} < \cdots < l'_i$.

All other "decompositions" of x as a face of y, $x = \partial^{u_1}_{v_1} \cdots \partial^{u_i}_{v_i} y$ *verify the following: let* $K_0(x, y)$ *(respectively,* $K_1(x, y)$*) be the cardinal of the set* $\{j/1 \leqslant j \leqslant i, k_j = 0\}$ *(respectively,* $\{j/1 \leqslant j \leqslant i, k_j = 1\}$*), then* $K_0(x, y)$ *(respectively,* $K_1(x, y)$*) is also the cardinal of* $\{j/1 \leqslant j \leqslant i, u_j = 0\}$ *(respectively,* $\{j/1 \leqslant j \leqslant i, u_j = 1\}$*). Hence* $K_k(x, y)$ *depends only on x and y.*

Proof. By induction on i (the length of the decomposition). The statement about K_0 and K_1 follows from the fact that these are invariant under commutation following the commutator rules. □

6.4. The face ordering

Lemma 6.13. *Let M be a non-self-linked cubical set, x and y two of its elements. The relation x is a face of y ("*$x <_F y$*") is a partial order.*

Proof. It is reflexive indeed. □

Now, if $x <_F y$ and $y <_F x$ then $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} y$ and $y = \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} x$ by definition. So $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} x$ and $K_0(x, x) = K_0(x, y) + K_0(y, x)$ and $K_1(x, x) = K_1(x, y) + K_1(y, x)$ by Lemma 6.12. But $K_0(x, x) = K_1(x, x) = 0$ again by Lemma 6.12 so are $K_0(x, y)$, $K_1(x, y)$, $K_0(y, x)$ and $K_1(y, x)$. Therefore $x = y$.

Finally, if $x <_F y$ and $y <_F z$ then $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} y$ and $y = \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} z$. So $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} z$ and $x <_F z$.

Lemma 6.14. *Let M be a non-self-linked cubical set, x, y and z three elements of M. Then* $x <_F y <_F z$ *implies* $K_0(x, z) = K_0(x, y) + K_0(y, z)$ *and* $K_1(x, z) = K_1(x, y) + K_1(y, z)$.

Proof. If $x <_F y$ and $y <_F z$ then $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} y$ and $y = \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} z$. So $x = \partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} \partial^{k'_1}_{l'_1} \cdots \partial^{k'_j}_{l'_j} z$ and the number of k_m or k'_m equal to 0 (respectively, 1) in the decomposition above is the number of k_m equal to 0 (respectively, 1) plus the number of k'_m equal to 0 (respectively, 1), hence the result. □

The face ordering has nice properties that we will exploit later on.

Lemma 6.15. *Let M be a non-self-linked cubical set and x, y be two elements of M such that* $x <_F y$*. Then there is a projection operator (which is a dimap)* p^y_x *of cube* $|y|$ *in* $|M|$ *whose carrier is y on the cube* $|x|$ *such that, if* $x = \partial^{k_1}_{l_1} \cdots \partial^{k_j}_{l_j} y$, $p^y_x(y, \delta^{k_j}_{l_j} \cdots \delta^{k_1}_{l_1} v) = (x, v)$.

Proof. In $|M|$, points (x, t) are identified with $(y, \delta^{k_j}_{l_j} \cdots \delta^{k_1}_{l_1}(t))$. Let p be the projection from $\mathbb{R}^n$ to $\mathbb{R}^m$ (if $y \in M_n$, $x \in M_m$, we have $t \in \mathbb{R}^m$ and $\delta^{k_j}_{l_j} \cdots \delta^{k_1}_{l_1}(t) \in R^n$) which projects out coordinates $l_1, \ldots, l_j$. Then set $p^y_x(y, u) = (x, p(u))$, then $p^y_x(y, \delta^{k_j}_{l_j} \cdots \delta^{k_1}_{l_1} v) = (x, v)$ thus $p^y_x(x, v) = (x, v)$. p is continuous and monotonic so p^y_x is a dimap. □

Remark 6.16. We actually only need M to be non-singular, i.e., that $\partial^k_l x = \partial^{k'}_{l'} x \Rightarrow k \neq k'$, for the above lemma to hold.

Lemma 6.17. *Let M be a non-self-linked cubical set. Then for all $y \in M$, for all faces b and b' of y in M, we have only two possibilities,*

- *b and b' have no face in common (we write $b \cap b' = \emptyset$),*
- *or b and b' have a maximal (with respect to the partial order $<_F$) face in common (that we write $b \cap b'$), which is a face of y.*

Proof. Since M is non-self-linked and we study iterated boundaries of $y \in M_n$, we can consider this a study of the n-cube $\Box_n$ with no identifications of boundaries except the ones given by the geometry.

Two boundaries b and b' are then considered as subsets of $\{(t_1, \dots, t_n) \in \Box_n\}$. We write $b = \partial_{l_1}^{k_1} \cdots \partial_{l_j}^{k_j} y$ and $b' = \partial_{l'_1}^{k'_1} \cdots \partial_{l'_{j'}}^{k'_{j'}} y$ on canonical form. Then $b = \{(t_1, \dots, t_n) \in \Box_n | t_{l_i} = k_i \text{ for } i = 1 \dots j\}$ and $b = \{(t_1, \dots, t_n) \in \Box_n | t_{l'_i} = k'_i \text{ for } i = 1 \dots j'\}$ and the intersection $b \cap b' = \{(t_1, \dots, t_n) \in \Box_n | t_{l_i} = k_i \text{ for } i = 1 \dots j \text{ and } t_{l'_i} = k'_i \text{ for } i = 1 \dots j'\}$.

If for some $i \in \{1, \dots, j\}$ and $i' \in \{1, \dots j'\}$ $l_i = l'_{i'}$ and $k_i \neq k'_{i'}$, then $b \cap b' = \emptyset$. Otherwise, the description as a subset of $\Box_n$ gives that $b \cap b'$ is a face. □

Another useful fact, which is geometrically quite clear is the following:

Lemma 6.18. *Suppose w is an upper face of an n-cube y and that x is a face in y which has a non-empty intersection with w. Then the upper vertex of x is in w.*

Proof. Let $w = \partial_{l(y,w)}^{\mathbf{1}} y$ and $x = \partial_{l(y,x)}^{k(y,x)} y$ be the canonical representations of w and x as faces in y. Since x and w have non-trivial intersection, whenever i is a coordinate in the vector/multi-index $l(y, x)$ (denoted $i \in l(y, x)$) and $k(y, x)_i = 0$, then $i \notin l(y, w)$. Now the upper vertex v of x considered in y is $v = \partial_{l(x,v)}^{\mathbf{1}} \partial_{l(y,x)}^{k(y,x)} y$. Hence v is the point $(t_1, \dots, t_n) \in \Box_n$ where $t_i = 0$ if $i \in l(y, x)$ and $k(y, x)_i = 0$ and $t_i = 1$ else. This is in w, since $t_i = 0$ only if $i \notin l(y, w)$. □

Remark 6.19. By symmetric arguments we have: if w is a lower face of z which intersects another face x non-trivially, then the lower vertex of x is in w.

6.5. A local partial order on the geometric realization of a non-self-linked precubical complex

In this section M is a non-self-linked precubical complex with geometric realization $|M|$. As remarked above, any p in $|M|$ has a unique representative (x, t) where $x =$ carrier (p) and $t \in \overset{\circ}{\Box}_n$ (for some n). When we choose a representative of a point in $|M|$, it will always be this one.

We give a partial order $\leqslant_x$ on the open neighbourhood $U^x = St(p, M)$ of p, for all p, and we prove that the open cover of $|M|$ $\mathcal{U} = \{St(v, M) | v \in M_0\}$ by stars of vertices with the partial order $\leqslant_v$ defines a local partial order. A local po-neighbourhood of p is given by $(St(x), \leqslant_x)$, where x is the carrier.

Let $(y, u) \in U^x$, then $x = \partial_{l_1}^{k_1} \dots \partial_{l_i}^{k_i}(y)$ because $y \in St(x, M)$, so (x, t) is identified with $(y, \delta_{l_i}^{k_i} \dots \delta_{l_1}^{k_1}(t))$. This identification is unique, since M is non-self-linked, so we can define

Definition 6.20. Let $U^x = St(p, M)$, where x is the carrier of p. We define a relation on U^x:

$$(x, t) \leqslant_{U^x} (y, u) \text{ if } \delta_{l_i}^{k_i} \dots \delta_{l_1}^{k_1}(t) \leqslant u \quad \text{in } \Box_{n+i},$$

$$(y, u) \leqslant_{U^x} (x, t) \text{ if } \delta_{l_i}^{k_i} \dots \delta_{l_1}^{k_1}(t) \geqslant u \quad \text{in } \Box_{n+i}.$$

And we give a partial order on U^x: let $x \in M$ and let (z, v) be a point in U^x with carrier z. We say

$$(z, v) \leqslant_x (y, u)$$

if there exists b in the star of x and t such that

$$(z, v) \leqslant_{U^b} (b, t) \leqslant_{U^b} (y, u).$$

To see that this is in fact a partial order, we need:

Lemma 6.21. • *Suppose* $(x, t) \leqslant_{U^x} (y, u)$ *and* $x \neq y$, *then necessarily*, $x = \partial^0_{l_1} \dots \partial^0_{l_j} y$ (*where* $j \geqslant 0$). *This implies that* $K_0(x, y) > 0$ *and* $K_1(x, y) = 0$.

• *Suppose* $(y, u) \leqslant_{U^x} (x, t)$ *and* $x \neq y$, *then necessarily*, $x = \partial^1_{l_1} \dots \partial^1_{l_j} y$ (*where* $j \geqslant 0$). *This implies that* $K_0(x, y) = 0$ *and* $K_1(x, y) > 0$.

Proof. We only prove the first statement since the proof of the other is similar. $y \in St(x, M)$ so there is a collection of indices such that $x = \partial^{k_0}_{l_0} \dots \partial^{k_j}_{l_j} y$ $(j \geqslant -1)$. j cannot be equal to -1 since $x \neq y$. Suppose now that there is an index k_i $(0 \leqslant i \leqslant j)$ which is equal to one. We have $u \in \overset{\circ}{\square}_n$ and $n \geqslant 1$ (since $n = 0$ is only possible when x is a vertex and $x = y$), therefore all coordinates u_i of u are strictly less than 1, so in particular $u_{l_i} < (\delta^{k_i}_{l_i}(t))_{l_i} = 1$ by definition of the operator $\delta^{k_i}_{l_i}$. This is a contradiction with the definition of $\leqslant_{U^x}$. □

And this gives the following characterization of the partial order:

Lemma 6.22. *Suppose* $(z, v) \leqslant_x (y, u)$, *that is*, $\exists (b, t)$, $b \in St(x, M)$, $(z, v) \leqslant_{U^b} (b, t) \leqslant_{U^b} (y, u)$. *Then we have the following cases*,

(a) $b = \partial^1_{l_1} \dots \partial^1_{l_j} z$, *for some collection of indices and* $K_1(b, z) = j \geqslant 1$ $(K_0(b, z) = 0)$; $b = \partial^0_{l'_1} \dots \partial^0_{l'_{j'}} y$, *for some collection of indices and* $K_0(b, y) = j' \geqslant 1$ $(K_1(b, y') = 0)$.

(b) $b = z$, $b = \partial^0_{l'_1} \dots \partial^0_{l'_j} y$, *for some collection of indices and* $K_0(b, y) = j \geqslant 1$ $(K_1(b, y') = 0)$, *and the relation above shrinks down to* $(z, v) \leqslant_{U^z} (y, u)$.

(c) $b = y$, $b = \partial^1_{l_1} \dots \partial^1_{l_j} z$, *for some collection of indices and* $K_1(b, z) = j \geqslant 1$ $(K_0(b, z) = 0)$, *and the relation above shrinks down to* $(z, v) \leqslant_{U^y} (y, u)$.

(d) $y = z$ *and the relation above shrinks down to* $v \leqslant u$.

Proof. This is entailed by Lemma 6.21. □

We will say in the sequel that (z, b, u) is in case (a), (b), (c) or (d) according to the criteria above. In order to simplify the proofs we will write in brief (for all cases (a), (b), (c) and (d)) $b = \partial^{\mathbf{1}}_* z$ and $b = \partial^{\mathbf{0}}_* y$ where $*$ means a multi-index, $\mathbf{1} = (1, 1, \dots, 1)$, $\mathbf{0} = (0, 0, \dots, 0)$ and they may all be empty.

Now we can state,

Theorem 6.23. *The geometric realization of a non-self-linked cubical set M defines a locally po-space with covering being* $\{St(x, M)/x \in M_0\}$ *and local partial order* $\leqslant_x$ *on* $St(x, M)$.

Proof. We check that $\leqslant_x$ is a partial order indeed for all x in M. First, reflexivity is obvious.

Antisymmetry. Suppose $(z, v) \leqslant_x (y, u)$ and $(y, u) \leqslant_x (z, v)$. This means there are (b, t) and (b', t') with $b \in St(x, M)$ and $b' \in St(x, M)$ such that

$$(z, v) \leqslant_{U^b} (b, t) \leqslant_{U^b} (y, u) \leqslant_{U^{b'}} (b', t') \leqslant_{U^{b'}} (z, v),$$

$$b = \partial^{\mathbf{1}}_* z, \quad b = \partial^{\mathbf{0}}_* y, \quad b' = \partial^{\mathbf{1}}_* y, \quad b' = \partial^{\mathbf{0}}_* z$$

and moreover

$$x = \partial^*_* b, \quad x = \partial^*_* b',$$

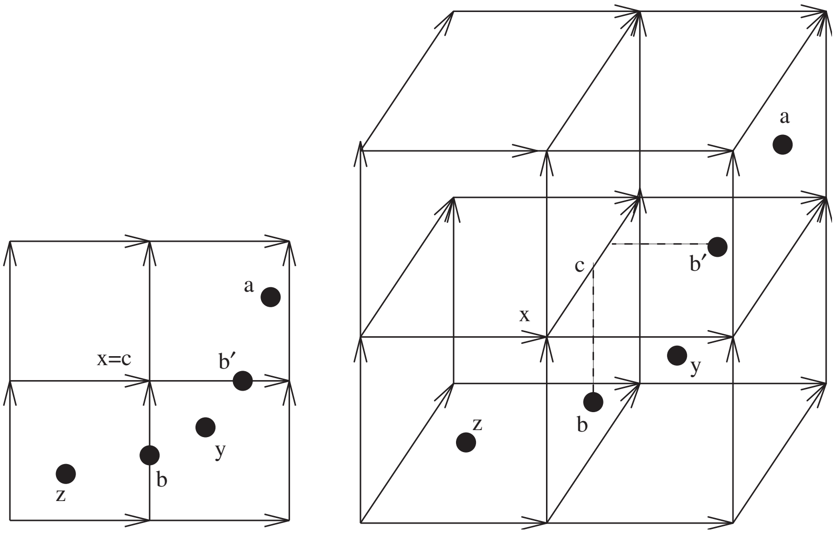


Fig. 11. Illustration of the proof.

where $*$ means a multi-index, $\mathbf{1} = (1, 1, \dots, 1)$, $\mathbf{0} = (0, 0, \dots, 0)$ and they may all be empty. Hence composing boundary maps we see that the following equalities hold:
(1) $K_0(x, z) = K_0(x, b)$ and $K_0(x, z) = K_0(x, b') + K_0(b', z)$,
(2) $K_0(x, y) = K_0(x, b) + K_0(b, y)$ and $K_0(x, y) = K_0(x, b')$,
(3) $K_1(x, y) = K_1(x, b)$ and $K_1(x, y) = K_1(x, b') + K_1(b', y)$,
(4) $K_1(x, z) = K_1(x, b) + K_1(b, z)$ and $K_1(x, z) = K_1(x, b')$.
Now 1 and 2 imply $K_0(b, y) + K_0(b', z) = 0$ and thus both are 0 which give $b = y$ and $b' = z$. Similarly 2 and 3 imply that $K_1(b', y) + K_1(b, z) = 0$ and thus $b' = y$ and $b = z$.

Transitivity. Suppose $(z, v) \leqslant_x (y, u) \leqslant_x (a, w)$. This means there are (b, t) and (b', t') with $b \in St(x, M)$ and $b' \in St(x, M)$ such that,

$$(z, v) \leqslant_{U^b} (b, t) \leqslant_{U^b} (y, u) \leqslant_{U^{b'}} (b', t') \leqslant_{U^{b'}} (a, w)$$

with

$$b = \partial^{\mathbf{1}}_* z, \quad b = \partial^{\mathbf{0}}_* y, \quad b' = \partial^{\mathbf{1}}_* y, \quad b' = \partial^{\mathbf{0}}_* a$$

and moreover

$$x = \partial^*_* b, \quad x = \partial^*_* b',$$

where $*$ means a multi-index, $\mathbf{1} = (1, 1, \dots, 1)$, $\mathbf{0} = (0, 0, \dots, 0)$ and they may all be empty. By Lemma 6.17, the intersection of b and b' is a face c of y containing x (see Fig. 11). The inclusion of c into b and b' (which are compositions of [dual] boundary operators δ^k_l) define projections p and p' of points in b and b' onto points of c by Lemma 6.15. Define $d = p(t)$. Necessarily (looking at the coordinates in y, as p is a dimap), $p(t') \geqslant d$. Now it is enough to see that

$$(z, v) \leqslant_{U^c} (c, d) \leqslant_{U^c} (a, w)$$

(hence the transitivity in that case). The first inequality is implied by the fact that $(b, t) \leqslant_{U^c} (c, d)$ since c is an upper boundary of b and d is the corresponding projection of t. The second is implied by the fact that $(c, d) \leqslant_{U^c} (b', t')$ since c is a lower boundary of b' and $p(t') \geqslant d$.

Po-neighbourhoods. To see that this defines a local partial order, we define po-neighbourhoods: let $p = (x, t) \in |M|$. Then we claim that $St(p, M)$ with partial order defined above, defines a po-neighbourhood of p:

Since $p \in St(v)$ implies $St(x) \subset St(v)$, we have to see, that when $(y, u), (z, v) \in St(x) \subseteq St(v_1) \cap St(v_2)$, then $(y, u) \leqslant_{v_1} (z, v) \Leftrightarrow (y, u) \leqslant_{v_2} (z, v)$. If x is a vertex, then $v_1 = v_2 = x$ and there is nothing to check. Suppose now, that

x is not a vertex and suppose $(y,u)\leqslant_{v_1}(z,v)$. Since $(y,u)\leqslant_{v_1}(z,v)$ if and only if $(y,(\frac{1}{2},\dots,\frac{1}{2}))\leqslant_{v_1}(z,(\frac{1}{2},\dots,\frac{1}{2}))$, we will suppose all coordinates are $\frac{1}{2}$ and write $y\leqslant_{v_1}z$.

(1) If $y\leqslant_{v_1}z$ is of type (b), (c) or (d), the inequality is induced by the Euclidean partial order on the cube in $|M|$ labelled y (in case b and d) or the cube labelled z (in case c and d), and the relation will clearly hold in $St(v_2)$ also.

(2) Suppose now $y\leqslant_{v_1}z$ is of type (a). Then there is a $w\in St(v_1)$ s.t. $y\leqslant_{U^w}w\leqslant_{U^w}z$. If $w\in St(v_2)$, we are done.
Now w is an upper face of y and a lower face of z, and v_1 is in the intersection of x,y,z and w. Hence, by 6.18 when we study x and w as faces in y, the upper vertex of x is in w. Similarly, by 6.19 from the inclusion in z, the lower vertex of x is in w, which implies that x is a face of w, so v_2 is a vertex in w. □

To give an idea of some of the locally po-spaces we can construct considering this:

Example 6.24. Let M be the cubical set

$$M_2=A,B,C,D,\quad M_1=a,b,c,d,e,f,g,h,\quad M_0=p,q,r,s,$$

$$d_1^0A=d_1^0C=a,\quad d_2^0A=d_2^0B=b,\quad d_1^0B=d_1^0D=c,\quad d_2^0D=d_2^0C=d,$$

$$d_1^1A=d_1^1D=e,\quad d_2^1A=d_2^1D=f,\quad d_1^1B=d_1^1C=g,\quad d_2^1B=d_2^1C=h,$$

$$d_1^0a=d_1^0b=d_1^0c=d_1^0d=p,\quad d_1^1a=d_1^1c=d_1^0f=d_1^0h=q,$$

$$d_1^1b=d_1^1d=d_1^0e=d_1^0g=r,\quad d_1^1e=d_1^1f=d_1^1g=d_1^1h=s.$$

Then $|M|$ is the projective plane, and one can give cubical models for projective spaces of all dimensions in the same way. Hence a local partial order does not induce an orientation in the manifold sense.

6.6. Subdivision and non-singular precubical sets

Even if the precubical set is self-linked, the above construction may give rise to a local partial order on the geometric realization: the circle constructed by identifying the vertices $\square_1$ could be subdivided into two intervals, which are then non-self-linked. The generalization of this is via barycentric subdivision which we explore in this subsection.

Definition 6.25. Let M be a pre-cubical set. M is a non-singular cubical set if for all its n-cubes x, $\partial_l^k(x)=\partial_{l'}^{k'}(x)$ implies $k\neq k'$.

So an upper face may be identified with a lower face.

Definition 6.26. Let K be a cubical set and let K' be another cubical set. Then K' is a subdivision of K if there is a dihomeomorphism $f:|K'|\to|K|$ (meaning that f and f^{-1} are dimaps) such that,

- $\forall x\in K'_n,\ \exists y\in K_n,\ f(x,\square_n)\subseteq(y,\square_m)$,
- $\forall y\in K,\ \exists x_1,\dots,x_k\in K',\ (y,\square_m)=\bigcup_{i=1,\dots,k}f(x_i,\square_{n_i})$.

Definition 6.27. The *standard n-dicube* is the topological space $\square_n$ with the covering $\mathbf{U}=\{\square_n\}$ and local partial order $\leqslant_{\square_n}$ induced from the pointwise ordering in $\mathbb{R}^n$. The n-dicube is then a local po-space.

Definition 6.28. Let M be a local po-space. A singular n-dicube is any dimap from the standard n-dicube to M.

Lemma 6.29. *Let K be a cubical set. The "barycentric subdivision" of K is defined as follows. Consider the singular n-dicubes of $|K|$, $\sigma_x:\square_n\to|K|$, $\sigma_x(t)=(x,t)$, where $x\in K$, and the 2^n functions,*

$$s_{b_1,\dots,b_n}:\square_n\to\square_n$$

for $(b_1,\dots,b_n)\in\{0,1\}^n$ with

$$s_{b_1,\dots,b_n}(t_1,\dots,t_n)=\left(\frac{t_1+b_1}{2},\dots,\frac{t_n+b_n}{2}\right).$$

Then the subset SdK of $|K|$ with,

$$(SdK)_n = \{\sigma_x \circ s_{b_1,\ldots,b_n}/x \in K, (b_1,\ldots,b_n) \in \{0,1\}^n\}$$

is a subdivision of K, called the barycentric subdivision of K.

Proof. Let $f : |SdK| \to |K|$ be defined as follows. Elements of $|SdK|$ are of the form $x = (u, v)$ with $u \in (SdK)_n$, i.e. $u = \sigma_y \circ s_{b_1,\ldots,b_n}$ $(b_i = 0, 1, y \in K_n)$, and $v \in \mathring{\square}_n$. We set $f(x) = (y, w)$ with $w \in \mathring{\square}_n$ and $w_i = \frac{v_i+b_i}{2}$ $(i = 1, \ldots, n)$.

For all such $x = (u, v) \in SdK$, $f(x) \in (y, \square_n)$ with the y defined above. Also for the same y, $(y, \square_n) = \bigcup_{b_1,\ldots,b_n=0,1} f(\sigma_y \circ s_{b_1,\ldots,b_n}, \square_n)$.

Let now $g : |K| \to |SdK|$ defined as follows. Points of $|K|$ are of the form $z = (y, w)$ with $y \in K_n$ and $w \in \mathring{\square}_n$. Set $g(z) = (\sigma_y \circ s_{b_1,\ldots,b_n}, 2w_i - b_i)$ with $b_i = \begin{cases} 0 & \text{if } 0 \leqslant w_i < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leqslant w_i \leqslant 1 \end{cases}$ $(i = 1, \ldots, n)$. Then f and g are continuous maps, $f \circ g = Id = g \circ f$. □

Lemma 6.30. *Let M be a non-singular cubical set. Then $Sd(M)$ is non-self-linked.*

Proof. Let $|M|$ be the realization of the cubical set $|M|$. Then for $x \in M_n$ the map $\sigma_x : \square_n \to |M|$ is injective on the interior of $\square_n$ by construction of $|M|$.

On the boundary of $\square_n$ there may be identifications corresponding to $\partial_i^k(x) = \partial_j^l(x)$. When $k \neq l$, this will identify boundaries of *different* cubes in the barycentric subdivision: $\partial_i^k(\sigma_x \circ s_{b_1,\ldots,b_{k-1},k,b_{k+1},\ldots,b_n}) = \partial_j^l(\sigma_x \circ s_{b_1,\ldots,b_{k-1},l,b_{k+1}\ldots,b_n})$. □

Now using both Lemma 6.30 and Theorem 6.23 we can give a local po-space structure to any non-singular cubical set.

6.7. The singular cube functor and locally po-spaces

We study the categorical properties of geometric realization: geometric realization is a functor—from non-self-linked precubical sets to local po-spaces.

We are going to construct a right-adjoint to it in this section.

Lemma 6.31. *Let $f : X \longrightarrow Y$ be a morphism between the two pre-cubical sets X and Y. Then f induces a continuous map $|f|$ from $|X|$ to $|Y|$.*

Proof. Define $\mathbf{R}(f) : \mathbf{R}(X) \longrightarrow \mathbf{R}(Y)$ by: $\mathbf{R}(f)((x,t)) = (f(x), t)$. It is obviously a continuous map.

Suppose $(x, t) \equiv (y, s)$. Then there exists $(y_1, s_1), \ldots, (y_u, s_u)$ such that $(y_1, s_1) = (x, t)$, $(y_u, s_u) = (y, s)$ and $\forall g, \exists k, j, d_j^k(y_g) = y_h$ and $s_g = \delta_j^k(s_h)$ with $h = g + 1$ or $h + 1 = g$.

We show by induction on u that $\mathbf{R}(f)((x,t)) \equiv \mathbf{R}(f)((y,s))$, thus inducing a map from $|X|$ to $|Y|$. It holds trivially for $u = 1$. To prove the induction step it suffices to see that $\mathbf{R}(f)((x,t)) \equiv \mathbf{R}(f)((y_2, s_2))$.

Suppose $\exists k, j, \partial_j^k(x) = y_2$ and $t = \delta_j^k(s_2)$. But $\partial_j^k(f(x)) = f(\partial_j^k(x))$. Thus, $\partial_j^k(f(x)) = f(y_2)$ and $t = \delta_j^k(s_2)$, which proves the result. □

When the precubical set in non-self-linked, we get a functor to local po-spaces:

Proposition 6.32. *Let $f : M \to N$ be a morphism of pre-cubical sets. Then $|f| : |M| \to |N|$ is a dimap.*

Proof. Recall that $|f|(x,t) = (f(x), t)$ for all $x \in M_n$ and $t \in \square_n$ (and for all n). Consider the "partial order" $\leqslant_{U^x}$ first. Suppose $(x,t) \leqslant_{U^x} (y,u)$. This means that $x = \partial_{l_1}^{k_1} \ldots \partial_{l_i}^{k_i}(y)$ and $\delta_{l_i}^{k_i} \ldots \delta_{l_1}^{k_1}(t) \leqslant u$. But f is a morphism of cubical sets so $f(x) = \partial_{l_1}^{k_1} \ldots \partial_{l_i}^{k_i}(f(y))$. So $(f(x), t) \leqslant_{U^{f(x)}} (f(y), u)$. Then it is easy to check that more generally $z \leqslant_x y$ implies $f(z) \leqslant_{f(x)} f(y)$. □

The combinatorial counterpart of the standard n-dicube is clearly the faces of such a cube with the boundary maps induced from the geometry.

Definition 6.33. Let $D_{[n]}$ be the free pre-cubical set generated by a unique n-cube $I_{[n]}$, i.e. the pre-cubical set of faces of $I_{[n]}$ which is formally,

- $(D_{[n]})_j = \{\partial^{k_1}_{l_1} \cdots \partial^{k_{n-j}}_{l_{n-j}}(I_{[n]}) | k_i = 0, 1, l_1 < \cdots < l_{n-j}\}$.
- The boundary operations are concatenations of the operator and of the element of $D_{[n]}$ itself.

Then,

Lemma 6.34. • *$|D_{[n]}|$ is homeomorphic to the space $\square_n$ with the usual topology in $\mathbb{R}^n$.*

• *Any morphism of pre-cubical sets $\sigma : D_{[n]} \to M$ such that $f \circ \delta^k_l = f \circ \delta^{k'}_{l'}$ implies $k = k'$ and $l = l'$ (we call this a singular cube) induces a continuous map $|\sigma| : \square_n \to |M|$.*

Not surprisingly, the geometric realization represents n-cubes by the standard n-dicube.

Proposition 6.35. • *$|D_{[n]}|$ is in fact (a refinement of) the po-space $\square_n$ with the componentwise partial order in $\mathbb{R}^n$.*

• *Any singular cube $\sigma_x : D_{[n]} \to M$ induces a dimap $|\sigma_x| : \square_n \to |M|$ (i.e. a singular n-dicube).*

Proof. The unique n-cube $I_{[n]}$ of $D_{[n]}$ is geometrically realized as the interior of $\square_n$ with the right partial order since inside the n-cube the local partial order is defined by case (d) of Lemma 6.22. Let us consider again the map

$$\begin{array}{rll} f_{I_{[n]}} : & |D_{[n]}| & \to \square_n \\ & (\partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i}(I_{[n]}), v) & \to \delta^{k_i}_{l_i} \cdots \delta^{k_1}_{l_1}(v). \end{array}$$

We know by Lemma 6.34 that f is an homeomorphism. We now have to see that f and f^{-1} are dimaps as well.

Let y be a face of $I_{[n]}$, i.e. any element of $D_{[n]}$ and x a face of y, $f_{I_{[n]}}(x,t) \leqslant f_{[n]}(y,u)$ (respectively, $f_{I_{[n]}}(y,u) \leqslant f_{I_{[n]}}(x,t)$) is equivalent to $(x,t) \leqslant_{U^x} (y,u)$ (respectively, $(y,u) \leqslant_{U^x} (x,t)$). To see this, let $y = \partial^{u_1}_{v_1} \cdots \partial^{u_j}_{v_j}(I_{[n]})$ and $x = \partial^{k_1}_{l_1} \ldots \partial^{k_i}_{l_i}(y)$. $(x,t) \leqslant_{U^x} (y,u)$ is equivalent to $\delta^{k_i}_{l_i} \ldots \delta^{k_1}_{l_1}(t) \leqslant u$ then to $f_{I_{[n]}}(y, \delta^{k_i}_{l_i} \cdots \delta^{k_1}_{l_1}(t)) \leqslant f_{I_{[n]}}(y,u)$ by monotony of f. But $f_{I_{[n]}}(x,t) = f_{I_{[n]}}(y, \delta^{k_i}_{l_i} \cdots \delta^{k_1}_{l_1}(t))$ so this is equivalent to $f_{I_{[n]}}(x,t) \leqslant f_{I_{[n]}}(y,u)$.

Let us consider now (z,v) and (y,u) be any element of $|D_{[n]}|$. Suppose $(z,v) \leqslant_x (y,u)$ for some vertex x of $D_{[n]}$. Then there exists $(b,t) \in |D_{[n]}|$ such that $(z,v) \leqslant_{U^b} (b,t) \leqslant_{U^b} (y,u)$, hence $f_{I_{[n]}}(z,v) \leqslant f_{I_{[n]}}(b,t) \leqslant f_{I_{[n]}}(y,u)$. Inversely, suppose that we have (z,v) and (y,u) such that $f_{I_{[n]}}(z,v) \leqslant f_{I_{[n]}}(y,u)$ and z and y belong to $St(x, D_{[n]})$ for some vertex x of $D_{[n]}$. Then by Lemma 6.17 there exists a maximal common face b between y and z since they are both faces of $I_{[n]}$. Furthermore this maximal face is such that $K_0(b,y) = 0$ and $K_1(b,z) = 0$, so b can be decomposed as $b = \partial^1_{f_1} \cdots \partial^1_{f_k}(y)$ and $b = \partial^0_{f'_1} \cdots \partial^0_{f'_{k'}}(z)$. So we have $u \leqslant \delta^1_{f_k} \cdots \delta^1_{f_1} p^y_b(u)$, i.e. $(y,u) \leqslant_{U^b} (b, p^y_b(u))$ (using the projection defined in Lemma 6.15).

We also have $\delta^0_{f'_{k'}} \cdots \delta^0_{f'_1} p^z_b v \leqslant v$ thus $(b, p^z_b(v)) \leqslant_{U^b} (z,v)$. Now, $p^y_b(u) \leqslant p^z_b(v)$ because otherwise, looking at the coordinates of $f(y,u)$ and $f(z,v)$ we cannot have $f(y,u) \leqslant f(z,v)$. Therefore $(y,u) \leqslant_{U^b} (b, p^y_b u) \leqslant_{U^b} (z,v)$ hence $(y,u) \leqslant_x (z,v)$.

The second statement is obvious since $|\sigma_x|$ is f_x. □

We are now ready for the definition of the singular cube functor.

Lemma 6.36. *Let $(M, \leqslant)$ be a locally partially ordered topological space. Define $S(M)$ to be the following graded set. For $n \in \mathbb{N}$, $S(M)_n$ is the set of singular n-dicubes of M together with the operators ∂^k_l such that $\partial^k_l(f) = f \circ \delta^k_l$. This gives $S(M)$ the structure of a pre-cubical set. Moreover, this is non-self-linked.*

Proof. This is mostly a standard proof [34]. We just have to check that faces of singular n-dicubes are still singular n-dicubes, i.e. that they verify the extra-condition that they are non-self-linked. This is straightforward. □

Similarly,

Lemma 6.37. *Let* $(M, \leqslant)$ *and* $(N, \leqslant)$ *be two locally partially ordered topological spaces and let* $f : M \to N$ *be a dimap. Then* $S(f) : S(M) \to S(N)$ *defined by, for all* $x : \square_n \to M \in S$, $f(x) = f \circ x : \square_n \to N$, *is a map of pre-cubical sets. S defines a functor from the category of locally partially ordered topological spaces to the (full sub-) category of (non-self-linked) pre-cubical sets.*

Proof. This is obvious (the composition of dimaps is a dimap). □

Proposition 6.38. $|.|$ *is left-adjoint to S.*

Proof. We prove that there exist two natural transformations

$$\eta : Id \to S(|.|),$$

$$\varepsilon : |S| \to Id$$

(respectively, the unit and counit of the adjunction) such that

$$S \xrightarrow{\eta S} S(|S|) \xrightarrow{S\varepsilon} S,$$

$$|\cdot| \xrightarrow{|\cdot|\eta} |S(|\cdot|)| \xrightarrow{\varepsilon|\cdot|} |\cdot|$$

are the identities.

We can first show that

$$(A): \quad M \hookrightarrow S(|M|),$$

$$(B): \quad |S(X)| \hookrightarrow X$$

in a natural manner for all M pre-cubical set and X any local po-space. We begin by (A). For all n, we have the identity arrows on $\square_n$ which induce the isomorphisms: for all x, $Id : \square_n \to (x, \square_n)$. These in turn induce injective morphisms $f_x : \square_n \to |M|$, because M is an amalgamated sum of the $(x, \square_n)$. The $(f_x)_x$ form a subset N of $S(|M|)$. It is an easy exercise to show that N is closed under the action of the δ_i^k. Thus N is a subpre-cubical set of $S(|M|)$. The naturality of the inclusion arrow $M \hookrightarrow S(|M|)$ is most obvious. This defines what is to be the unit of the adjunction.

Now, we come to (B). Elements of $S(X)_n$ are $f : \square_n \to X$. Now, $|S(X)|$ is an amalgamated sum of $(x, \square_n)$, $x \in S(X)_n$. The x induce on $\coprod_x (x, \square_n)$ and then on $|S(X)|$ an injective morphism in the category of local po-spaces. It is an easy exercise to show that these arrows are natural in X. This defines what is to be the counit of the adjunction.

Then, we have to verify that two compositions of natural transformations are the identity. This is easy verification. □

6.8. Combinatorial dihomotopy

We introduce combinatorial dihomotopy and prove that a combinatorial dihomotopy of dipaths gives rise to an ordinary dihomotopy in the geometric realization. The other implication that dihomotopic geometric realizations of combinatorial dipaths are in fact combinatorially dihomotopic is also true, but is not proved in this paper.

Definition 6.39. Let N be a cubical set. A dipath in N is any sequence $p = (p_1, \ldots, p_k)$ of elements of N_1 such that for all i, $1 \leqslant i < k$, $\partial_1^1(p_i) = \partial_1^0(p_{i+1})$. $\partial_1^0(p_1)$ is the initial point of p. $\partial_1^1(p_k)$ is the final point of p.

Definition 6.40. Let N be a cubical set and p, q two dipaths in N with the same initial and final points. We say that p and q are elementary dihomotopic if there exists A in N_2, k and j in $\mathbb{N}$ such that,

(1) $p = (p_1, \ldots, p_k)$ and $q = (q_1, \ldots, q_k)$,

(2) for all i, $1 \leqslant i < j$, $p_i = q_i$, and for all i, $j+1 < i \leqslant k$, $p_i = q_i$,
(3) $p_j = \partial_1^0(A)$, $p_{j+1} = \partial_2^1(A)$, $q_j = \partial_2^0(A)$, $q_{j+1} = \partial_1^1(A)$.

Dihomotopy of (cubical) dipaths is the reflexive and transitive closure of elementary dihomotopy.

Using the local po-space structure defined in the previous theorems, we have also the following link with the combinatorial structure of M.

Proposition 6.41. • *Any combinatorial dipath p in M induces a* (*topological*) *dipath $|p|$ in $|M|$.*

• *Any combinatorial dihomotopy between two paths p and q in M induces a* (*topological*) *dihomotopy between $|p|$ and $|q|$.*

Proof. Let $p = (p_1, \ldots, p_k)$ be a dipath in N. Let $x : I \to |N|$ defined by, $\forall i$, $0 \leqslant i \leqslant k-1$, $\forall t \in I$, $\frac{i}{k} \leqslant t \leqslant \frac{i+1}{k}$, $x(t) = (p_{i+1}, k(t - \frac{i}{k})) \in |N|$.

Then x is a dimap (the local partial order in $|N|$ being defined as in Theorem 6.23),

- around each point in $(p_i, \mathring{\square}_1)$ (for some i), the local partial order is the same as the one in I,
- around each "glueing point" $\partial_0^1(p_i)$, the partial order is the same as the one in I since $\partial_0^0(p_{i+1}) = \partial_0^1(p_i)$.
- also, for all t such that $\frac{i}{k} < t < \frac{i+2}{k}$, $x(t) \in St(\partial_1^0(p_{i+1}), \square_0) = St(\partial_1^1(p_i), \square_0)$.

We set $|p| = x$.

It is enough to prove now that if p and q are elementary dihomotopic, then $|p|$ and $|q|$ are dihomotopic. We suppose that we have $A \in N_2, k, l \in \mathbb{N}$ such that $p = (p_1, \ldots, p_k), q = (p_1, \ldots, p_{l-1}, q_l, q_{l+1}, p_{l+2}, \ldots, p_k)$ and $\partial_0^0(A) = p_l$, $\partial_1^0(A) = q_l$, $\partial_0^1(A) = q_{l+1}$, $\partial_1^1(A) = p_{l+1}$. Now, define $H(\alpha, t)$ for $\alpha \in I$, $t \in I$ to be the map,

- for $0 \leqslant t \leqslant \frac{l-1}{k}$ and $\frac{l+1}{k} \leqslant t \leqslant 1$, $H(\alpha, t) = |p|(t) = |q|(t)$,
- for $\frac{l-1}{k} \leqslant t \leqslant \frac{l}{k}$, $H(\alpha, t) = (A, (\alpha(kt - l + 1), (1-\alpha)(kt - l + 1))) \in |N|$,
- for $\frac{l}{k} \leqslant t \leqslant \frac{l+1}{k}$, $H(\alpha, t) = (A, ((1-\alpha)(kt - l) + \alpha, \alpha(kt - l) + 1 - \alpha)) \in |N|$.

Then H is the desired dihomotopy between $|p|$ and $|q|$. □

For the converse to this proposition, see [19].

7. Two-phase locking is safe: a dihomotopy proof

7.1. Introduction

Let us take a simple example, first given in [48]. Consider a distributed database in which transactions T_1 to T_n access shared variables a, b, ... using locks: *Pa* to lock the exclusive access to a and *Va* to unlock a so that other transactions can use a. This is the same language that we used up to now.

Let us now have a look at the so-called "serialization" problem. Consider the following two transactions `R` and `S` put in parallel, where we have put actual assignments of the shared variables, when properly locked:

```
R: P A; P B;              S: P B;
   A:= B+1;                  B:= B+2;
   V A; V B;                 V B;
   P B;                      P A; P B;
   B:= 3;                    A:= 2*B;
   V B;                      P A; P B;
```

When beginning with the initial values `A=0`, `B=0` we can get the following values:

```
R: A=1     R: A=1    R: A=1    S: B=2    S: B=2    S: B=2
R: B=3     S: B=2    S: B=2    R: A=3    R: A=3    S: A=6
S: B=5-    R: B=3-   S: A=4-   S: A=4-   R: B=3-   R: A=3-
S: A=10-   S: A=6-   R: B=3-   R: B=3-   S: A=6-   R: B=3-
```

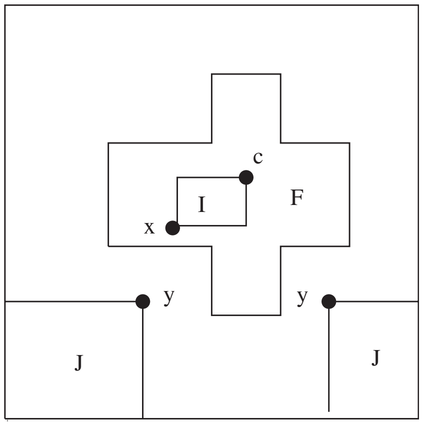


Fig. 12. Sets F, $I(\mathbf{x}, \mathbf{c})$, $K(\mathbf{y}, \mathbf{c})$.

Only the first trace (with result `A=10, B=5`) and the last trace (with result `A=3, B=3`) are correct. The other traces are interferences: as a matter of fact we want that all the execution traces give the same result as a sequential trace, i.e., `R` then `S` or `S` then `R` in their totality. It is the property called *serializability*.

A solution is given by the two-phase locking protocol: all processes P accessing to a database should first do all the lock operations then the computation then all the unlock operations. The same operations programmed using this protocol are the following:

```
R:                       S:
   P A;                     P A;
   P B;                     P B;
   A:= B+1;                 B:= 2;
   B:= 3;                   A:= 2*B;
   V B;                     V B;
   V A;                     V A;
```

It is well-known, that such a "two-phase locking" protocol is a scheduling strategy that ensures that a concurrent program has the same effect as a serial execution of the individual programs. We will present a geometric proof of the serializability of the two-phase protocol in the spirit of Gunawardena [48], but directly based on the directed homotopy theory concepts we have developed.

The "two-phase locking" protocol is a scheduling strategy that ensures that a concurrent program has the same effect as a serial execution of the individual programs as explained above.

The aim of this section is twofold. First of all, we want to give a modification of Gunawardena's reasoning in the framework of the present paper. Our proof is certainly more technical, but it seems to have several advantages: first of all, we avoid Gunawardena's "wobbling" problems (cf. [48, p. 189]): in his construction, he has to consider intermediate paths that are *not* dipaths—and to replace them by such. Secondly, our proof does not only work in the case of semaphore programs, but for general "mutual exclusion" programs—a fixed number $a \geqslant 1$ of transactions can acquire a lock to the same shared object at the same time.

After this paper had been written, Grandis proved in [46] an oriented version of the Seifert–van Kampen theorem (see also a weaker version, proven by the third author in [41]). This theorem allows, at least in principle, to calculate the dihomotopy sets $\pi_1(X; x_0, x_1)$ between certain start and end points from the dihomotopy sets of pieces of X and to their intersections. Our proof relies on a clever choice of such pieces to make these calculations work in practice.

7.2. Blockwise starshaped sets

First, we modify the concept of a "star-shaped" set in a vector space (used in [48]) in the presence of a partial order:

Definition 7.1. (1) For $x, c \in \mathbf{R}$ let $I(x, c)$ denote the interval $[x, c] \cup [c, x]$ between x and c.

(2) For $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbf{R}^n$ let $I(\mathbf{x}, \mathbf{c}) = \prod I(x_i, c_i)$, cf. Fig. 12.

(3) Let $\mathbf{c} \in F \subset \mathbf{R}^n$. The set F is called *blockwise starshaped with respect to* $\mathbf{c}$ if and only if $I(\mathbf{x}, \mathbf{c}) \subset F$ for every $\mathbf{x} \in F$.

Since the "block" $I(\mathbf{x}, \mathbf{c})$ is convex, a set that is blockwise starshaped with respect to $\mathbf{c}$ is also starshaped with respect to $\mathbf{c}$ in the classical sense.

Example 7.2. (1) An n-cube R (n-rectangle in [21]) is blockwise starshaped with respect to every point in R. A Euclidean ball $\{\mathbf{x} \in \mathbf{R}^n \mid |\mathbf{x} - \mathbf{c}| \leqslant r\}, r > 0$, is blockwise starshaped only with respect to its centre $\mathbf{c}$. The triangle $T = \{(x_1, x_2) \in \mathbf{R}^2 \mid x_i \geqslant 0, x_1 + x_2 \leqslant 1\}$ is starshaped with respect to every of its points. It is blockwise starshaped only with respect $(0, 0)$.

(2) A *union* F of n-cubes is starshaped with respect to every point in their intersection. The forbidden region in a process graph is modelled by such a union of n-cubes. It has a non-empty "central" intersection if it is a model of a two-phase locked transaction system.

What are the properties of *complements* of blockwise starshaped sets? Let $I = [a, b],\ F \subset I^n$ and $X = I^n \setminus F$.

Definition 7.3. (1) For $x, c \in I = [a, b]$ define the interval $K(x, c) \subset I$ by

$$K(x, c) = \begin{cases} [a, x], & x < c, \\ [x, b], & x > c, \\ [a, b], & x = c. \end{cases}$$

(2) For $\mathbf{x} = [x_1, \ldots, x_n], \mathbf{c} = [c_1, \ldots, c_n] \in \mathbf{R}^n$ let $K(\mathbf{x}, \mathbf{c}) = \prod K(x_i, c_i)$, cf. Fig. 12.

Lemma 7.4. *Let $F \subset I^n = [a, b]^n$ be blockwise starshaped with respect to $\mathbf{c} \in F$. Let $X = I^n \setminus F$. Then $K(\mathbf{y}, \mathbf{c}) \subset X$ for every $\mathbf{y} \in X$.*

Proof. Assume $\mathbf{x} \in K(\mathbf{y}, \mathbf{c}) \cap F$. Then $\mathbf{y} \in I(\mathbf{x}, \mathbf{c}) \subset F$. Contradiction! $\square$

7.3. Partitions and contractions

Suppose $\mathbf{c} \in \overset{\circ}{F} \subset I^n$ and F is blockwise starshaped with respect to $\mathbf{c}$. We want to study dipaths in $X = I^n \setminus F$ from $\mathbf{a} = (a, \ldots, a)$ to $\mathbf{b} = (b, \ldots, b)$. In order to get formulas that are easy to verify and to overlook, we apply a dihomeomorphism $\Psi : [a, b]^n \to [-1, 1]^n$ with $\Psi(\mathbf{c}) = \mathbf{0}$. This dihomeomorphism should be chosen as a product of maps that are increasing in each coordinate. This ensures that $\Psi(F)$ is blockwise starshaped with respect to $\mathbf{0}$. For $\mathbf{c} = \mathbf{0}$, we can describe the I- and K-sets above as follows:

$$I(\mathbf{x}, \mathbf{0}) = \{(y_1, \ldots, y_n) \in I^n \mid \operatorname{sgn}(x_i) = \operatorname{sgn}(y_i) = \operatorname{sgn}(x_i - y_i)\};$$

$$K(\mathbf{x}, \mathbf{0}) = \{(y_1, \ldots, y_n) \in I^n \mid x_i \neq 0 \Rightarrow \operatorname{sgn}(x_i) = \operatorname{sgn}(y_i) = \operatorname{sgn}(y_i - x_i)\}$$

using a subdivision of I^n with respect to $\mathbf{0}$. The decomposition of the interval $I^i = I^i_{-1} \cup I^i_1 := [-1, 0] \cup [0, 1]$ induces a decomposition of I^n into 2^n sub-n-cubes $I_\Delta = I_{\delta_1 \ldots \delta_n} = \prod I^i_{\delta_i}$. There is an obvious partial order between those subcubes giving rise to $n!$ directed paths from $I_{-1 \ldots -1}$ to $I_{1 \ldots 1}$. We need the following subsets of I_Δ:

Definition 7.5. Given $\Delta \in \{-1, 1\}^n$ with $\delta_i = 1$ and $\delta_j = -1$. Then
(1) $I^{i,j}_\Delta = \{\mathbf{x} \in I_\Delta \mid (x_i, x_j) \neq (0, 0)\}$.
(2) $I^{i,j}_{\Delta 0} = \{\mathbf{x} \in I^{i,j}_\Delta \mid x_i = 0\},\ I^{i,j}_{\Delta 1} = \{\mathbf{x} \in I^{i,j}_\Delta \mid x_j = 0\}$.
(3) $J^{i,j}_\Delta = \{\mathbf{x} \in I_\Delta \mid x_k = \delta_k,\ k \neq i, j;\ x_j = -1 \text{ or } x_i = 1\}$.
(4) The latter is a 1-complex with endpoints $\mathbf{p}^{i,j}_{\Delta 0}$ and $\mathbf{p}^{i,j}_{\Delta 1}$ with
$(\mathbf{p}^{i,j}_{\Delta *})_k = \delta_k,\ k \neq i, j,\ (\mathbf{p}^{i,j}_{\Delta 0})_i = 0,\ (\mathbf{p}^{i,j}_{\Delta 0})_j = -1,\ (\mathbf{p}^{i,j}_{\Delta 1})_i = 1,\ (\mathbf{p}^{i,j}_{\Delta 1})_j = 0.$

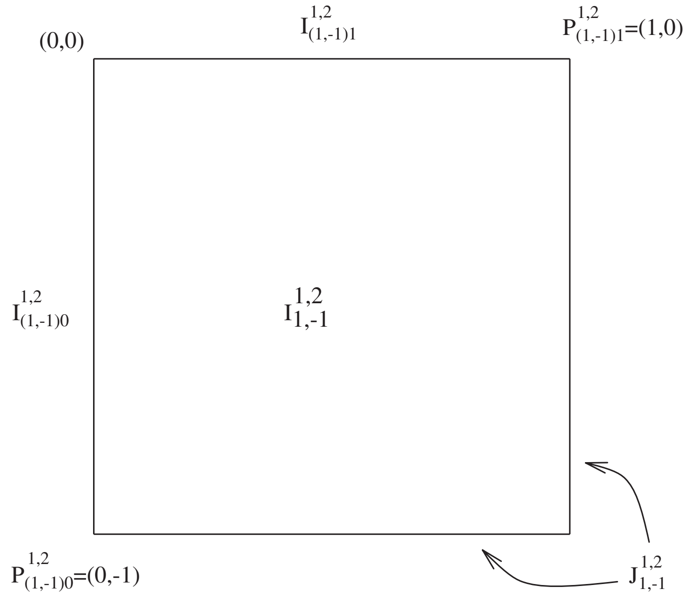


Fig. 13. Subrectangle with faces.

In fact, $I_\Delta^{i,j}$ is one of the 2^n sub-n-cubes mentioned above with a 2-codimensional face removed; $I_{\Delta 0}^{i,j}$ and $I_{\Delta 1}^{i,j}$ represent two of its faces; $J_\Delta^{i,j}$ represents a (totally ordered) 1-complex in its boundary from $\mathbf{p}_{\Delta 0}^{i,j}$ to $\mathbf{p}_{\Delta 1}^{i,j}$ (Fig. 13).

Next comes a definition of a *dimap*, cf. Definition 3.7,

$$\Phi^{\Delta,i,j} = (\Phi_1^{\Delta,i,j}, \dots \Phi_n^{\Delta,i,j}) : (I_\Delta^{i,j}; I_{\Delta 0}^{i,j}, I_{\Delta 1}^{i,j}) \to (J_\Delta^{i,j}, \mathbf{p}_{\Delta 0}^{i,j}, \mathbf{p}_{\Delta 1}^{i,j}).$$

Definition 7.6. (1) $\Phi_k^{\Delta,i,j}(\mathbf{x}) = \delta_k$ for $k \neq i, j$.

(2) $\Phi_i^{\Delta,i,j}(\mathbf{x}) = \frac{x_i}{-x_j}$ and $\Phi_j^{\Delta,i,j}(\mathbf{x}) = -1$ for $x_i \leqslant -x_j$.

(3) $\Phi_i^{\Delta,i,j}(\mathbf{x}) = 1$ and $\Phi_j^{\Delta,i,j}(\mathbf{x}) = \frac{x_j}{x_i}$ for $x_i \geqslant -x_j$.

There is a directed version of deformation retracts (cf. [16]) in the po-space environment:

Definition 7.7. Let $A \subset X$ denote an inclusion of two po-spaces. The subspace A is called a *strong deformation di-retract* of X if there exists a dimap $\Phi : X \to A$ restricting to the identity on A and a dihomotopy $H : X \times I \to X$ between Φ and the identity map on X which restricts to the trivial homotopy on A: $H_t|A = id,\ t \in I$.

Considering the map $\Phi^{\Delta,i,j}$ above, we can then show:

Proposition 7.8. (1) *The* 1-*dimensional subcomplex* $J_\Delta^{i,j}$ *is a strong deformation di-retract of* $I_\Delta^{i,j}$.

(2) *Moreover*, $J_\Delta^{i,j} \cap X = J_\Delta^{i,j} \setminus F$ *is a strong deformation di-retract of* $I_\Delta^{i,j} \cap X = I_\Delta^{i,j} \setminus F$.

(3) *Let* $\Delta = (\delta_1, \dots, \delta_n)$ *and* $\Delta' = (\delta'_1, \dots, \delta'_n)$ *with* $\delta'_k = \delta_k$ *for* $k \neq j,\ \delta_j = -1,\ \delta'_j = 1,\ \delta_l = \delta'_l = -1$ *denote two successive subcubes. Then, the maps* $\Phi^{\Delta,i,j}$ *and* $\Phi^{\Delta',j,l}$ *agree on the intersection of their domains of definition. Likewise for the homotopies in the deformation retractions from* (1) *above*.

Proof. (1) The dimap property depends only on the (x_i, x_j)-coordinates. In the projection on the (x_i, x_j)-plane, the map $\Phi^{\Delta,i,j}$ "stretches" every oriented wedge (like the stipled one in Fig. 14) out on the boundary. In particular, points under/over the "antidiagonal" $\{x_i = -x_j\}$ are mapped to points in subsequent 1-simplices. It is elementary to see that the restrictions of $\Phi^{\Delta,i,j}$ to points "under", resp. "over" that antidiagonal are dimaps and that $\Phi^{\Delta,i,j}$ fixes $J^{\Delta,i,j}$.

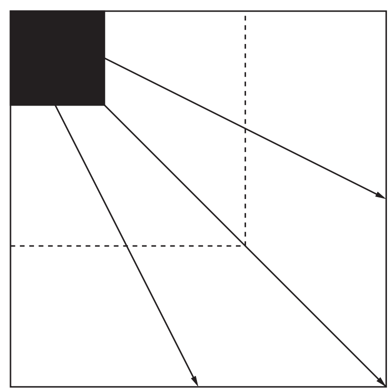


Fig. 14. The dimap $\Phi^{\Delta,i,j}$.

The (linear) self-dihomotopy $H^{i,j}$ on $I_{\Delta}^{i,j}$ given by $H^{i,j}(\mathbf{x},t) := (1-t)\mathbf{x}+t\Phi^{\Delta,i,j}(x)$ connects $\mathbf{x}$ to $\Phi^{\Delta,i,j}(x)$ along a line; in particular, all of the maps H_t are di-maps.

(2) We have to show that the $\Phi^{\Delta,i,j}$ preserves $I_{\Delta}^{i,j} \setminus F$. By definition, $\Phi_k^{\Delta,i,j}(\mathbf{x}) \geqslant x_k$ for $\delta_k = 1$ and $\Phi_k^{\Delta,i,j}(\mathbf{x}) \leqslant x_k$ for $\delta_k = -1$. Use Lemma 7.4. Similarly for the dihomotopy $H^{i,j}$.

(3) On the intersection $I_{\Delta 1}^{i,j} \cap I_{\Delta 0}^{j,k}$, the maps $\Phi^{\Delta,i,j}$ and $\Phi^{\Delta',j,l}$ are *constant* with value $\mathbf{p}_{\Delta 1}^{i,j} = \mathbf{p}_{\Delta 0}^{j,l}$. $\square$

We need special care for the minimal and maximal subrectangles I_{Δ_-}, resp. I_{Δ_+} corresponding to $\Delta_- = (-1, \dots, -1)$ and $\Delta_+ = (1, \dots, 1)$. Let $I_{\Delta_\pm}^i = \{\mathbf{x} \in I_{\Delta_\pm} | x_k = \pm 1,\ k \neq i;\ \ x_i = 0\}$, $J_{\Delta_\pm}^i = \{\mathbf{x} \in I_{\Delta_\pm} | x_k = \pm 1,\ k \neq i\}$, and $(\mathbf{p}_{\Delta_\pm}^i)_k = \pm 1,\ k \neq i,\ (\mathbf{p}_{\Delta_\pm}^i)_i = 0$. Then, we define

$$\Phi^{\Delta_\pm,i} : (I_{\Delta_\pm}; I_{\Delta_\pm}^i) \to (J_{\Delta_\pm}^i, \mathbf{p}_{\Delta_\pm}^i), \quad \Phi^{\Delta_\pm,i}(\mathbf{x}) = \begin{cases} \pm 1, & k \neq i, \\ x_i, & k = i. \end{cases}$$

Proposition 7.9. *The analogue of Proposition* 7.8 *holds in these two cases as well.*

To formulate the corollary, we need some notation about sequences, resp. unions of the subcubes I_Δ above: let $\sigma \in \Sigma_n$ denote a permutation of the integers $\{1, \dots, n\}$. Let $\Delta_\sigma(k) = (\delta_1, \dots, \delta_n)$ be given by $\delta_i = 1$ if $i \in \{\sigma(1), \dots \sigma(k)\}$ and $\delta_i = -1$ otherwise. Then $\Delta_\sigma(0), \dots, \Delta_\sigma(n)$ is an ascending *chain* of subcubes from Δ_- to Δ_+. Let

$$I_\sigma = I_{\Delta -} \cup \bigcup_{k=1}^{n-1} I_{\Delta_\sigma(k)}^{\sigma(k),\sigma(k+1)} \cup I_{\Delta +}, \quad J_\sigma = J_{\Delta -}^{\sigma(1)} \cup \bigcup_{k=1}^{n-1} J_{\Delta_\sigma(k)}^{\sigma(k),\sigma(k+1)} \cup J_{\Delta +}^{\sigma}(n) \tag{2}$$

denote a union of $(n+1)$ subcubes (without certain 2-dimensional faces) in an ascending chain, resp. a *totally ordered* 1-dimensional subcomplex in the boundary. Glueing the dimaps and dihomotopies from Propositions 7.8 and 7.9 together, we obtain

Corollary 7.10. (1) *The* 1*-complex J_σ is a strong deformation di-retract of the complex I_σ.*

(2) *If* $\mathbf{0} \in \mathring{F}$ *and F is blockwise starshaped with respect to* $\mathbf{0}$, *then J_σ is a strong deformation di-retract of the complex $I_\sigma \setminus F$.*

Remark 7.11. A po-space with a totally ordered 1-subcomplex (dihomeomorphic to an interval) as a strong deformation di-retract is the analogue to a *contractible* space in ordinary topology. Hence, one might call such a po-space *dicontractible*. The strategy of the proof was thus to subdivide the underlying po-space into dicontractible pieces with control on the intersection. Applying Grandis oriented version [46] of the Seifert–van Kampen theorem (with the chosen subdivision) makes it in fact possible to calculate the *fundamental category* [46,43] of the space $X = I^n \setminus F$.

7.4. Application to dipaths and serializability

As an application, we obtain the following result about dipaths generalizing Gunawardena's result [48]. Remark that the forbidden region in a process graph is blockwise starshaped with respect to a central point according to Example 7.2. Remark also, that a dipath on the 1-skeleton of the boundary of a hypercube I^n corresponds exactly to a serial execution.

Theorem 7.12. *Let* $F \subset I^n = [-1, 1]^n$ *such that* $\mathbf{0} \in \overset{\circ}{F}$ *and F is blockwise starshaped with respect to* $\mathbf{0}$. *Every dipath in* $X = I^n \setminus F$ *from* $-\mathbf{1} = (-1, \ldots, -1)$ *to* $\mathbf{1} = (1, \ldots, 1)$ *is dihomotopic to a dipath on the* 1-*skeleton* $(\partial I^n)_1$ *of the boundary* $\partial I^n = \{(x_1, \ldots, x_n) | \exists k : x_i = \pm 1,\ i \neq k\}$ *of* I^n.

Proof. Let $\Sigma = \{(x_1, \ldots, x_n) \in I^n | \ x_i = x_j = 0 \text{ for some } 1 \leqslant i < j \leqslant n\} \subset I^n$ denote the "singular set". Every dipath in $I^n \setminus F$ *avoiding* Σ is contained in one of the complexes I_σ for a permutation $\sigma \in \Sigma_n$ and thus dihomotopic in $X = I^n \setminus F$ to a dipath in $J_\sigma \subset (\partial I^n)_1$ from $-\mathbf{1}$ to $\mathbf{1}$ by Corollary 7.10.

It remains to handle dipaths intersecting Σ. Since a dipath has dimension 1 and the singular set Σ has codimension 2, we can apply a (locally linear) transversality argument to see that every dipath in *X* is dihomotopic to one avoiding Σ. Alternatively, We may give *X* the structure of a cubical complex in such a way that no 1-cube intersects Σ, and then argue that every dipath is dihomotopic to a dipath on the 1-skeleton of *X*, cf. Section 6. If *X* is the "forbidden region" corresponding to "mutual exclusion" in a process graph (cf. [48,21]), the subdivision will have that non-intersection property by construction. □

8. Open mathematical problems

This paper has the purpose of stating some of the most basic definitions and properties in directed homotopy theory. Since a preprint version of this paper appeared in 1999, a lot of progress has been made, but there are still a lot of challenging problems. Let us mention some of them, with pointers to the literature that has appeared meanwhile:

(M1) (Local) po-spaces do not right away have nice categorical properties. Cartesian closedness is important. Several alternative ways have been proposed, notably the category of flows by Gaucher [28] and the category of *d*-spaces by Grandis [46]. Gaucher has also defined and investigated relations between flows that capture the analogue of homotopy equivalence in the classical setting.

(M2) The dicomponents of this paper have obtained a more solid foundation starting from the fundamental *category* (dipaths modulo dihomotopy) of a po-space in which a certain system of morphisms gets inverted. In particular, the dicomponents are interpreted as the objects of a category, which can be viewed as a graph with commutation relations. For definitions and properties, cf. [43,68,20]. The investigation and exploitation of naturality properties for these categories presents still a major challenge.

(M3) Compositionality is also a major mathematical issue: suppose *X* is a po-space and that the fundamental categories (cf. (M2)) are well-understood for the pieces A, B and their intersection $A \cap B$. What can one infer about the fundamental category of *X*? What about the categories of dicomponents? The first question has been solved by Grandis [46] in a Seifert–van Kampen theorem for *d*-spaces, cf. (M1) above, and by Goubault, under some restricted circumstances, for local po-spaces in [41]. Consequences for the dicomponents (M2) still have to be sorted out.

(M4) How can one define directed *homology* and exploit it for calculations? What are the relations between dihomotopy and dihomology? Here also, a categorical framework has to be established. This is work in progress; for a very first attempt cf. [34]. More refined theories have been mostly defined by Gaucher, see for instance [27] or [30,29].

(M5) *Higher* dihomotopy (with a suitable structure) has to be investigated more closely, having a potential of giving more subtle invariants of lpo-spaces. For some first attempts, cf. [76,43,68].

(M6) Many po-spaces arising from concurrency set-ups are equipped with a certain degree of *symmetry*. Hence, a group equivariant dihomotopy theory ought to be established.

9. Open computer-scientific problems

We tried in this paper to motivate the mathematics by some examples and concepts taken from several areas of computer science. Some new applications, or new results might be derived from this theory, e.g.:

(CS1) How can we exploit (M2) so that we can derive the "essential schedules" of a concurrent system? A nice application has already been made for a small subset of the set of diconnected components, namely the unsafe region and the unreachable region in simple cases, see [21]. A generalization would be important since these schedules describe fine (safety) properties of concurrent systems (about the possible orderings on accesses to shared resources for instance) that for instance encompass serializability issues. (M3) would make it possible to attack the difficult problem of reasoning about schedules of a system compositionally, i.e. inductively on the knowledge of its subparts. This would be of a great algorithmic value for program analysis for instance. Since we wrote this paper, hints about how to do this have now been published, see in particular [44,20,68].

(CS2) (M4) and (M5) would make it possible to consider more refined properties of fault-tolerant systems and make a direct link with the work of Herlihy, Rajsbaum and Shavit. Basically the aim is to give the semantic foundations to a computability and complexity theory for fault-tolerant distributed systems. (M1) would help to describe the basic fundamental synchronization models that one can imagine for concurrent systems.

(CS3) (M4) would make it possible to have good algorithms (from linear algebra) giving compatible information about schedules of concurrent programs. This was already hinted in [35].

(CS4) It is an old question [77] as to which monoids can be represented by finite canonical term rewriting systems, which is quite related to our notions and techniques. More generally, we could ask what can be computed in more general structures [62]. This is very much linked to (CS2).

(CS5) We think that ideas from this framework might help to design better algorithms for distributed databases schedulers (like better "path-pushing" algorithms) and for micro-instructions schedulers.

10. Related work

A typical problem in concurrency theory and in distributed systems theory is to distinguish non-determinism from "true" concurrency. In some cases, we would like to be able to specify the actual use of shared resources of a parallel program, like, how many processors are busy or idle, or should a process wait for a shared variable? As can be seen in Fig. 15, the parallel execution of a with b is identified with the non-deterministic choice between $a.b$ and $b.a$, called interleaving of a and b, in ordinary transition systems semantics. These two should denote entirely different behaviours in fact. The former should indicate that actions a and b can overlap in time, whereas the latter should prescribe that a and b are conflicting operations and that one has to be executed before the other. This is central to the discussion of mutual exclusion properties for instance. Our solution is to explicitly fill in the boundary of a square if we want to specify true concurrency instead of mutual exclusion, or to fill in the boundary of a cube when we want to specify that three processes can be active at the same time and not just two etc.

A first solution has been proposed in slightly different forms, asynchronous transition systems [4], concurrent automata [78], transition systems with independence [81], etc. These solutions very often consist in adding an independence relation between atomic actions involved in an ordinary transition system. In these semantics, the interleaving of two independent actions means their execution in parallel, whereas the interleaving of two non-independent actions means their execution in mutual exclusion. Unfortunately, in these models, it is difficult to speak in a natural manner about more complex mutual exclusion properties, like shared resources that one can access in parallel n times but not $n+1$ ($n \geqslant 2$), nor of the number of busy processes at some instant in a distributed system. For instance, given three actions a, b and c, should we understand a and b independent, b and c independent and c and a independent as the same as a, b and c are independent? This probably is not true if you are considering a, b and c as the requests to print a (different) file on a printer addressed to a server of printers. If the server controls two printers then, on the program side, all pairs of actions are independent, whereas three requests cannot be treated at the same time. If the server controls three printers, all three requests are independent. This is dealt with in the notion of a counting semaphore, initialized to 2 (or 2-semaphore, i.e. that can be shared by 2 but not by 3 processes at a time). The geometric semantics of such an object is pictured in Fig. 16.

These points are actually crucial in a number of applications. In particular, concerning the proof of parallel programs on constrained architectures, or the proof of fault-tolerant distributed protocols in which the number of

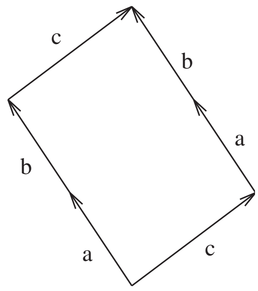


Fig. 15. Interleaving of *A* and *c*.

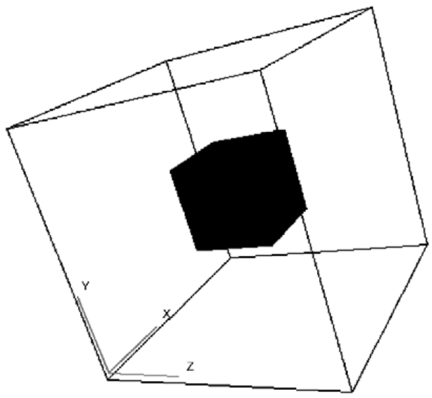


Fig. 16. A 2-semaphore.

busy processors is of primary importance (see Sections 9 and 10), or for optimization of the use of shared resources.

This was first proposed in [65,32], and further treated in [34]. Now we can understand the problem of pools of printers explained briefly above as follows. If we have three printers in the pool then we allow traces (i.e. paths) that are inside the cube delimited by the three actions a, b and c whereas if we have only two printers, we do not allow them, but only those which are on the boundary of the cube (which is a closed surface).

Most of these aspects are dealt with traditionally by resorting to Petri nets. But, even if the operational meaning of Petri nets is simple, it is not of the same nature as for transition systems. For instance, Petri nets are difficult to use in a compositional way, which is not the case for transition systems [2].

Slightly different geometric models have been used in the work of Herlihy, Shavit and Rajsbaum on fault-tolerant protocols for distributed systems. This has given numerous results in the field of distributed systems. It is proved for instance in [53] that in a shared-memory model with single reader/single writer registers providing atomic read and write operations, k-set agreement requires at least $\lfloor f/k \rfloor + 1$ rounds where f is the number of processes that can fail. General tests are also given for solving t-resilient problems. Not only impossibility results can be given but also constructive means for finding algorithms derive from this work (see for instance [54]). We refer the reader to other articles in this area, in particular, the book on distributed algorithms [60], other articles by Herlihy et al. [50–52], some slightly different methods, still geometric in nature, in [3,5,6,11,24] (which originated this field of research, starting with graph theoretical arguments), [69]. Several ideas about classifying data structures according to what protocols they manage to solve are described in [56,57,71]. This should be related to problem (CS2). Some links with directed homotopy have been hinted by one of the co-authors, see [37–39].

In concurrent databases, we refer the reader to the very nice introductory paper of Gunawardena [48]. An application of this theory to the problem of deadlock detection has been made by the authors in [21].

There are other scheduling problems of interest for the theory we developed in this paper, like the ones appearing due to the architecture of recent microprocessors. Let us take an example from [49,66]. Many modern CPUs like SPARCs or MIPS pipeline instructions. Of course, their functional units, registers or bus are all used in mutual exclusion. Unfortunately the pipelined instructions overlap in time as they use more than one clock cycle and some of them cannot

instructions/cycle	0	1	2	3	4	5
add.s	U	S+A	A+R	R+S	Ø	Ø
add.s		U	S+A	A+R	R+S	Ø

where U is unpack, S is shift, A is adder and R is round.

Fig. 17. MIPS R4000 floating point unit.

be executed (otherwise "structural hazards" occur) within a certain number of cycles after some others (see Fig. 17 [1]). We do not want to use the pipeline in mutual exclusion since we would have to empty it after every instruction. The problem addressed in [66] is to verify that schedulers for a single process ensure that structural hazards will not occur (safety). In a concurrent framework, if there are more processes than processors, we can address the new problem of finding a way to interleave actions from different processes executed on the same processor, that verify the constraints while using the pipeline at the best of its capabilities (this is a view formalized in [1], see below). Some processors (like INTEL's Pentium) are even more complex to deal with since some resources may be used by at most two processes in parallel but not three. [2]

We see from Fig. 17 that if we suppose that we want to execute two instructions add.s one after the other on the MIPS R4000 floating point unit, at cycle 2, the adder A has to be used by both instructions (coming from the same thread). The same holds at cycle 3 for the round unit R. We say in that case that there is a hazard on A at cycle 2 and a hazard on R at cycle 3. A good scheduler should have prevented us from this situation by interleaving the two threads after the first add.s and continue with non-conflicting instructions of the second thread for the pipeline to be emptied a bit before executing the second add.s.

If we translate this problem to P and V operations on the shared resources, we see that our problem is to find schedules that do not deadlock.

On the semantic side, there is a strong link between po-spaces and progress graphs [10,15]. The model presented in [65,32] originated part of this work, and most of the work of one of the co-authors, [42,33,34] and also [59]. There are also links with the homological considerations of [25,26]. Some potentially related semantic models are the n-categorical formulations of [9] and some other combinatorial or categorical formulations in [23,70,73–75].

In program analysis, some proposals have been made to use the scheduling information that one can extract from the geometry of executions, to derive automatic parallelization algorithms. This has been hinted in [35] (where some other ideas for program analysis are exemplified) and also fully treated for CCS in [79,80]. A prototype Parallel Pascal Analyser has been implemented by Cridlig (http://www.dmi.ens.fr/∼cridlig); its principles are described in [12]. Another application on CML has been designed by the same author in [13] and to Linda-based languages in [14].

On the more mathematical side, homology of monoids has been studied in [22,47] (with the specific use of cubical sets), [58,77]. An extension to homology of categories has been proposed in [62]. Also cubical sets as such have been used in [72] and studied combinatorially and categorically in [7,8].

References

[1] H. Attiya, R. Friedman, A correctness condition for high-performance multiprocessors, in: Proc. 24th STOC, ACM Press, New York, 1992.
[2] A. Arnold, Systèmes de transitions finis et sémantique des processus communicants, Masson, 1992.
[3] H. Attiya, A. Bar-Noy, D. Dolev, D. Peleg, R. Reischuk, Renaming in an asynchronous environment, J. ACM 37 (3) (1990) 524–548.
[4] M.A. Bednarczyk, Categories of asynchronous systems, Ph.D. Thesis, University of Sussex, 1988.
[5] E. Borowsky, Capturing the power of resiliency and set consensus in distributed systems, Technical Report, University of California in Los Angeles, 1995.
[6] E. Borowsky, E. Gafni, Generalized FLP impossibility result for t-resilient asynchronous computations, in: Proc. 25th STOC, ACM Press, New York, 1993.
[7] R. Brown, P.J. Higgins, On the algebra of cubes, J. Pure Appl. Algebra 21 (1981) 233–260.
[8] R. Brown, P.J. Higgins, Colimit theorems for relative homotopy groups, J. Pure Appl. Algebra 22 (1981) 11–41.
[9] R. Buckland, M. Johnson, ECHIDNA: a system for manipulating explicit choice higher dimensional automata, in: AMAST'96: Fifth Internat. Conf. on Algebraic Methodology and Software Technology, Munich, 1996.
[10] S.D. Carson, P.F. Reynolds Jr., The geometry of semaphore programs, ACM Trans. Programming Languages Systems 9 (1) (1987) 25–53.

[1] Taken from [66].
[2] It has two integer arithmetic units.

[11] S. Chaudhuri, Agreement is harder than consensus: set consensus problems in totally asynchronous systems, in: Proc. Ninth Annu. ACM Symp. on Principles of Distributed Computing, ACM Press, New York, 1990, pp. 311–334.
[12] R. Cridlig, Semantic analysis of shared-memory concurrent languages using abstract model-checking, in: Proc. PEPM'95, ACM Press, La Jolla, 1995.
[13] R. Cridlig, Semantic analysis of concurrent ML by abstract model-checking, in: Proc. LOMAPS Workshop, 1996.
[14] R. Cridlig, E. Goubault, Semantics and analyses of Linda-based languages, in: Proc. WSA'93, Lecture Notes in Computer Science, Vol. 724, Springer, Berlin, 1993.
[15] E. Dijkstra, Cooperating Sequential Processes, Academic Press, New York, 1968.
[16] A. Dold, Lectures on Algebraic Topology, second ed., Grundlehren der Matematischen Wissenschaften, Vol. 200, Springer, Berlin, 1980.
[17] U. Fahrenberg, The geometry of timed PV-programs, Electronic Notes in Computer Science, Vol. 81, 2003, pp. 1–14.
[18] L. Fajstrup, Dicovering spaces, Homology Homotopy Appl. 5 (2) (2003) 1–17.
[19] L. Fajstrup, Dipaths and dihomotopies in a cubical complex, Adv. Appl. Math. 35 (2) (2005) R–2003–22.
[20] L. Fajstrup, E. Goubault, E. Haucourt, M. Raussen, Components of the fundamental category, Appl. Categ. Structures 12 (1) (2004) 81–108.
[21] L. Fajstrup, E. Goubault, M. Raussen, Detecting deadlocks in concurrent systems, in: D. Sangiorgi, R. de Simone (Eds.), CONCUR '98; Concurrency Theory, Lecture Notes in Computer Science, Vol. 1466, Springer, Nice, France, Ninth Internat. Conf., Proceedings, 1998, pp. 332–347.
[22] D.R. Farkas, The Anick resolution, J. Pure Appl. Algebra 79 (1992) 159–168.
[23] M. Fiore, G. Plotkin, J. Power, Complete cuboidal sets in axiomatic domain theory (extended abstract), in: Proc. 12th Annu. IEEE Symp. on Logic in Computer Science, IEEE Computer Society Press, Warsaw, Poland, 1997, pp. 268–279.
[24] M. Fisher, N.A. Lynch, M.S. Paterson, Impossibility of distributed commit with one faulty process, J. ACM 32 (2) (1985) 374–382.
[25] P. Gaucher, Connexion de flux d'information en algèbre homologique, Technical Report, IRMA, Strasbourg, available at ⟨http://irmasrv1.u-strasbg.fr/∼gaucher/activite.html⟩, 1997.
[26] P. Gaucher, Etude homologique des chemins de dimension 1 d'un automate, Technical Report, IRMA, Strasbourg, available at ⟨http://irmasrv1.u-strasbg.fr/∼gaucher/activite.html⟩, 1997.
[27] P. Gaucher, Homotopy invariants of higher dimensional categories and concurrency in computer science, Math. Structures Comput. Sci. 10 (4) (2000) 481–524.
[28] P. Gaucher, A convenient category for the homotopy theory of concurrency, Technical Report, math.AT/0201252, 2002, 138pp.
[29] P. Gaucher, From concurrency to algebraic topology, Electronic Notes in Computer Science, Vol. 39 (2), 2001, pp. 1–19.
[30] P. Gaucher, E. Goubault, Topological deformation of higher dimensional automata, Technical Report, arXiv:math.AT/010760, 2001.
[31] G. Gierz, K.H. Hofmann, K. Keimel, J.D. Lawson, M. Mislove, D.S. Scott, A Compendium of Continuous Lattices, Springer, Berlin, Heidelberg, New York, 1980.
[32] R. van Glabbeek, Bisimulation semantics for higher dimensional automata, Technical Report, Stanford University, Manuscript available on the web as ⟨http://theory.stanford.edu/∼rvg/hda⟩ 1991.
[33] E. Goubault, Domains of higher-dimensional automata, in: Proc. CONCUR'93, Springer, Hildesheim, 1993.
[34] E. Goubault, The geometry of concurrency, Ph.D. Thesis, Ecole Normale Supérieure Paris, available at ⟨http://www.dmi.ens.fr/∼goubault⟩, 1995.
[35] E. Goubault, Schedulers as abstract interpretations of HDA, in: Proc. PEPM'95, ACM Press, New York, also available at ⟨http://www.dmi.ens.fr/∼goubault⟩, La Jolla, 1995.
[36] E. Goubault, Durations for truly-concurrent actions, in: Proc. ESOP'96, no. 1058, Springer, Berlin, 1996, pp. 173–187.
[37] E. Goubault, The dynamics of wait-free distributed computations, Technical Report, Research Report LIENS-96-26, December 1996.
[38] E. Goubault, A semantic view on distributed computability and complexity, in: Proc. Third Theory and Formal Methods Section Workshop, Imperial College Press, also available at ⟨http://www.dmi.ens.fr/∼goubault⟩, 1996.
[39] E. Goubault, Optimal implementation of wait-free binary relations, in: Proc. 22nd CAAP, Springer, New York, 1997.
[40] E. Goubault, Cubical sets are generalized transition systems, Technical Report, Pre-proceedings of CMCIM'02, also available at ⟨http://www.di.ens.fr/∼goubault⟩, 2001.
[41] E. Goubault, Some geometric perspectives in concurrency theory, Homology Homotopy Appl. 5 (2) (2003) 95–136.
[42] E. Goubault, T.P. Jensen, Homology of higher-dimensional automata, in: Proc. CONCUR'92, Springer, Stonybrook, New York, 1992.
[43] E. Goubault, M. Raussen, Dihomotopy as a tool in state space analysis, in: S. Rajsbaum (Ed.), LATIN 2002: Theoretical informatics, Lecture Notes in Computer Science, Vol. 2286, Springer, Cancun, Mexico, 2002, pp. 16–37.
[44] E. Goubault, M. Raussen, Dihomotopy as a tool in state space analysis, in: Proc. LATIN'02, 2002.
[45] M. Grandis, Directed homotopy theory II. Homotopy constructs, Theory Appl. Categ. 10 (14) (2002) 369–391.
[46] M. Grandis, Directed homotopy theory I. The fundamental category, Cahiers Topologie Géom. Différentielle Catégoriques (2003) 281–316.
[47] J.R.J. Groves, Rewriting systems and homology of groups, in: L.G. Kovacs (Ed.), Groups—Canberra 1989, Lecture Notes in Mathematics, Vol. 1456, Springer, Berlin, 1991, pp. 114–141.
[48] J. Gunawardena, Homotopy and concurrency, Bull. EATCS 54 (1994) 184–193.
[49] E. Harcourt, J. Mauney, T. Cook, From processor timing specifications to static instruction scheduling, in: Proc. Static Analysis Symp.'94, Lecture Notes in Computer Science, Springer, Berlin, 1994.
[50] M. Herlihy, A tutorial on algebraic topology and distributed computation, Technical Report, Presented at UCLA, 1994.
[51] M. Herlihy, S. Rajsbaum, Set consensus using arbitrary objects, in: Proc. 13th Annu. ACM Symp. on Principles of Distributed Computing, ACM Press, New York, 1994.
[52] M. Herlihy, S. Rajsbaum, Algebraic topology and distributed computing, a primer, Technical Report, Brown University, 1995.
[53] M. Herlihy, N. Shavit, The asynchronous computability theorem for t-resilient tasks, in: Proc. 25th STOC, ACM Press, New York, 1993.

[54] M. Herlihy, N. Shavit, A simple constructive computability theorem for wait-free computation, in: Proc. STOC'94, ACM Press, New York, 1994.
[55] J. Jardine, Cubical homotopy theory: a beginning, Technical Report NI02030-NST, Isaac Newton Institute, Cambridge, UK, 2002.
[56] P. Jayanti, On the robustness of Herlihy's hierarchy, in: Proc. 12th Annu. ACM Symp. on Principles of Distributed Computing, Ithaca, New York, USA, 1993, pp. 145–157.
[57] P. Jayanti, Robust wait-free hierarchies, J. ACM 44 (4) (1997) 592–614.
[58] Y. Kobayashi, Complete rewriting systems and homology of monoid algebras, J. Pure Appl. Algebra 65 (1990) 263–275.
[59] E. Lanzmann, Automates d'ordre supérieur, Master's Thesis, Université d'Orsay, 1993.
[60] N. Lynch, Distributed Algorithms, Morgan Kaufmann, Los Altos, CA, 1996.
[61] A. Mazurkiewicz, Basic notions of trace theory, in: Lecture Notes for the REX Summer School in Temporal Logic, Springer, Berlin, 1988.
[62] F. Morace, Finitely presented categories and homology, Technical Report, Université Joseph Fourier, 1995.
[63] L. Nachbin, Topology and Order, Van Nostrand, Princeton, 1965.
[64] R. Penrose, Techniques of Differential Topology in Relativity, Conference Board of the Mathematical Sciences, Vol. 7, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, USA, 1972.
[65] V. Pratt, Modeling concurrency with geometry, in: Proc. 18th ACM Symp. on Principles of Programming Languages, ACM Press, New York, 1991.
[66] T.A. Proebsting, C.W. Fraser, Detecting pipeline structural hazards quickly, in: Proc. Symp. on Principles of Programming Languages, ACM Press, New York, 1994.
[67] M. Raussen, On the classification of dipaths in geometric models for concurrency, Math. Structures Comput. Sci. 10 (4) (2000) 427–457.
[68] M. Raussen, State spaces and dipaths up to dihomotopy, Homology Homotopy Appl. 5 (2) (2003) 257–280.
[69] M. Saks, F. Zaharoglou, Wait-free k-set agreement is impossible: the topology of public knowledge, in: Proc. 25th STOC, ACM Press, New York, 1993.
[70] V. Sassone, G.L. Cattani, Higher-dimensional transition systems, in: Proc. LICS'96, 1996.
[71] E. Schenk, The consensus hierarchy is not robust, in: Proc. 16th Annu. ACM Symp. on Principles of Distributed Computing, Santa Barbara, California, 1997, p. 279.
[72] J. Serre, Homologie singulière des espaces fibrés applications, Ph.D. Thesis, Ecole Normale Supérieure, 1951.
[73] S. Sokolowski, Homotopy in concurrent processes, Technical Report, Institute of Computer Science, Gdansk Division, 1998.
[74] S. Sokolowski, Investigation of concurrent processes by means of homotopy functors, Technical Report, Institute of Computer Science, Gdansk Division, 1998.
[75] S. Sokolowski, Point glueing in cpo-s, Technical Report, Institute of Computer Science, Gdansk Division, 1998.
[76] S. Sokołowski, Classifying holes of arbitrary dimension in partially ordered cubes, Kansas State University, February 2000, Manuscript.
[77] C.C. Squier, F. Otto, Y. Kobayashi, A finiteness condition for rewriting systems, Theoret. Comput. Sci. 131 (1994) 271–294.
[78] A. Stark, Concurrent transition systems, Theoret. Comput. Sci. 64 (1989) 221–269.
[79] Y. Takayama, Towards Cycle Filling as Parallelization, RIMS Workshop on Concurrency Theory and Applications 96, RIMS Kyoto University, 1998.
[80] Y. Takayama, Extraction of concurrent processes from higher-dimensional automata, in: Proc. CAAP'96, 1996, pp. 72–85.
[81] G. Winskel, M. Nielsen, Models for concurrency, Handbook of Logic in Computer Science, Vol. 3, Oxford University Press, Oxford, 1994, pp. 100–200.

Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 365 (2006) 247 – 257
Theoretical Computer Science
www.elsevier.com/locate/tcs

Deadlocks and dihomotopy in mutual exclusion models

Martin Raußen

Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg Øst, Denmark

Abstract

Higher dimensional automata (HDA) represent a promising tool for modelling ("true") concurrency in a both combinatorial and topological framework. Within these models, fast algorithms investigating deadlocks and unreachable regions have been devised previously on a background of easily understandable "directed" geometric ideas. In this article, we modify notions and methods from homotopy theory to define and investigate "essentially different" schedules in a HDA and to detect whether two given runs are essentially different using an algorithm again based on "directed geometry".

Keywords: Mutual exclusion; Deadlock detection; Dihomotopy

1. Introduction: mutual exclusion models

Already back in 1968, Dijkstra [2] proposed to apply a geometric point of view in the consideration of coordination situations in *concurrency*. His *progress graphs* were at the basis of the *Higher Dimensional Automata* (HDA) introduced by Pratt [16] and developed in the thesis of Goubault [9] and in later research [7].

In this article, we stick to Dijkstra's simple continuous geometric model. A system of n concurrent processes will be represented as a subset of Euclidean space $\mathbf{R}^n$ with the usual partial order. Each coordinate axis represents an idealized program counter of one of the processes performing a linear program; [1] the sequence of discrete increasing values of the counter is replaced by a continuous line. A point $(x_1, x_2, \ldots, x_n)$ in $\mathbf{R}^n$ represents the compound state in which the programs' counters simultaneously have the values $x_1, x_2, \ldots,$ resp. x_n.

A run of a concurrent program is modelled by a *continuous increasing* path between two states. A path in $\mathbf{R}^n$ is increasing if the projection to *each* of the axes increases; this condition has to be added to the picture since time in a program cannot flow backward.

Shared resources can often only be used by one or a limited number of processors at the same time. As a consequence, certain *hyperrectangles*—corresponding to sets of values of the program counters that lead to conflict in the access to such a resource—cannot be entered by any allowable path; together, they form the *forbidden region*, which will be neglected in the state space.

The resulting *mutual exclusion models* are more general than those modelling only *semaphore* programs. They allow us to consider also k-semaphores, where a shared object may be accessed by k, but not by $k+1$ processors; cf. [1] and Example 2.1. For an introduction of the more general notion of HDA and some of their properties, we refer to [9,7,11].

E-mail address: raussen@math.aau.dk.

[1] The methods can be adapted to more general programs by replacing an axis by a graph and the state space by a product of graphs, cf. [8].

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.07.052

In this note, we attempt to use methods from geometry and homotopy theory (after suitable modification) to analyse properties of the resulting state spaces and, in particular, of the allowable runs on them. It has been noticed [15,7], that increasing paths (schedules) between two given states, that are homotopic in a specific way (dihomotopic, i.e., there exists a 1-parameter family of directed paths interpolating between them; cf. Definition 3.1), will always yield to the same results in a concurrent computation.

Therefore, it is important to obtain a better understanding of the dihomotopy relation between increasing paths and to use and to further develop the existing machinery from elementary algebraic topology (homotopy, homology, etc.) in this direction. It cannot be straightforward to do so: Fajstrup et al. [7] give examples of two directed paths in a certain state space that are homotopic but not dihomotopic; i.e., in every interpolation there will occur non-directed paths. For several attempts to extend topological methods and results to the directed situation related to the present one, consult e.g. [7,10,17,13,3,12,18].

We address two questions in this article: How can one use the geometric/combinatorial description of the forbidden region to

(1) detect *deadlocks* and associated *unsafe*, resp. *unreachable* regions? In Section 2, we give a survey of the results obtained in [6], mainly as a background for (2) below.
(2) obtain information on the number of "essentially different" schedules between two given states? These results are new and will be developed and explained in Sections 4 and 5.

Somewhat surprisingly, the two questions above turn out to be related: It can be detected that certain schedules cannot be equivalent because a related model exhibits a combination of an unreachable and a deadlock state; cf. Section 4. On the other hand, it will be shown, that a certain restriction of the complexity of the synchronization yields state spaces in which all runs are dihomotopic, cf. Section 5; a classical homological result (Remark 5.2.1) made this plausible.

In a sense, the number of "essentially different" schedules from 2 above should be seen as a measure for the complexity of the synchronization: in a program with trivial synchronization, all directed paths are dihomotopic. In data base theory, Gunawardena showed [15,7] that all dipaths in models corresponding to the "2-phase locking strategy" are homotopic to a serial one, and thus a concurrent execution will always have the same result as one of the $n!$ possible executions with "one process at a time" (serializability). In general, the more complex the synchronization (or rather its effects on concurrent computations), the larger the number of essentially different (mutually non-dihomotopic) directed paths.

2. Deadlock detection in mutual exclusion models

To get more formal, let $I = [0, 1]$ denote the unit interval, and let $I^n \subset \mathbf{R}^n$ denote the unit hypercube. An (open) isothetic hyperrectangle is a subset

$$R = (a_1, b_1) \times \cdots \times (a_n, b_n) \subset I^n;$$

closed or half-open coordinate intervals are exceptionally allowed at the boundaries in the forms $[0, b)$, $(a, 1]$, resp. $[0, 1]$. The *forbidden region* $F = \bigcup_1^r R^i$ is then a finite union of n-hyperrectangles $R^i = (a_1^i, b_1^i) \times \cdots \times (a_n^i, b_n^i)$, and the *state space* is its complement $X = I^n \setminus F$. We assume that $\mathbf{0} = (0, \ldots, 0)$ and $\mathbf{1} = (1, \ldots, 1)$ are *not* contained in the forbidden region F; they represent the initial, resp. the final state of the concurrent program.

Example 2.1. Forbidden regions allow to model all sorts of k-semaphore. Assume for example, that 3 processors P_i have to access a shared resource while their counter $x_i \in (a_i, b_i)$. If only one processor can have access to the resource at the same time (2-semaphore), then the forbidden region includes the union of 3 cubes

$$(a_1, b_1) \times (a_2, b_2) \times [0, 1] \cup (a_1, b_1) \times [0, 1] \times (a_3, b_3) \cup [0, 1] \times (a_2, b_2) \times (a_3, b_3).$$

If two (but not all three) processors can access it concurrently, then only the cube $(a_1, b_1) \times (a_2, b_2) \times (a_3, b_3)$ is forbidden.

The "Swiss flag" example [6] from Fig. 1 (the forbidden region consisting of two forbidden rectangles is dashed) conveys the idea, that deadlocks—with no possible legal move—in such mutual exclusion models are associated to n-dimensional "lower corners" below the forbidden region.

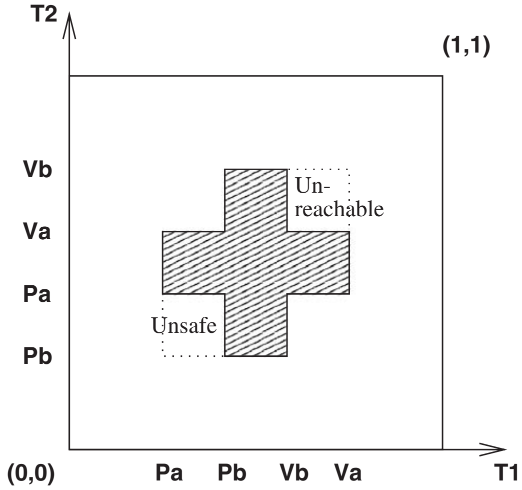


Fig. 1. "Swiss flag".

To make this formal, we call a continuous path[2] $\alpha : \vec{I} \to X \subset \vec{I}^n$ from $\mathbf{x} = \alpha(0)$ to $\mathbf{y} = \alpha(1)$ a *dipath* (directed path) if every composition $pr_i \circ \alpha$ is increasing. We introduce a new partial order $\preccurlyeq$ on X by

$$\mathbf{x} \preccurlyeq \mathbf{y} \Leftrightarrow \text{ there is a dipath } \alpha \text{ from } \mathbf{x} \text{ to } \mathbf{y} \text{ in } X.$$

As can be seen e.g. in Fig. 1, this partial order is in general finer than the one X inherits from the usual partial order on $\mathbf{R}^n$.

An element $\mathbf{x} \in X$ is called *admissible* if $\mathbf{x} \preccurlyeq \mathbf{1}$ and *unsafe* else. An element $\mathbf{y} \in X$ is called *reachable* if $\mathbf{0} \preccurlyeq y$ and *unreachable* else. An element $\mathbf{1} \neq \mathbf{x} \in X$ is called a *deadlock* if $\mathbf{x} \preccurlyeq \mathbf{y} \Rightarrow \mathbf{y} = \mathbf{x}$; cf. Fig. 1.

To formulate results, we need to introduce k-element intersections of the hyperrectangles R_i forming part of the forbidden region $F = \bigcup_1^r R^i$, cf. Section 1: For any k-element index set

$$J = \{i_1, \ldots, i_k\} \subset \{1, \ldots, r\}, k > 0, \text{ let } R^J = R^{i_1} \cap \cdots \cap R^{i_k}.$$

Unless $R^J = \emptyset$, it is again a hyperrectangle $R^J = (a_1^J, b_1^J) \times \cdots \times (a_n^J, b_n^J)$ with $a_j^J = \max\{a_j^i \mid i \in J\}$ and $b_j^J = \min\{b_j^i \mid i \in J\}$. The minimal vertex of R^J is given by $\mathbf{a}^J = (a_1^J, \ldots, a_n^J)$. Moreover, let $\tilde{a}_j^J$ denote the "second largest" of the jth coordinates a_j^i; we need also consider the "unsafe corner" $Us^J =]\tilde{\mathbf{a}}^J, \mathbf{a}^J] =]\tilde{a}_1^J, a_1^J] \times \cdots]\tilde{a}_n^J, a_n^J] \subset X$.

Proposition 2.2. (1) *An element $\mathbf{1} \neq \mathbf{x} \in X$ in the interior of $\vec{I}^n$ is a deadlock if and only if there is an n-element index set $J = \{i_1, \ldots, i_n\}$ with $R^J \neq \emptyset$ and $\mathbf{x} = \mathbf{a}^J = \min R^J$.*

(2) *If $\mathbf{x} = \mathbf{a}^J = \min R^J$ is a deadlock, then all elements of the n-hyperrectangle Us^J are unsafe.*

Remark 2.3. (1) In a similar way, one can find an "unreachable corner" Ur^J "above" the maximal element of an n-intersection R^J.

(2) A simple trick allows to detect deadlock points that are contained in the boundary of I^n as well; cf. [6] and also Section 5.3.

In [6], we describe a fast incremental algorithm, that detects the *entire* unsafe region (consisting of *all* unsafe elements in X) in few steps—usually, many (discrete) states are detected in one single step. One has to take into account the (order) combinatorics of intersections of forbidden hyperrectangles and of those hyperrectangles that have found to be unsafe in previous steps. An implementation of this algorithm can be found on the URL http://www.ens.fr/~goubault.

[2] We distinguish between the interval $\vec{I}$ with the *usual* order as (partial) order and the interval I neglecting order (or rather, with equality as the partial order relation).

3. The dihomotopy concept

An execution of a concurrent process corresponds to a dipath (cf. Section 2) in the state space X. The most interesting dipaths are those starting at **0** and terminating at **1** (a complete run), but also dipaths starting and/or terminating at other elements need to be considered; both for practical purposes in state space analysis and as intermediate steps in theoretical calculations.

Many executions will "automatically" be equivalent; this means that all conceivable concurrent calculations along the corresponding schedules/paths yield the same result. In geometric language, this is the case when the dipaths corresponding to the executions are *dihomotopic*, cf. [7] dihomotopy is a modification of the notion *homotopy*—which is fundamental and well studied in Algebraic Topology. With dihomotopy we take into account not only continuity but also partial order. There are several definitions for dihomotopy, all of which are equivalent in the case of our simple partially ordered state space; cf. Proposition 3.4, or in greater generality [4]. We need to work with three of these definitions:

Definition 3.1. A continuous 1-parameter deformation (*dihomotopy*) $H : \vec{I} \times I \to X$ with $H(0, t) = \mathbf{x}$, $H(1, t) = \mathbf{y}$ for all $t \in I$ is called

(1) a *dihomotopy* [7] if, for all t, the "interpolating" paths $\alpha_s : t \mapsto H(s, t)$ are *dipaths*.

(2) an *elementary d-homotopy* [14] if, for all $s \in \vec{I}$ and such that for all s and t, the "interpolating" paths $\alpha_t : s \mapsto H(s, t)$ and $\alpha_s : t \mapsto H(s, t)$ are *dipaths*.

Two continuous dipaths $\alpha, \beta : \vec{I} \to X$ from $\mathbf{x} \in X$ to $\mathbf{y} \in X$ are called

(1) *dihomotopic* [7] if there exists a dihomotopy H with $H(s, 0) = \alpha(s)$, $H(s, 1) = \beta(s)$ for all $s \in \vec{I}$.

(2) *d-homotopic* [14] if there exist dipaths $\alpha_0 = \alpha, \alpha_1, \ldots, \alpha_{2s} = \beta$ and elementary d-homotopies from α_{2k} to α_{2k+1} and from α_{2k+2} to α_{2k+1}—i.e., a "zig-zag homotopy" between α and β.

Remark 3.2. Both notions are defined in far more general situations for maps between locally partially ordered spaces (dihomotopy), resp. for maps between d-spaces (d-homotopy). The latter notion is preferable for homotopy theoretic purposes.

Only the order requirement for the interpolating paths is characteristic for a di/d-homotopy compared to a homotopy (with fixed ends). Examples (cf. Example 1 or [7]) show that dihomotopy in general is a finer relation than homotopy of dipaths. Moreover, it is important to notice that, unlike for homotopy, dihomotopy in general does *not* satisfy a cancellation property: $\alpha * \beta_1$ dihomotopic to $\alpha * \beta_2$ does not always imply that β_1 is dihomotopic to β_2. Examples for non-cancellation are given in [7]; it also occurs in Example 4.6.

In the case of the state space of a mutual exclusion model (more generally, for a cubical complex), one may restrict attention to dipaths on the 1-*skeleton* of $X \subseteq \vec{I}^n$ and to *combinatorial* dihomotopies [7]. To explain these notions in our simple case, one considers the projections of *all* hyperrectangles within the forbidden region to the coordinate axes. This gives rise to a subdivision of the axes $[0, 1]$ into subintervals—at requests for shared resources or terminations of such. The 1-skeleton corresponding to that subdivision consists of the *line sections* parallel to one of the axes and constant at one of the subdivision points for all other directions. A (locally serial) dipath along this 1-skeleton proceeds at every time along one of these line sections. An *elementary* dipath proceeding with "unit speed and one step" parallel to the x_i-axis will be denoted σ_i—the ith process proceeds one step forward while the others wait. (This notation is not unambiguous, but good enough for our purposes.) Two such elementary dipaths σ_i, σ_j can be concatenated to yield the dipath $\sigma_i * \sigma_j$ if the target of the first agrees with the source of the second.

Definition 3.3. (1) Two dipaths $\alpha = \sigma_i * \sigma_j$ and $\beta = \sigma_j * \sigma_i$ in X with the same source $\mathbf{x}$ and target $\mathbf{y}$ are called *elementarily dihomotopic* if the two-dimensional rectangle with lower vertex in $\mathbf{x}$ and upper vertex in $\mathbf{y}$ is contained in X.

(2) The (combinatorial) dihomotopy relation is obtained from elementary dihomotopy as the closure under concatenation, reflexivity and transitivity. We write $\alpha \vec{\sim} \beta$ to denote that α (combinatorially) dihomotopic to β.

More general definitions for combinatorial dihomotopy are given in [7,4].

An elementary dihomotopy (given by such a rectangle in the state space) reflects the fact that the result of the compound execution of σ_i and σ_j is independent of the order in which these are performed (even after possible subdivisions into smaller pieces).

Within the state space I^n—no mutual exclusion—any such skeletal dipath can be obtained from any other (in I^n) by permutations and thus by a succession of transpositions and hence elementary dihomotopies. In particular, any two dipaths with the same source and target are combinatorially dihomotopic in I^n. In a (smaller) state space $X \subset I^n$ however—with mutual exclusion, e.g., $X = I^n \setminus F$—a chain of elementary dihomotopies within I^n might contain a particular elementary dihomotopy along a two-dimensional rectangle that is *not* contained in the state space *X* although its boundary is. If this is the case *for all* such chains between two given dipaths, these two dipaths are *not* combinatorially dihomotopic in *X*.

In the "Swiss flag" example from Fig. 1 in Section 2, there are two dihomotopy classes of dipaths connecting **0** and **1**. A complete classification algorithm for dipaths up to dihomotopy in two-dimensional models had previously been given in [17]. It is the aim of this article to pave the way for a generalization of those results to the general *n*-dimensional case.

In this article, we will allow ourselves to use whatever notion of dihomotopy, *d*-homotopy or combinatorial dihomotopy is most suitable for our purposes. We may do so because of the following result, which applies in particular to the geometric cubical complex $X = \vec{I}^n \setminus F$:

Proposition 3.4 (*[4, Theorems 5.1 and 5.6]*). *All three notions are equivalent in geometric cubical complexes.*

4. Dihomotopy and deadlocks in mutual exclusion models

The purpose of this section is to make a link between the detection of deadlocks and unsafe regions in mutual exclusion models and the occurrence of non-dihomotopic dipaths in such models. It had been conjectured for a long time, that, just as *n* intersecting *n*-rectangles give rise to deadlocks, unsafe and unreachable regions, so should likewise $(n-1)$ intersecting *n*-rectangles give rise to non-trivial non-local dihomotopy. [3] We discuss here when and why this in fact is the case.

Forgetting about the last coordinate (processor) amounts to projecting the forbidden hyperrectangles and the forbidden region under $\pi : \mathbf{R}^n \to \mathbf{R}^{n-1}, \mathbf{x} = (x_1, \ldots, x_n) \mapsto \pi\mathbf{x} = (x_1, \ldots, x_{n-1})$, arriving at a forbidden region $\bar{F} = \pi(F)$ and a state space $\bar{X} = \vec{I}^{n-1} \setminus \bar{F}$ (different from $\pi(X)$, in general!).

Let us compare the forbidden regions in X and in $\bar{X}$: consider an $(n-1)$-element index set J with non-empty intersection hyperrectangle $R^J \subset F$. If the participating hyperrectangles intersect *generically*—in particular, if $R^J \neq R^K$ for every smaller index set $K \subset J$—then every of the $(n-1)$ hyperrectangles R_i will "contribute" at least one coordinate to the minimum $\mathbf{a}^J = [a_1^J, \ldots, a_n^J]$ of R^J—and similarly to its maximum $\mathbf{b}^J$. We may then suppose without restriction, that

$$a_1^J = a_1^1, \ldots, a_{n-2}^J = a_{n-2}^{n-2}, \quad a_{n-1}^J = a_{n-1}^{n-1}, a_n^J = a_n^{n-1}.$$

The $(n-1)$ hyperrectangles $\pi(R^i)$ in $\vec{I}^{n-1}$ intersect in $\pi(R^J) = \pi(R)^J$—for short πR^J—a hyperrectangle with minimal vertex $\pi\mathbf{a}^J = (a_1^J, \ldots, a_{n-1}^J)$, which is a *deadlock* for the model space $\bar{X}$. The intersection $\pi R^J = [\pi\mathbf{a}^J, \pi\mathbf{b}^J]$ gives furthermore rise to an unsafe region $Us(\pi R^J) =]\pi\tilde{\mathbf{a}}^J, \pi\mathbf{a}^J] \subset \bar{X}$. As in Section 2, the point $\tilde{\mathbf{a}}^J$ has the "second" largest coordinates among the a_j^i as its coordinates.

In a similar way, we can consider the projection $\pi' : \mathbf{R}^n \to \mathbf{R}^{n-1}, (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-2}, x_n)$, giving rise to the deadlock $\pi'\mathbf{a}^J = (a_1^J, \ldots, a_{n-2}^J, a_n^J)$ and the unsafe region $Us(\pi' R^J) \subset \vec{I}^{n-1} \setminus \pi' F$.

Lemma 4.1. (1) *Let* $\mathbf{x}, \mathbf{y} \in X$ *satisfy*

$$\pi\mathbf{x} \in Us(\pi R^J) \text{ or } \pi'\mathbf{x} \in Us(\pi' R^J), \mathbf{x} \preccurlyeq \mathbf{a}^J, \mathbf{b}^J \preccurlyeq \mathbf{y}.$$

[3] Even *a single n*-rectangle in the forbidden region creates dihomotopy, but only between points that are "sufficiently close" to that *n*-rectangle, cf. the discussion in dimension 3 in [5].

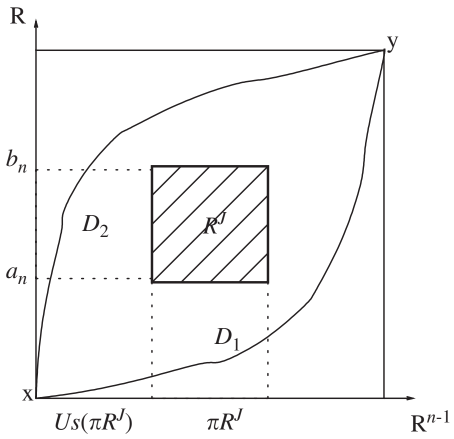


Fig. 2. Non-dihomotopic dipaths.

A dipath $\alpha = (\alpha_1, \dots, \alpha_n)$ *from* **x** *to* **y** *satisfies either*
(P1) $\alpha_n(t) \leqslant b_n \Rightarrow \pi\alpha(t) \in Us(\pi R^J)$ *or*
(P2) $\alpha_n(t) > a_n \Rightarrow \alpha^n(t) \notin Us(\pi R^J)$.
Two dihomotopic dipaths satisfy either both (P1) *or* both (P2).
(2) *Let* **u**, **v** $\in X$ *chosen such that*

$$\pi\mathbf{v} \in Ur(\pi R^J) \text{ or } \pi'\mathbf{v} \in Ur(\pi R^J),\ \mathbf{u} \preccurlyeq \mathbf{a}^J,\ \mathbf{b} \preccurlyeq \mathbf{v}^J.$$

A dipath $\beta = (\beta_1, \dots, \beta_n)$ *from* **u** *to* **v** *satisfies either*
(P3) $\beta_n(t) \geqslant a_n \Rightarrow \pi\beta(t) \in Ur(\pi R^J)$ *or*
(P4) $\beta_n(t) < b_n \Rightarrow \pi\beta(t) \notin Ur(\pi R^J)$.
Two dihomotopic dipaths satisfy either both (P3) *or* both (P4).

Corollary 4.2. *Two dipaths* α, β *from* **x** *to* **y** *with* α *satisfying* (P1) *and* β *satisfying* (P2), cannot *be dihomotopic*.

An instructive example is given by two dipaths α, β from $\mathbf{a}^J$ to $\mathbf{b}^J$: while all the other coordinates remain fixed, we let α_{n-1} grow from a^J_{n-1} to b^J_{n-1} before α_n grows from a_n to b_n; for β, the nth coordinate grows before the $(n-1)$st. Remark for later use that there are no *upward* restrictions for the *end* point **y** (Fig. 2).

Proof of Lemma 4.1.
(1) The crucial property is:

$$\pi R^J \times]a_n, b_n[) \subset F.$$

A dipath in X from **x** to **y** has thus to pass through

$$D = D_1 \cup D_2 = (\pi R^J \times [x_n, a_n]) \cup (Us(\pi R^J) \times]a_n + \varepsilon, b_n - \varepsilon[) \text{ for a small } \varepsilon > 0,$$

since adding D to F disconnects **x** from **y**. Since D_1 and D_2 are *disconnected* from each other, any dipath from **x** to **y** has to pass through one and only one of those sets. There cannot be a dihomotopy between a dipath intersecting the first and a dipath intersecting the second, since this would yield a division of the connected parameter interval I into two open non-empty sets, cf. [17].
(2) Is proved by a symmetric argument. □

A *single* arrangement of $(n-1)$ intersecting hyperrectangles will in general *not* lead to non-dihomotopic dipaths from **0** to **1**. This can be seen e.g. for the state space with a *single* wedge (cf. Example 1) as the forbidden region. We have to consider (at least) *two disjoint* arrangements J, K consisting of $(n-1)$ intersecting n-rectangles each within the forbidden region F; as usual, $X = \vec{I}^n \setminus F$. The two intersection n-rectangles and their projections will be called

$$R^J = [\mathbf{a}^J, \mathbf{b}^J], \quad R^K = [\mathbf{a}^K, \mathbf{b}^K], \quad \pi R^J = [\pi\mathbf{a}^J, \pi\mathbf{b}^J], \quad \pi R^K = [\pi\mathbf{a}^K, \pi\mathbf{b}^K]$$

with unsafe, resp. unreachable regions

$$Us(\pi R^K) =]\pi\tilde{\mathbf{a}}^K, \quad \pi\mathbf{a}^K], Ur(\pi R^J) = [\pi\mathbf{b}^J, \pi\tilde{\mathbf{b}}^J[.$$

We suppose that $a_n^J < b_n^K$.

A dipath $\alpha = (\alpha_1, \ldots, \alpha_n) : \vec{I} \to X$ from $\mathbf{0}$ to $\mathbf{1}$ is called *inter-JK* if it satisfies

$$a_n^J < \alpha_n(t) < b_n^K \Rightarrow \pi\mathbf{b}^J < \pi\alpha(t) < \pi\mathbf{a}^K. \tag{1}$$

—where the <-relation on the right-hand side is understood for all $n - 1$ coordinates.

Proposition 4.3. *Let $F = \bigcup_{i=1}^r R_i \subset \vec{I}^n$ denote the forbidden region. Let $J, K \subset \{1, \ldots, k\}$ denote two disjoint subsets indexing $(n-1)$ intersecting hyperrectangles R_i each and such that $\pi\tilde{\mathbf{a}}^K < \pi\mathbf{b}^J < \pi\mathbf{a}^K < \pi\tilde{\mathbf{b}}^J$. Any dipath $\beta : \vec{I} \to X$ from $\mathbf{0}$ to $\mathbf{1}$ that is dihomotopic to an inter-JK-dipath α is then an inter-JK-dipath itself.*

Corollary 4.4. *Under the assumptions of Proposition* 4.3, *a dipath in X from* **0** *to* **1** *that is not inter-JK*, *e.g.*, *a dipath on the boundary of $\vec{I}^n$, is not dihomotopic to an inter-JK-dipath. In particular*, *if there exist both an inter-JK-dipath from* **0** *to* **1** *and another one that is not inter-JK*, *then these two are not dihomotopic to each other*.

Remark 4.5. (1) From an application point of view, Corollary 4.4 implies the existence of different terminating schedules that can possibly yield different results of distributed calculations according to different schedules.

(2) Corollary 4.4 applies in particular to the example of a three-dimensional PV-program given in [1].

Example 4.6. The situation from Proposition 4.3 arises in three dimensions, when the forbidden region is a *cylinder* (with a "thick" rectangle as cross-section). More strikingly, there are state spaces with *trivial fundamental group*, that allow non-dihomotopic dipaths: it suffices to consider a forbidden region consisting of two "wedges", one behind the other and not connected to each other; one of them yields a deadlock after projection (to the "front") and the other unreachable points; cf. Fig. 3. A dipath (from lower left to upper right) through the area between the wedges is homotopic (relative to the end points) but not dihomotopic to a dipath avoiding it.

Proof of Proposition 4.3. By Proposition 3.4, it is enough to show that α and β are not d-homotopic. To this end, we apply Marco Grandis' van Kampen theorem [14, Theorem 3.6], to the decomposition $X = X_1 \cup X_2$ with $X_1 = X \cap (\vec{I}^{n-1} \times [0, b_n^K[)$, $X_2 = X \cap (\vec{I}^{n-1} \times]a_n^J, 1])$. Given a decomposition of the dipath $\alpha = \alpha^1 * \alpha^2$ with α_i in X_i and division point $\mathbf{u} = \alpha^1(1) = \alpha^2(0) \in [\pi\mathbf{b}^J, \pi\mathbf{a}^K] \times]a_n^J, b_n^K[$ by assumption. It is sufficient to prove, for β_i dihomotopic to α_i *in* X_i (with fixed end points $\mathbf{0}$ and $\mathbf{u}$, resp. $\mathbf{u}$ and $\mathbf{1}$), that $\beta = \beta^1 * \beta^2$ is an inter-*JK* dipath, as well.

The assumption of Proposition 4.3 has the following consequence:

$$Ur(\pi(R)^J) \cap Us(\pi(R)^K) = [\pi\mathbf{b}^J, \pi\mathbf{a}^K] = Ur(\pi(R)^J) \cap \downarrow \pi\mathbf{a}^K = Us(\pi(R)^K) \cap \uparrow \pi\mathbf{b}^J. \tag{2}$$

Hence, $\alpha_n^1(t) > a_n^J$ implies $\pi\alpha(t) \in Ur(\pi(R)^J) \cap Us(\pi(R)^K)$. From Lemma 4.1, we may conclude, that $\beta_n^1(t) > a_n^J$ implies $\pi\beta^1(t) \in Ur(\pi(R)^J)$, as well. Moreover, $\pi\beta^1(t) \leqslant \pi\mathbf{u} \leqslant \pi\mathbf{a}^K$ for all t, and we can conclude from (2): $\pi\beta^1(t) \in [\pi\mathbf{b}^J, \pi\mathbf{a}^K]$. In the same way, it can be shown that $\beta_n^2(t) < b_n^K$ implies $\pi\beta^2(t) \in [\pi\mathbf{b}^J, \pi\mathbf{a}^K]$. □

5. Trivial dihomotopy for models with less complicated constraints

In contrast, for a model space with a less complicated forbidden region, we can show by a simple essentially combinatorial argument and using the characterization of dihomotopy from Definition 3.3:

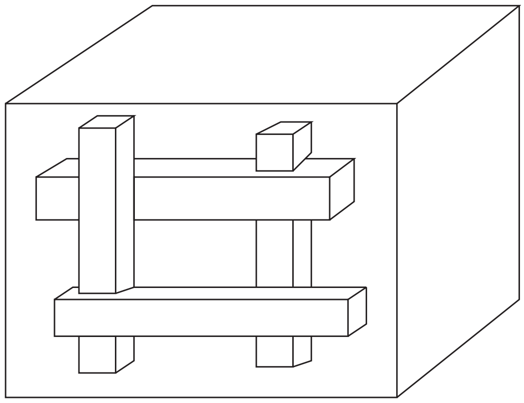


Fig. 3. Two wedges.

Proposition 5.1. *For a model space X with the property that* $R^J = \emptyset$ *for all index sets J of cardinality* $n-1$, *every two dipaths from* **0** *to* **1** *are dihomotopic to each other.*

Remark 5.2. A similar result holds also in the classical non-directed case: using duality and Čech-type cohomology, it is easy to see that the complement of a forbidden region with $R^J = \emptyset$ for all index sets J of cardinality $n-1$ has a trivial first *homology* group.

From an application point of view, the criterion from Proposition 5.1 is easy to check and ensures that *all* runs in such a distributed calculation will yield the *same* result. This should also be interesting for data base scheduling; compare [15] and [7], Section 8. On the other hand, it cannot be applied for traditional semaphore scheduling, since those will always generate lots of intersections. One would have to restrict to k-semaphores with $k > 1$.

5.1. Local futures

Given a point $\mathbf{x} \in X = \vec{I}^n \setminus F$ and $F = \bigcup_{i=1}^n R_i$ intersecting in general position. We describe the local future $\uparrow_l \mathbf{x}$ of such a point, i.e., the intersection of a small cube with lower vertex $\mathbf{x}$ in I^n within the state space X (reversing inequalities yields similar results for the local past $\downarrow_l \mathbf{x}$ of $\mathbf{x}$):

For a hyperrectangle $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, let

$$\partial_-^j R = \{\mathbf{x} = (x_1, \ldots, x_n) \mid x_j \leqslant b_j \wedge \exists 1 \leqslant i_1 < \cdots < i_j \leqslant n : x_{i_k} = a_{i_k}\},$$

i.e., the intersection of j of its lower faces. In particular, the lower boundary of a standard s-cube $\vec{I}^S$ is $\partial_- \vec{I}^s = \partial_-^1 \vec{I}^s = \{(a_1, \ldots, a_s) \mid \exists i : a_i = 0\}$.

Lemma 5.3. *Assume* $\mathbf{x} \in X$ *is contained in* $\partial_-^{j_1} R_1 \cap \cdots \cap \partial_-^{j_k} R_k$*—with* $k \geqslant 0$ *and* $j_s \geqslant 1$ *maximal—and* $j := j_1 + \cdots + j_k \leqslant n$. *Then* $\uparrow_l \mathbf{x}$ *is dihomeomorphic to* $\partial_- \vec{I}^{j_1} \times \cdots \times \partial_- \vec{I}^{j_k} \times \vec{I}^{n-j}$.

(In the case that $\mathbf{x}$ is *not* contained in the *lower boundary* $\partial_- R$ of any forbidden hyperrectangle R, then $j = 0$ and $\uparrow_l \mathbf{x}$ is dihomeomorphic to $\vec{I}^n$.)

Proof. An element $\mathbf{y} \in \vec{I}^n$ with $\mathbf{x} \leqslant \mathbf{y}$ close to $\mathbf{x}$ is contained in the state space X if and only if $\mathbf{y} \notin (R_1 \cup \cdots \cup R_k)$, i.e., if and only if at least one of the j_i "critical" coordinates in $\partial_-^{j_i} R_i$ of $\mathbf{y}$ coincides with the respective coordinate of $\mathbf{x}$—which gives rise to a factor homeomorphic to $\partial_- \vec{I}^{j_i}$. □

Remark 5.4. Another way to phrase Lemma 5.3 is, that the local building blocks of a mutual exclusion state space are of type $\partial_- \vec{I}^{j_1} \times \cdots \times \partial_- \vec{I}^{j_k} \times \vec{I}^{n-j}$. These simple ingredients can thus be seen as the building blocks of the state space for any mutual exclusion model. This should be of independent interest!

Lemma 5.5. *Let σ_j, σ_k denote two elementary dipaths* (*cf. Section* 3) *starting at* **x**—*as above. There is a elementary dipath σ_l commuting with both σ_k and σ_l up to dihomotopy if at least one of the following conditions is satisfied*:
(1) $j := j_1 + \cdots + j_k < n$;
(2) $j_s \geqslant 3$ *for at least one index* s;
(3) $j_{s_1}, j_{s_2} \geqslant 2$ *for two different indices* s_1, s_2.

Example 5.6. In the case $n = 2$, $j_1 = 2$, the local future of such a point x is of the form $\partial^2_- \vec{I}^2$, which is a one-dimensional wedge (like a letter L). In this case the two dipaths σ_1, σ_2 along the legs of the wedge do *not* commute locally up to dihomotopy. In the case $n = 3$, $j_1 = 2$, the local future of a point is the product of a wedge and an interval. In this case, the dipaths σ_1, σ_2 commute both with the dipath σ_3 in the third direction.

Proof. By Lemma 5.3, the local future $\uparrow_l \mathbf{x}$ is of the from $\partial\_\vec{I}^{j_1} \times \cdots \times \partial\_\vec{I}^{j_k} \times \vec{I}^{n-j}$. Remark that a factor $\partial\_\vec{I}^1$ consists of a single element and thus has no effect on the product.
(1) Let σ_l denote a elementary dipath within $\vec{I}^{n-j}$. Then, for $j \neq l$, the rectangle "spanned" by σ_j and σ_l is contained in the state space X. In particular, σ_l commutes with σ_j.
(2) For given indices j, k choose an index $j \neq l \neq k$ referring to one of the axes in the cube $\vec{I}^3$ the lower boundary $\partial\_\vec{I}^3$ of which is a factor in $\uparrow_l x$. Same argument as in 1 above.
(3) The only case not yet covered is a product including a factor $\partial\_\vec{I}^2 \times \partial\_\vec{I}^2$. Two dipaths in *different* factors of this product span a rectangle in X and thus do commute with each other. □

Lemma 5.7. *The conditions of Proposition* 5.1 *ensure that Lemma* 5.5 *is applicable*.

Proof. By assumption $j := j_1 + \cdots + j_k \leqslant n$ with $k \leqslant n - 2$. Hence $j < n$ or $j_i \geqslant 3$ for at least one i or $j_{i_1} = j_{i_2} = 2$ for at least two indices i_1 and i_2. □

5.2. Proof of Proposition 5.1

As explained in Section 3, within the state space $X \subset \vec{I}^n$ we need only consider dipaths of the form $\sigma = \sigma_{i_1} * \cdots * \sigma_{i_N}$, $1 \leqslant i_j \leqslant n$ from **0** to **1**. By Proposition 3.4, we are done if we can show that all those are *combinatorially* dihomotopic.

Proof. The proof is by induction on the length l of dipaths ending at **1**—and thus starting at an (arbitrary) element **x** at "taxi cab distance" l from **1**. For $l = 0$ and 1, there is nothing to prove. Assume inductively that, for all elements **x** at distance k from **1**, all dipaths starting at **x** and ending at **1** are combinatorially dihomotopic to each other.

Let **y** denote a vertex of X at distance $k + 1$ from **1** and let $\sigma = \sigma_{i_{n-k}} * \sigma_{i_{n-k-1}} * \cdots * \sigma_n =: \sigma_{i_{n-k}} * \bar{\sigma}$ and $\sigma' = \sigma'_{i_{n-k}} * \sigma'_{i_{n-k-1}} * \cdots * \sigma'_n =: \sigma'_{i_{n-k}} * \bar{\sigma}'$ denote two elementary dipaths from **y** to **1**. By Lemmas 5.5 and 5.7, there exists an elementary dipath σ_l with source **y** that commutes with both $\sigma_{i_{n-k}}$ and $\sigma'_{i_{n-k}}$.

We denote by **z** the target of $\sigma_{n-k} * \sigma_l \overset{\sim}{\to} \sigma_l * \sigma_{n-k}$. The condition of Proposition 5.1 assures also that the future $\uparrow \mathbf{z}$ of **z** is *deadlock-free* [6]. In particular, there exists a dipath $\hat{\sigma}$ from **z** to **1**. By induction, $\sigma_l * \hat{\sigma} \overset{\sim}{\to} \bar{\sigma}$. Likewise, $\bar{\sigma}'$ is dihomotopic to a dipath $\sigma_l * \hat{\sigma}'$. But then

$$\sigma = \sigma_{i_{n-k}} * \bar{\sigma} \overset{\sim}{\to} \sigma_{i_{n-k}} * \sigma_l * \hat{\sigma} \overset{\sim}{\to} \sigma_l * \sigma_{i_{n-k}} * \hat{\sigma} \overset{\sim}{\to} \sigma_l * \sigma'_{i_{n-k}} * \hat{\sigma}' \overset{\sim}{\to} \sigma'_{i_{n-k}} * \sigma_l * \hat{\sigma}' \overset{\sim}{\to} \sigma'_{i_{n-k}} * \bar{\sigma}';$$

the combinatorial dihomotopy in the middle exists by induction. □

5.3. Dipaths up to dihomotopy between arbitrary points

As mentioned in the introduction to Section 4, Footnote 3, dihomotopy between intermediate states may be non-trivial although dihomotopy between the initial and the terminal state is trivial. A simple example for this phenomenon occurs for $X = \vec{I}^3 \setminus \vec{\mathbf{J}}^3$ with $\vec{\mathbf{J}} \subset \vec{I}$ an open subinterval: all dipaths in X from **0** to **1** are dihomotopic to each other, but there are two dihomotopy classes of dipaths between **x** and **y** in $(\vec{I} \setminus \mathbf{J}) \times (\vec{I} \setminus \mathbf{J}) \times \mathbf{J}$ whenever $x_1, x_2 \leqslant a \leqslant y_1, y_2$ for all $a \in \mathbf{J}$. This example is studied in detail in [5] which determines the *components of the fundamental category* of that state space X.

To study dipaths up to dihomotopy between $\mathbf{x}$ and $\mathbf{y}$ in $X = \vec{I}^n \setminus F$, we have to work with the state space $X_{\mathbf{xy}} = \{\mathbf{z} \in X \mid x_i \leqslant z_i \leqslant y_i,\ 1 \leqslant i \leqslant n\}$. Similar to the techniques in [6,17], it can be regarded as $X_{\mathbf{xy}} = \vec{I}^n \setminus (F \cup F_{\mathbf{xy}})$ with

$$F_{\mathbf{xy}} = \bigcup_{1 \leqslant i \leqslant n} (\vec{I} \times \cdots \times \vec{I} \times [0, x_i[\times \vec{I} \times \cdots \times \vec{I}) \cup (\vec{I} \times \cdots \times \vec{I} \times]y_i, 1[\times \vec{I} \times \cdots \times \vec{I}).$$

This means that $2n$ additional (outer) hyperrectangles are added to the forbidden region.

The techniques from Sections 4 and 5 apply. In particular, if $(n-1)$ of the rectangles in $F \cup F_{\mathbf{xy}}$ have a non-empty intersection (apart from the trivial intersections among the hyperrectangles in $F_{\mathbf{xy}}$), then Proposition 4.3 and Corollary 4.4 will in many cases show that there exist non-dihomotopic dipaths from $\mathbf{x}$ to $\mathbf{y}$. This applies e.g., to the example $X = \vec{I}^3 \setminus \vec{J}^3$ discussed above. If, on the other hand, all relevant intersections of $(n-1)$ hyperrectangles in $F \cup F_{\mathbf{xy}}$ are empty, then Proposition 5.1 shows that all dipaths from $\mathbf{x}$ to $\mathbf{y}$ are dihomotopic to each other.

6. Concluding remarks. Future work

The main results of [6] as described in Proposition 2.2 and the end of Section 2 show that the "ordered combinatorics" of intersections of hyperrectangles in the forbidden region associated to a mutual exclusion model can be applied to yield a very efficient algorithm determining deadlocks, unsafe and unreachable regions for such a model space. We have modified these techniques to attack a more difficult problem, i.e., to determine the (number of) essentially different computation paths in such a model. The results indicate that the ordered combinatorics of intersections of hyperrectangles in the forbidden region (at one level lower) again will play a key role.

The ultimate goal for the work initiated in this paper is the construction of an algorithm determining the set of dihomotopy classes between two given states, building on the deadlock algorithm from [6] and generalizing the algorithm given in [17] in the two-dimensional case. To this end, one has to investigate the "directed combinatorics" between situations as they arise in Proposition 4.3 more closely. Several unreachable and unsafe regions (associated to a projection of the forbidden region) can have an interplay that is not that easy to analyse, as you can see in Fig. 4. Moreover, one has to get to grips with situations where projections along various different axes have to be combined.

It should also be interesting to see how the components of the fundamental category of X from [5] relate to this approach.

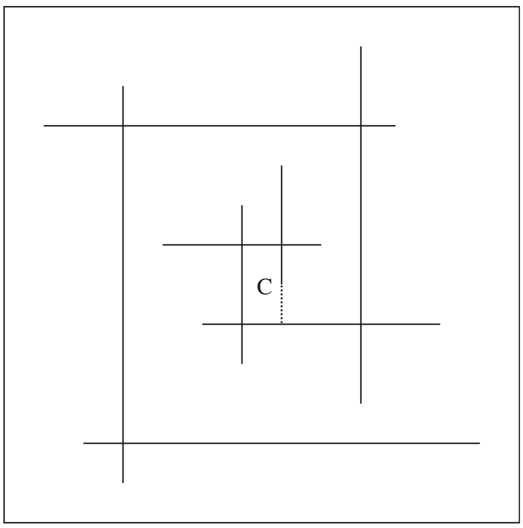


Fig. 4. A labyrinth state space.

References

[1] S.D. Carson, P.F. Reynolds, The geometry of semaphore programs, ACM TOPLAS 9 (1) (1987) 25–53.
[2] E.W. Dijkstra, Co-operating sequential processes, in: F. Genuys (Ed.), Programming Languages, Academic Press, New York, 1968, pp. 43–110.
[3] L. Fajstrup, Discovering spaces, Homology, Homotopy, Appl. 5 (2) (2003) 1–17.
[4] L. Fajstrup, Dipaths and dihomotopies in a cubical complex, Adv. in Appl. Math. 35 (2005) 188–206.
[5] L. Fajstrup, É. Goubault, E. Haucourt, M. Raussen, Components of the fundamental category, Appl. Categ. Structures 12 (1) (2004) 81–108.
[6] L. Fajstrup, É. Goubault, M. Raussen, Detecting deadlocks in concurrent systems, in: D. Sangiorgi, R. de Simone (Eds.), CONCUR '98; Concurrency Theory (Nice, France), Lecture Notes in Computer Science, Vol. 1466, Springer, Berlin, September 1998, Ninth Internat. Conf., Proceedings, pp. 332–347.
[7] L. Fajstrup, É. Goubault, M. Raussen, Algebraic topology and concurrency, Theoret. Comput. Sci. 357 (2006) 241–278.
[8] L. Fajstrup, S. Sokołowski, Infinitely running concurrent processes with loops from a geometric viewpoint, Electronic Notes Theoretical Computer Science, Vol. 39, Springer, Berlin, no. 2, 2000, 19pp., URL: http://www.elsevier.nl/locate/entcs/volume39.html.
[9] É. Goubault, The geometry of concurrency, Ph.D. Thesis, Ecole Normale Superieure, Paris, 1995.
[10] É. Goubault, Geometry and concurrency: a user's guide, Math. Structures Comput. Sci. 10 (4) (2000) 411–425.
[11] É. Goubault, Géométrie du Parallelisme, Analyse Statique et Autres Applications, Mémoire d' habilitation à diriger les recherches, DEIN/SLA, CEA/Saclay, F-91191 Gif-sur-Yvette, France, 2001.
[12] É. Goubault, Some geometric perspectives in concurrency theory, Homology, Homotopy, Appl. 5 (2) (2003) 95–136.
[13] É. Goubault, M. Raussen, Dihomotopy as a tool in state space analysis, in: S. Rajsbaum (Ed.), LATIN 2002: Theoretical Informatics, Cancun, Mexico, April 2002, Lecture Notes in Computer Science, Vol. 2286, Springer, Berlin, Proceedings, pp. 16–37.
[14] M. Grandis, Directed directed homotopy theory I. The fundamental category, Cahiers Topologie Géom. Différentielle Catégoriques 44 (2003) 281–316.
[15] J. Gunawardena, Homotopy and concurrency, Bull. EATCS 54 (1994) 184–193.
[16] V. Pratt, Modelling concurrency with geometry, Proc. 18th ACM Symp. on Principles of Programming Languages, 1991, pp. 311–322.
[17] M. Raussen, On the classification of dipaths in geometric models for concurrency, Math. Struct. Comput. Sci. 10 (4) (2000) 427–457.
[18] M. Raussen, State spaces and dipaths up to dihomotopy, Homology, Homotopy, Appl. 5 (2) (2003) 257–280.

Acta Informatica (2007) 43:501–519
DOI 10.1007/s00236-006-0037-5

ORIGINAL ARTICLE

Geometric analysis of nondeterminacy in dynamical systems

Towards a geometric analysis of concurrent systems

Rafael Wisniewski · Martin Raussen

Received: 9 January 2006 / Revised: 7 November 2006 /
Published online: 13 January 2007

Abstract This article intends to provide some new insights into concurrency using ideas from the theory of dynamical systems. Inherently discrete concurrency corresponds to a parallel continuous concept: a discrete state space corresponds to a differential manifold, an execution path corresponds to a flow line of a dynamical system. To model non-determinacy within dynamical systems, we introduce a new geometrical object, a section cone. A section cone is a convex set in the space of vector fields, all elements having the same singular points. We show that it is enough to consider flow lines of a *single* vector field in order to capture the behavior of all flow lines in the section cone up to homotopy (corresponding to equivalence of executions).

1 Introduction

A major problem in concurrency is connected with the so-called state space explosion making it impossible to evaluate all possible interleaved schedules of a concurrent programme. Many methods have been proposed to reduce the numbers of schedules to be considered. Some of these suggestion are geometrically founded, e.g. [7–9,20,24]. In this article, we analyze a situation that deals with analogous problem in a continuous environment. The discrete state space is replaced by a differential manifold (with infinitely many points/states). A dynamical system on such a manifold gives rise to a *collection* of flow (solution) lines, that are directed as the schedules in a concurrent programme. Flow lines themselves are determined by an initial condition and have thus not the same flexibility as schedules have (if interleaving and subdivision is

R. Wisniewski ( )
Department of Electronic Systems, Aalborg University, Aalborg, Denmark
e-mail: raf@control.aau.dk

M. Raussen
Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark
e-mail: raussen@math.aau.dk

considered). To model an analogous amount of flexibility, we investigate the solutions corresponding to a section cone. More fuzzy than a dynamical system, it allows choices of paths in different directions, but at the same time it gives rise to a partial order on the state space.

The prize is an even higher number of flow lines to be considered. We then show, that it is enough to consider the flow lines of a "core" dynamical system in order to capture the behaviour of all flow lines in the section cone up to homotopy. This is already a dramatic reduction of the solutions to be considered, although we still end up with infinitely many of them. But now, choices can only be made at "critical" points of the dynamical system. Furthermore, it is anticipated, that sophisticated algebraic topological methods will allow to achieve a reduction to a *finite* and *computable* number of flow lines up to homotopy. Although the state spaces considered in the references above are not differential manifolds on the nose, we believe that insights gained from the differential approach (in particular focusing at critical points) might help to the further analysis of concurrency problems from a geometric viewpoint.

1.1 Continuous parallel to concurrency

Due to interleaving the execution of the same program may result in different execution paths. It is essential that small variations of a concurrent program do not change its performance qualitatively. Consider two transition systems both with two vertices a and b and a single edge $a \to b$. V. Pratt in [20] pointed out that with true concurrency of these two transition systems a trace is contained in a filled) square whose four sides are $(a,a) \to (a,b)$, $(a,b) \to (b,b)$, $(a,a) \to (b,a)$, and $(b,a) \to (b,b)$, see the diagram below,

$$\begin{array}{ccc} (a,b) & \longrightarrow & (b,b) \\ \uparrow & & \uparrow \\ (a,a) & \longrightarrow & (b,a). \end{array}$$

A generalization of this observation led to the definition of higher-dimensional automata, cf. [20], in which squares, cubes and hypercubes are pasted together possibly in a complicated pattern. General higher-dimensional automata allow models of state spaces, in which certain forbidden regions are removed with the effect that execution paths are restricted, since shared resources can only be used by a limited number of processes at the same time. Loops, branching and mergings in programs can result in a complex geometry. Geometric properties were investigated in [6–8].

We return to the example above and observe that if subdivision and different interleavings are allowed, any trace from (a,a) to (b,b) might be modelled as a flow line of some vector field belonging to the cone spanned by $(a,a) \to (a,b)$ and $(a,a) \to (b,a)$, Fig. 1. With this motivation we shall introduce a new geometric concept, a section cone that allows a continuum of directed flow lines through every point.

Figure 2 illustrates two processes T_1 and T_2 running in parallel. The white hole represents a forbidden region. Each flow line illustrates different scheduling. We remark that due to causality (time does not flow backwards) the vector fields at each point are locally encapsulated in a cone.

We simplify the picture in this article by assuming that the underlying space is a closed smooth manifold, say M. Executions correspond to smooth curves on M.

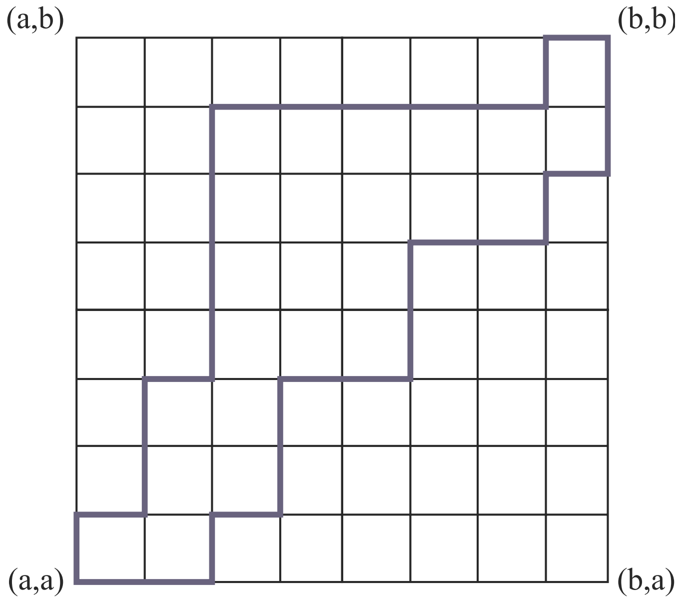


Fig. 1 Due to subdivision and interleaving the execution paths belong to a cone spanned by $(a, a) \rightarrow (a, b)$ and $(a, a) \rightarrow (b, a)$

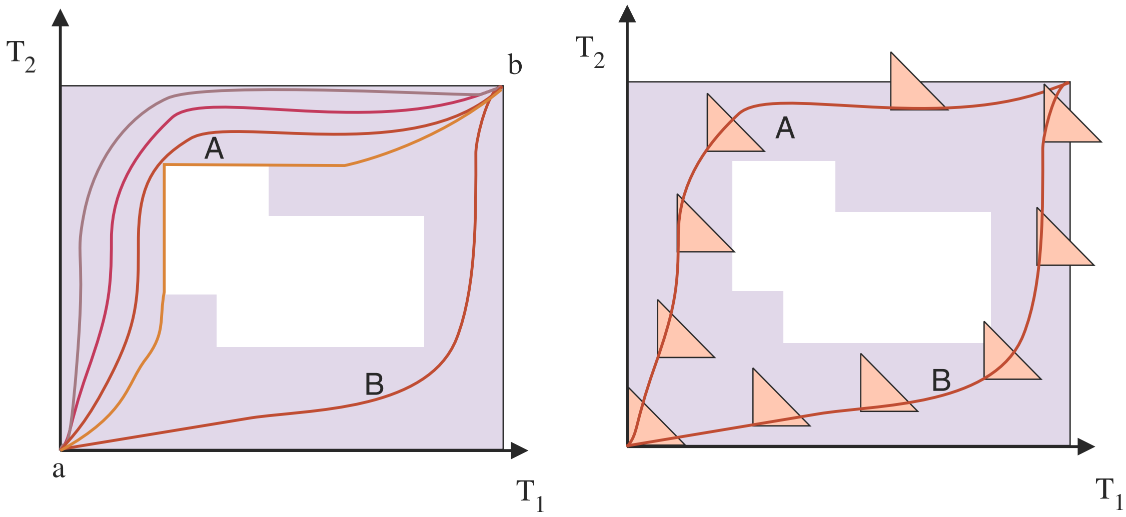


Fig. 2 A concurrent program can be modelled as a flow line corresponding to a vector field lying within a section cone

Due to causality, only certain curves can be allowed: We define a section cone as a convex subset $\mathcal{K}$ of the vector space of smooth vector fields on M characterized by the property that if a is a singular point for some vector field in $\mathcal{K}$ then this is the case for all members of $\mathcal{K}$, Section 3. Allowable curves are then flow lines tangent to a vector field in the section cone $\mathcal{K}$. We denote a flow line of ξ by $\phi_x^\xi(t)$, that is

$$\frac{d}{dt}\phi_x^\xi(t) = \xi\left(\phi_x^\xi(t)\right) \quad \text{with } \phi_x^\xi(0) = x.$$

The manifold M is compact thus the vector field ξ generates a 1-parameter group $\phi_t^\xi : M \rightarrow M$, $t \in \mathbb{R}$, of diffeomorphisms and the smooth flow map $\Phi^\xi : \mathbb{R} \times M \rightarrow M$ related in the following way to each other

$$\Phi^\xi(t, x) = \phi_t^\xi(x) = \phi_x^\xi(t).$$

Flow lines of a particular vector field ξ are completely characterized by an initial condition x. Flow lines within a section cone allow more flexibility corresponding

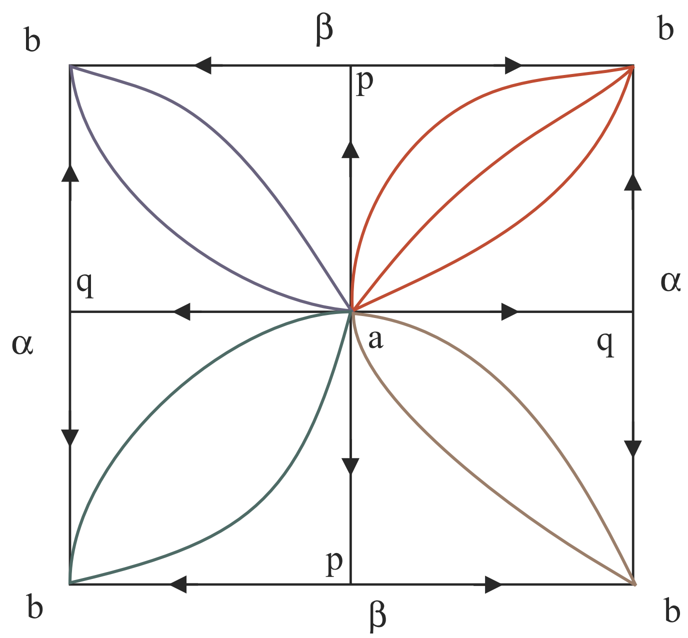


Fig. 3 Flow lines on a torus. The sides and points of the square indicated by the same letters are identified. The set of flow lines from a to b has four components contained in the four "squares" that constitute the torus

to non-determinacy. In order to reduce the number of scenarios in a validation task we consider two types of deformations of flow lines with fixed endpoints: homotopy by a vector field (a one-parameter deformation of flow lines tangent to a specific vector field) and homotopy by a section cone (a one-parameter deformation of flow lines, with flow lines tangent to arbitrary vector fields in the section cone). Executions corresponding to homotopic flow lines have the same global behaviour.

- Two flow-lines γ_0, γ_1 of a vector field ξ joining two singular points a and b are said to be homotopic by ξ if there is a homotopy H such that H_t is a flow line of ξ joining a and b, and $H_0 = \gamma_0$, $H_1 = \gamma_1$.
- Two flow lines: α_0 of a vector field ξ_0 and α_1 of a vector field ξ_1, both joining singular points a and b are di-homotopic if there exist a path $\sigma : I \to \mathcal{K}$ and a path $\beta : I \to M$ such that for any $t \in I$, we have $\beta(t)$ is a flow line from a to b corresponding to the vector field $\sigma(t)$ and $\gamma_0(t) = \phi^{\sigma(0)}_{\beta(0)}(t)$, $\gamma_1(t) = \phi^{\sigma(1)}_{\beta(1)}(t)$.

Both homotopy notions define equivalence relations on the set of flow lines. In Fig. 2, the set of the flow lines up to homotopy by the indicated section cone consists of two classes that cannot be deformed into each other by any of the homotopy notions considered. The two classes have also different semantic interpretations that are equivalent to:

1. Execute first the process T_1 then T_2;
2. Execute first T_2 then T_1.

The set of equivalence classes of flow lines from a to b of a Morse–Smale vector field ξ, Sect. 2, up to homotopy by ξ corresponds to the set of connected components of the space of flow lines joining a with b. Figure 3 presents a torus with a vector field given by the gradient of the height function. The set of flow lines from a to b has four connected components. The article shows that for a Lyapunov–Smale section cone, Sect. 3, the classes of flow lines joining a and b up to homotopy by $\xi \in \mathcal{K}$ are in one-to-one correspondence with the classes of flow lines connecting a and b up to homotopy by $\mathcal{K}$. This is the subject of Sect. 5.

It seems possible at least for gradient like vector fields, cf. [5,21], to do algebraic calculations allowing to determine and characterize flow-lines of a particular vector field up to homotopy with respect to a particular vector field ξ. By the result above, these calculations then extend to the more realistic scenario characterizing flow-lines within a section cone $\mathcal{K}$.

Ultimately, knowledge of the connected components of the space of execution paths would allow to test the possible outcome of a concurrent program on just one representative "schedule" in every connected component. An effective calculation of the set of these equivalence classes can thus be viewed as an antidote to the state space explosion problem. In the article, however, we make a number of assumptions which simplify the problem in hand. We assume that the state spaces is a smooth manifold and the execution paths are flow lines of a section cone. A more accurate, however considerably more complex, model of concurrency consists of a state space composed of hypercubes attached to each other by a pasting scheme. Then the execution paths correspond to absolutely continuous functions—solutions of a differential inclusion, say $\dot{x} \in F(x)$, where F is a cone-valued set function, cf. [2,23]. It is then anticipated that the problem of reducing the solutions can be reformulated in terms of viability theory, e.g. [2]. The concepts of a differential inclusion and a section cone are related, c.f. Remark 1 in Section 3. An important distinction, however, is that a flow line of any vector field belonging to the inclusion at any time is a solution of the corresponding differential inclusion; while the flow line of a section cone must follow a single vector field chosen arbitrarily from among those available in the section cone.

2 Gradient like vector fields

We recall some basic facts from the geometric theory of dynamical systems.

Let M^n be a closed smooth manifold, where n stands for the dimension of M. Let $\mathfrak{X}^r(M)$ be the space of C^r vector fields (having continuous derivatives up to r) on M. Suppose $\xi \in \mathfrak{X}^r(M)$ $(r \geq 1)$ and a is a *singular point* of ξ, that is $\xi(a) = 0$. Consider a local chart (ψ, U) with $a \in U$ and $\psi(a) = 0$. In these local coordinates, ξ is represented by

$$\hat{\xi} = d\psi \xi \circ \psi^{-1}.$$

Definition 1 (Section 2.3 in [17]) Suppose $\xi \in \mathfrak{X}^r(M)$. A singular point $a \in M$ is called *hyperbolic* if and only if the differential of $\hat{\xi}$ at 0, $d\hat{\xi}_0 : \mathbb{R}^n \to \mathbb{R}^n$ is *hyperbolic*, i.e. $d\hat{\xi}_0$ does not have any complex eigenvalues whose real part is zero.

The significance of the notion of a hyperbolic singular point stems from the following observation. If $L \in \mathcal{L}(\mathbb{R}^n)$, that is $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map, which is hyperbolic, then there is a direct sum decomposition

$$\mathbb{R}^n = E^s \oplus E^u,$$

where E^s and E^u are invariant subspaces for L. Moreover, the eigenvalues of $L^s \equiv L|_{E^s}$ have negative real part and the eigenvalues of $L^u \equiv L|_{E^u}$ have positive real part, c.f. Ch. 7 in [12].

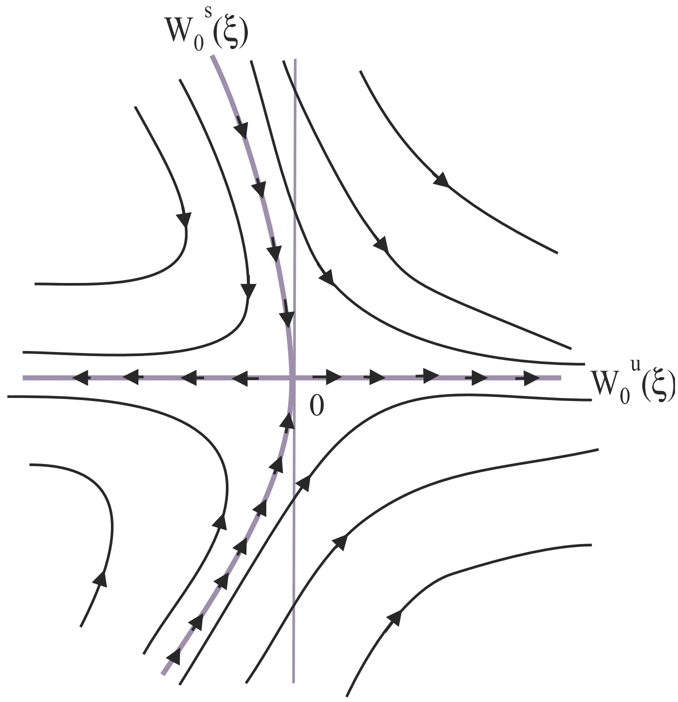


Fig. 4 Stable and unstable manifolds for $\dot{x} = x + y^2$, $\dot{y} = -y$

The *stable manifold* of ξ at a singular point a, cf. Sect. 2.6 in [17], is the set of all initial values $x \in M$ such that the flow $\phi_x^{\xi}(t)$ converges to a with t going to infinity,

$$W_a^s(\xi) = \Big\{x \in M| \lim_{t\to+\infty} \phi_x^{\xi}(t) = a\Big\}.$$

The *unstable manifold* of ξ at a is

$$W_a^u(\xi) = \Big\{x \in M| \lim_{t\to-\infty} \phi_x^{\xi}(t) = a\Big\}.$$

In particular if $L \in \mathcal{L}(\mathbb{R}^n)$ then $W_0^s(L) = E^s$ and $W_0^u(L) = E^u$. The stable and unstable manifolds of the vector field

$$\xi(x, y) = \Big(x + y^2, -y\Big)$$

at the origin are depicted in Fig. 4.

The next theorem states that if a is a hyperbolic singular point of $\xi \in \mathfrak{X}^r(M)$ then the stable (unstable) manifold is indeed a C^r manifold with dimension (co-dimension) corresponding to the index of ξ. Recall that *index* of the vector field ξ is the number of negative eigenvalues of the differential of its local representation at 0, $d\hat{\xi}_0$.

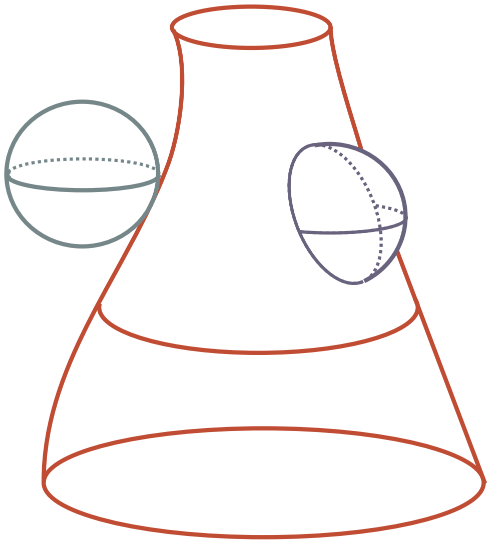
Theorem 1 (Global stable manifold theorem 4.15 in [3]) *Suppose $\xi \in \mathfrak{X}^r(M)$, $r \geq 1$, a is a hyperbolic singular point and λ is the index of ξ. Then $W_a^s(\xi)$ is the surjective image of a C^r injective immersion*

$$\alpha^s : \mathbb{R}^\lambda \to W_a^s(\xi) \subset M.$$

Hence, $W_a^s(\xi)$ is an injectively immersed open disk in M. Furthermore,

$$T_a(W_a^s(\xi)) = T_a(M)^s. \tag{1}$$

Fig. 5 The left sphere does not intersect the "deformed" cylinder transversally in $\mathbb{R}^3$



The equality (1) states that the tangent space of the stable manifold $W_a^s(\xi)$ at the singular point a coincides with the stable subspace of the tangent space of the manifold M at a,

As mentioned in the Introduction, small variations of a concurrent program do not change essentially its performance. Therefore, in the continuous model we focus on a subset $\mathfrak{G}^r(M)$ of vector fields $\mathfrak{X}^r(M)$ having the following stability property: If $\xi \in \mathfrak{G}^r(M)$ then there is an open set U in $\mathfrak{X}^r(M)$ such that any vector field in U is in a certain sense equivalent to ξ. To make this precise, we need the definition of a transversal intersection.

Definition 2 (Section 1.3 in [17], Chap. 3 in [11]) Suppose that N_1 and N_2 are embedded submanifolds of M. We say that N_1 intersects N_2 *transversally* if, whenever $p \in N_1 \cap N_2$, we have $T_p(N_1) + T_p(N_2) = T_p(M)$.

Definition 2 is illustrated in Fig. 5. Only the right sphere intersects the "deformed" cylinder transversally in $\mathbb{R}^3$. Notice that if the right sphere is slightly perturbed, the intersection is still homeomorphic to a circle. Whereas perturbing the left sphere arbitrarily little gives a circle or an empty set (the left sphere and the cylinder are disjoint).

We are ready to define elements of the set $\mathfrak{G}^r(M)$ discussed above.

Definition 3 A vector field ξ on M will be called *gradient-like* provided it satisfies the following three conditions:

1. The vector field ξ has a finite number of singular points, say $\beta_1, ..., \beta_k$, each hyperbolic.
2. Let

$$\alpha(x) = \bigcap_{\tau \le 0} \overline{\bigcup_{t \le \tau} \phi_t^\xi(x)}$$

and

$$\omega(x) = \bigcap_{\tau \ge 0} \overline{\bigcup_{t \ge \tau} \phi_t^\xi(x)},$$

where $\overline{A}$ stands for the closure of A.

Then for each $x \in M$, $\alpha(x) = \{\beta_i\}$ and $\omega(x) = \{\beta_j\}$ for some i and j.

3. Let $\Omega(\xi)$ be the set of nonwandering points[1] for ξ, then $\Omega(\xi) = \{\beta_1, ..., \beta_k\}$.
4. The stable and unstable manifolds associated with the β_i have transversal intersection.

The set of gradient-like vector fields on M is denoted by $\mathfrak{G}^r(M)$.

We remark that a gradient-like vector field is a Morse–Smale vector field, cf. Sect. 4.1 in [17], which does not have any closed orbits.

We conclude the section by stating important results on the structural stability of gradient-like vector fields.

Definition 4 Two vector fields $\xi, \eta \in \mathfrak{X}^r(M)$ are *topologically equivalent* if there exists a homeomorphism $h : M \to M$ (h is continuous and has continuous inverse) such that

1. $h \circ \Phi^{\xi}(\mathbb{R}, x) = \Phi^{\eta}(\mathbb{R}, h(x))$ for each $x \in M$,
2. h preserves the orientation, that is if $x \in M$ and $\delta > 0$ there exists $\epsilon > 0$ such that, for $0 < t < \delta$, $h \circ \Phi^{\xi}(t, x) = \Phi^{\eta}(\tau, h(x))$ for some $0 < \tau < \epsilon$.

The map h is called a topological equivalence.

The first condition of the definition states that the homeomorphism h takes orbits into orbits. The second states that a stable manifold of ξ goes to a stable manifold of η. Specifically, for a pair of topologically equivalent vector fields ξ and η via a homeomorphism $h : M \to M$ and a singular point p we have $W^s_{\xi}(p) = h^{-1}(W^s_{\eta}(h(p)))$.

Example 1 Consider two vector fields ξ, η on $\mathbb{R}^2$ given by

$$\xi(x, y) = (x, y) \quad \text{and} \quad \eta(x, y) = (x + y, -x + y).$$

The origin corresponds to a node source for the vector field ξ, and a spiral source for η. By the Grobman-Hartman Theorem, Proposition 2.14 in [17], any two linear hyperbolic vector fields with the same indices are topologically equivalent; in particular, ξ and η have index 0 and are thus topologically equivalent.

As mentioned before, we are interested in the behaviour of a vector field whose orbits do not change qualitatively under small perturbations.

Definition 5 A vector field $\xi \in \mathfrak{X}^r(M)$ is structurally stable if there exists an open neighborhood U of ξ in $\mathfrak{X}^r(M)$ such that every $\eta \in U$ is topologically equivalent to ξ.

If $\xi \in \mathfrak{X}^r(M)$ is a gradient-like (or more generally Morse-Smale) vector field then ξ is structurally stable, cf. Theorem 4.1 in [18]. This means that small perturbations of a gradient-like vector field behave qualitatively the same.

It is often easier to analyze a vector field using an associated function defined on M. This idea has been successfully used in both classical mechanics and control engineering for stability analysis.

A function $f : M \to \mathbb{R}$ is *Morse* if all its critical points (p is a *critical point* for f if $df_p = 0$) are isolated, cf. [14]. We denote the set of all critical points of f by $\mathcal{C}r(f)$,

$$\mathcal{C}r(f) \equiv \{p \in M |\ df_p = 0\}.$$

[1] We say that $p \in M$ is a wandering point for ξ if there exists a neighborhood V of p and a number t_0 such that $\phi^{\xi}_t(V) \cap V = \emptyset$ for $|t| > t_0$. Otherwise we say that p is nonwandering.

Definition 6 (Lyapunov function for a gradient-like vector field, [13]) Let ξ be a gradient-like vector field on M. Then a Morse function f for M will be called a *Lyapunov function* for ξ provided

1. $\xi(f)(x) < 0$ for all $x \in M - \mathcal{C}r(f)$;
2. There exists a real number $\kappa > 0$ and neighborhoods V_i of the critical points δ_i of f, $i = 1, ..., l$ such that

$$-\xi(f)(y) \geq \kappa d(y, \delta_i)^2 \quad \text{for } y \in V_i.$$

   Here, d is the distance induced by a Riemannian metric g on M (c.f. Example 3 in the next section) and $\xi(f)$ is the Lie derivative of the function f along the vector field ξ.

Condition 1 of Definition 6 states that f decreases along any non-trivial flow line of ξ. By Theorem 1 in [13], every gradient-like (Morse–Smale) vector field admits a Lyapunov function.

3 Section cones

We define a new mathematical object—a section cone—a generalization of a vector field to a field of cones. Intuitively, a section cone is characterized by: At every point $p \in M$, a section cone localizes to a cone in the tangent space $T_p(M)$. This cone varies smoothly with p. Given a section cone, one can define the di-paths (directed paths) associated to it. A di-path is a curve which is a finite concatenation of integral arcs of the vector fields within the section cone. Later in this section we will see that a certain section cone - a Lyapunov section cone - defines a partial order relation $\succeq$ on M by $p \succeq q$ if and only if there is a di-path from p to q. For motivation, we refer to Fig. 2 in the Introduction, a concurrent program flows within a field of cones. In the figure the cones are constructed in a way that the induced partial order is the “natural one” in the concurrency model.

Let us make precise the notion of a cone in a vector space.

Definition 7 ([4]) Let V be a real vector space. A *cone* K in V is a subset of V satisfying

1. If $\alpha, \beta \geq 0$ and $x, y \in K$, then $\alpha x + \beta y \in K$;
2. $K \cap (-K) = \{0\}$.

It follows from the definition that a cone is a convex set containing 0, that is

$$\alpha x + (1-\alpha) y \in K \quad \text{for all } x, y \in K \quad \text{and} \quad 0 \leq \alpha \leq 1.$$

If x and $-x$ are in K then $x = 0$. For a finite dimensional vector space V the dimension of a cone K, $\dim(K)$ is by definition the dimension of the subspace $K - K = \{x - y \mid x, y \in K\}$. The cone K in V is *reproducing* if and only if $K - K = V$.

Example 2 If $\{v_i\}_{i=1,...,n}$ is a basis and n is the dimension of V, then

$$\mathrm{span}^+\{v_1, ..., v_n\} \equiv \{w \in V \mid w = \alpha_1 v_1 + \cdots + \alpha_n v_n,\ \alpha_i \geq 0\}$$

is a reproducing cone. In particular the quadrant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0\}$ is a reproducing cone.

We denote the set of singularities of a vector field ξ by

$$\mathcal{C}r(\xi) \equiv \{p \in M |\, \xi(p) = 0\}.$$

Definition 8 (Section Cone) Let M be a smooth manifold. A C^r *section cone* $\mathcal{K}$ on M is a subset of $\mathfrak{X}^r(M)$ satisfying the following two conditions:

1. For every pair $\xi, \eta \in \mathcal{K}$, $\mathcal{C}r(\xi) = \mathcal{C}r(\eta)$.
2. If ξ and η are in $\mathcal{K}$ and $\alpha, \beta > 0$ then $\alpha\xi + \beta\eta \in \mathcal{K}$.

The first condition says that all vector fields in a section cone have the same singularities. Also if the zero section 0_M is in $\mathcal{K}$ then $\mathcal{K} = \{0_M\}$. The second condition imposes convexity on the subset $\mathcal{K}$. Particularly, if $\xi \in \mathcal{K}$ then $\alpha\xi \in \mathcal{K}$ for $\alpha > 0$.

Condition 1 allows to speak about singular points of a section cone.

Definition 9 A point p is a *singular point* of a section cone $\mathcal{K}$ if $p \in \mathcal{C}r(\xi)$ for some, thus for all, $\xi \in \mathcal{K}$. We denote the set of singular points of $\mathcal{K}$ by $\mathcal{C}r(\mathcal{K})$.

Example 3 Let g be a Riemannian metric on M, i.e. a pointwise inner product on T_pM varying smoothly in p. We pick $\eta \in \mathfrak{X}^r(M)$ and define the set $\mathcal{K}(\eta) \subset \mathfrak{X}^r(M)$ by

$$\mathcal{K}(\eta) = \{\alpha(\eta + \xi) \in \mathfrak{X}^r(M) |\, \xi \in \mathfrak{X}^r(M), \alpha > 0,\ g(\xi, \eta) = 0, g(\eta, \eta) \geq g(\xi, \xi)\}.$$

Note that for $\eta+\xi \in \mathcal{K}(\eta)$ we have $\eta(p) = 0$ for some $p \in M$ if and only if $(\eta+\xi)(p) = 0$. Furthermore, if $\vartheta_i = \alpha_i(\eta + \xi_i) \in \mathcal{K}(\eta)$ for $\alpha_i > 0$ and $i \in \{1, 2\}$ then

$$\|\alpha_1\xi_1 + \alpha_2\xi_2\|^2 \leq (\alpha_1 + \alpha_2)^2\|\eta\|^2, \quad \text{where } \|\cdot\|^2 \equiv g(\cdot, \cdot).$$

Hence $\vartheta_1 + \vartheta_2 = (\alpha_1 + \alpha_2)\eta + \alpha_1\xi_1 + \alpha_2\xi_2 \in \mathcal{K}(\eta)$, and $\mathcal{K}(\eta)$ is a section cone.

We shall use the notation $\mathcal{K}(p) \equiv \{s(p) |\, s \in \mathcal{K}\} \subset T_p(M)$. In particular, $p \in \mathrm{Cr}(\mathcal{K})$ if and only if $\mathcal{K}(p) = \{0\}$.

Proposition 1 *Let $\mathcal{K}$ be a section cone. If $\xi, \eta \in \mathcal{K}$ and $\xi(p) = -\eta(p)$ for some $p \in M$ then $p \in \mathcal{C}r(\xi)$. As a consequence, for each $x \in M$, the set $\mathcal{K}(x) \cup \{0\}$ is a cone in the vector space $T_x(M)$.*

Proof Since $\xi, \eta \in \mathcal{K}$, $\vartheta = \xi + \eta \in \mathcal{K}$. In particular

$$\vartheta(p) = \xi(p) + \eta(p) = -\eta(p) + \eta(p) = 0.$$

We conclude that $p \in \mathcal{C}r(\vartheta)$. By condition 1 the point p is also a singular point of ξ. □

Since $\mathfrak{X}^r(M)$ has the structure of a real vector space we may consider a cone in $\mathfrak{X}^r(M)$. Proposition 1 implies that if $\mathcal{K}$ is a section cone then $\mathcal{K} \cup \{0_M\}$ (0_M is the zero section) is a cone in $\mathfrak{X}^r(M)$.

Remark 1 An alternative tool for modelling non-determinacy within dynamical systems is a differential inclusion, e.g. [2,23]. The concepts of a section cone and a differential inclusion are related but not equivalent.

Let $\mathcal{K}$ be a C^r section cone on an open set U of $\mathbb{R}^n$. The evaluation map $e_{\mathcal{K}}$: $U \times \mathcal{K} \to \mathbb{R}^n$ given by

$$e_{\mathcal{K}}(x, \xi) = \xi(x)$$

is C^r, cf. Corollary B.16 in [1]. We define a set-valued map $F : U \to \mathbb{R}^n$ by

$$F(x) = \{e_{\mathcal{K}}(x, \xi) | \xi \in \mathcal{K}\} = \mathcal{K}(x).$$

Since in particular $e_{\mathcal{K}}$ is continuous, the set-valued map F is lower semi-continuous.

A flow line of the section cone $\mathcal{K}$, that is a flow line of a vector field $\xi \in \mathcal{K}$ is also a solution of the differential inclusion

$$\dot{x} \in F(x). \tag{2}$$

The converse, however, does not hold, as can be seen in the following example. Let U be an open neighborhood of 0 in $\mathbb{R}^2$ and consider two vector fields

$$\xi_1 : x \mapsto \begin{bmatrix} a & 0 \\ \lambda_1 & b \end{bmatrix} x \quad \text{and} \quad \xi_2 : x \mapsto \begin{bmatrix} a & 0 \\ \lambda_2 & b \end{bmatrix} x,$$

with $a, b, \lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1 < \lambda_2$. Suppose $\mathcal{K}$ is the section cone spanned by ξ_1 and ξ_2, i.e. $\eta \in \mathcal{K}$ if and only if there are positive reals α, β such that $\eta = \alpha\xi_1 + \beta\xi_2$. Let $\lambda : U \to [\lambda_1, \lambda_2]$ be a non-constant smooth function. Then the solution of the differential equation

$$\dot{x} = \begin{bmatrix} a & 0 \\ \lambda(x) & b \end{bmatrix} x, \; x(0) \neq 0 \tag{3}$$

is also a solution of the differential inclusion (2), but is not a flow line of $\mathcal{K}$.

In conclusion, we have defined in the equation (3) the vector field on U, say ξ, such that $\xi(x) \in F(x)$ for all $x \in U$. Then the flow line ϕ_p^{ξ} of ξ ($p \in U$) is a solution of the differential inclusion (2). On the other hand $\xi \notin \mathcal{K}$, thus ϕ_p^{ξ} is not a flow line of the section cone $\mathcal{K}$. Paraphrasing, the solution of a differential inclusion follows any vector field in the inclusion at any time. While a solution of a section cone must follow a single vector field picked from among those in the section cone.

In the following definitions we impose more structure on the section cone.

Definition 10 A C^r section cone $\mathcal{K}$, $r \geq 1$, on a closed smooth manifold M is *gradient-like* if and only if the section cone $\mathcal{K} \subset \mathfrak{G}^r(M)$, i.e. all vector fields in $\mathcal{K}$ are gradient-like.

If $\mathcal{K}$ is a gradient-like section cone then the singular points are isolated. Since any two vector fields $\xi, \eta \in \mathcal{K}$ can be connected by a path in $\mathcal{K} \subset \mathfrak{G}^r(M)$, ξ and η are topologically equivalent (there exists a homeomorphism $h : M \to M$ taking orbits of ξ to orbits of η). Another consequence is that the indices of ξ and η at the same singular point p agree. This will be shown in Theorem 2.

Define a set $\mathfrak{X}_p^r(M) = \{\xi \in \mathfrak{X}^r(M) | \xi(p) = 0 \text{ and } p \text{ is hyperbolic}\}$. The index function on $\mathfrak{X}_p^r(M)$ is

$$\mathrm{Ind}_p : \mathfrak{X}_p^r(M) \to \mathbb{N}, \; \xi \mapsto \mathrm{index}_{\xi}(p),$$

where $\mathrm{index}_{\xi}(p)$ is the index of the vector field ξ at the singular point p.

Theorem 2 *Suppose $\mathcal{K}$ is a gradient-like section cone and $p \in \mathcal{C}r(\mathcal{K})$. The index function on $\mathcal{K}$ at p, $\mathrm{Ind}_p^{\mathcal{K}} \equiv \mathrm{Ind}_p|_{\mathcal{K}}$, is continuous and thus constant.*

In the proof we make use of the following proposition.

Proposition 2 (Proposition 2.2.18, [17]) *The eigenvalues of an operator $L \in \mathcal{L}(\mathbb{R}^n)$ depend continuously on L.*

Proof (of Theorem 2) We shall denote the map in Proposition 2 by θ

$$\theta : \mathcal{L}(\mathbb{R}^n) \to \mathbb{C}^n/S_n,\ L \mapsto [(\lambda_1, ..., \lambda_n)],$$

where S_n is the symmetric group of degree n, and $\lambda_1, ..., \lambda_n$ are eigenvalues of L, possibly with multiplicities.

Let (ψ, V) be a coordinate neighborhood of the point p with $U \equiv \psi(V)$. Consider the set $\mathfrak{X}^r(U)$ of C^r vector fields on U, that is C^r maps $U \to \mathbb{R}^n$. Define the composition $\kappa = \theta \circ e_p \circ d$, where $d : \mathfrak{X}^r(U) \to C^r(U, \mathcal{L}(\mathbb{R}^n))$ is the derivative, $e : U \times C^r(U, \mathcal{L}(\mathbb{R}^n)) \to \mathcal{L}(\mathbb{R}^n)$ is the evaluation map and $e_p(\cdot) = e(p, \cdot)$. By Proposition 2, θ is continuous, so is the composition κ. We denote by $\varrho : (\mathbb{C} - i\mathbb{R})^n/S_n \to \mathbb{N}$ the map assigning the number of complex numbers with negative real part in the n-tuple of complex numbers. The map ϱ is a continuous discrete valued map. The local representation of Ind_p is the composition $\text{Ind}_p = \varrho \circ \kappa$. Thus we conclude that Ind_p is continuous.

A topological space is connected if and only if every discrete valued map defined on it is constant. The space $\mathfrak{X}^r_p(M)$ has $n+1$ connected component corresponding to index 0 to n. On the other hand any pair $\xi,\ \eta \in \mathcal{K}$ can be connected by a path in $\mathcal{K}$. Hence $\mathcal{K}$ is a subset of one and only one connected component of $\mathfrak{X}^r_p(M)$. Therefore $\text{Ind}^{\mathcal{K}}_p$ is constant. □

Theorem 2 allows to define the index of a singular point $p \in \mathcal{C}r(\mathcal{K})$ as follows.

Definition 11 Suppose $\mathcal{K}$ is a gradient-like section cone and p is a singular point. Then the *index* of $\mathcal{K}$ at p, $\text{index}_{\mathcal{K}}(p) = \text{index}_{\xi}(p)$, for some (thus all) $\xi \in \mathcal{K}$.

The objective of this chapter is to set up a notion of a section cone which induces a partial order relation on M. The candidates are those section cones which do not allow closed orbits. More specifically, we define a Lyapunov section cone.

Definition 12 A C^r section cone $\mathcal{K}$, $r \geq 1$, on a smooth manifold M is *Lyapunov* if and only if there exists a C^r Morse function $f : M \to \mathbb{R}$ and a Riemannian metric on M such that for any $\xi \in \mathcal{K}$ we have

1. $\xi(f)(x) < 0$ for all $x \in M - \text{C}r(\mathcal{K})$,
2. there exist a constant $\kappa > 0$ and open neighborhoods $\{U_q\}_{q \in \text{C}r(\mathcal{K})}$ of the singular points such that

$$-\xi(f)(x) \geq \kappa d(x,p)^2 \text{ for } p \in U_p, \quad \text{where } d \text{ is the Riemannian distance.}$$

It follows from Definition 12 that the function f is a Lyapunov function for each vector field in the Lyapunov section cone $\mathcal{K}$. This means in particular that the function f is monotonically decreasing along all non-trivial flow lines of the vector fields belonging to $\mathcal{K}$.

By a theorem of Peixoto, cf. [19], if M is a two dimensional compact manifold then any Lyapunov section cone on M is also gradient-like. In general, for dimension greater than 2 the above result is not known. Therefore, we define a new type of a section cone that is gradient like and Lyapunov at the same time.

Definition 13 A *Lyapunov–Smale* section cone is a Lyapunov section cone which is gradient-like.

4 Partial orders

A partial order on a set is a reflexive, transitive and antisymmetric relation. It is demonstrated in [7] that partial orders can be appropriate tools for modelling of causality in computer programs. To make an explicit relation with [7] we define a partial order on M by means of flow lines of a section cone, cf. Definition 17. We say that a point q is greater than or equal to a point p if there exists a flow line from p to q corresponding to some vector field in a Lyapunov section cone $\mathcal{K}$. The partial order that—under a certain condition—arises from the transitive closure of that relation—gives rise to (the concept of) a *di-path* (directed path). That is a continuous map from the closed unit interval with the natural partial order inherited from the order of the real numbers to the manifold with the partial order defined as above, which furthermore preserves the partial orders.

An *integral arc* from a point p to a point q on M is a segment of the flow line ϕ_x^ξ for a $\xi \in \mathfrak{X}^r(M)$ and an $x \in M$ joining p with q. If $\alpha > 0$ and $\xi \in \mathfrak{X}^r(M)$, then the orbits of ξ and of $\alpha\xi$ coincide.

Definition 14 $\gamma : I \to M$ is an integral arc of a vector field ξ if there exists an $\alpha > 0$ and an $x \in M$ such that $\phi_x^{\alpha\xi}(t) = \gamma(t)$ for all $0 < t < 1$.

This definition allows to re-parameterize a flow line. Given a flow line ϕ_x^ξ for a vector field ξ and an $x \in M$ let $\phi_x^\xi(t_1) = p$ and $\phi_x^\xi(t_2) = q$ with $t_1 < t_2$. We permit a re-parameterization of the flow ϕ_x^ξ by a function $\beta : \mathbb{R} \to \mathbb{R}$, $\beta(t) = (t_2 - t_1)t + t_1$. Then $\phi_p^{(t_2-t_1)\xi}(t) = \phi_x^\xi(\beta(t))$ with $\phi_p^{(t_2-t_1)\xi}(0) = p$ and $\phi_p^{(t_2-t_1)\xi}(1) = q$. We notice that if ξ is in a section cone $\mathcal{K}$ and $\alpha > 0$, then also $\alpha\xi \in \mathcal{K}$.

We will study paths consisting of a concatenation of a finite number of integral arcs of vector fields in $\mathcal{K}$.

Definition 15 If γ is an integral arc from x_0 to x_1, and μ is an integral arc from x_1 to x_2, then the product $\gamma * \mu$ is the path σ defined by the equation

$$\sigma(t) = \begin{cases} \gamma(2t) & \text{for } t \in [0, \frac{1}{2}], \\ \mu(2t-1) & \text{for } t \in [\frac{1}{2}, 1]. \end{cases}$$

The function σ is well defined, continuous (by the pasting lemma, cf. Theorem 18.3 in [15]) and piecewise C^r.

Given a section cone, we define a di-path as a concatenation of finite number of integral arcs corresponding to the vector fields belonging to this cone.

Definition 16 Suppose $\mathcal{K}$ is a C^r section cone. We call a piecewise C^r path $\sigma : I \to M$ a *di-path* of $\mathcal{K}$ if it is a constant path or there exists a finite set of integral arcs $\{\gamma_1, ..., \gamma_k\}$, for $i = 1, \ldots, k$ where γ_i is an integral arc of the vector field ξ_i satisfying

1. $\{\xi_1,\ldots,\xi_k\} \subset \mathcal{K}$ and
2. $\sigma = \gamma_1 * ... * \gamma_k$, in particular $\gamma_i(1) = \gamma_{i+1}(0)$.

Example 4 We define two C^r section cones ($r \geq 1$) on $\mathbb{R}^2$:

$$\mathcal{K} \equiv \left\{ \alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y} \,\middle|\, \alpha, \beta \geq 0 \right\}$$

and

$$\mathcal{C} \equiv \left\{ \alpha(x,y)\frac{\partial}{\partial x} + \beta(x,y)\frac{\partial}{\partial y} \,\middle|\, \alpha,\beta : \mathbb{R}^2 \to \mathbb{R} \text{ are nonnegative } C^r \text{ functions} \right\}.$$

The flow lines of the section cone $\mathcal{K}$ are monotone straight lines (monotone means here monotone in each coordinate). The flow lines of $\mathcal{C}$ are monotone C^r curves. The di-paths of $\mathcal{K}$ are monotone piecewise straight lines. Lastly, the di-paths of the section cone $\mathcal{C}$ are monotone piecewise C^r curves.

Definition 17 Suppose $\mathcal{K}$ is a section cone on the manifold M. For a pair of points p and q in M we say that $p \succeq_{\mathcal{K}} q$ if and only if there exists a di-path $\sigma : I \to M$ of $\mathcal{K}$ with $\sigma(0) = p$ and $\sigma(1) = q$.

The next theorem shows that if the section cone $\mathcal{K}$ in Definition 17 is Lyapunov then the relation $\succeq_{\mathcal{K}}$ is a partial order relation.

Theorem 3 *Let $\mathcal{K}$ be a Lyapunov section cone. Then the relation $\succeq_{\mathcal{K}}$ is a partial order relation.*

Proof Reflexivity and transitivity follow directly from the definition of the di-path.

We show antisymmetry. Suppose $x \succeq_{\mathcal{K}} y$ and $y \succeq_{\mathcal{K}} x$. Let f be the Lyapunov function associated to the Lyapunov section cone. Then

$$\begin{aligned} x \succeq_{\mathcal{K}} y &\Rightarrow f(x) \geq f(y) \\ y \succeq_{\mathcal{K}} x &\Rightarrow f(x) \leq f(y), \end{aligned}$$

thus $f(x) = f(y)$. Assume $x \neq y$ and let $\sigma = \gamma_1 * \cdots * \gamma_k$ be a di-path joining x with y. However, f is strictly decreasing along any of the γ_i. In particular,

$$f(y) = f(\sigma(1)) < f(\sigma(0)) = f(x).$$

This is a contradiction. □

5 The central vector field theorem

In order to reduce the number of scenarios in a validation task we consider and compare two types of deformations of flow lines (concurrent programs): first, homotopy by a vector field, which is a one-parameter deformation of flow lines tangent to a *specific* vector field; second, homotopy by a section cone, which is a one-parameter deformation of flow lines, with flow lines tangent to *arbitrary* vector fields in the section cone. Both deformations give rise to equivalence relations. Theorem 4 below shows that classes of these two equivalence relations are in one to one correspondence.

For a gradient-like vector field ξ the intersection $W(a,b;\xi) \equiv W^u_a(\xi) \cap W^s_b(\xi)$ is transversal, and $W(a,b;\xi)$ is a manifold of dimension equal to $\text{index}_\xi(\text{b}) - \text{index}_\xi(\text{a})$. Let $P(a,b;\eta)$ be the set of flow lines of η from the singular point a to the singular point b. Furthermore, let the set of flow lines of the vector fields in a section cone $\mathcal{K}$ born in a and dying in b be denoted by $P(a,b;\mathcal{K})$.

Definition 18 Let M be a closed smooth manifold. For $r \geq 1$, let $\xi \in \mathfrak{X}^r(M)$ and $\mathcal{K}$ be a C^r section cone on M.

1. $\gamma_0,\ \gamma_1 \in P(a,b;\xi)$. We say that γ_0 is *homotopic to* γ_1 *by* ξ, write $\gamma_0 \sim_\xi \gamma_1$, if and only if there is a path $\beta : I \to M$ such that $\beta(t) \in W(a,b;\xi)$, $\gamma_0(t) = \phi^{\xi}_{\beta(0)}(t)$ and $\gamma_1(t) = \phi^{\xi}_{\beta(1)}(t)$.
2. Suppose $\gamma_0,\ \gamma_1 \in P(a,b;\mathcal{K})$. We say that γ_0 is *homotopic to* γ_1 *by* $\mathcal{K}$ and write $\gamma_0 \sim_{\mathcal{K}} \gamma_1$ if and only if there exist a path $\sigma : I \to \mathcal{K}$ and a path $\beta : I \to M$ such that $\beta(t) \in W(a,b;\sigma(t))$, $\gamma_0(t) = \phi^{\sigma(0)}_{\beta(0)}(t)$ and $\gamma_1(t) = \phi^{\sigma(1)}_{\beta(1)}(t)$.

We remark that the relations $\sim_{\mathcal{K}}$ and $\sim_\xi$ are equivalence relations.

Theorem 4 *Let M be a closed smooth manifold of dimension n. Suppose $\mathcal{K}$ is a Lyapunov-Smale C^r section cone on M, $r \geq 1$, and $\xi \in \mathcal{K}$. If a, b are singular points of $\mathcal{K}$ with indices* 0 *and n, respectively. Then the map*

$$\Pi : P(a,b;\xi)/_{\sim_\xi} \to P(a,b;\mathcal{K})/_{\sim_{\mathcal{K}}},$$

induced by the inclusion $P(a,b;\xi) \hookrightarrow P(a,b;\mathcal{K})$, is a bijection.

Theorem 4 follows from the following two propositions.

Proposition 3 *Let M be a closed smooth manifold of dimension n, let $\mathcal{K}$ be a gradient-like C^r section cone $r \geq 1$. Suppose that a, b are singular points of $\mathcal{K}$ with $a \succeq_{\mathcal{K}} b$. If $\xi \in \mathcal{K}$ then for any $\eta \in \mathcal{K}$ and any $\gamma_0 \in P(a,b;\eta)$ there is some $\gamma_1 \in P(a,b;\xi)$ such that $\gamma_0 \sim_{\mathcal{K}} \gamma_1$. That is the following composition*

$$P(a,b;\xi) \hookrightarrow P(a,b;\mathcal{K}) \to P(a,b;\mathcal{K})/\sim_{\mathcal{K}}$$

is surjective.

Proof Since $\xi, \eta \in \mathcal{K}$ we can define a path $\sigma \in C^r(I,\mathcal{K})$ by $\sigma(t) = t\xi + (1-t)\eta$ with $\sigma(0) = \eta$ and $\sigma(1) = \xi$. Let c_η be a constant path $c_\eta(t) = \eta$. We define a C^r map $g : I \times I \to \mathcal{K}$ by $g(s,t) = (1-s)c_\eta(t) + s\sigma(t)$. The map g gives rise to a map $G : I \to \mathcal{R}^r(M)$. Pick $\tau > 0$ then $G(s)(t) \equiv G^\tau(s)(t) = \left(t, \phi^{g(s,t)}_\tau\right)$. We shall use the notation $G_s(t) = G(s)(t)$. We note that $G_0(t) = (t, \phi^{\eta}_\tau)$ and $G_1(t) = \left(t, \phi^{\sigma(t)}_\tau\right)$.

By Corollary 3 in Appendix 6, there exists a homeomorphism $h : I \to I$ and there is a map $H : I \times M \to M$ with $H_t : M \to M$, where $H_t(z) \equiv H(t,z)$ is a conjugacy between $G_0(t)$ and $G_1(h(t))$ for all $t \in I$.

Since $\gamma_0 \in P(a,b;\eta)$ there is a point $x \in W(a,b;\eta)$ such that $\gamma_0(t) = \phi^{\eta}_x(t)$ for all $t \in \mathbb{R}$. Let c_x be a constant path in M given by $c_x : t \mapsto x$. Then $c_x(t) \in W(a,b,c_\eta(t))$.

We define a path $\beta : I \to M$ by $\beta(t) = H_0^{-1} \circ H_t \circ c_x \circ h(t)$ if $h(0) = 0$. If $h(0) = 1$ then we define β by $\beta(t) = \beta'(1-t)$ where $\beta'(t) = H_1^{-1} \circ H_t \circ c_x \circ h(t)$. We observe that $\beta(t) \in W(a,b;\sigma(t))$, $\beta(0) = x$, $\sigma(0) = \eta$, and $\sigma(1) = \xi$. If $\gamma_1(t) = \phi^{\xi}_{\beta(1)}(t)$ then $\gamma_1 \in P(a,b;\xi)$ and $\gamma_0 \sim_{\mathcal{K}} \gamma_1$. □

The next proposition shows that the map Π in Theorem 4 is injective.

Proposition 4 *Let M be a closed smooth manifold and $\mathcal{K}$ be a Lyapunov-Smale C^r section cone on M, $r \geq 1$. Let a, b be singular points of $\mathcal{K}$ with indices* 0 *and n, respectively. Suppose $\xi \in \mathcal{K}$. If $\gamma_1, \gamma_2 \in P(a,b;\xi)$ and $\gamma_1 \sim_{\mathcal{K}} \gamma_2$ then $\gamma_1 \sim_\xi \gamma_2$.*

Proof Suppose there is a map $\sigma \in C^0(I,\mathcal{K})$ such that $\sigma(0) = \sigma(1) = \xi$ and a path $\beta : I \to M$ with $\beta(t) \in W(a,b;\sigma(t))$ for $t \in I$. We shall show that there is a path $\beta' : I \to M$ with $\beta'(0) = \beta(0)$ and $\beta'(1) = \beta(1)$ such that $\beta'(t) \in W(a,b;\xi)$.

There is $\sigma' \in C^r(I,\mathcal{K})$ such that $\sigma'(0) = \sigma'(1) = \xi$ and $\beta(t) \in W(a,b;\sigma'(t))$ for $t \in I$, c.f. Proposition 7.4.1 in [25].

Let c_ξ be a constant path $c_\xi(t) = \xi$. We define a C^r map $g : I \times I \to \mathcal{K}$ by $g(s,t) = (1-s)\sigma'(t)+sc_\xi(t)$. We use the same argument as in the proof of Proposition 3. The map g gives rise to a map $G : I \to \mathcal{R}^r(M)$, where $\mathcal{R}^r(M)$ is the set of arcs of Morse-Smale diffeomorphisms on M. Pick $\tau > 0$ then $G(s)(t) \equiv G^\tau(s)(t) = \left(t, \phi_\tau^{g(s,t)}\right)$. We shall use the notation $G_s(t) = G(s)(t)$. Note that $G_0(t) = \left(t, \phi_\tau^{\sigma'(t)}\right)$ and $G_1(t) = (t, \phi_\tau^\xi)$.

By Corollary 3, Appendix 6, there exist a homeomorphism $h : I \to I$ and a map $H : I \times M \to M$, where H_t is a conjugacy between $G_0(t)$ and $G_1(h(t))$ for $t \in I$. We observe that H_0 and H_1 are both selfconjugacies[2] of ϕ_τ^ξ.

The homeomorphism h takes 0 to 0 or 0 to 1. Without loss of generality we assume that $h(0) = 1$. We define a path $\gamma : I \to M$ by $\gamma(t) = H_1^{-1} \circ H_{(1-t)} \circ \beta \circ h^{-1}(1-t)$. Since $H_t(\cdot) \equiv H(t,\cdot)$ is a conjugacy between $G_0(t)$ and $G_1(h(t))$ (orbits go to orbits) we have that $H_\tau \circ \beta \circ h^{-1}(\tau) \in W(a,b;\xi)$ for each $\tau \in I$. Thereby, $\gamma(t) \in W(a,b;\xi)$. We see that $\gamma(0) = \beta(0)$. Moreover, since $\mathcal{K}$ is a Lyapunov-Smale section cone, the points $\gamma(1)$ and $\beta(1)$ are in the same connected component of $W(a,b;\xi)$, c.f Propositions 7.4.9 and 7.4.6 in [25]. It follows that there is a path $\gamma' : I \to M$ connecting $\gamma(1)$ with $\beta(1)$ such that $\gamma'(t) \in W(a,b;\xi)$ for all $t \in I$. The desired path β' is then the concatenation of γ and γ':

$$\beta'(t) = \begin{cases} \gamma(2t) & \text{for } t \in [0, \frac{1}{2}], \\ \gamma'(2t-1) & \text{for } t \in [\frac{1}{2}, 1]. \end{cases}$$

□

We have the following interpretation of Theorem 4. If a and b are singular points of the section cone $\mathcal{K}$ such that the index of a is 0 and the index of b is maximal, then the study of the connected components of the space of flow lines of the vector fields within the section cone $\mathcal{K}$ can be reduced to the study of the connected components of $W(a,b;\xi)$ for an arbitrary $\xi \in \mathcal{K}$.

The following corollaries are easy consequences of Theorem 4.

Corollary 1 *Let M be a closed smooth manifold and $\mathcal{K}$ be a Lyapunov–Smale C^r section cone on M, $r \geq 1$. Let a, b be singular points of $\mathcal{K}$ with indices 0 and n, respectively. If $\xi, \eta \in \mathcal{K}$ then there is a bijection $\Theta : P(a,b;\xi)/_{\sim_\xi} \to P(a,b;\eta)/_{\sim_\eta}$.*

For $p, q \in \mathcal{C}r(\mathcal{K})$, $p \succ_\xi q$ means that there is a non-trivial orbit of ξ from p to q.

Corollary 2 *Let M be a closed smooth manifold and $\mathcal{K}$ be a Lyapunov–Smale C^r section cone on M, $r \geq 1$. Let a, b be singular points of $\mathcal{K}$ with indices 0 and n, respectively. If $\xi, \eta \in \mathcal{K}$ then*

$$a \succ_\xi b \Leftrightarrow a \succ_\eta b.$$

[2] A homeomorphism $G : M \to M$ is called a selfconjugacy of a diffeomorphism $\Psi : M \to M$ if an only if $G \circ \Psi = \Psi \circ G$.

Corollary 2 means that the partial order $a \succ_\xi b$ between critical points a, b does not depend on a particular choice of ξ in $\mathcal{K}$.

6 Conclusion

To model non-determinacy within dynamical systems we introduced a new geometrical object—a section cone. This is a convex subset $\mathcal{K}$ of $\mathfrak{X}^r(M)$ characterized by the property that all vector fields in $\mathcal{K}$ have coinciding singular points. The section cone specifies the set of flow lines of the vector fields belonging to $\mathcal{K}$. We introduced two equivalence relations on the set of flow lines with fixed end points: homotopy by a vector field $\xi \in \mathcal{K}$ and homotopy by the section cone $\mathcal{K}$. The first relation has a potential of being algebraically computable since choices are restricted to finitely many critical points; and the latter is a scenario in a continuous counterpart of concurrency with executions corresponding to flow lines within a section cone. If $\mathcal{K}$ is a Lyapunov–Smale section cone then the classes of flow lines joining two singular points a and b up to homotopy by ξ are in one-to-one correspondence with the classes of flow lines connecting a and b up to homotopy by $\mathcal{K}$.

The result of this paper can be seen as a promising tool for decreasing the number of considered scheduling scenarios in validation. So far, we have only developed the notion of a section cone on a closed smooth manifold. This is a limitation, since a true state space is *not* a differential manifold but "almost"; it can be modelled as a manifold with corners, cf. [10]. There is expectation that the result will be extended to such a true state space.

Appendix: A stability of one-parameter families of diffeomorphisms

We study a one-parameter family $\{\kappa_t\}_{t\in I}$ of diffeomorphisms starting at a Morse–Smale diffeomorphism. For a bifurcation point $t_0 \in I$ the diffeomorphism κ_{t_0} ceases to be Morse–Smale, that is κ_{t_0} have a nonhyperbolic singular point or its stable and unstable manifolds do not intersect transversally. We shall recall that if p is a hyperbolic singularity for a diffeomorphism κ_t, then the stable manifold $W^s_p(\kappa_t)$ is an injectively immersed open disk in M. The same is true for the strong stable and strong unstable manifolds at a nonhyperbolic singular point, cf. Appendix III [22].

Let M^n be a closed smooth manifold. We follow [16] and consider the set of C^r diffeomorphisms on M denoted by $\mathrm{Diff}^r(M)$.

Definition 19 ([16]) We call a C^r map $\kappa : I \times M \to I \times M$ an arc of diffeomorphisms on M if and only if $\kappa(t,x) = (t, \kappa_t(x))$, where $x \mapsto \kappa_t(x)$ is a C^r diffeomorphism for each $t \in I$. The space of arcs of C^r diffeomorphisms on M will be denoted by $\mathcal{P}^r(M)$.

We give $\mathrm{Diff}^r(M)$ and $\mathcal{P}^r(M)$ the C^r topology.

Suppose an arc $\kappa \in \mathcal{P}^r(M)$ with $\kappa_0 \in \mathcal{MS}^r$, where $\mathcal{MS}^r$ is the set of Morse-Smale C^r diffeomorphisms on M. Let $b = b(\kappa) = \inf\{t \in I \mid \kappa(t) \notin \mathcal{MS}^r\}$. As for diffeomorphisms also for arcs (of diffeomorphisms) we can introduce a notion of conjugacy.

Definition 20 ([16]) If $\kappa, \kappa' \in \mathcal{P}^r(M)$, then we say that (h, H) is a conjugacy if $h : I \to I$ is a homeomorphism with $h(b(\kappa)) = b(\kappa')$, $H : I \times M \to M$ is a map with H_t being a conjugacy between κ_t and $\kappa'_{h(t)}$ for all t in some neighborhood of $[0, b(\kappa)]$.

The definition of conjugacies gives rise to the concept of structural stability for arcs of diffeomorphisms.

Definition 21 ([16]) An arc $\kappa \in \mathcal{P}^r(M)$ is stable if there is an open neighborhood U of κ in $\mathcal{P}^r(M)$ such that any $\kappa' \in U$ is conjugate to κ.

The necessary and sufficient conditions for structural stability of arcs of diffeomorphisms have been formulated and proven in [16].

Definition 22 ([16]) The subset $\mathcal{S}^r(M) \subset \mathcal{P}^r(M)$ is the set of arcs κ that satisfy:

1. The limit set of each κ_t has finitely many orbits, $t \in I$;
2. κ has only finitely many bifurcation values, say b_1 to b_s in $(0, 1)$;
3. All stable, strong stable, unstable, and strong unstable manifolds intersects transversally;
4. For each $i \in \{1, ..., s\}$, κ_{b_i} has no cycles and has exactly one non-hyperbolic periodic orbit which is either a noncritical saddle-node, cf. Sec. 3, [16], or a flip, cf. Sec. 4, [16]; this non-hyperbolic orbit unfolds generically.

We shall not explain the meaning of Definition 22, instead we refer to [16] for details and remark merely that any arc of diffeomorphisms κ, such that $\kappa_t \in \mathcal{MS}^r$ for $t \in I$ belongs to the set $\mathcal{S}^r(M)$.

Theorem 5 (Theorem 4.4 in [16]) *For $r \geq 1$, the arcs in $\mathcal{S}^r(M)$ are stable.*

We define the subset $\mathcal{R}^r(M) \subset \mathcal{S}^r(M)$ of arcs κ that satisfy $\kappa_t \in \mathcal{MS}^r$.

Corollary 3 *Let $G : I \to \mathcal{R}^r(M)$, $r \geq 1$, be a map. Then there is a conjugacy between the arc G_0 and G_1. In particular there exist a homeomorphism $h : I \to I$ and a map $H : I \times M \to M$, where H_t is a conjugacy between $G_0(t)$ and $G_1(h(t))$ for $t \in I$.*

Proof The conclusion follows from compactness of the unit interval. We cover I be finite number of open intervals $\{U_i\}_{i\in\{1,...,l\}}$ and propagate the conjugacy from the neighborhood of 0 to 1. □

We shall relate the results on conjugacy of arcs of diffeomorphism with arcs of vector fields, which is the primary object of the study in this article. Below, we show that the Stable Manifold Theorem for Vector Fields follows from for the Stable Manifold Theorem for Diffeomorphisms.

Let M be a compact smooth manifold. A vector field $\xi \in \mathfrak{X}^r(M)$, $r \geq 1$, determines a one-parameter family of C^r diffeomorphisms $\phi_t : M \to M$ for $t \in \mathbb{R}$ given by

$$\phi_t^\xi(x) = \phi_x^\xi(t).$$

Suppose Ψ is a C^r diffeomorphism and a is a fixed point, then we define the stable manifold for Ψ at a by

$$W_a^s(\Psi) = \left\{x \in M | \lim_{n\to+\infty} \Psi^n(x) = a\right\}.$$

Since $(\phi_t^\xi)^n = \phi_{nt}^\xi$ for all $n \in \mathbb{N}$, for any singular point a of ξ and any fixed $t > 0$ we have

$$W_a^s(\phi_t^\xi) = \left\{x \in M | \lim_{n\to+\infty} \phi_{tn}^\xi(x) = a\right\} = \left\{x \in M | \lim_{\lambda\to+\infty} \phi_x^\xi(\lambda) = a\right\} = W_a^s(\xi).$$

References

1. Arwin, M.: Smooth Dynamical Systems. World Scientific, Singapore (1980)
2. Aubin, J.-P.: Viability Theory (Systems and Control: Foundations and Applications). Birkhauser (1991)
3. Banyaga, A., Hurtubise, D.: Lectures on Morse Homology. Kluwer, Dordrecht (2004)
4. Barker, G.: Theory of cones. Linear Algebra and Its Applications. **39**, 263–291 (1981)
5. Cohen, R.: Topics in Morse Theory: Lecture Notes. Stanford University (1991)
6. Fahrenberg, U.: Higher-dimensional automata from a topological viewpoint. Ph.D. Thesis, Aalborg University (2005)
7. Fajstrup, L., Goubault, E., Raussen, M.: Algebraic Topology and Concurrency. Theoretical Computer Science **357**, 241–278 (2006)
8. Goubault, E.: The geometry of concurrency. Ph.D. Thesis, Ecole Normale Superieure, Paris (1995)
9. Goubault, E., Jensen, T.: Homology of higher dimensional automata. In: CONCUR'92, Lecture Notes in Computer Science, vol. 630. Springer, Heidelberg (1992)
10. Handron, D.: Generalized billiard paths and Morse theory for manifolds with corners. Topology Appl. **126**, 83–118 (2002)
11. Hirsch, M.W.: Differential Topology. Springer, Heidelberg (1976)
12. Hirsch, M.W., Smale, S.: Differential Equations, Dynamical Systems, and Linear Algebra. Academic, New York (1974)
13. Meyer, K.R.: Energy functions for Morse-Smale systems. Am. J. Math. **90**(4), 1031–1040 (1968)
14. Milnor, J.: Morse Theory. Annals of Mathematics Studies. Princeton University Press, Princeton (1973)
15. Munkres, J.: Topology. Prentice-Hall, Englewood Chiffs (2000)
16. Newhouse, S.E., Palis, J., Takens, F.: Bifurcations and stability of family of diffeomorphisms. Publications mathématiques de l'I.H.É.S. **57**, 5–71 (1983)
17. Palis, J., de Melo, W.: Geometric Theory of Dynamical Systems. Springer, Heidelberg (1982)
18. Palis, J., Smale, S.: Structural stability theorems. In: Global Analysis. Proceedings of Symposium in Pure, Math. vol. XIV. Americal Math. Soc., pp. 223–231 (1970)
19. Peixoto, M.: Structural stability on two-dimensional manifolds. Topology **1**, 101–120 (1962)
20. Pratt, V.: Modeling concurrency with geometry. In: 18th ACM Symposium on Principles of Programming Languages, pp. 311–322. ACM Press, New York (1991)
21. Schwarz, M.: Morse Homology. Birkhauser (1993)
22. Shub, M.: Global Stability of Dynamical Systems. Springer, Heidelberg (1986)
23. Smirnov, G.V.: Introduction to the Theory of Differential Inclusion. AMS (2002)
24. van Glabbeek, R.: On the expressiveness of higher dimensional automata. Theoret. Comput. Sci. **356**, 265–290 (2006)
25. Wisniewski, R.: Flow Lines under Perturbations within Section Cones. Ph.D. Thesis, Aalborg University (2005)

Appl Categor Struct (2007) 15:355–386
DOI 10.1007/s10485-007-9085-4

Invariants of Directed Spaces

Martin Raussen

Received: 28 April 2006 / Accepted: 3 November 2006 / Published online: 1 June 2007

Abstract Directed spaces are the objects of study within directed algebraic topology. They are characterised by spaces of directed paths associated to a source and a target, both elements of an underlying topological space. The algebraic topology of these path spaces and their connections are studied from a categorical perspective. In particular, we study the preorder category associated to a directed space and various "quotient" categories arising from algebraic topological functors. Furthermore, we propose and study a new notion of directed homotopy equivalence between directed spaces.

Keywords d-space · Preorder category · Component category · Automorphic homotopy flow · Dihomotopy equivalence

Mathematics Subject Classifications (2000) 18A32 · 55Pxx · 55Q05 · 55U40 · 68N30 · 68Q85

1 Introduction

1.1 Background

With motivations arising originally from concurrency theory within Computer Science, a new field of research, directed algebraic topology, has emerged. The main characteristic is, that it involves spaces of "directed paths" (or timed paths, executions): these directed paths can be concatenated, but in general *not* reversed; time is not reversable. These executions can be viewed as objects themselves (this is the point of view of Gaucher, c.f., e.g. [11, 12]) or as elements of subspaces of spaces

M. Raussen ( )
Department of Mathematical Sciences, Aalborg University,
Fredrik Bajers Vej 7G, 9220 Aalborg Øst, Denmark
e-mail: raussen@math.aau.dk

of paths in an underlying topological space (with concatenation as a partial binary operation). We will apply the latter approach, compare also [9, 16].

A particular model in the investigation of concurrency phenomena leads to Higher Dimensional Automata (HDA); for these consult e.g. the recent [24]. The underlying space in these models is then – instead of a directed graph – the geometric realization of a cubical complex; like a simplicial complex, but with cubes as building blocks; cf. e.g. [4, 5]. Every cube carries a natural partial order, and directed paths have to respect the partial orders in their range. In the models, directed paths correspond to executions (calculations); they have two crucial properties:

(1) The reverse of a directed path is, in general, *not* directed;
(2) Directed paths that are *dihomotopic*, i.e., that can be connected through a one-parameter family of directed paths, are equivalent; in the sense that HDA – calculations along these "schedules" will always lead to the same result.

The primary motivation to study directed spaces (i.e., topological spaces with preferred directed paths) stems from the investigation and use of HDAs; a different area, in which directed paths play a natural role, is relativity theory in physics; the directed paths are the time-like or the causal paths in a space – time, cf. [21].

Since one cannot reverse directed paths, one can no longer expect invariants in algebraic structures with inverses like (fundamental) groups or groupoids etc.; they will rather live in categories like the fundamental category and others discussed further on in this article.

A nice and flexible framework for directed paths was introduced by Marco Grandis with the notion of d-spaces and, in particular, of d-paths (cf. Definition 2.1) on a topological space X. In [16] and the subsequent [15], he developed a framework for directed homotopy theory. In particular, he proved a van Kampen theorem for fundamental categories that allows to do calculations on the (refined) d-homotopy sets (cf. Definition 2.1) of directed paths. Moreover, in these and more recent papers by Grandis, the methodology was extended to directed homotopy on general categories.

The success of the use of algebraic invariants like the fundamental group in algebraic topology resides both in the fact that even complicated spaces often have calculable "small" invariant groups and in the exploitation of the algebraic structure in the comparison of invariants of two spaces.

It is far "heavier" to handle fundamental categories: The hom-sets between different pairs of elements are in general not isomorphic to each other (as sets). Moreover, they lack further algebraic structure. The objects of the fundamental category of a topological space X correspond to the elements of the underlying set X; only the morphism sets are "compressed" by the dihomotopy relations. It is in general a difficult task to use the van Kampen theorem for calculations of a fundamental category.

It was the aim of [9], to divide the underlying space into "components" in a systematic way, so that one only needs to investigate the (dihomotopy classes of) directed paths between components. The outcome of that paper was the *component category* of the fundamental category of a space; with certain generalizations to component categories of other categories.

The approach chosen in [9] has the drawback, that it involves *choices* of certain subcategories and thus does not yield *the* category of components. In [13], the authors overcome this dilemma by an additional pullback/pushout requirement (cf. Definition 4.3) giving rise to unique well-defined component categories. On the other hand, it is not always clear, how natural this extra requirement is. Moreover, if the directed space allows non-trivial *directed loops*, this method (and also those of the previous [9]) does not give satisfactory results.

1.2 Aims of This Paper

This paper contains several contributions that are linked to each other. First of all, we begin a systematic study of the topology of spaces of directed paths and not only of their path components (which are studied via the fundamental category). To get organized, we need to introduce a zoo of indexing categories, most prominently the so-called *preorder category* of a directed space, cf. Section 3.2: Its objects are *pairs* of elements, reflecting the fact that a directed space is mainly characterized by the directed paths between a given source *and* a given target (and the relations between different such path spaces). The indexing categories come along with functors to well-known categories such as the (homotopy) category of topological spaces, homology functors to abelian groups and so on. Morphisms in the indexing category are then regarded as (weakly) invertible if they are mapped into *isomorphisms* by the functor. As a result, one can apply the localization method leading to components suggested in [9] and modified in [13] *to the indexing category*.

Other possibilities of compressing the size of categories and obtaining minor (or even minimal) models of categories and retaining the essential "directed information" were pursued in more recent work of Grandis [17, 18] and of Krishnan [20]. It is certainly desirable to compare the different methods and their results in more detail.

In ordinary algebraic topology, topological spaces are considered to have the same shape if they are homotopy equivalent. It is not clear what the corresponding notion ought to be for directed spaces, and the obvious generalization has very weak properties with respect to path spaces: corresponding spaces of directed paths are in general not homotopy equivalent to each other. The "missing link" is formalized by the notion of an automorphic homotopy flow (Section 5) on a given directed space; such a homotopy flow produces a family of directed flow lines, but it need not give rise to self-homeomorphisms. These automorphic homotopy flows can usually only flow within limited regions; these regions can then be used to give an alternative way to obtain components and component categories.

Components and component categories are usually not at all preserved under directed continuous maps. We investigate under which conditions at least a directed homotopy equivalence gives rise to equivalent component categories. If a computation of the component categories of two spaces yields inequivalent results, then they cannot be directed homotopy equivalent (under an additional coherence condition).

1.3 Outline of the Paper

In Sections 2 and 3, we extend the toolbox associated to a directed space from a single tool (the fundamental category) to a kit consisting of several related categories and functors reflecting both the underlying path spaces and their (higher) homotopy and homology. In order to make concatenation associative, we use the notion of directed traces (directed paths up to directed reparametrization) investigated in [7]. Path spaces are *di*pointed (source and target); this is reflected in the (indexing) preorder category of a directed space, on which we can define the most important functors (organising the path spaces, their homology or homotopy groups etc.). In particular, we "reorganize" the information contained in the fundamental category through a functor from the preorder category to *Sets* as a special case.

Section 4 studies methods of inverting (a collection of) morphisms in the preorder category that induce isomorphisms under the functor under investigation. The preorder setting allows, e.g., to invert a directed path from x to y seen as a morphism between certain pairs, but not between others; this is quintessential for directed spaces with non-trivial directed loops; cf. Examples 4.6 and 7.4. Systematic inversion requires certain extension properties which allow to arrive at component *categories*.

In Section 5, we introduce automorphic homotopy flows on a given directed space and investigate their main properties. This is first applied in Section 6 to the definition of a dihomotopy equivalence between two given directed spaces; these dihomotopy equivalences are then shown to induce homotopy equivalences on related spaces of directed paths.

In Section 7, we describe an alternative way to arrive at components in an organized way by exploiting automorphic homotopy flows. These homotopy flows lead to invertibles in an extended preorder category that satisfy the extension properties right away. We show in Section 8, that the component categories from this approach are equivalent under dihomotopy equivalences satisfying a coherence condition. Finally, we sketch how these ideas might be generalized from functors on preorder categories to functors on more general categories.

2 Reparametrizations and Traces

2.1 Review on d-spaces

We start with a review of the basic notions, mainly taken from [16]: Let X denote a Hausdorff topological space, let $P(X) = C(I, X) = X^I$ denote the space of all paths in X, i.e., of all continous maps from the unit interval I into X equipped with the compact-open topology.

Definition 2.1 [16]

(1) A d-*space* is a topological space X together with a set $\vec{P}(X) \subseteq P(X)$ of continuous paths $I \to X$ such that
   - (a) $\vec{P}(X)$ contains all constant paths;
   - (b) $\vec{P}(X)$ is closed under concatenation;
   - (c) $p \circ \varphi \in \vec{P}(X)$ for any $p \in \vec{P}(X)$ and any continuous increasing (not necessarily surjective, not necessarily strictly increasing) map $\varphi : I \to I$;

(2) Elements of $\vec{P}(X)$ are called d-*paths*. For $x, y \in X$, $\vec{P}(X)(x, y)$ consists of all d-*paths* $p \in \vec{P}(X)$ with source x and target y ($p(0) = x$, $p(1) = y$).

(3) A continuous map $f : X \to Y$ between two d-spaces is called a d-map if $f(\vec{P}(X)) \subseteq \vec{P}(Y)$.

For the *oriented* unit interval $\vec{I}$, $\vec{P}(\vec{I})$ consists of the continuous increasing maps $\varphi : I \to I$. A d-path on X is then a d-map from $\vec{I}$ to X.

The product of two d-spaces is a d-space in a natural way. For homotopy purposes, we will in particular be interested in the products $X \times I$ and $X \times \vec{I}$ of a d-space X with the unoriented interval I ($\vec{P}(I) = P(I)$) and the oriented interval $\vec{I}$.

Definition 2.2

(1) A dihomotopy (between d-maps $f = H_0, g = H_1 : X \to Y$) is a d-map $H : X \times I \to Y$ (i.e., each map H_t is a $d - map$);
(2) A d-homotopy $f \xrightarrow{H} g$ (from $f = H_0$ to $g = H_1$) in X is a d-map $H : X \times \vec{I} \to Y$ (i.e., additionally all paths $H(x, t), x \in X$ are d-paths in Y);
(3) Dihomotopy is the equivalence relation defined by (1); d-homotopy is the equivalence relation generated by (2) (as transitive and symmetric closure).

Obviously, d-homotopic maps are dihomotopic; the opposite is in general not true. It is true, though, for paths in certain cubical complexes [8].

2.2 Reparametrization Equivalence

This section reviews the main definitions and results from [7] concerning continuous *reparametrizations* of d-paths. A *reparametrization* $\varphi : I \to I$ is a *continuous surjective increasing* self-map of the interval I; in particular: $s \leq t \Rightarrow \varphi(s) \leq \varphi(t)$ and $\varphi(0) = 0, \varphi(1) = 1$. A *regular reparametrization* $\varphi : I \to I$ is a reparametrization satisfying $s < t \Rightarrow \varphi(s) < \varphi(t)$; in other words, it is a self-homeomorphism of the interval respecting end-points.

Let $Rep_+(I)$ denote the *topological monoid* of all reparametrizations. A topology on $Rep_+(I)$ is induced from the compact-open topology on the space $C(I, I)$ (inherited from the supremum metric) of all self-maps of the interval. Let $Homeo_+(I) \subset Rep_+(I)$ denote the *topological group* of all regular reparametrizations.

The monoid $Rep_+(I)$ acts, for all $x, y \in X$, continuously on $\vec{P}(X)(x, y)$ by composition (on the right). Note that this action preserves d- and di-homotopy classes.

Definition 2.3 [7] Two d-paths $p, q \in \vec{P}(X)$ are called *reparametrization equivalent* if there exist reparametrizations φ, ψ such that $p \circ \varphi = q \circ \psi$.

Proposition 2.4 [7] *Reparametrization equivalence is an equivalence relation, in particular it is transitive.*

Taking quotients with respect to reparametrization equivalence yields a quotient map $q_P : \vec{P}(X)(x, y) \to \vec{T}_P(X)(x, y) := \vec{P}(X)(x, y)/_{\simeq}$; we endow this quotient space with the quotient topology.

Definition 2.5 [7]

(1) A d-path $\alpha : I \to X$ is called *regular* if it is either constant or if there is no non-trivial subinterval $J \subset I$ on which it is constant.
(2) Let $\vec{R}(X)(x, y) \subset \vec{P}(X)(x, y)$ denote the space of all regular d-paths from x to y with the induced topology.

Proposition 2.6 [7] *Let $x \neq y$ be elements of a d-space X. The action of $Homeo_+(I)$ on $\vec{R}(X)(x, y)$ is* free.

The group action of $Homeo_+(I)$ defines an (orbit) equivalence relation $\simeq$ and a quotient map $q_R : \vec{R}(X)(x, y) \to \vec{T}_R(X)(x, y) := \vec{P}(X)(x, y)/_{\simeq}$. We arrive at a commutative diagram of inclusions and quotient spaces

$$\begin{array}{ccc} \vec{P}(X)(x, y) & \hookrightarrow & \vec{R}(X)(x, y) \\ \downarrow{\scriptstyle q_P} & & \downarrow{\scriptstyle q_R} \\ \vec{T}_P(X)(x, y) & \xrightarrow{i} & \vec{T}_R(X)(x, y). \end{array}$$

Proposition 2.7 [7] *The map $i : \vec{T}_P(X)(x, y) \to \vec{T}_R(X)(x, y)$ is a* homeomorphism.

The proof in [7] proceeds in three steps: First, we show that every d-path is a reparametrization of a regular d-path. This yields surjectivity. Injectivity relies on a factorization property for reparametrizations. It is obvious that i is continuous. To see that it is also open relies on the fact that $Homeo_+(I)$ is dense in $Rep_+(I)$, cf. [7].

Lemma 2.8 [7] *For $x \neq y$, the quotient map $q_R : \vec{R}(X)(x, y) \to \vec{T}_R(X)(x, y)$ is a weak homotopy equivalence.*

Proof The free group action yields a fibration with contractible fiber $Homeo_+(I)$. □

Remark 2.9 It would be interesting to know whether (or under which conditions) the quotient map q_R from Lemma 2.8 is a genuine homotopy equivalence.

In conclusion, for calculations of homotopy or homology invariants, we can use any of the spaces $\vec{R}(X)(x, y)$, $\vec{T}_P(X)(x, y)$ or $\vec{T}_R(X)(x, y)$. In many cases, the conclusion of Lemma 2.8 holds also for $x = y$; cf. [7]. In the following, we will use the notation $\vec{T}(X)(x, y)$ for both $\vec{T}_P(X)(x, y)$ and $\vec{T}_R(X)(x, y)$.

3 The Trace Category and Its Relatives

In this section, we describe various indexing categories that can be used to organize the spaces of traces with given source and target. One may note a certain analogy to various categories (i.e., orbit categories) organising G-spaces in equivariant topology, cf. e.g. [23].

3.1 The Trace Category

The *trace category* $\vec{T}(X)$ of a d-space has the elements of X as objects; the morphisms from x to y are given by $\vec{T}(X)(x, y)$ – with the topology as a quotient space of $\vec{R}(X)(x, y)$. Composition on $\vec{T}(X)$ is inherited from concatenation on $\vec{R}(X)$. The latter is only associative up to reparametrization, just enough to make composition on $\vec{T}(X)$ associative!

A d-map $f : X \to Y$ between two d-spaces X and Y induces a functor $\vec{T}(f) : \vec{T}(X) \to \vec{T}(Y)$ by composition on the left on morphisms, i.e., $\vec{T}(f) : \vec{T}(X)(x, y) \to \vec{T}(Y)(fx, fy)$ is given by $\vec{T}(f)[\psi] = [f \circ \psi]$.

The *fundamental category* $\vec{\pi}_1(X)$ [9, 22] arises from the trace category $\vec{T}(X)$ as the category of path components, with the *di*homotopy relation. Concatenation on the trace spaces is homotopy invariant and factors over the fundamental category. In particular, left and right concatenation define maps

$$C_l : \vec{\pi}_1(X)(x, y) \to [\vec{T}(X)(y, z), \vec{T}(X)(x, z)] \text{ and} \tag{3.1}$$

$$C_r : \vec{\pi}_1(X)(y, z) \to [\vec{T}(X)(x, y), \vec{T}(X)(x, z)] \tag{3.2}$$

with $[-, -]$ denoting morphisms in the homotopy category $Ho - Top$.

Remark 3.1 The fundamental category in the sense of Grandis [16] is different, in general, since a d-homotopy between d-paths is not just a path in the space of d-paths; compare Remark 3.2.1.

An algebraic counterpart of the trace category is the homology category $\vec{H}(X)$ with elements as objects and with $\vec{H}_*(X)(x, y) = \bigoplus_{n\geq 0} H_n(\vec{T}(X)(x, y); R)$, the *total* homology with coefficients in a ring R. A composition law is then given by a generalization of the Pontryagin product for H-spaces, i.e.,

$$H_*(\vec{T}(X)(x, y)) \times H_*(\vec{T}(X)(y, z)) \xrightarrow{\times} H_*(\vec{T}(X)(x, y) \times \vec{T}(X)(y, z)) \xrightarrow{*} H_*(\vec{T}(X)(x, z)),$$

where the first map is given by the homological cross-product and the second is induced by concatenation.

Taking homology of the trace spaces corresponds to functors $\vec{H}_*(X) : \vec{T}(X) \to \vec{H}_*(X)$. A d-map $f : X \to Y$ between two d-spaces X and Y induces natural transformations $f_* : \vec{H}_*(X) \to \vec{H}_*(Y)$ – with group homomorphisms between the morphism groups.

A similar construction can be done for homotopy groups. Define homotopy categories $\vec{\Pi}_{n+1}(X)$ with elements of X as objects and with $\vec{\Pi}_{n+1}(X)(x, y) := \coprod_{\sigma\in\vec{T}(X)(x,y)} \vec{\pi}_n(\vec{T}(X)(x, y); \sigma)$ with composition law given by

$$\vec{\pi}_n(\vec{T}(X)(x, y); \sigma_1) \times \vec{\pi}_n(\vec{T}(X)(y, z); \sigma_2) \xrightarrow{\cong} \vec{\pi}_n(\vec{T}(X)(x, y) \times \vec{T}(X)(y, z); (\sigma_1, \sigma_2))$$
$$\xrightarrow{*} \vec{\pi}_n(\vec{T}(X)(x, z); \sigma_1 * \sigma_2);$$

the second map is again induced by concatenation. Note that these homotopy groups – up to (non-canonical) isomorphism – only depend on the connected component of the base points σ_j, i.e., on the classes they represent in $\vec{\pi}_1(X)(-, -)$.

3.2 Preorder Categories and Homology

More useful indexing devices are several variants of *preorder categories* $\vec{D}(X)$ of a d-space X. The categories to be discussed differ in their morphism sets, but they have all the same objects: A d-space X comes equipped with a natural preorder $x \preceq y \Leftrightarrow \vec{P}(X)(x, y) \neq \emptyset$. For all the preorder categories below, the objects are the pairs $(x, y) \in X \times X$ with $x \preceq y$.

The morphisms in $\vec{D}(X)$ are $\vec{D}(X)((x, y), (x', y')) := \vec{T}(X)(x', x) \times \vec{T}(X)(y, y')$ with composition given by pairwise contra-, resp. covariant concatenation. Hence, $\vec{D}(X)$ is a full subcategory of the category $\vec{T}(X)^{op} \times \vec{T}(X)$. Note that every morphism $(\sigma_x, \sigma_y) \in \vec{T}(X)(x', x) \times \vec{T}(X)(y, y')$ decomposes as follows: $(\sigma_x, \sigma_y) = (c_{x'}, \sigma_y) \circ (\sigma_x, c_y) = (\sigma_x, c_{y'}) \circ (c_x, \sigma_y)$ with $c_u \in \vec{T}(X)(u, u)$ the constant trace at $u \in X$.

Trace spaces can be organised by the trace space functor $\vec{T}^X : \vec{D}(X) \to Top$ given by $\vec{T}^X(x, y) = \vec{T}(X)(x, y)$ and $\vec{T}^X(\sigma_x, \sigma_y)(\sigma) := \sigma_x \circ \sigma \circ \sigma_y \in \vec{T}(X)(x', y')$ for $\sigma \in \vec{T}(X)(x, y)$. This functor can be viewed as (a restriction of) the Top-enriched hom-functor of $\vec{T}(X)$.

A d-map $f : X \to Y$ induces a functor $\vec{D}(f) : \vec{D}(X) \to \vec{D}(Y)$ with $\vec{D}(f)(x, y) = (fx, fy)$ and $\vec{D}(f)(\sigma_x, \sigma_y) = (\vec{T}(f)(\sigma_x), \vec{T}(f)(\sigma_y)) = (f \circ \sigma_x, f \circ \sigma_y)$; moreover, it induces a natural transformation $\vec{T}(f)$ from $\vec{T}^X$ to $\vec{T}^Y$.

A homotopical variant is given by the category $\vec{D}_\pi(X)$ with the same objects as above and with $\vec{D}_\pi(X)((x, y), (x', y')) := \vec{\pi}_1(X)(x', x) \times \vec{\pi}_1(X)(y, y')$. Hence, this category is a full subcategory of the category $\vec{\pi}_1(X)^{op} \times \vec{\pi}_1(X)$ where $\vec{\pi}_1(X)$ denotes the fundamental category (cf. Section 3). It comes with a functor $\vec{T}_\pi^X : \vec{D}_\pi(X) \to Ho - Top$ into the homotopy category; a d-map $f : X \to Y$ induces a natural transformation $\vec{T}_\pi(f)$ from $\vec{T}_\pi^X$ to $\vec{T}_\pi^Y$. Together with the (vertical) forgetful functors, we obtain a commutative diagram

$$\begin{array}{ccc} \vec{D}(X) & \xrightarrow{\vec{T}^X} & Top \\ \downarrow & & \downarrow \\ \vec{D}_\pi(X) & \xrightarrow{\vec{T}_\pi^X} & Ho - Top. \end{array} \tag{3.3}$$

The functors $\vec{T}_X$ and $\vec{T}_\pi^X$ may be composed with homology functors into categories of (graded) abelian groups, R-modules or graded rings. In particular, we obtain, for $n \geq 0$, functors $\vec{H}_{n+1}(X) : \vec{D}(X) \to Ab$ with $(x, y) \mapsto H_n(\vec{T}(X)(x, y))$ and $(\sigma_x, \sigma_y)_*$ given by the map induced on n-th homology groups by concatenation with those two traces on trace space $\vec{T}(X)(x, y)$. This functor factors obviously over $\vec{D}_\pi(X)$. In the same spirit, one can define homology with coefficients and cohomology; a d-map $f : X \to Y$ induces a natural transformation $\vec{H}_{n+1}(f) : \vec{H}_{n+1}(X) \to \vec{H}_{n+1}(Y), n \geq 0$.

Remark 3.2

(1) Composing with the functor $\pi_0 : Top \to Sets$ that associates to a topological space its set of path components, yields a functor $\vec{\Pi}_1 : \vec{D}(X) \to Sets$ with $\vec{\Pi}_1(X)(x, y) = \vec{\pi}_1(X)(x, y)$, the set of morphisms in the fundamental category; this functor factors over $\vec{D}_\pi(X)$. Also the information in the fundamental category in the sense of Grandis can be recovered by a functor from $\vec{D}(X)$ to *Sets*; that functor does in general not factor over $\vec{T}^X$; it involves "directed" information.

(2) It would be desirable to have better theoretical and computational tools concerning the homology of trace (or path) spaces. First steps in this direction have been taken for spaces arising from cubical sets in an unpublished master's thesis from Aarhus University.

3.2.1 Preorder Endomorphism Categories

The effect of self-d-maps $f : X \to X$ (from a d-space X to itself) on the trace spaces is reflected by the preorder endomorphism category $\vec{D}E(X)$: It has the same objects as the category $\vec{D}(X)$, whereas $\vec{D}E(X)((x, y)(x', y')) := \{f : X \to X | f \text{ a dimap with } f(x) = x', f(y) = y'\}$ and composition is composition of d-maps. This category organises trace spaces through the functor $\vec{T}E^X : \vec{D}E(X) \to Top$ with $\vec{T}E^X(x, y) := \vec{T}(X)(x, y)$ and $\vec{T}E^X(f) := \vec{T}(f) : \vec{T}(X)(x, y) \to \vec{T}(X)(x', y')$ for $f \in \vec{D}E(X)((x, y)(x', y'))$.

The quotient functor into $Ho - Top$ factors over a category, where the morphisms are homotopy classes of d-maps with homotopies fixing "endpoints." Also these functors can be composed with homology functors. Note that these preorder endomorphism categories are *not* functorial with respect to d-maps between different d-spaces. We do not go into details, since it seems to be difficult to obtain localizations (cf. Section 7) with good properties within these categories.

It is sometimes more significant to restrict the preorder endomorphism category to one with the same objects, but with fewer endomorphisms: $\vec{D}E_0(X)((x, y)(x', y')) := \{f : X \to X | f \text{ a dimap } d-\text{homotopic to } id_X \text{ with } f(x) = x', f(y) = y'\}$.

3.3 Factorization Categories and Higher Homotopy

For indexing purposes, that are finer than those in Section 3.2, it is convenient to consider the *factorization category* $F\vec{T}(X)$[1] of the trace category $\vec{T}(X)$: The objects of $F\vec{T}(X)$ are just the morphisms of the trace category $\vec{T}(X)$. Moreover, we define

$$F\vec{T}(X)(\sigma_{xy}, \sigma'_{x'y'}) := \{(\varphi_{x'x}, \varphi_{yy'}) \in \vec{T}(X)(x', x) \times \vec{T}(X)(y, y') \,|\, \sigma'_{x'y'} = \varphi_{yy'} \circ \sigma_{xy} \circ \varphi_{x'x}\};$$

composition is defined as in Section 3.2 above, by restriction.

This category comes with a functor $F\vec{T}^X : F\vec{T}(X) \to Top_*$ into the category of *pointed* topological spaces. It associates to σ_{xy} the pointed topological space $(\vec{T}(X)(x, y); \sigma_{xy})$. As in Section 3.2, one may consider homotopical variants, a category $F_\pi\vec{T}(X)$ with pairs of *dihomotopy classes* of commuting traces as morphisms and a functor $F_\pi\vec{T}^X : F_\pi\vec{T}(X) \to Ho - Top_*$.

The main interest in these categories and functors arises after composition with homotopy functors from either Top_* or from $Ho - Top_*$ and to arrive at homotopy functors $\vec{\pi}_2(X) : F\vec{T}(X) \to Groups$ and $\vec{\pi}_{n+1}(X)(\sigma_{xy}) : F\vec{T}(X) \to Ab,\ n > 1$, given by $\vec{\pi}_{n+1}(X) := \pi_n(\vec{T}(X)(x, y); \sigma_{xy})$ and the obvious induced maps. These homotopy functors factor over $F_\pi\vec{T}(X)$. The constructions are again functorial with respect to d-maps.

A dihomotopy between dipaths σ_1 and σ_2 corresponds to a change of base point in the same component of the trace space. In particular, the associated homotopy groups $\vec{\pi}_{n+1}(X)(\sigma_1)$ and $\vec{\pi}_{n+1}(X)(\sigma_2)$ are isomorphic; the isomorphism depends on the homotopy class of the dihomotopy, in general.

3.3.1 Endomorphism Trace Categories

Looking at the effect of d-self maps leads to the consideration of yet another category related to a d-space X. The endomorphism trace category $E\vec{T}(X)$ has the same objects as $F\vec{T}(X)$, whereas $E\vec{T}(X)(\sigma_{xy}, \sigma'_{x'y'}) := \{f : X \to X|\ f \text{ a d-map with}$

$f(\sigma_{xy}) = \sigma'_{x'y'}\}$; equality means of course reparametrization equivalence of representatives. Endomorphism trace categories are not functorial with respect to d-maps between different d-spaces; they do not seem to enjoy good enough properties for localization purposes (cf. Section 7).

The significance of endomorphism trace categories arises from the functor $E\vec{T}^X : E\vec{T}(X) \to Top_*$, which to a d–map f seen as morphism from σ_{xy} to $\sigma'_{x'y'}$ associates the pointed map $\vec{T}(f) : (\vec{T}(X)(x, y); \sigma_{xy}) \to (\vec{T}(X)(x', y'); \sigma'_{x'y'})$. Also this functor can be composed with homotopy functors.

As before, one may restrict attention to the category $E_0\vec{T}(X)$ with $E_0\vec{T}(X)(\sigma_{xy}, \sigma'_{x'y'}) := \{f : X \to X \mid f$ a dimap $d-$homotopic to id_X with $f(\sigma_{xy}) = \sigma'_{x'y'}\}$.

It is also possible to view these endomorphism trace categories as 2-categories with elements as objects, traces as 1-morphisms and d-self maps as 2-morphisms; it is easy to check that the interchange law holds.

4 Weakly Invertible Systems and Component Categories

4.1 Motivation. An Example

The article [9] describes a method of "compressing" information in a small category by quotienting out a subcategory of so-called weakly invertible morphisms satisfying certain properties. This method has since been refined in [13] and related to a quotient approach based on general congruences on small categories that was described earlier in [2].

Definition 4.1 [13] A morphism $\sigma \in \mathcal{C}(x, y)$ in a small category $\mathcal{C}$ is called *Yoneda invertible* if, for every object z of $\mathcal{C}$ with $\mathcal{C}(y, z) \neq \emptyset$, resp. $\mathcal{C}(z, x) \neq \emptyset$ the maps $\mathcal{C}(y, z) \xrightarrow{\sigma\circ} \mathcal{C}(x, z)$ and $\mathcal{C}(z, x) \xrightarrow{\circ\sigma} \mathcal{C}(z, y)$ are bijections.

Yoneda invertible morphisms are at the base of the quotienting process described and applied to the fundamental category of a d-space in [9, 13]. While these seem to be very adequate for categories where all isomorphisms are identities (as in d-spaces arising from a partial order), the presence of loops causes serious problems.

Example 4.2 The simplest example is that of the oriented circle $\vec{S}^1$ with $\vec{P}(\vec{S}^1)$ consisting of the counterclockwise paths (arising from non-decreasing paths on the reals under the exponential map). In this case, for every pair of angles $\alpha, \beta \in S^1$, the *trace* space (after dividing out reparametrizations) is the discrete space represented by rotations by angles $\beta - \alpha + 2k\pi, k \in \mathbf{Z}, k \geq 0$. The trace category and the fundamental category (cf. Section 3.1) agree: $\vec{\pi}_1(\vec{S}^1)(\alpha, \beta) = \vec{T}(\vec{S}^1)(\alpha, \beta) = \{\beta - \alpha + 2k\pi, k \in \mathbf{Z}, k \geq 0\}$. Concatenation with a rotation by an angle $\rho, 0 < \rho < 2\pi$, corresponds to addition of angles and yields thus maps

$$\begin{array}{cccccccc} \vec{T}(\vec{S}^1)(\alpha,\beta) & \xrightarrow{+\rho} & \vec{T}(\vec{S}^1)(\alpha, \beta+\rho \mod 2\pi) & \vec{T}(\vec{S}^1)(\alpha,\beta) & \xrightarrow{\rho+} & \vec{T}(\vec{S}^1)(\alpha-\rho \mod 2\pi, \beta) \\ \beta-\alpha+2k\pi & \mapsto & \beta-\alpha+\rho+2k\pi & \beta-\alpha+2k\pi & \mapsto & \beta-\alpha+\rho+2k\pi. \end{array}$$

Note that the addition of ρ is $\mod 2\pi$ on the objects and "on the nose" on the morphisms. In particular, for $\beta - \alpha + \rho > 2\pi$, neither of the maps $+\rho$ or $\rho+$ is surjective. Given any $\beta \in S^1$ and $\rho \neq 0$, the morphism $+\rho$ is certainly *not* a bijection for $\alpha = \beta + \frac{\rho}{2}$ and the morphism $\rho+$ is *not* a bijection for $\beta = \alpha - \frac{\rho}{2}$. In conclusion, the only Yoneda invertible morphisms in the fundamental category of $\vec{S}^1$ are the identities; nothing is gained by going over to a quotient category.

As a cure, we will consider the homotopy preorder category $\vec{D}_\pi(\vec{S}^1)$, cf. Section 3.2 of the trace-fundamental category. In this category, the morphism (using analogous notation) $\vec{D}_\pi(\vec{S}^1)(\alpha,\beta) \stackrel{\rho+,+\sigma}{\longrightarrow} \vec{D}_\pi(\vec{S}^1)(\alpha-\rho,\beta+\sigma)$ will be considered as weakly invertible if and only if $\beta - \alpha + \rho + \sigma < 2\pi$. Note that these weakly invertible morphisms form a closed wide subcategory in the homotopy preorder category of $\vec{S}^1$. To be continued in Example 4.6.

4.2 Weakly Invertible Morphisms with Respect to a Functor

We now present a very general method of pointing out subcategories of weakly invertible morphisms: Consider a functor $F : \mathcal{C} \to \mathcal{D}$ between two small categories. A morphism $\sigma \in \mathcal{C}(x,y)$ will be called *F-invertible* if and only if $T(\sigma) \in \mathcal{D}(Fx, Fy)$ is an *isomorphism* in $\mathcal{D}$. Let $\mathcal{C}_F(x,y) \subseteq \mathcal{C}(x,y)$ denote the set of *all* F-invertible morphisms from x to y. The collection of all $\mathcal{C}_F(x,y)$ form a wide subcategory $\mathcal{C}_F$ of $\mathcal{C}$ since the composition of two F-invertible morphisms obviously is F-invertible again; remark that $\mathcal{C}_F(x,y)$ contains the $\mathcal{C}$-isomorphisms.

For example, consider the functor $\vec{T}^X : \vec{D}_\pi(X) \to Ho - Top$ or the functors $\vec{H}_{n+1}(X) : \mathcal{D}_\pi(X) \to Ab$ from Section 3.2. A morphism $(\sigma_x, \sigma_y) \in \vec{D}_\pi(X)((x,y), (x',y'))$ is $\vec{T}^X$-invertible if and only if $\vec{T}(X)(\sigma_x,\sigma_y) : \vec{T}(X)(x,y) \to \vec{T}(X)(x',y')$ is a homotopy equivalence; it is $\vec{H}_{n+1}$-invertible if $(\sigma_x,\sigma_y)_* : H_n(\vec{T}(X)(x,y)) \to H_n(\vec{T}(X)(x',y'))$ is an isomorphism.

In Example 4.2, the weakly invertible morphisms on $\vec{S}^1$ correspond exactly to the $T^{\vec{S}^1}$-invertible morphisms, which again are the same as the $\vec{H}_1(\vec{S}^1)$-invertible morphisms.

4.3 Component Categories

Having defined the wide category $\mathcal{C}_F$ of $\mathcal{C}$, we can proceed along the lines of [9] or of [13] to arrive at a quotient category identifying objects and morphisms that are linked to each other by $\mathcal{C}_F$-morphisms. In order to get a consistent construction, it is usually necessary to restrict the morphisms furthermore to a wide subcategory $\Sigma \subseteq \mathcal{C}_F \subseteq \mathcal{C}$:

Definition 4.3 [3, 9, 13] Let $\Sigma \subseteq \mathcal{C}$ denote a wide subcategory of a small category $\mathcal{C}$. The pair $(\mathcal{C}, \Sigma)$

LEP/REP Satisfies the *left/right extension property* LEP/REP if and only if the diagrams

$$\begin{array}{ccc} x' & \stackrel{\tau'}{\dashrightarrow} & y' \\ {\scriptstyle\sigma}\uparrow & & \uparrow{\scriptstyle\sigma'} \\ x & \stackrel{\tau}{\longrightarrow} & y \end{array} \qquad\qquad \begin{array}{ccc} x & \stackrel{\tau}{\longrightarrow} & y \\ {\scriptstyle\sigma'}\uparrow & & \uparrow{\scriptstyle\sigma} \\ x' & \stackrel{\tau'}{\dashrightarrow} & y' \end{array} \tag{4.1}$$

can be filled in with $\tau' \in \mathcal{C}, \sigma' \in \Sigma$ given any $\tau \in \mathcal{C}, \sigma \in \Sigma$.

pure Is *pure* if and only if $\sigma \circ \tau \in \Sigma \Leftrightarrow \sigma, \tau \in \Sigma$; (and left pure, resp. right pure, if one can only conclude $\sigma \in \Sigma$, resp. $\tau \in \Sigma$).

SLEP/SREP Has the *strong left/right extension property* SLEP/SREP if the diagrams (4.1) can be filled in to yield pushout, resp. pullback squares in $\mathcal{C}$; c.f. [13] for details.

Proposition 4.4 [13] *Let $\mathcal{B}$ denote a wide subcategory of $\mathcal{C}$ and suppose that the pair $(\mathcal{C}, Iso(\mathcal{C}))$ is pure.*

(1) *If $(\mathcal{C}, B)$ satisfies SLEP and SREP, then $(\mathcal{C}, B)$ is pure.*
(2) *The family of all wide subcategories $Iso(\mathcal{C}) \subseteq \mathcal{D} \subseteq \mathcal{B} \subseteq \mathcal{C}$ such that $(\mathcal{C}, \mathcal{D})$ satisfies SLEP and SREP is a* complete lattice*; in particular, there is a wide subcategory $\Sigma_{\mathcal{B}} \subseteq \mathcal{B}$ such that $(\mathcal{C}, \Sigma_{\mathcal{B}})$ satisfies SLEP and SREP and such that $\mathcal{D} \subseteq \Sigma_{\mathcal{B}}$ for all $\mathcal{D}$ above.*

There are now two possible points of departure from which to obtain a component category from a wide subcategory $Iso(\mathcal{C}) \subseteq \Sigma \subseteq \mathcal{C}_F \subseteq \mathcal{C}$:

(1) Σ is *a* subcategory such that $(\mathcal{C}, \Sigma)$ satisfies LEP/REP and purity; no maximality ensured.
(2) $\Sigma = \Sigma_{\mathcal{C}_F}$, the maximal subcategory satisfying SLEP and SREP.

For the convenience of the reader, we include a brief description of the construction of the component category of $\mathcal{C}$ with respect to Σ from [9] starting with the category of fractions $\mathcal{C}[\Sigma^{-1}]$ [3, 9]. The exposition can be simplified since the extension properties from Definition 4.3 imply that the subcategory Σ admits a left, resp. right calculus of fractions [3, 9] on $\mathcal{C}$. A morphism in this category (with the same objects as those of $\mathcal{C}$) is a ("*zig-zag*")-morphisms of the form $\sigma^{-1} \circ \tau$, resp. $\tau \circ \sigma^{-1}$, $\sigma \in \Sigma, \tau \in \mathcal{C}$; the extension properties make sure that the composite of two morphisms can be written in this "standard form" again. Morphisms of the form $\sigma_1^{-1} \circ \sigma_2$, resp. $\sigma_1 \circ \sigma_2^{-1}$, $\sigma_i \in \Sigma$ are the Σ-*zig-zag morphisms* [9]. The functor F maps $\Sigma \subseteq \mathcal{C}_F$ into $Iso(\mathcal{D})$ and can therefore be extended to a functor $F[\Sigma^{-1}] : \mathcal{C}[\Sigma^{-1}] \to \mathcal{D}, \sigma^{-1} \circ \tau \mapsto (F(\sigma))^{-1} \circ F(\tau)$; likewise for $F[\Sigma^{-1}]$.

Two objects x, y of $\mathcal{C}$ are called Σ-*equivalent* ($x \simeq_\Sigma y$) if there exists a Σ-zig-zag-morphism between them. The equivalence classes are called the Σ-*components* of $\mathcal{C}$; they are the path components *with respect to the Σ-zig-zag morphisms*. Moreover, we generate an equivalence relation on the morphisms of $\mathcal{C}[\Sigma^{-1}]$ by requiring that $\tau \simeq \tau \circ \sigma, \tau \simeq \sigma \circ \tau$ whenever $\sigma \in \Sigma$ and the composition is defined.

The *component category* $\pi_0(\mathcal{C}; \Sigma)$ of the preorder category has the Σ-components as objects; the morphisms from $[x]$ to $[y]$ are the equivalence classes of morphisms in $\bigcup_{x' \simeq x, y' \simeq y} \mathcal{C}[\Sigma^{-1}]$. Two morphisms in $\pi_0(\mathcal{C}; \Sigma)$ represented by $\tau_i \in \mathcal{C}(x_i, y_i), 1 \leq i \leq 2$, with $y_1 \simeq_\Sigma x_2$ can be composed by inserting any Σ-zig-zag-morphism connecting y_1 and x_2, cf. [9] for details.

Taking equivalence classes results in a functor $q_\Sigma : \mathcal{C} \to \pi_0(\mathcal{C}; \Sigma)$. By construction, the functor $F : \mathcal{C} \to \mathcal{D}$ factors over q_Σ and the component category to yield a functor $F : \pi_0(\mathcal{C}; \Sigma) \to \mathcal{D}$.

The extension and purity properties have the following consequences:

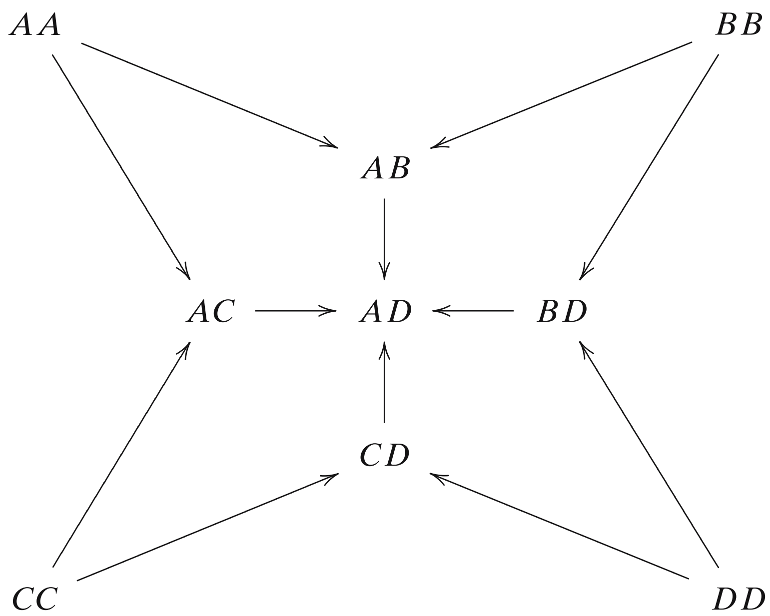
Proposition 4.5

(1) ([9], *Proposition* 5): (Σ, Σ) *satisfies LEP/REP.*
(2) ([9], *Proposition* 3): *Given a component* $C \subseteq ob(\mathcal{C})$ *and elements* $x, y \in C$. *Every morphism* $\tau' \in \mathcal{C}(x', y')$ *with* $x' \in C$ *(resp.* $y' \in C$*) is* Σ*-equivalent to a morphism* $\tau \in \mathcal{C}(x, -)$ (resp. $\tau \in \mathcal{C}(-, y)$).
(3) *Every isomorphism in* $\pi_0(\mathcal{C}; \Sigma)$ *is an endomorphism.*
(4) *If* $\tau_1 \circ \tau_2 \in \pi_0(\mathcal{C}; \Sigma)(C, C)$ *is an isomorphism, then the* $\tau_i, 1 \le i \le 2$. *are isomorphisms.*

The last property is particularly important: it makes it impossible to leave and then reenter a component.

Example 4.6

(1) Let X denote the subspace of $\vec{I}^2$ obtained by removing an (open) isothetic subsquare $\vec{J}^2$, cf. Fig. 2. This space is divided into four components A (below the hole), B, C, D (above the hole). It is easy to see that the trace spaces $\vec{T}(X)(x, y), x \preceq y$, are homotopy equivalent to a discrete space consisting of *two* elements if and only if $x \in A$ and $y \in D$ and of *one* element else. Concatenation with a d-path is a homotopy equivalence unless the d-path crosses one of the boarders of A, resp. D (the stipled lines in Fig. 2).
It is not difficult to see that the subcategory $\Sigma(\vec{X}) \subseteq \vec{D}_\pi(X)$ consisting of pairs (σ_x, σ_y) with source and target in the same of the blocks A, B, C, D maps into homotopy equivalences under T^X and it satisfies SLEP/SREP. As a result, the component category $\pi_0(\vec{D}_\pi(X); \vec{\Sigma}(X))$ can be depicted as follows (e.g., AD is represented by a trace from an element in A to an element in D):

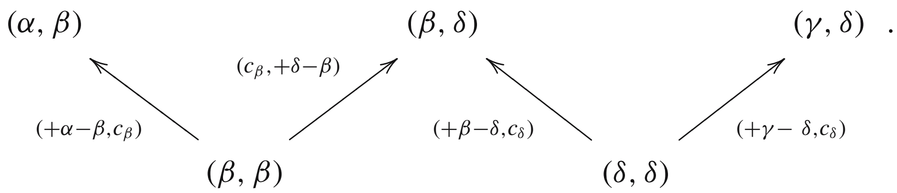


The quadrilaterals from BB and from CC to AD commute, the other two quadrilaterals do not: there are two essentially different ways to extend a d-path starting and ending in A to a d-path starting in A and ending in D.

(2) Let $\vec{S}^1$ denote the oriented circle from Example 4.2. Let $\Sigma(\vec{S}^1) \subseteq \vec{D}_\pi(\vec{S}^1)$ denote the wide subcategory consisting of all morphisms

$$\vec{D}_\pi \vec{S}^1(\alpha, \beta) \overset{\rho+,+\sigma}{\longrightarrow} \vec{D}_\pi \vec{S}^1(\alpha - \rho, \beta + \sigma) \text{ with } \beta - \alpha + \rho + \sigma < 2\pi;$$

these are exactly the morphisms that are mapped to homotopy equivalences under $\vec{T}^{\vec{S}^1}$. Also in this case, it is easy to see that the pair of categories satisfies SLEP/SREP – by extending to the maximum of the targets/minimum of the sources.
The following diagram yields a zig-zag of $\vec{\Sigma}(\vec{S}^1)$-paths between arbitrary pairs of angles (α, β), (γ, δ):



As a result, the component category $\pi_0(\vec{D}_\pi(\vec{S}^1); \vec{\Sigma}(\vec{S}^1))$ has a single object and a morphism set in bijective correspondence with the non-negative integers generated by a single loop $+2\pi$. This underpins our assertion that it is commendable to use the preorder category as indexing category rather than the fundamental category; compare Example 4.2.

4.4 Component Categories with Respect to Homotopy Equivalences, Homology, Homotopy

We will now apply the general construction of a component category from Section 4.3 to the case where $\mathcal{C}$ is the homotopy preorder category or the homotopy factorization category of a d-space X and F is one of the functors T^X, resp. the homology and homotopy functors considered in Section 3). In both the preorder and the factorization category, only the identities are isomorphisms. Moreover, the factorization category is loopfree (This property is essential for the later parts of [13].)

In these cases, C_F will be the subcategory consisting of concatenation morphisms that induce homotopy equivalences, resp. induce isomorphisms on certain homology or homotopy groups.

We will indicate later in Section 5 a topologically natural way to ensure LEP/REP. The delicate point is then to choose a suitable subcategory $\Sigma \subseteq C_F$ satisfying purity, in addition. It is not so clear to me how natural it is to require SLEP/SREP (with purity as a consequence), and case studies (in particular in spaces/categories with non-trivial loops) should be performed.

Having chosen a suitable such subcategory $\Sigma_F \subseteq \mathcal{C}_F$, we can construct the component categories $\pi_0(\vec{D}_\pi(X), \Sigma_F)$. This leads to a consistent way of identifying pairs (x, y), (x', y') in the preorder category with $\vec{T}(X)(x, y)$, $\vec{T}(X)(x', y')$ being homotopy equivalent, homology equivalent in certain dimensions etc. We obtain factorizations of the original functors

$$\vec{T}_{\pi}^{X} : \vec{D}_{\pi}(X) \stackrel{q_{\Sigma_T}}{\longrightarrow} \pi_0(\vec{D}_{\pi}(X), \Sigma_T) \stackrel{\pi_0(\vec{T}_{\pi}^{X})}{\longrightarrow} Ho - Top,$$

$$\vec{H}_{n+1}(X) : \vec{D}_{\pi}(X) \stackrel{q_{\Sigma_H}}{\longrightarrow} \pi_0(\vec{D}_{\pi}(X), \Sigma_H) \stackrel{\pi_0(\vec{H}_{n+1}(X))}{\longrightarrow} Ab,$$

$$\vec{\pi}_{n+1}(X) : F\vec{T}_{\pi}(X) \stackrel{q_{\Sigma_\pi}}{\longrightarrow} \pi_0(F\vec{T}_{\pi}(X), \Sigma_\pi) \stackrel{\pi_0(\vec{\pi}_{n+1}(X))}{\longrightarrow} Sets/Grps/Ab.$$

Remark 4.7 A similar and somewhat simpler situation has been investigated under the heading "persistent homology" by G. Carlsson and collaborators, in particular for applications in the analysis of statistical point cloud data, c.f. e.g. [6, 25]. The indexing category corresponds to the preorder category of ordinals $1 \to 2 \to 3 \to \cdots$. A (homology) functor from this category into k-modules, k a field, is a $k[t]$-module and splits as such into irreducible modules of the form $\Sigma^i k[t]$ or $\Sigma^i k[t]/(t^j)$ of classes that are "born at step" i and possibly annihilated after j steps. These are denoted by "barcodes" with bars from i to ∞ or to $i + j$. In our case, we can have homology classes that are born at (x, y) (in the cokernel of maps into $\vec{H}_*(X)(x, y)$) that survive to (x', y') and are annihilated at (x'', y'') (in the kernel of the map into $\vec{H}_*(X)(x'', y'')$). These births and deaths are obstructions to invertibility with respect to the functors $\vec{H}_*$.

5 Automorphic Homotopy Flows

This section prepares a new definition for directed homotopy equivalences and an investigation of their properties. Which requirements should a d-map $f : X \to Y$ satisfy in order to qualify as a directed homotopy equivalence? Obviously, there should be a reverse d-map $g : Y \to X$ such that both $g \circ f$ and $f \circ g$ are d-homotopic to the resp. identity maps. But this is not enough: The (d-path) structures on X and Y ought to be homotopically related, i.e., the maps $\vec{T}(f) : \vec{T}(X)(x, y) \to \vec{T}(Y; fx, fy)$ should be ordinary homotopy equivalences – for all x, y with $\vec{T}(X)(x, y) \neq \emptyset$ – and that in a natural way.

Remark 5.1

(1) For the lpo-spaces of [10] one might instead ask, that the "intervals" $[x, y]$, resp. $[fx, fy]$ containing all elements between x and y, resp. between fx and fy are homotopy equivalent.
(2) Compare with the future, resp. past homotopy equivalence from [18]; note also the coherence requirements there.

This is in general not the case: A first indication for this is that a self-d-map $h : X \to X$ that is d-or di-homotopic to the identity map $id_X : X \to X$ does *not* always yield homotopy equivalences $\vec{T}(X)(x, y) \to \vec{T}(X)(hx, hy)$, cf. Example 5.10 below.

5.1 Homotopy Flows

To get an idea of what a dihomotopy equivalence should satisfy – and also for a suggestion on subcategories related to components, one needs to understand d-homotopies of the identity id_X of a d-space X.

Definition 5.2

(1) A d-map $H : X \times \vec{I} \to X$ is called a *future homotopy flow* if $H_0 = id_X$ and a *past homotopy flow* if $H_1 = id_X$.
(2) The sets consisting of all future homotopy flows, resp. of all past homotopy flows will be denoted by $\vec{P}_+C_0(X, X)$, resp. by $\vec{P}_-C_0(X, X)$.

Homotopy flows (and the later refinements) generalise the concept of a *flow* on a differentiable manifold. We do not require that the maps $H(-, t) : X \to X$ are homeomorphisms. The orbits of a flow have the following counterpart: For every $x \in X$, the map $H_x : \vec{I} \to X,\ t \mapsto H(x, t)$ is a d-path (with $H_x(0) = x$, resp. $H_x(1) = x$). Evaluation at $x \in X$ defines maps

$$ev_+^x : \vec{P}_+C_0(X, X) \to \vec{T}(X)(x, -),\ \text{resp. } ev_-^x : \vec{P}_-C_0(X, X) \to \vec{T}(X)(-, x). \quad (5.1)$$

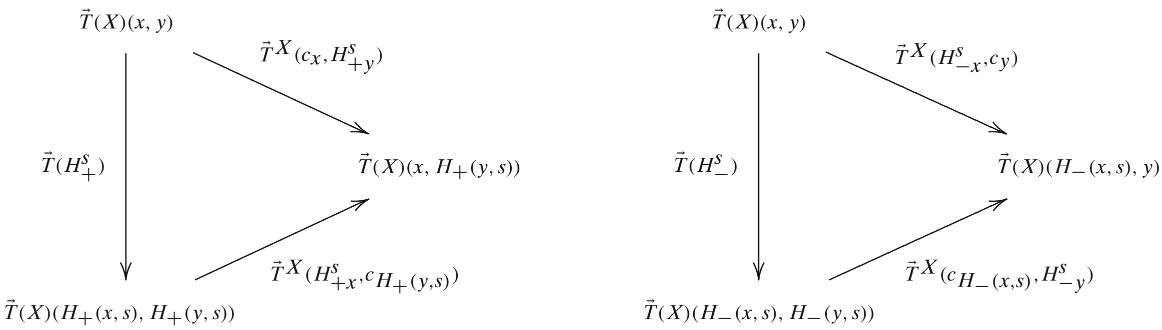
A maximal element $x_+ \in X$ – with the constant path as the only d-path with source x_+ – will be fixed under a future homotopy flow, likewise a minimal element under a past homotopy flow.

Note that future homotopy flows can be pieced together in various natural ways:

$$(H_1 * H_2)(x, t) = \begin{cases} H_1(x, 2t), & t \leq \frac{1}{2} \\ H_2(H_1(x, 1), 2t - 1), & t \geq \frac{1}{2}, \end{cases} \quad \text{resp. } (H_1 \Box H_2)(x, t) = H_2(H_1(x, t), t). \quad (5.2)$$

Similarly for past homotopy flows. In particular, if $f, g : X \to X$ are future, resp. past d-homotopic to id_X, then their compositions $f \circ g : X \to X$ and $g \circ f : X \to X$ are so, as well; they form thus sub-monoids of the monoid $\vec{C}(X, X)$ of all self-d-maps of X. Likewise for *traces* of past/future homotopy flows (quotienting out "global" reparametrizations).

Homotopy flows induce several interesting maps on trace spaces: Let $H_+, H_- : X \times \vec{I} \to X$ denote d-homotopies $id_X \to f$, resp. $g \to id_X$. These d-homotopies define, for every $x \in X$ the d-paths $H_{\pm x}$ from x to fx, resp. from gx to x. For reasons to be explained in Remark 5.7 below, we need moreover to consider restrictions of the homotopies and their effect on trace spaces: for every $s \in I$, there is a (restricted) d-homotopy $H_\pm^s : X \times \vec{I} \to X$, $H_\pm^s(x, t) = H_\pm(x, st)$ with associated d-paths $H_{\pm x}^s$ from x to $H_+(x, s)$, resp. from $H_-(x, s)$ to x. As explained in Section 3, the maps $H^s_{\pm}$, H^s_{+y}, H^s_{-x} induce maps on trace spaces, that are linked to each other by the following homotopy commutative diagrams:



(5.3)

(The maps $\vec{T}^X(-, -)$ are defined in Section 3.2 and arise from concatenation). Commutativity in the diagram is a consequence of a more general

Lemma 5.3 *A d-homotopy $H : X \times \vec{I} \to Y$ induces, for every object (x, y) in $\vec{D}(X)$, a d-homotopy $\bar{H}(x, y) : \vec{T}(X)(x, y) \times I \to \vec{T}(Y)(H(x, 0), H(y, 1))$ between $\bar{H}_0(x, y)(\sigma) = H(\sigma, 0) * H(y, t)$ and $\bar{H}_1(x, y)(\sigma) = H(x, t) * H(\sigma, 1)$.*

A similar construction has been used in [8].

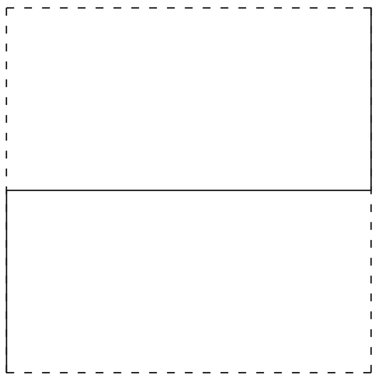
Proof Consider the d-map $K : \vec{I}^2 \to \vec{I}^2$ given by

$$K(s, t) = \begin{cases} (0, 3st) & \text{for } t \leq \frac{1}{3} \\ (3t - 1, s) & \text{for } \frac{1}{3} \leq t \leq \frac{2}{3} \\ (1, 3(t + s - st)) - 2 & \text{for } \frac{2}{3} \leq t, \end{cases}$$

cf. Fig. 1. Note that $K(s, 0) = (0, 0)$, $K(s, 1) = (1, 1)$, $s \in I$, and that $K(0, t)$, $K(1, t)$ are parametrizations of the two d-paths connecting $(0, 0)$ with $(1, 1)$ along the boundary $\partial\vec{I}^2$ of $\vec{I}^2$.

Composition $H \circ K$ defines a d-homotopy (with fixed boundary) connecting a reparametrization of $\bar{H}_0(x, y)(\sigma)$ to a reparametrization of $\bar{H}_1(x, y)$. □

Fig. 1 The path $K(\frac{1}{2}, t)$



5.2 Automorphic Homotopy Flows

Definition 5.4 A future/past homotopy flow $H : X \times \vec{I} \to X$ is called

(1) *Automorphic* if, for all $x, y \in X$ with $\vec{T}(X)(x, y) \neq \emptyset$ and all $s \in I$, the maps $\vec{T}(H^s_+)$, resp. $\vec{T}(H^s_-)$ (vertical in Eq. 5.3) are *homotopy equivalences*;
(2) The sets consisting of all automorphic future/past homotopy flows will be denoted by $\vec{P}_\pm Aut(X)$.
(3) A self-d-map $f : X \to X$ is called a future/past-automorphism if there exists an automorphic future/past homotopy flow between f and the identity on X. The set of all future/past-automorphisms on X will be denoted $Aut_*(X) \subseteq \vec{C}(X, X)$, $* = +, -$.

Remark 5.5

(1) In particular, the maps

$$\vec{T}(f) : \vec{T}(X)(x, y) \to \vec{T}(X)(fx, fy) \text{ resp. } \vec{T}(g) : \vec{T}(X)(gx, gy) \to \vec{T}(X)(x, y)$$

are homotopy equivalences.
(2) Using the concatenation $*$ of homotopy flows from Eq. 5.2, it is obvious that *automorphisms* form sub*monoids* of $\vec{C}(X, X)$.
(3) The definitions above come close to that of a flow on a manifold. But remark again, that the maps $H(-, t) :\to X$ are not required to be homeormorphic; in particular, they will in general not be invertible.

Lemma 5.6 *Let H denote a future/past homotopy flow on X.*

(1) *If all (skew) concatenation maps in Eq. 5.3 are homotopy equivalences, then H is automorphic.*
(2) *Let H be automorphic. If one of the (skew) concatenation maps in Eq. 5.3 is a homotopy equivalence, then the other is as well.*

Proof Immediate from Eq. 5.3. □

Remark 5.7 It is in general not enough to ask that the maps induced by the entire d-homotopy H in Eq. 5.3 are homotopy equivalences. In general, one cannot conclude that the maps induced by H^s are homotopy equivalences, as well. We need that finer requirement crucially in the discussion of components in Section 7.

Remark 5.8 In the undirected case, it is unnecessary to ask homotopies to be automorphic : if $x, y, x', y' \in X$ are in the same path-component of a topological space, then the sets of paths $P(X)(x, y)$ and $P(X)(x', y')$ are always homotopy equivalent to each other.

One may also study the effects of the maps induced by a homotopy flow under the homology, resp. homotopy functors from Section 3 ($n > 0$ and $\vec{T}(X)(x, y) \neq \emptyset$):

$$
\begin{array}{ccc}
\vec{H}_{n+1}(X)(x,y) & & \\
\Big\downarrow \vec{T}(H^s_+)_* & \searrow^{\vec{T}^X(c_x, H^s_{+y})_*} & \vec{H}_{n+1}(X)(x, H_+(y,s)) \\
\vec{H}_{n+1}(X)(H_+(x,s), H_+(y,s)) & \nearrow_{\vec{T}^X(H^s_{+x}, c_y)_*} &
\end{array}
\qquad
\begin{array}{ccc}
\vec{H}_{n+1}(X)(x,y) & & \\
\Big\downarrow \vec{T}(H^s_-)_* & \searrow^{\vec{T}^X(H^s_{-x}, c_y)_*} & \vec{H}_{n+1}(X)(H_-(x,s), y) \\
\vec{H}_{n+1}(X)(H_-(x,s), H_-(y,s)) & \nearrow_{\vec{T}^X(c_x, H^s_{-y})_*} &
\end{array}
\tag{5.4}
$$

and

$$
\begin{array}{ccc}
\vec{\pi}_{n+1}(X)(\sigma xy) & & \\
\Big\downarrow \vec{T}(H^s_+)_\# & \searrow^{\vec{T}^X(c_x, H^s_{+y})_\#} & \vec{\pi}_{n+1}(X)(\sigma xy * H^s_{+y}) \\
\vec{\pi}_{n+1}(X)(H^s_+ \sigma xy) & \nearrow_{\vec{T}^X(H^s_{+x}, c_y)_\#} &
\end{array}
\qquad
\begin{array}{ccc}
\vec{\pi}_{n+1}(X)(\sigma xy) & & \\
\Big\downarrow \vec{T}(H^s_-)_\# & \searrow^{\vec{T}^X(H^s_{-x}, c_y)_\#} & \vec{\pi}_{n+1}(X)(H^s_{-x} * \sigma xy) \\
\vec{\pi}_{n+1}(X)(H^s_- \sigma xy) & \nearrow_{\vec{T}^X(c_x, H^s_{-y})_\#} &
\end{array}
\tag{5.5}
$$

An automorphic homotopy flow induces bijections/group isomorphisms $\vec{T}(H^s)_*$ on $\vec{H}_*(X)$, resp. $\vec{T}(H^s)_\#$ on $\vec{\pi}_*(X)$. Sometimes, a weaker requirement can do (and will be used in Section 7):

Definition 5.9

(1) A homotopy flow is said to be *automorphic up to homology*, if it induces isomorphisms on all relevant homology groups in Eq. 5.4.
(2) A homotopy flow is said to be automorphic up to homotopy/homology *in or up to a fixed dimension k* if it induces isomorphisms on all homotopy/homology sets/groups in Eq. 5.5/5.4 *in or up to dimension k*.
(3) Spaces of such homotopy flows are denoted by an additional decoration, we write e.g. $\vec{P}_{\pm}Aut_{H_{\leq k}}(X)$.

Again, maps that are future/past d-homotopic to the identity by a d-homotopy satisfying one of the requirements above, form a submonoid of $\vec{C}(X, X)$.

Example 5.10

(1) Let X denote the d-space (square with a hole) from Example 4.6. A future homotopy flow will always preserve A and D, but it may move elements of both B, resp. C into D. An *automorphic* future homotopy flow does *not* allow this:

Let $f : X \to X$ denote a d-map, $x,\ fx \in A,\ y \in B,\ fy \in D$. Then $\vec{T}(X)(x, y)$ is contractable whereas $\vec{T}(X)(fx, fy)$ consists of two path-components. Hence, there cannot exist an automorphic future homotopy flow $H : id_X \to f$.

(2) Let S denote the "Swiss flag" po-space [10], cf. the drawing in the middle of Fig. 2. By a combination of a future and a past homotopy flow, the identity on S is d-homotopic to a map that sends X to the 1-skeleton of the outer square; the area Y_1 and in particular the "deadlock" d will be sent to the minimal element x_0. But the dihomotopies involved are *not* automorphic. It is important for applications in concurrency theory that the Swiss-flag space and its outer boundary should *not* be considered as equivalent: Deadlocks should not disappear under an equivalence!

6 Dihomotopy Equivalences

6.1 Definitions and Remarks

Definition 6.1

(1) A d-map $f : X \to Y$ is called a *future dihomotopy equivalence* if there exist d-maps $f_+ : X \to Y, g_+ : Y \to X$ such that f, f_+ are d-homotopic and *automorphic* d-homotopies $H^X : id_X \to g_+ \circ f_+$ on X and $H^Y : id_Y \to f_+ \circ g_+$ on Y.

(2) The d-map $f : X \to Y$ is called a *past dihomotopy equivalence* if there exist d-maps $f_- : X \to Y, g_- : Y \to X$ such that f, f_- are d-homotopic and *automorphic* d-homotopies $H^X : g_- \circ f_- \to id_X$ on X and $H^Y : f_- \circ g_- \to id_Y$ on Y.

(3) The d-map f is called a *dihomotopy equivalence* if it is *both* a future and a past dihomotopy equivalence.

Remark 6.2

(1) The requirements in Definition 6.1 above should be seen as a requirement to a d-homotopy *class* of dimaps from X to Y.

(2) Example 5.10 shows that we ask for more than just the existence of an inverse up to d-homotopy.

(3) Only the essential extra (automorphism) requirement allows to arrive at valuable conclusions regarding effects on the trace, fundamental, homotopy and homology categories; cf. Proposition 6.4 below.

(4) A similar requirement is unnecessary in the classical undirected case; cf. Remark 5.8.

(5) Our definition above is related to the definition of a *faithful* future, resp. past equivalence in the work of Grandis [18] between (general) categories. For a category like the fundamental category $\mathcal{C} = \vec{\pi}_1(X)$ or the trace category $\mathcal{C} = \vec{T}(X)$, faithfulness amounts to asking that the map induced by right concatenation with H^X_y from $\mathcal{C}(x, y)$ to $\mathcal{C}(x, gfy)$ is *epi* and *mono* within the category $\mathcal{C}$. This requirement has nice consequences (cf. the Cancellation Lemma 2.2 in [18], which requires an extra coherence condition), but it is restricted to what can be seen from within the category $\mathcal{C}$. Our requirements are expressed with respect to a functor (e.g., $\vec{T}^X$) to a *different* category (e.g. $Ho - Top$).

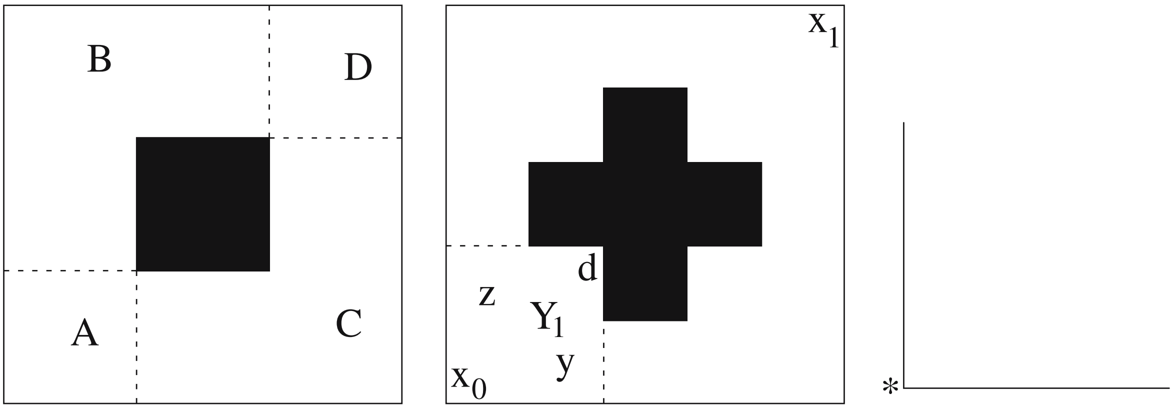


Fig. 2 The po-spaces X and S from Example 5.10 and L from Example 6.3

(6) An argument similar to that used for pf-equivalences in [18], Section 3, shows that the "homotopy inverses" g_+, g_- of a dihomotopy equivalence are d-homotopic to each other: $g_- \mapsto g_+ \circ f_+ \circ g_-$ (by H_+^X), $g_+ \circ f_+ \circ g_-$ is d-homotopic to $g_+ \circ f_- \circ g_-$, and $g_+ \circ f_- \circ g_- \mapsto g_+$ (by H_-^Y). Another d-homotopy is given by $g_- \mapsto g_- \circ f_+ \circ g_+ \simeq g_- \circ f_- \circ g_+ \mapsto g_+$. Grandis requires $f = f_+ = f_-$ and, for coherence, that the two compositions above agree.

Example 6.3 Here is a simple example of a past dihomotopy equivalence that is *not* future: Let L denote a "branching" po-space (a subspace of Euclidean space with induced partial order and hence d-space structure) in the shape of the letter L with base element $*$ as its lower left vertex, cf. drawing in Fig. 2 on the right.

Inclusion $i : \{*\} \to L$ and the constant map $c : L \to \{*\}$ satisfy $c \circ i = id_*$; moreover, there is an (increasing) dihomotopy $i \circ c \to id_L$. But for *no* map $i_+ : \{*\} \to L$ does there exist an (increasing) dihomotopy $id_L \to i_+ \circ c$. Another way to phrase this is: The space L is *past contractible*, but *not future contractible*. Compare [16]. It is crucial for applications in concurrency that dihomotopy equivalence distinguish between a branching and a non-branching space.

In general, a d-space with more than one local maximum cannot be future contractible; if it has more than one local minimum, it cannot be past contractible.

6.2 Properties of Dihomotopy Equivalences

Proposition 6.4 *The natural transformation $\vec{T}_\pi(f) : \vec{T}_\pi^X \to \vec{T}_\pi^Y$ induced by a (past or future) dihomotopy equivalence $f : X \to Y$ between d-spaces X and Y is an equivalence, i.e., the induced maps $\vec{T}(f)(x, y) : \vec{T}(X)(x, y) \to \vec{T}(Y)(fx, fy)$ are homotopy equivalences.*

Proof By abuse of notation, we write f, g instead of f^+, g^+, resp. f^-, g^- in the following. In the diagram

$$\vec{T}(X)(x, y) \xrightarrow{\vec{T}(f)} \vec{T}(Y)(fx, fy) \xrightarrow{\vec{T}(g)} \vec{T}(X)(gfx, gfy) \xrightarrow{\vec{T}(f)} \vec{T}(Y)(fgfx, fgfy),$$

let I denote a homotopy inverse to $\vec{T}(g) \circ \vec{T}(f)$ and let J denote a homotopy inverse to $\vec{T}(f) \circ \vec{T}(g)$. Then $\vec{T}(g)$ has a homotopy right inverse $\vec{T}(f) \circ I$ and a homotopy left inverse $J \circ \vec{T}(f)$. By general nonsense, the right homotopy inverse and the left homotopy inverse are homotopic to each other, and thus $\vec{T}(g)$ is a homotopy equivalence. Since $\vec{T}(g \circ f) = \vec{T}(g) \circ \vec{T}(f)$ is a homotopy equivalence by definition, the map $\vec{T}(f)$ is a homotopy equivalence, as well. □

Example 6.5 The "Swiss flag" space S can*not* be dihomotopy equivalent to a *graph* G: With reference to the middle drawing in Fig. 2, assume $f : S \to G$ is a dihomotopy equivalence. By Proposition 6.4, there are unique directed paths $\rho_y : fx_0 \to fy$ (from fx_0 to fy), $\rho_z : fx_0 \to fz, \sigma_y : fy \to fd, \sigma_z : fy \to fd, \tau_y : fy \to fx_1, \tau_z : fz \to fx_1$ and

$$\rho_y * \sigma_y = \rho_z * \sigma_z, \ \ \rho_y * \tau_y \neq \rho_z * \tau_z. \tag{6.1}$$

There is only one directed connection from fx_0 to fd in G. Hence, we can assume without restriction of generality, that ρy is a "prefix" of ρ_z, i.e., there exists $\rho_{yz} : fy \to fz$ such that $\rho_z = \rho_y * \rho_{yz}$. But then $\rho_z * \tau_z = \rho_y * \rho_{yz} * \tau_z = \rho_y * \tau_y$. This contradicts Eq. 6.1!

As in [18], a future dihomotopy equivalence gives rise to "adjunction" maps induced by the homotopy flows H^X, H^Y from Definition 6.1 above; in the following, we shall write f, g instead of f_+, g_+, f_-, g_-:

$$\mathcal{H}^X : \vec{T}(Y)(fx, y) \xrightarrow{\vec{T}(g)} \vec{T}(X)(gfx, gy) \xrightarrow{\vec{T}^X(H_{+x}, c_{gy})} \vec{T}(X)(x, gy), \ \alpha \mapsto H_{+x} * g\alpha, \tag{6.2}$$

$$\mathcal{H}^Y : \vec{T}(X)(gy, x) \xrightarrow{\vec{T}(f)} \vec{T}(Y)(fgy, fx) \xrightarrow{\vec{T}^Y(H_{+y}, c_{fx})} \vec{T}(Y)(y, fx), \ \beta \mapsto H_{+y} * f\beta. \tag{6.3}$$

For past equivalences, the respective maps are given by $\alpha \mapsto g\alpha * H_{-x}$, resp. $\beta \mapsto f\beta * H_{-y}$. These maps are homotopy equivalences if and only if the maps $\vec{T}^X(H_{+x}, c_{gy})$ are homotopy equivalences.

Future and past dihomotopy equivalences behave well under composition:

Proposition 6.6 *The composition $g \circ f : X \to Z$ of (future or past) dihomotopy equivalences $X \xrightarrow{f} Y \xrightarrow{g} Z$ is again a (f/p) dihomotopy equivalence.*

Proof Let $Z \xrightarrow{g'} Y \xrightarrow{f'} X$ denote homotopy inverses to f, resp. g, and let $H^X : id_X \to f' \circ f$ and $H^Y : id_Y \to g' \circ g$ denote automorphic homotopy flows. By a slight abuse of notation, let $f' \circ H^Y \circ f : f' \circ f \to f' \circ g' \circ g \circ f$ denote the induced d-homotopy on X. The composition $\bar{H} : id_X \xrightarrow{H^X} f' \circ f \xrightarrow{f' \circ H^Y \circ f} f' \circ g' \circ g \circ f$ is a homotopy flow; we have to show that it is automorphic. The "levels" $\bar{H}_s$ of $\bar{H}$ are either of the form H^X_s – which induce homotopy equivalences by definition – or of the form $f' \circ H^Y_s \circ f$.

Let $I : \vec{T}(X)(f'fx, f'fy) \to \vec{T}(X)(x, y)$ denote a homotopy inverse to $\vec{T}(f') \circ \vec{T}(f)$, let $J : \vec{T}(Y)(ff'fx, ff'fy) \to \vec{T}(Y)(fx, fy)$ denote a homotopy inverse to

$\vec{T}(f) \circ \vec{T}(f')$, and let $K_s : \vec{T}(Y)(H^Y_s(fx), fH^Y_s(fy)) \to \vec{T}(Y)(fx, fy)$ denote a homotopy inverse to $\vec{T}(H^Y_s)$. Then $\vec{T}(f') \circ \vec{T}(H^Y_s) \circ \vec{T}(f)$ has homotopy right and left inverses:

- $(I \circ \vec{T}(f') \circ K_s \circ J \circ \vec{T}(f)) \circ (\vec{T}(f') \circ \vec{T}(H^Y_s) \circ \vec{T}(f)) \simeq id$ on $\vec{T}(X)(x, y)$;
- $(\vec{T}(f') \circ \vec{T}(H^Y_s) \circ \vec{T}(f)) \circ (\vec{T}(f') \circ J \circ K_s \circ \vec{T}(f) \circ I) \simeq id$ on $\vec{T}(X)(x, y)$,

which have to agree by general nonsense. □

7 Automorphic Homotopy Flows and Components

7.1 Motivation

The localization construction from Section 4 has a drawback: It is in general difficult to get hold on a (preferably large) subcategory Σ of the original category that both maps into isomorphisms and satisfies the extension properties (and hopefully also purity). The construction in [13] focussing on the strong extension properties is categorically very satisfactory, but it is not clear that the resulting category will be "large enough" to yield satisfactory compression.

The construction below is of a more "geometric nature" and uses the automorphic homotopy flows introduced in Section 5. On the positive side, the (non-strong) extension properties follow right away. Moreover, the set-up is very much related to the dihomotopy equivalences from Section 6, a fact that will be exploited in the final Section 8. On the negative side, it seems not to be possible to prove that the resulting subcategory is pure and Proposition 4.5.(4) need not always be satisfied: the associated component category can therefore have isomorphisms that split into non-isomorphisms (which may leave the component of the start element).

7.2 The Extended Preorder Category

We extend the preorder category $\vec{D}_\pi(X)$ to a category $\vec{D}^{Aut_+}_\pi(X)$ with the same objects but with more morphisms. The morphisms in this new category are generated by those from the previous and additionally by morphisms $f(x, y)$ from (x, y) to (fx, fy) for every $f \in Aut_+(X)$ and every $x \preceq y$ subject to the following relations (compare Eq. 5.3 for every automorphic future homotopy flow H_+, resp. every past homotopy flow H_-):

$$\begin{array}{ccc} (fx, fy) & \xrightarrow{(c_{fx}, f\sigma)} & (fx, fz) \\ {\scriptstyle f(x,y)}\uparrow & & \uparrow{\scriptstyle f(x,z)} \\ (x, y) & \xrightarrow{(c_x, \sigma)} & (x, z) \end{array} \qquad \begin{array}{ccc} (fx, fy) & \xrightarrow{(f\tau, c_{fy})} & (fu, fy) \\ {\scriptstyle f(x,y)}\uparrow & & \uparrow{\scriptstyle f(u,y)} \\ (x, y) & \xrightarrow{(\tau, cy)} & (u, y) \end{array} \tag{7.1}$$

for $\sigma \in \vec{T}(X)(y,z)$, $\tau \in \vec{T}(u,x)$, and

$$\begin{array}{ccc} (x,y) & \xrightarrow{(c_x, H_{+y})} & (x, H_+(y)) \\ {\scriptstyle H_+(x,y)}\downarrow & \nearrow{\scriptstyle (H_{+x}, c_{H_+(y)})} & \\ (H_+(x), H_+(y)) & & \end{array} \qquad \begin{array}{ccc} (x,y) & \xrightarrow{(H_{-x}, c_y)} & (H_-(x), y) \\ {\scriptstyle H_-(x,y)}\downarrow & \nearrow{\scriptstyle (c_{H_-(x)}, H_{-y})} & \\ (H_-(x), H_-(y)) & & \end{array} \tag{7.2}$$

for a an automorphic future homotopry flow H_+, resp. a past homotopy flow H_-.

By abuse of notation, we let $Aut_+(X)$ denote the wide subcategory of $\vec{D}_\pi^{Aut_+}(X)$ with morphisms stemming from automorphisms alone; this is in fact a subcategory, since $Aut_+(X)$ is closed under concatenation. There is an obvious variant $Aut_-(X)$ giving rise to an extended preorder category $\vec{D}_\pi^{Aut_-}(X)$. Considering future and past homotopy flows simultaneously leads to the category $Aut(X)$ with morphisms $(f_+, f_-)(x,y) : (x,y) \to (x', y')$ with $f_+(x) = x'$, $f_+(y) = y'$, $f_-(x') = x$, $f_-(y') = y$ such that $\vec{T}(f_- f_+) \simeq id$ on $\vec{T}(X)(x,y)$ and $\vec{T}(f_+ f_-) \simeq id$ on $\vec{T}(X)(x', y')$. The remaining part of this section are formulated for $Aut(X)$, but most of it applies also for the future, resp. past versions.

Proposition 7.1

(1) *Every element of* $\vec{D}_\pi^{Aut_+}(X)$ *can be written in the form* $c \circ f$ *with* c *a morphism in the preorder category and* $f \in Aut(X)$.
(2) *Within* $\vec{D}_\pi^{Aut}(X)$*, the subcategory* $Aut(X)$ *satisfies LEP/REP* with respect to $\vec{D}_\pi(X)$ *– explanation in the proof.*
(3) *The functor* $\vec{T}^X : \vec{D}_\pi(X) \to Ho - Top$ *extends to* $\vec{D}_\pi^{Aut}(X)$ *and maps* $Aut(X)$ *into isomorphisms.*

Proof

(1) Successive application of Eq. 7.1 applied to a "mixed" morphism.
(2) For $f_+ \in Aut_+(X)$, $f_- \in Aut_-(X)$, $\tau \in \vec{T}(X)(x,y)$, we can fill in extension diagrams as follows:

$$\begin{array}{ccc} (f_+x, f_+y) & \overset{(c_{f_+x}, f_+\tau)}{\dashrightarrow} & (f_+x, f_+y') \\ {\scriptstyle f_+(x,y)}\uparrow & & \uparrow{\scriptstyle f_+(x,y')} \\ (x,y) & \xrightarrow{(c_x, \tau)} & (x, y') \end{array} \qquad \begin{array}{ccc} (x,y) & \xrightarrow{\tau} & (x, y') \\ {\scriptstyle f_-(x,y')}\uparrow & & \uparrow{\scriptstyle f_-(x,y)} \\ (f_-x, f_-y) & \overset{(c_{f_-x}, f_-\tau)}{\dashrightarrow} & (f_-x, f_-y'). \end{array}$$

(3) $\vec{T}^X$ maps $f(x,y) : (x,y) \to (fx, fy)$ into the homotopy equivalence $\vec{T}(f)(x,y) : \vec{T}(X)(x,y) \to \vec{T}(X)(fx, fy)$. The relations in Eqs. 7.1 and 7.2 are respected. □

Remark 7.2

(1) The coherent automorphisms (cf. Definition 7.1) satisfy LEP/REP with respect to all of $\vec{D}_\pi^{Aut}(X)$.
(2) An alternative treatment, letting automorphic homotopy flows "act" on the preorder category (not extended) is outlined in the final Section 8.4.

7.3 Category of Fractions and Components

Now we construct a component category from the pair $(\vec{D}_\pi^{Aut}(X), Aut(X))$ along the outline from Section 4.3. Since Proposition 7.1(2) only leads to solutions of extension properties with respect to the subcategory $\vec{D}_\pi(X)$, elements of the category of fractions $\vec{D}_\pi^{Aut}(X)[Aut(X)^{-1}]$ will in general have a normal form of the type $c \circ g$ with $g = g_1 \circ h_1^{-1} \circ \cdots \circ g_n \circ h_n^{-1} \in Aut(X)^{-1}$. Still, the component category $\pi_0(\vec{D}_\pi^{Aut}(X), Aut(X))$ is well-defined, coming with a factorization

$$\vec{T}_\pi(X) : \vec{D}_\pi^{Aut}(X) \xrightarrow{q_{Aut}} \pi_0(\vec{D}_\pi^{Aut}(X), Aut(X)) \xrightarrow{\pi_0(\vec{T}_\pi(X))} Ho - Top.$$

More serious is the following fact: Morphisms in the original preorder category $\vec{D}_\pi(X)$ can contribute to the isomorphisms in the category of fractions: For example, assume that $h = g \circ f \in Aut(X)$, $H : id_X \to g$ with $gfx = fx$ and $\tau = H_{fy} \in \vec{T}(fy, gfy = hy)$. Then $h(x, y) = (c_{fx}, \tau) \circ f(x, y)$ and hence $(c_{fx}, \tau) = h(x, y) \circ f(x, y)^{-1}$ is an isomorphism with inverse $f(x, y) \circ h(x, y)^{-1}$.

The purity condition can be decided within the preorder category $\vec{D}_\pi(X)$: Let $Inv(D_\pi(X))$ denote the pullback subcategory in the diagram

$$\begin{array}{ccc} & & Aut(X)^{-1} \\ & & \downarrow \\ \vec{D}_\pi(X) & \longrightarrow & \vec{D}_\pi^{Aut}(X)[Aut(X)^{-1}]. \end{array}$$

It is easy to check that

Lemma 7.3 $(\vec{D}_\pi^{Aut}(X)[Aut(X)^{-1}], Aut(X)^{-1})$ *is pure if and only if* $(\vec{D}_\pi(X), Inv(D_\pi(X)))$ *is pure. If this is the case, isomorphisms in the component category can only split up into isomorphisms.*

If the construction above does not result in a pure subcategory, the biggest subcategory of $Aut(X)^{-1}$ satisfying SLEP/SREP will yield one, but there will in general be no control of its size.

The approach above will (for $Aut(X)$, not necessarily for $Aut_\pm(X)$) lead to particular components containing only elements in the diagonal as objects; both a future flow and a past flow preserve the diagonal.

Example 7.4 Using the approach described above yields almost the same component categories for the spaces X and $\vec{S}^1$ from Example 4.6 as depicted previously.

X Every future homotopy flow preserves A, since its maximum is a fixed point; likewise, every past homotopy flow preserves D. No flow line can ever connect an element of B to an element of C. For $x \preceq y$, the trace spaces $\vec{T}(X)(x, y)$ are contractable unless $x \in A$ and $y \in B$, in which case the trace space has two contractable components. It is not difficult to construct (piecewise linear) automorphic homotopy flows with flow lines connecting any two $x \preceq y$ from a given "component." As a result, we get additional initial components $\Delta_A, \Delta_B, \Delta_C, \Delta_D$, that have to be taken from the components AA etc.

$\vec{S}^1$ The component category has two objects: the diagonal Δ in the torus $S^1 \times S^1$ and its complement T', the torus with the diagonal deleted. To see that pairs $(\alpha_i, \beta_i), i = 1, 2$ are connected within $Aut(X)$, note that a rotation (in $Aut(X)$) connects (α_1, β_1) to $(\alpha_2, \beta_1 + (\alpha_2 - \alpha_1))$, which is connected to (α_2, β_2) by an automorphism that inflates/deflates the arc starting at α_2 and its complement. The component category has two objects Δ and T' and is freely generated by two morphisms $c : \Delta \to T', d : T' \to \Delta$.

Remark 7.5 The set-up goes through without major changes if one replaces automorphisms by automorphisms up to homology/homotopy in a given range of dimensions.

7.4 Naturality Issues

As discussed earlier [9, 14, 22], a general d-map does in general *not* preserve components. Some coherence with the automorphic flows on the two spaces is needed.

Definition 7.6

(1) A d-map $f : X \to Y$ is called *coherent* if, for every pair $x \preceq y$ and every morphism $F(x, y) \in Aut(X)((x, y), (Fx, Fy))$ there exists a morphism $G \in Aut(Y)^{-1}((fx, fy)(fFx, fFy))$.

(2) f is called *strongly coherent* if, for every strictly automorphic past/future homotopy flow H on X, there is a strictly automorphic homotopy flow $\bar{H}$ on Y solving the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \uparrow{\scriptstyle H} & & \uparrow{\scriptstyle \bar{H}} \\ X \times \vec{I} & \xrightarrow[f \times id_I]{} & Y \times \vec{I}. \end{array}$$

It is in general not easy to check for coherence unless one already knows the components of the two spaces. On the other hand, strong coherence (which implies coherence) is more rare, but it is, e.g., satisfied for certain inclusion maps. As a direct consequence of the definitions, we obtain

Proposition 7.7 *A coherent d-map $f : X \to Y$ satisfies $\vec{T}^X(f)(Aut(X)) \subseteq Aut(Y)^{-1}$. Hence it maps components of X into components of Y and induces a functor*

$$f_\# : \pi_0(\vec{D}_\pi^{Aut}(X); Aut(X)) \to \pi_0(\vec{D}_\pi^{Aut}(Y); Aut(Y)).$$

8 Components and Dihomotopy Equivalences

In the following, we collect what we know about the behaviour of dihomotopy equivalences on components. There are many examples of d-maps (e.g. inclusions, cf. [14]) that do not respect components (with respect to whatever equivalence relation). It is not clear from the definitions either whether a dihomotopy equivalence always does; one needs to impose weak coherence (Definition 8.6) as an extra requirement; we will show that weakly coherent dihomotopy equivalences give rise to isomorphic component categories.

8.1 Homotopy Flows and Components

We start by considering the effects of automorphic homotopy flows $H_1 : id_X \to h_1$, resp. $H_2 : h_2 \to id_X$ on the components of a d-space X with respect to the subcategory $Aut(X)$. The following result follows immediately from the definitions:

Lemma 8.1

(1) *The pairs* $(h_2(x), h_2(y)) \mapsto (x, y) \mapsto (h_1(x), h_1(y))$ *are contained in the* same *component for every object* (x, y) *in* $\vec{D}_\pi(X)$.
(2) *For* $h = h_i, i = 1, 2$, *the morphisms* $\tau \in \vec{D}_\pi(X)((x, y), (x', y'))$ *and* $(h \circ \tau) \in \vec{D}_\pi(X)((hx, hy), (hx', hy'))$ *are equivalent for every morphism* τ *in the preorder category.*
(3) *h induces the identity on the component category* $\pi_0(\vec{D}_\pi^{Aut}(X); Aut(X))$.

Analogous results hold for the other component categories considered in Section 7.

8.2 Dihomotopy Equivalences and Components

In order to phrase our first result concerning dihomotopy equivalences, we make use of

Definition 8.2 Let $Y' \subset Y$ denote a (non-empty) subset of a d-space Y. Y' is called *dense* if for every pair (x, x') in $\vec{D}(Y)$ there is a pair (y, y') in $\vec{D}(Y')$ with $(x, x') \preceq_D (y, y')$.

Remark 8.3 The definition corresponds in fact to denseness in an appropriate order topology.

In the following, we consider components with respect to $Aut(X) \subseteq \vec{D}_\pi^{Aut}(X)$.

Proposition 8.4 *Let* $f : X \to Y$ *denote a dihomotopy equivalence. Then*

(1) *The intersection of the image* $f(\vec{D}_\pi(X)) \subset \vec{D}_\pi(Y)$ *with* every *component in* $\vec{D}_\pi(Y)$ *is* dense.
(2) *If* $f(x_1) = f(x_2)$, $f(x_1') = f(x_2')$, $x_i, x_i' \in X$, *then* (x_1, x_1') *and* (x_2, x_2') *are contained in the same component of* $\vec{D}_\pi(X)$.

(3) *Every morphism β in $\vec{D}_\pi(Y)$ is $Aut(Y)$-equivalent to a morphism $f \circ \alpha$ with α a morphism in $\vec{D}_\pi(X)$.*
(4) *If $f \circ \alpha_1 = f \circ \alpha_2$ for morphisms α_i in $\vec{D}_\pi(X)$, then α_1 and α_2 are $Aut(X)$-equivalent.*

Similar results hold also for the other component categories considered previously.

Proof Let $g : Y \to X$ denote a dihomotopy inverse to f.

(1) By Lemma 8.1(1), (y, y') and (fgy, fgy') are contained in the same component in $\vec{D}(Y)$.
(2) (x_i, x_i') is contained in the same component as $(gfx_1, gfx_1') = (gfx_2, gfx_2')$.
(3) β is $Aut(Y)$-equivalent to $fg(\beta)$.
(4) α_i is $Aut(X)$-equivalent to $gf(\alpha_1) = gf(\alpha_2)$. □

Dihomotopy equivalences preserve the topology of trace spaces in the following sense:

Proposition 8.5 *Let $f : X \to Y$ denote a dihomotopy equivalence. Let $(x, y), (x', y')$ be objects of $\vec{D}_\pi(X)$ such that $\vec{T}(x, y)$ and $\vec{T}(x', y')$ are homotopy equivalent. Then the trace spaces $\vec{T}(fx, fy)$ and $\vec{T}(fx', fy')$ are homotopy equivalent, too. Similarly for trace spaces that induce isomorphisms in homology or homotopy in a range of dimensions.*

Proof The diagram

$$\begin{array}{ccccc}
\vec{T}(Y)(fx, fy) & \xrightarrow[\simeq]{\vec{T}(g)} & \vec{T}(X)(gfx, gfy) & \xleftarrow[\simeq]{\vec{T}(gf)} & \vec{T}(X)(x, y) \\
 & & & & \downarrow \simeq \\
\vec{T}(Y)(fx', fy') & \xrightarrow[\vec{T}(g)]{\simeq} & \vec{T}(X)(gfx', gfy') & \xleftarrow[\vec{T}(gf)]{\simeq} & \vec{T}(X)(x', y')
\end{array}$$

yields the homotopy equivalence asked for. □

8.3 Weakly Coherent Dihomotopy Equivalences

8.3.1 Definition

Unfortunately, Proposition 8.5 does not imply that a dihomotopy equivalence induces isomorphisms of component categories. In order to get functoriality, a coherence condition has to be imposed. Note that the condition below is weaker than asking the map f to be coherent itself in the sense of Definition 7.6.

Definition 8.6 Let $f : X \to Y$ denote a (future/past) dihomotopy equivalence with homotopy inverse $g : Y \to X$. The pair (f, g) is called *weakly coherent*, if the d-self maps $g \circ f : X \to X$ and $f \circ g : Y \to Y$ are coherent (cf. Definition 7.6). The map f

is called *weakly coherent* itself, if there exists a homotopy inverse such that (f, g) is weakly coherent.

The extra coherence requirement will be crucial in the investigation of the effect of a dihomotopy equivalence from X to Y on component categories associated to these two d-spaces. One may ask how natural weak coherence is:

Remark 8.7

(1) In an analogous situation, a self-diffeomorphism $F : X \to X$ conjugates a flow diffeomorphism $h_1 : X \to X$ to the flow diffeomorphism $h_2 = F \circ h_1 \circ F^{-1}$ – of a different dynamical system.
(2) The coherence condition corresponds to a weaker form of the coherence requirement in the definition of future and past equivalences in [18].

8.3.2 Coherent Dihomotopy Equivalences Induce Isomorphisms of Component Categories

Proposition 8.8 *A weakly coherent dihomotopy equivalence $f : X \to Y$ induces an* isomorphism $\pi_0\vec{T}(f) : \pi_0(\vec{D}_\pi(X), Aut(X)) \to \pi_0(\vec{D}_\pi(Y), Aut(Y))$ *of component categories.*

Proof The only substantial difficulty arises in proving the *existence* of a functor $\pi_0\vec{T}(f)$ as above; it is here that weak coherence is needed. Let $h_1 \in Aut_+(X)$ and $(x, y) \in \vec{D}_\pi(X)$. We have to show, that there is an $Aut(Y)^{-1}$-morphism connecting (fx, fy) and $fh_1x, fh_1y)$.

Since $gf : X \to X$ is coherent, there exists $h_2 \in Aut(X)^{-1}$ with $(gfh_1x, gfh_1y) = (h_2gfx, h_2gfy)$. Then $fh_2g : Y \to Y$ is an automorphism (see the proof of Proposition 6.6) that connects (fx, fy) to (fh_2gfx, fh_2gfy), whereas the automorphism fg connects (fh_1x, fh_1y) to $(fgfx, fgfh_1y) = (fgfx, fh_2gfy)$:

$$
\begin{array}{ccc}
(fh_2gfx,\ fh_2gfy) & \overset{=}{\longleftrightarrow} & (gfgh_1x,\ fgfh_1y) \\
\uparrow{\scriptstyle fh_2g(x,y)} & & \uparrow{\scriptstyle fg(fh_1x,\ fh_1y)} \\
(fx,\ fy) & & (fh_1x,\ fh_2y)
\end{array}
$$

Iterating the argument above, (for a zig-zag path in $\vec{D}_\pi^{Aut}(X)$) shows that equivalent pairs in $\vec{D}_\pi^{Aut}(X)$ are mapped to equivalent pairs in $\vec{D}_\pi^{Aut}(Y)$ under $\vec{T}_\pi(f)$. The construction is well-behaved with respect to concatenation maps and yields therefore a functor $\pi_0\vec{T}(f) : \pi_0(\vec{D}_\pi^{Aut}(X), Aut(X)) \to \pi_0(\vec{D}_\pi^{Aut}(Y), Aut(Y))$.

Using the coherence of $fg : Y \to Y$, one obtains a reverse functor $\pi_0\vec{T}(g) : \pi_0\vec{D}_\pi(Y) \to \pi_0\vec{D}_\pi(X)$. By Lemma 8.1(3), the compositions of the two functors yield identity functors on the component categories of X, resp. Y. □

In conclusion, if two d-spaces have non-isomorphic component categories, then there cannot exist a weakly coherent dihomotopy equivalence between them.

8.4 A More General Perspective?

The preceding sections dealt only with categories arising from path spaces of (topological) d-spaces – as quotients etc. A generalization of the approach to more general categories can be described as follows:

We consider small categories *over* a given small category $\mathcal{D}$, i.e., every category $\mathcal{C}$ is endowed with a functor $F : \mathcal{C} \to \mathcal{D}$ into $\mathcal{D}$; typically, $\mathcal{D} = Ho - Top$, *Sets*, *Grps*, *Ab* etc. Consider an equivalence relation $\equiv$ on the objects of $\mathcal{C}$ such that

- For every $x \equiv x'$, there is a subset $\emptyset \neq I(F(x), F(x')) \subset Iso(F(x), F(x'))$ such that $I(F(x), F(x')) = \varphi \circ I(F(x), F(x))$ for every $\varphi \in I(F(x), F(x'))$;
- For every $x \equiv x', \varphi \in I(F(x), F(x')), \sigma \in Mor_{\mathcal{C}}(x, y)$, there exist $y \equiv y', \varphi' \in I(F(y), F(y'))$ and $\sigma' \in Mor_{\mathcal{C}}(x', y')$ such that the diagrams

$$\begin{array}{ccc} F(x) & \xrightarrow{F(\sigma)} & F(y) \\ {\scriptstyle \varphi}\downarrow & & \downarrow{\scriptstyle \varphi'} \\ F(x') & \overset{F(\sigma')}{\dashrightarrow} & F(y') \end{array} \qquad\qquad \begin{array}{ccc} F(x) & \overset{F(\tau')}{\dashrightarrow} & F(y) \\ {\scriptstyle \psi'}\downarrow & & \downarrow{\scriptstyle \psi} \\ F(x') & \xrightarrow{F(\tau)} & F(y') \end{array}$$

commute.

The relation $\equiv$ extends to morphisms ($\sigma \equiv \sigma', \tau \equiv \tau'$) and sequences of morphisms and generates *generalized congruence* on $\mathcal{C}$ in the sense of [2] and a quotient functor $T : \mathcal{C} \to \mathcal{C}/_{\equiv}$; compare [13, 19]. We consider then $\mathcal{C}$ and $\mathcal{C}/_{\equiv}$ as *similar* categories with respect to $\mathcal{D}$.

In order to make similarity an equivalence relation, consider the transitive symmetric closure of this relation arising from zig-zags

$$\mathcal{C}_1 \to \mathcal{C}_1/_{\equiv_1} \cong \mathcal{C}_2/_{\equiv_2} \leftarrow \mathcal{C}_2 \to \cdots,$$

where the quotient categories are assumed to be equivalent. This defines the notion of categories (or rather functors) that are *similar over* $\mathcal{D}$.

A method to obtain an equivalence relation $\equiv$ as above generalising automorphic homotopy flows, can be described as follows: Given a category $F : \mathcal{C} \to \mathcal{D}$ over $\mathcal{D}$, consider an endofunctor $\Phi : \mathcal{C} \to \mathcal{C}$ together with a "directed homotopy" $\varphi : 1_{\mathcal{C}} \to \Phi$ in the sense of [16], i.e., a natural transformation. Such a pair (Φ, φ) is called a *future auto-equivalence over* $\mathcal{D}$ if the morphisms $F(\varphi(x)) : F(x) \to F(\Phi(x))$ are $\mathcal{D}$-isomorphisms for *all* objects x in $\mathcal{C}$. Note the following consequence for every $\mathcal{C}$-morphism σ: $F(\sigma)$ is a $\mathcal{D}$-isomorphism if and only if $F(\Phi(\sigma))$ is a $\mathcal{D}$-isomorphism.

Past auto-equivalences are defined in the same way using directed homotopies $\varphi : \Phi \to 1_{\mathcal{C}}$. Both future and past auto-equivalences over $\mathcal{D}$ form monoids $M^{\pm}_{\mathcal{C}}$. An equivalence relation, $M_{\mathcal{C}}$-equivalence, on objects x, x' arises from the requirement that there are a future auto-equivalence Φ and a past auto-equivalence Ψ with $\Phi x = x'$ and $\Psi x' = x$. Similarly for morphisms.

Also the notion of dihomotopy equivalence can be generalised to this setting: Two functors $F_i : \mathcal{C}_i \to \mathcal{D}$ can be related by a functor pair $(\Psi : \mathcal{C}_1 \to \mathcal{C}_2, \psi : F_1(\mathcal{C}_1) \to \mathcal{D})$ such that the diagram

$$\begin{array}{ccc} \mathcal{C}_1 & \xrightarrow{\Psi} & \mathcal{C}_2 \\ {\scriptstyle F_1}\downarrow & & \downarrow{\scriptstyle F_2} \\ F_1(\mathcal{C}_1) & \xrightarrow{\psi} & \mathcal{D} \end{array}$$

commutes on both objects and morphisms. Let $M_{\mathcal{C}_1}$, resp. $M_{\mathcal{C}_2}$ denote the monoids of auto-equivalences over $\mathcal{D}$. A functor pair $(\Psi, \psi) : \mathcal{C}_1 \to \mathcal{C}_2$ over $\mathcal{D}$ is then called a *future homotopy equivalence* if there is a functor pair $(\Gamma, \gamma) : \mathcal{C}_2 \to \mathcal{C}_1$ over $\mathcal{D}$ such that $F_1 \circ \Gamma \circ \Psi = \gamma \circ \psi \circ F_1$, $F_2 \circ \Psi \circ \Gamma = \psi \circ \gamma \circ F_2$, and $\Gamma \circ \Psi, \Psi \circ \Gamma$ are $\mathcal{M}^+_{\mathcal{C}_i}$-equivalences. Arguing formally as in the proof of Proposition 6.4, it can then be seen that $\psi : F_1(x) \to F_2(\Psi(x))$ is a $\mathcal{D}$-isomorphism for every object x in $\mathcal{C}_1$.

Not all meaningful equivalence relations and generalised congruences arise from (automorphic) homotopy flows, resp. auto-equivalences; those do not always identify "enough" objects:

Example 8.9 Let $X = \vec{I}^2 \setminus \vec{J}^2$ denote the square with a hole studied in Examples 4.6 and 7.4. Every (continuous) future auto-equivalence leaves the region A in Fig. 2 invariant; likewise, every past auto-equivalence leaves D invariant. Regardless the chosen functor on $\vec{D}(X)$, it will thus only be possible to identify objects in A with each other etc.

The functor $\vec{H}_2(X)$ maps every object (x, y) of $\vec{D}(X)$ to the first homology of the trace space from x to y which is trivial: $\vec{H}_2(X)(x, y) = 0$ for all $x \preceq y$. Hence, there is a generalised congruence over Ab arising from $\vec{H}_2$ that identifies all objects with each other, and also all morphisms with each other. With respect to $\vec{H}_2$, the category $\vec{D}(X)$ and the category with one object and only one identity morphism are similar.

Acknowledgements The author thanks Institut Mittag-Leffler, Djursholm, Sweden, for its hospitality during his participation, in April 2006, in its Algebraic Topology programme, that allowed him to organise the material in this article. He also thanks the referee for very valuable suggestions on how to improve the presentation.

References

1. Baues, H.J.: Algebraic homotopy. In: Cambridge Studies in Advanced Mathematics, no. 15. Cambridge University Press, Cambridge (1989)
2. Bednarczyk, M., Borzyskowski, A., Pawlowski, W.: Generalized congruences – epimorphisms in $\mathcal{C}at$. Theory Appl. Categ. **5**(11), 266–280 (1999)
3. Borceux, T.: Handbook of categorial algebra I: basic category theory. In: Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge (1994)
4. Brown, R., Higgins, P.J.: Colimit theorems for relative homotopy groups. J. Pure Appl. Algebra **22**, 11–41 (1981)
5. Brown, R., Higgins, P.J.: On the algebra of cubes. J. Pure Appl. Algebra **21**, 233–260 (1981)
6. Carlsson, G., Zomorodian, A., Collins, A., Guibas, L.: Persistence barcodes for shapes. Int. J. Shape Model. **11**, 149–187 (2005)

7. Fahrenberg, U., Raussen, M.: Reparametrizations of continuous paths. Tech. Report R-2006-22, Department of Mathematical Sciences, Aalborg University (2006) (to appear in J. Homotopy Relat. Struct. **2**(2), 2007)
8. Fajstrup, L.: Dipaths and Dihomotopies in a cubical complex. Adv. Appl. Math **35**(2), 188–206 (2005)
9. Fajstrup, L., Goubault, É., Haucourt, E., Raussen, M.: Components of the fundamental category. Appl. Categ. Structures **12**(1), 81–108 (2004)
10. Fajstrup, L., Goubault, É., Raussen, M.: Algebraic topology and concurrency. Theor. Comput. Sci. **357**, 241–278 (2006)
11. Gaucher, P.: A model category for the homotopy theory of concurrency. Homology Homotopy Appl. **5**(2), 549–599 (2003)
12. Gaucher, P.: Inverting weak dihomotopy equivalence using homotopy continuous flow. Theory Appl. Categ. **16**(3), 59–83 (2006)
13. Goubault, É., Haucourt, E.: Components of the fundamental category II. Appl. Categ. Structures (2007) (this volume)
14. Goubault, É., Raussen, M.: Dihomotopy as a tool in state space analysis. In: Rajsbaum, S. (ed.) LATIN 2002: Theoretical Informatics (Cancun, Mexico). Lecture Notes in Computer Science, vol. 2286, pp. 16–37. Springer, Berlin (2002)
15. Grandis, M.: Directed homotopy theory II. Homotopy constructs. Theory Appl. Categ. **10**(14), 369–391 (2002)
16. Grandis, M.: Directed homotopy theory I. The fundamental category. Cahiers Topologie Géom. Différentielle Catég. **44**, 281–316 (2003)
17. Grandis, M.: Modelling fundamental 2-categories for directed homotopy. Homology Homotopy Appl. **8**, 31–70 (2006)
18. Grandis, M.: The shape of a category up to directed homotopy. Theory Appl. Categ. **15**(4), 95–146 (2005)
19. Haucourt, E.: Topologie algé. Ph.D. thesis, Univeristé Paris 7, UER Informatique (2005)
20. Krishnan, S.: A convenient category of locally ordered spaces. In: GETCO '05, Preliminary Proceedings (San Francisco), vol. NS-05, BRICS, (Full version to appear in ENTCS Proceedings, p. 33) (2005)
21. Penrose, R.: Techniques of differential topology in relativity. In: Conference Board of the Mathematical Sciences. Regional Conference Series in Applied Mathematics, vol. 7. SIAM, Philadelphia, PA (1972)
22. Raussen, M.: State spaces and dipaths up to dihomotopy. Homology Homotopy Appl. **5**(2), 257–280 (2003)
23. tom Dieck, T.: Transformation groups. In: Studies in Mathematics, vol. 8. de Gruyter, Berlin (1987)
24. van Glabbeek, R.J.: On the expressiveness of higher dimensional automata, express 2004. In: Proceedings of the 11th International Workshop on Expressiveness in Concurrency, pp. 5–34 (2005)
25. Zomorodian, A.: Topology for computing. In: Cambridge Monographs on Applied and Computational Mathematics, vol. 16. Cambridge University Press, Cambridge (2005)

Journal of Homotopy and Related Structures, vol. 2(1), 2007, pp.1–24

REPARAMETRIZATIONS OF CONTINUOUS PATHS

ULRICH FAHRENBERG AND MARTIN RAUSSEN

(*communicated by Ronnie Brown*)

Abstract

A reparametrization (of a continuous path) is given by a surjective weakly increasing self-map of the unit interval. We show that the monoid of reparametrizations (with respect to compositions) can be understood via "stop-maps" that allow to investigate compositions and factorizations, and we compare it to the distributive lattice of countable subsets of the unit interval. The results obtained are used to analyse the space of traces in a topological space, i.e., the space of continuous paths up to reparametrization equivalence. This space is shown to be homeomorphic to the space of regular paths (without stops) up to increasing reparametrizations. Directed versions of the results are important in directed homotopy theory.

1. Introduction and Outline

1.1. Introduction

In elementary *differential geometry*, the most basic objects studied (after points perhaps) are *paths*, i.e., *differentiable* maps $p : I \to \mathbf{R}^n$ defined on the closed interval $I = [0, 1]$. Such a path is called *regular* if $p'(t) \neq \mathbf{0}$ for all $t \in]0, 1[$. A *reparametrization* of the unit interval I is a surjective differentiable map $\varphi : I \to I$ with $\varphi'(t) > 0$ for all $t \in]0, 1[$, i.e. a (strictly increasing) self-diffeomorphism of the unit interval.

Given a path $p : I \to \mathbf{R}^n$ and a reparametrization $\varphi : I \to I$, the paths p and $p \circ \varphi$ represent the same geometric object. In differential geometry one investigates equivalence classes (identifying p with $p \circ \varphi$ for any reparametrization φ) and their invariants, like curvature and torsion.

Motivated by applications in *concurrency theory*, a branch of theoretical Computer Science trying to model and to understand the coordination between many different processors working on a common task, we are interested in *continuous* paths $p : I \to X$ in more general topological spaces up to more general reparametrizations $\varphi : I \to I$. When the *state space* of a concurrent program is viewed as

The authors would like to thank the referee for many useful comments and in particular for pointing out an inaccuracy in the original proof of Theorem 3.6.
Received June 9, 2006, revised March 15, 2007; published on June 30, 2007.
2000 Mathematics Subject Classification: 55,68
Key words and phrases: path, regular path, reparametrization, reparametrization equivalence, trace, stop map, d-space, concurrency

a topological space (typically a cubical complex; cf. [**6**]), "directed" paths in that space respecting certain "monotonicity" properties correspond to *executions*. A nice framework to handle *directed* topological spaces (with an eye to homotopy properties) is the concept of a d-space proposed and investigated by Marco Grandis in [**9**]. Essentially, a topological space comes equipped with a subset of preferred *d-paths* in the set of all paths in X, cf. Definition 4.1. Note in particular, that the reverse of a directed path in general is *not* directed; the slogan is "breaking symmetries".

We do not try to capture the *quantitative* behaviour of executions, corresponding to particular parametrizations of paths, but merely the *qualitative* behaviour, such as the order of shared resources used, or the result of a computation. Hence the objects of study are paths up to certain reparametrizations which

1. do not alter the *image* of a path, and
2. do not alter the *order of events*.

We are thus interested in general paths in topological spaces, up to *surjective* reparametrizations $\varphi : I \to I$ which are *increasing* (and thus continuous!—cf. Lemma 2.7), but not necessarily strictly increasing. Two paths are considered to have the same behaviour if they are *reparametrization equivalent*, cf. Definition 1.2.

To understand this equivalence relation, we have to investigate the space of all reparametrizations which includes strange (e.g. nowhere differentiable) elements. Nevertheless, it enjoys remarkable properties: It is a monoid, in which compositions and factorizations can be completely analysed through an investigation of *stop intervals* and of *stop values*. The quotient space after dividing out the self-homeomorphisms has nice algebraic lattice properties.

A path is called *regular* if it does not "stop"; and we are able to show that the space of general paths *modulo reparametrizations* is homeomorphic to the space of *regular* paths *modulo increasing auto-homeomorphisms* of the interval. Hence to investigate properties of the former, it suffices to consider the latter. This is one of the starting points in the homotopy theoretical and categorical investigation of invariants of d-spaces in [**15**]. Further possible areas of application of the results (and of higher-dimensionsal generalisations still to be investigated) include categorical homotopy theory as in [**8**], categorified gauge theory as in [**1**] and n-transport theory; cf. the blog "The n-category café" at `http://golem.ph.utexas.edu/category`.

This article does not build on any sophisticated machinery. Most of the concepts and proofs can be understood with an undergraduate mathematical background. There are certain parallels to the elementary theory of distribution functions in probability theory, cf. e.g. [**13**]. The flavour is nevertheless different, since continuity (no jumps, i.e., surjectivity) is essential for us. For the sake of completeness, we have chosen to include also elementary results and their proofs (some of which may be well-known).

Marco Grandis has studied piecewise linear reparametrizations in [**10**] for different purposes, but also in the framework of "directed algebraic topology".

1.2. Basic definitions

Let always X denote a Hausdorff topological space and $I = [0, 1]$ the unit interval. The set of all (nondegenerate) closed subintervals of I will be denoted by $\mathfrak{P}_{[\,]}(I) =$

$\{[a, b] \mid 0 \leqslant a < b \leqslant 1\}$. Let $p : I \to X$ denote a continuous map (a path), and remark that the pre-image $p^{-1}(x)$ of any element $x \in X$ is a closed set.

Definition 1.1. 1. An interval $J \in \mathfrak{P}_{[\,]}(I)$ is called a *p-stop interval* if the restriction $p|_J$ is constant and if J is a maximal interval with that property.

2. The set of all p-stop intervals will be denoted as $\Delta_p \subseteq \mathfrak{P}_{[\,]}(I)$. Remark that the intervals in Δ_p are disjoint and that Δ_p carries a natural total order. We let $D_p := \bigcup_{J \in \Delta_p} J \subset I$ denote the *stop set* of p.
3. A path $p : I \to X$ is called *regular* if $\Delta_p = \emptyset$ or if $\Delta_p = \{I\}$ (no stop or constant).
4. A continuous map $\varphi : I \to I$ is called a *reparametrization* if $\varphi(0) = 0, \varphi(1) = 1$ and if φ is *increasing*, i.e. if $s \leqslant t \in I$ implies $\varphi(s) \leqslant \varphi(t)$.

Remark that neither a regular path nor a reparametrization need be injective.

Definition 1.2. Two paths $p, q : I \to X$ are called *reparametrization equivalent* if there exist reparametrizations φ, ψ such that $p \circ \varphi = q \circ \psi$.

We will show later (Corollary 3.3) that reparametrization equivalence is indeed an equivalence relation. As in differential geometry, we are interested in equivalence classes of paths modulo reparametrization equivalence. We call these equivalence classes *traces*[1] in the space X. In particular, we would like to know *whether every trace can be represented by a regular path.* The (positive) answer to this question in Proposition 3.7 is based on a closer look at the space of reparametrizations of the unit interval.

1.3. Outline of the article

Section 2 contains a detailed study of reparametrizations (in their own right) and characterizes their behaviour essentially by an order-preserving bijection between the set of stop intervals and the set of stop values (Definition 1.1 and Proposition 2.13). This pattern analysis allows to study compositions, and in particular, factorizations in the monoid of reparametrizations from an algebraic point of view. In particular, Proposition 2.18 shows that the space of all reparametrizations "up to homeomorphisms" is a distributive lattice isomorphic to the lattice of countable subsets of the unit interval.

Section 3 investigates the *space* of all paths in a Hausdorff space up to reparametrization equivalence. The main result (Theorem 3.6) states that two quotient spaces are in fact homeomorphic: the orbit space arising from the action of the group of all oriented homeomorphisms of the unit interval on the space of *regular* paths (with given end points, cf. Definition 1.1) on the one side, and the space of *all* paths with given end points up to reparametrization equivalence (Definition 1.2); in particular, *every trace can be represented by a regular path.* It might be a bit surprising that the proof makes essential use of the results on factorizations of reparametrizations from Section 2.

---

[1] with a geometric meaning; the notion has nothing to do with algebraic traces.

The final Section 4 deals with spaces of *directed* traces (directed paths up to reparametrization equivalence) on a d-space (cf. Section 1.1 and Definition 4.1). Corollary 4.5 confirms that the result of Theorem 3.6 has an analogue for directed paths in *saturated* (cf. Definition 4.3) d-spaces. This result is one of the starting points for the (categorical) investigations into invariants of directed spaces in [**15**]. Furthermore, it is shown how to relate reparametrization equivalence of directed paths to *thin dihomotopies*; this result is needed in the study [**5**] of directed squares ("two-dimensional paths").

2. Reparametrizations

2.1. Stop and move intervals, stop values, stop maps

The following definitions (extending Definition 1.1) and elementary results will mainly be used for reparametrizations. For the sake of generality, we will state and prove them for general paths $p : I \to X$ in a Hausdorff space X.

Definition 2.1. 1. An element $c \in X$ is called a *p-stop value* if there is a p-stop interval $J \in \Delta_p$ with $p(J) = \{c\}$. We let $C_p \subseteq X$ denote the set of all p-stop values.

2. The map p induces the *p-stop map* $F_p : \Delta_p \to C_p$ with $F_p(J) = c \Leftrightarrow p(J) = \{c\}$.
3. An interval $J \in \mathfrak{P}_{[\,]}(I)$ is called a *p-move interval* if it does not contain any p-stop interval and if it is maximal with that property.
4. The set of all p-move intervals will be denoted $\Gamma_p \subseteq \mathfrak{P}_{[\,]}(I)$, a collection of disjoint closed intervals. We let $O_p := \bigcup_{J\in\Gamma_p} \operatorname{int} J \subseteq I$ denote the *p-move set.*

Lemma 2.2. *For any path $p : I \to X$, the sets of p-stop intervals Δ_p, of p-move intervals Γ_P and of p-stop values C_p are at most countable.*

Proof. The set O_p and $\bigcup_{J\in\Delta_p} \operatorname{int} J$ of *interior points* in move, resp. stop intervals are open subsets of I and thus unions of at most *countably* many maximal open intervals. Their closures constitute Γ_p, resp. Δ_p. The stop value set C_p is at most countable as image of Δ_p under the p-stop map F_p. □

Remark 2.3. This result is similar in spirit to the assertion (relevant for distribution functions in probability theory) that a nondecreasing function to an interval has at most countably many discontinuity points, cf. e.g. [**13**, Sec. 11].

It is important to analyse the *boundary* ∂D_p of the p-stop set: It can be decomposed as $\partial D_p = \partial_1 D_p \cup \partial_2 D_p$ as follows:

- $\partial_1 D_p = \partial D_p \cap D_p$ – the set of all boundary points of intervals in D_p, an at most countable set;
- $\partial_2 D_p = \partial D_p \setminus D_p$ – the set of all (honest) accumulation points of these boundary points. $\partial_2 D_p$ can be uncountable; compare Ex. 2.11.

The move set O_p is the complement $O_p = I \setminus \overline{D_p} \subset I$ of the closure of D_p. It does occur that O_p is empty; compare Ex. 2.11.

The following elementary technical lemma concerning stop sets will be needed in the proof of Proposition 3.7.

Lemma 2.4. *Let $p : I \to X$ denote a path and $U \subseteq X$ an open subspace. Then $p^{-1}(U)$ is a union of (at most) countably many disjoint open intervals, and for any maximal open interval $]a,b[$ in $p^{-1}(U)$, $[a,c] \not\subseteq D_p$ and $[c,b] \not\subseteq D_p$ for every $c \in]a,b[$.*

Proof. As an open subset of I, $p^{-1}(U)$ is a union of (at most) countably many disjoint open intervals. Let $]a,b[$ be one of these, and let $c \in]a,b[$. Then $p(c) \in U, p(a) \notin U, p(b) \notin U$. In particular, p is neither constant on $[a,c]$ nor on $[c,b]$. □

2.2. Spaces of reparametrizations

Within the set of all self-maps of the unit interval I fixing its boundary points, we study the following subsets:

Definition 2.5.
- $\mathrm{Mon}_+(I) := \{\varphi : I \to I \mid \varphi \text{ increasing}, \varphi(0) = 0, \varphi(1) = 1\}$;
- $\mathrm{Rep}_+(I) := \{\varphi \in \mathrm{Mon}_+(I) \mid \varphi \text{ continuous}\}$ – the set of all increasing reparametrizations;
- $\mathrm{Homeo}_+(I) = \{\rho \in \mathrm{Rep}_+(I) \mid \rho \text{ strictly increasing}\}$ – the set of all increasing auto-homeomorphisms of the interval.

Note that $\mathrm{Homeo}_+(I) \subset \mathrm{Rep}_+(I) \subset \mathrm{Mon}_+(I)$. The compact-open topology on the space of *all* continuous maps $C(I,I)$ induces topologies on the latter two spaces $\mathrm{Rep}_+(I)$ and $\mathrm{Homeo}_+(I)$. Composition $\circ$ of maps turns $\mathrm{Mon}_+(I)$ into a monoid, $\mathrm{Rep}_+(I)$ into a topological monoid and $\mathrm{Homeo}_+(I)$ into a topological group (consisting of the units in $\mathrm{Rep}_+(I)$).

All three mapping sets come equipped with a natural partial order: $\varphi \leqslant \psi$ if and only if $\varphi(t) \leqslant \psi(t)$ for all $t \in I$, and they form complete lattices with respect to $\leqslant$. Least upper bounds, resp. greatest lower bounds are given by the max, resp. min of the functions involved:

$$(\varphi \vee \psi)(t) := \max\{\varphi(t), \psi(t)\} \qquad (\varphi \wedge \psi)(t) := \min\{\varphi(t), \psi(t)\}$$

Lemma 2.6.
1. *All three sets $\mathrm{Mon}_+(I), \mathrm{Rep}_+(I), \mathrm{Homeo}_+(I)$ are convex. In particular, the latter two spaces are contractible.*
2. *Any two reparametrizations $\varphi, \psi \in \mathrm{Rep}_+(I)$ are d-homotopic (cf. Definition 4.7 for the general definition), i.e. there exists a reparametrization $\varphi, \psi \leqslant \eta \in \mathrm{Rep}_+(I)$ and increasing paths $G, H : \vec{I} \to \mathrm{Rep}_+(I)$ with $G(0) = \varphi, H(0) = \psi, G(1) = H(1) = \eta$.*

Proof.
1. The sets are closed under convex combinations $(1-s)\varphi + s\psi$.
2. For $\eta = \varphi \vee \psi$, define $G(s) = (1-s)\varphi + s\eta$ and $H(s) = (1-s)\psi + s\eta$.

□

A characterization of the elements of $\mathrm{Mon}_+(I)$, $\mathrm{Rep}_+(I)$, and $\mathrm{Homeo}_+(I)$ is achieved in the elementary

Lemma 2.7. *Let $\varphi \in \mathrm{Mon}_+(I)$.*

1. *For every interval $J \subseteq I$, the pre-image $\varphi^{-1}(J) \subseteq I$ is an interval, as well. In particular, $\varphi^{-1}(a)$ is an interval (possibly degenerate) for every $a \in I$.*
2. *$\varphi \in \mathrm{Rep}_+(I)$ if and only if φ is* surjective.
3. *$\varphi \in \mathrm{Homeo}_+(I)$ if and only if φ is* bijective.

Proof. The only non-obvious statement is that surjectivity of $\varphi \in \mathrm{Mon}_+(I)$ implies continuity; we show that the pre-image $\varphi^{-1}(J)$ of an open interval $J \subset I$ is open:

Let $d \in \varphi^{-1}(J)$ and $\varphi(d) = c \in J$. Then there exist $\varepsilon > 0$ such that $[c-\varepsilon, c+\varepsilon] \subseteq J$ and $d_1, d_2 \in I$ such that $\varphi(d_1) = c - \varepsilon, \varphi(d_2) = c + \varepsilon$. Monotonicity implies: $\varphi([d_1, d_2]) \subseteq [c - \varepsilon, c + \varepsilon]$ and $d' \notin [d_1, d_2] \Rightarrow \varphi(d') \notin]c - \varepsilon, c + \varepsilon[$. Surjectivity implies: $\varphi([d_1, d_2]) = [c - \varepsilon, c + \varepsilon] \subseteq J$; hence d has an open neighbourhood in $\varphi^{-1}(J)$. □

The following information about images of intervals under reparametrizations is needed in the proof of Proposition 3.7:

Lemma 2.8. *Let $a, b \in I$ and $\varphi \in \mathrm{Rep}_+(I)$. Then*

1. *$\varphi([a, b]) = [\varphi(a), \varphi(b)]$.*
2. *$]\varphi(a), \varphi(b)[\subseteq \varphi(]a, b[) \subseteq [\varphi(a), \varphi(b)]$.*
3. *$\varphi(]a, b[) \neq]\varphi(a), \varphi(b)[$ if and only if there is $c \in]a, b[$ such that $[a, c] \subseteq D_\varphi$ or $[c, b] \subseteq D_\varphi$.*

Proof. Only the last assertion requires proof. If $[a, c] \subseteq D_\varphi$ for some $c \in]a, b[$, then $\varphi(a) = \varphi(c) \in \varphi(]a, b[)$; similarly, if $[c, b] \subseteq D_\varphi$, then $\varphi(b) \in \varphi(]a, b[)$. For the reverse direction, assume $\varphi(a) \in \varphi(]a, b[)$, and let $c \in]a, b[$ such that $\varphi(a) = \varphi(c)$. Then $\varphi(a) \leqslant \varphi(t) \leqslant \varphi(c) = \varphi(a)$ for any $t \in [a, c]$, hence $[a, c] \subseteq D_\varphi$. The other implication is similar. □

The following result deals with the relative size of the homeomorphisms within the reparametrizations. It will be needed in the proof of the main result in Section 3.

Lemma 2.9. *In the topology induced from the compact-open topology, both* $\mathrm{Homeo}_+(I)$ *and its complement are dense in* $\mathrm{Rep}_+(I)$.

Proof. The compact-open topology is induced by the supremum metric on the space $C(I, I)$ of all self-maps of the interval. Hence, for a given $\varphi \in \mathrm{Rep}_+(I)$ and $n \in \mathbf{N}$, we need to construct $\rho \in \mathrm{Homeo}_+(I)$ such that $\|\varphi - \rho\| \leqslant \frac{1}{n}$: Choose $c_k, 0 \leqslant k \leqslant n$, such that $c_0 = 0, c_n = 1$ and $\varphi(c_k) = \frac{k}{n}$; clearly c_k is strictly increasing with k. Hence the piecewise linear map ρ given by $\rho(c_k) = \frac{k}{n}$ is contained in $\mathrm{Homeo}_+(I)$. Furthermore, for $x \in [c_k, c_{k+1}], k < n$, we have $\frac{k}{n} \leqslant \rho(x), \varphi(x) \leqslant \frac{k+1}{n}$, and thus $\|\varphi - \rho\| \leqslant \frac{1}{n}$.

For the same $\varphi \in \mathrm{Rep}_+(I)$ and the same definition for c_0 and c_1 as above, let $\psi \in \mathrm{Rep}_+(I) \setminus \mathrm{Homeo}_+(I)$ be given by $\psi(x) = \varphi(x), x \geqslant c_1$; on the interval $[0, c_1]$, we let ψ be the piecewise linear map with $\psi(0) = \psi(\frac{c_1}{2}) = 0$ and $\psi(c_1) = \varphi(c_1)$. See also Figure 1. □

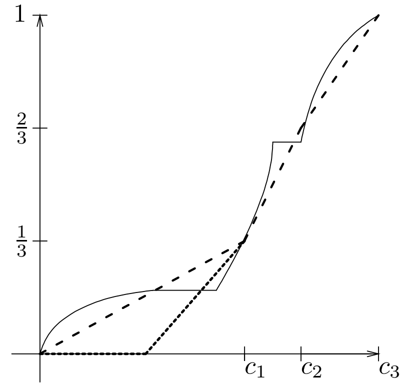


Figure 1: Reparametrization φ (full line), homeomorphism ρ (broken line), and reparametrization ψ (dotted line).

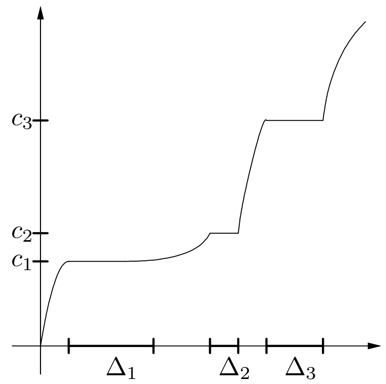


Figure 2: Stop intervals and stop values

2.3. Classification of reparametrizations

In the following, we are mainly interested in an investigation of the algebraic monoid structure on $\text{Rep}_+(I)$ induced by *composition* $\circ$ of maps. Note that there is another structure on the sets (spaces) $\text{Mon}_+(I)$, $\text{Rep}_+(I)$, and $\text{Homeo}_+(I)$, induced by *concatenation* of paths

$$(\varphi, \psi) \mapsto \varphi * \psi; \quad (\varphi * \psi)(t) = \begin{cases} \varphi(2t) & \text{for } t \leqslant \frac{1}{2} \\ \psi(2t-1) & \text{for } t > \frac{1}{2} \end{cases} \tag{2.1}$$

This composition does not induce a monoidal structure on these sets, as concatenation is not associative and does not have units "on the nose".

We wish to describe a reparametrization $\varphi \in \text{Rep}_+(I)$ by its φ-stop map $F_\varphi : \Delta_\varphi \to C_\varphi$ illustrated in Figure 2 and by the restriction of φ to its φ-move set $O_\varphi \subseteq I$; cf. Definition 2.1.

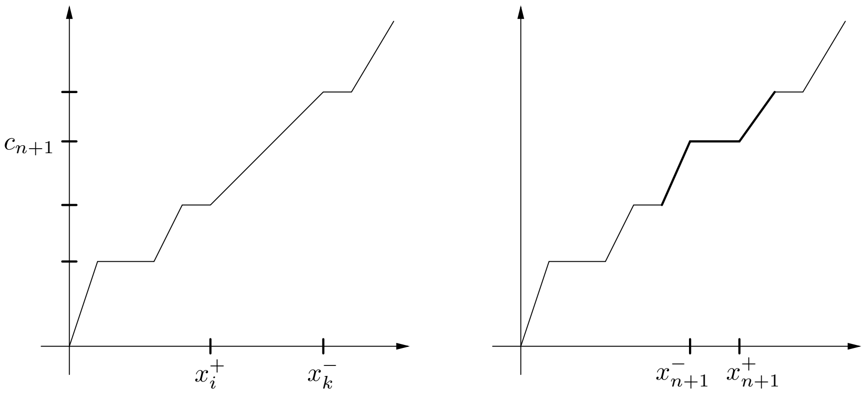


Figure 3: Inserting the stop value c_{n+1}

2.3.1. All countable sets in the interval are stop value sets

Lemma 2.2 tells us that the φ-stop value set $C_\varphi \subset I$ of a reparametrization φ is an at most countable ordered subset of I. Remark also that automorphisms $\rho \in \mathrm{Homeo}_+(I)$ are characterized by the properties $\Delta_\rho = C_\rho = \emptyset$, resp. $O_\rho = I$.

Which (countable) subsets of the unit interval can be realized as φ-stop sets of some $\varphi \in \mathrm{Rep}_+(I)$? It is easy to construct (piecewise linear) reparametrizations with a *finite* set of stop values. Rather surprisingly, this construction can be extended to arbitrary (at most) *countable* sets of stop values:

Lemma 2.10. *For every countable set $C \subset I$, there is a reparametrization $\varphi \in \mathrm{Rep}_+(I)$ with $C_\varphi = C$.*

Proof. Let $C = \{c_1, c_2, \dots\} \subset I$ denote an injective enumeration of the countable set C. We shall first construct a uniformly convergent sequence of piecewise linear maps $\varphi_n \in \mathrm{Rep}_+(I), n \geqslant 0$ with $C(\varphi_n) = \{c_1, \dots, c_n\}$ and thus $\Delta_{\varphi_n} = \{\varphi_n^{-1}(c_i) | 1 \leqslant i \leqslant n\}$. Let $[x_i^-, x_i^+] = \varphi_n^{-1}(c_i), 1 \leqslant i$; moreover, $x_0^- = 0, x_0^+ = 1$.

We start with $\varphi_0 = \mathrm{id}_I$. Inductively, assume φ_n given as above. Among the $x_j^\pm, 1 \leqslant j \leqslant n$, choose x_i^+, x_k^- such that $c_{n+1} \in \varphi(]x_i^+, x_k^-[)$ and such that the restriction of φ_n on that interval is strictly increasing (and linear). The map φ_{n+1} will differ from φ_n only on (the interior of) that subinterval $[c_i, c_k]$. The linear map on that interval is replaced by a piecewise linear map, which comes in three pieces. The middle one takes the constant value c_{n+1} on a subinterval $[x_{n+1}^-, x_{n+1}^+]$. On the left and right subinterval, we connect linearly to the values c_i, c_k on the boundaries. See also Figure 3.

The interval $[x_{n+1}^-, x_{n+1}^+]$ is chosen so small that $||\varphi_{n+1} - \varphi_n||_\infty < \frac{1}{2^n}$ ensuring uniform convergence of the maps φ_n to a continuous map $\varphi \in \mathrm{Rep}_+(I)$. For this map φ, we have $\Delta_\varphi = \{[x_n^-, x_n^+] | n > 0\}$ and $C_\varphi = C$. □

Example 2.11. If one chooses a *dense* countable subset $C \subset I$, e.g., $C = \mathbf{Q} \cap I$, then φ cannot be injective on *any* non-trivial interval; hence $O_\varphi = \emptyset$ and $\overline{D_\varphi} = I$.

The (uncountable) complement $I \setminus D_\varphi = \partial_2 D_\varphi$ does not contain any non-trivial interval.

2.3.2. Classification

What are the essential data to describe a reparametrization in terms of stop maps and move sets?

Proposition 2.12. *Let $\varphi \in \mathrm{Rep}_+(I)$ denote a reparametrization.*

1. *The reparametrization φ induces an* order-preserving bijection $F_\varphi : \Delta_\varphi \to C_\varphi$. *The restriction $\varphi|_J : J \to \varphi(J)$ to every move interval $J \in \Gamma_\varphi$ (Definition 2.1) is an (increasing) homeomorphism onto its image.*
2. *The restriction $\varphi|_{\overline{D_\varphi}} : \overline{D_\varphi} \to \overline{C_\varphi}$ of φ to $\overline{D_\varphi}$ is* onto.
3. *Two reparametrizations $\varphi, \psi \in \mathrm{Rep}_+(I)$ with $\Delta_\varphi = \Delta_\psi, C_\varphi = C_\psi, F_\varphi = F_\psi$ agree on $\overline{D_\varphi}$; if, moreover, $\varphi|_{O_\varphi} = \psi|_{O_\psi}$, then φ and ψ agree on all of I.*

Proof.

1. The first statement is obvious from the definitions. For the second, note that $J \cap D_\varphi = \partial J$ (or empty) for every such interval J.
2. Every element $b \in \overline{C_\varphi}$ is the limit of a monotone sequence of elements in C_φ which is the image of a monotone and bounded sequence of elements in D_φ; the limit of such a sequence exists and maps to b under φ.
3. By definition, φ and ψ agree on D_φ; by continuity, they have to agree on its closure $\overline{D_\varphi}$, as well. The last statement is obvious.

□

A reparametrization $\varphi \in \mathrm{Rep}_+(I)$ is thus *uniquely* characterized by its stop map $F_\varphi : \Delta_\varphi \to C_\varphi$ and by a (fitting) collection of homeomorphisms $\varphi|_{\overline{J}} : \overline{J} \to \varphi(\overline{J}),\ J \in \Gamma_\varphi$. Now we ask which conditions an "abstract" stop map has to satisfy in order to arise from a genuine reparametrization. We start with the following data:

- $\Delta \subseteq \mathfrak{P}_{[\,]}(I)$ denotes an (at most) countable subset of *disjoint closed* intervals – with a natural total order.
- $C \subseteq I$ denotes a subset with the same cardinality as Δ.
- $F : \Delta \to C$ denotes an order-preserving bijection.

Let $\Delta_-, \Delta_+ \subseteq I$ denote the set of lower, resp. upper boundaries of intervals in Δ. Define $D := \bigcup_{J\in\Delta} \Delta \subset I$.

Let $O = I \setminus \overline{D}$. Since O is open, it is a disjoint union $O = \bigcup_{J\in\Gamma} J$ of maximal open intervals indexed by an (at most) countable set Γ – possibly empty.

For every map $G : \Delta \to C$ we define a map $\varphi_G : D \to C$ by $\varphi_G(t) = G(J) \Leftrightarrow t \in J$. If F is order-preserving, then φ_F is increasing. Moreover:

Proposition 2.13. 1. *A reparametrization $\varphi \in \mathrm{Rep}_+(I)$ satisfies the following for every pair of strictly monotonely converging sequences $x_n \uparrow x,\ y_n \downarrow y$ for* *which $x, x_n \in (\Delta_\varphi)_+$ and $y, y_n \in (\Delta_\varphi)_-$:*

$$x = y \quad \Rightarrow \quad \lim \varphi(x_n) = \lim \varphi(y_n), \tag{2.2}$$
$$x < y \quad \Rightarrow \quad \lim \varphi(x_n) < \lim \varphi(y_n), \tag{2.3}$$
$$x = 1 \quad \Rightarrow \quad \lim \varphi(x_n) = 1, \tag{2.4}$$
$$y = 0 \quad \Rightarrow \quad \lim \varphi(y_n) = 0, \tag{2.5}$$

2. *For every order preserving bijection $F : \Delta \to C$ with φ_F satisfying* (1)–(4) *above for every pair of strictly monotonely converging sequences $x_n \uparrow x$, $y_n \downarrow y$ for which $x, x_n \in (\Delta_\varphi)_+$ and $y, y_n \in (\Delta_\varphi)_-$, there exists a reparametrization $\psi \in \mathrm{Rep}_+(I)$ with $\Delta_\psi = \Delta, C_\psi = C$ and $F_\psi = F$. The set of all such reparametrizations is in one-to-one correspondence with $\prod_\Gamma \mathrm{Homeo}_+(I)$.*

Proof. 1. By continuity, $\lim \varphi(x_n) = \varphi(x)$ and $\lim \varphi(y_n) = \varphi(y)$; this settles all but (2.3). Suppose $\varphi(x) = \varphi(y)$ in (2). Then $[x, y]$ is contained in a stop interval, hence $x \notin \Delta_+$ and $y \notin \Delta_-$.

2. First, we extend φ_F to $\overline{D}$: there is a *unique continuous* (and increasing!) extension of φ_F from D to $\overline{D}$: $\lim_{x_n \uparrow x} \varphi_F(x_n)$ exists and is independent of the sequence x_n by monotonicity and agrees with $\lim_{y_n \downarrow x} \varphi_F(y_n)$ by condition (2.2) of Proposition 2.13. Moreover, we let $\varphi_F(0) = 0$ and $\varphi_F(1) = 1$, in accordance with (3) and (4) above.
Let $J =]a_-^J, a_+^J[\in \Gamma$ denote a maximal open interval. Its boundary points a_-^J, a_+^J are contained in ∂D unless possibly if $a_-^J = 0$ and/or $a_+^J = 1$, in which case we are covered by (2.4) and/or (2.5) above. In conclusion, φ_F is defined on ∂J. Moreover, $\varphi_F(a_-^J) < \varphi_F(a_+^J)$, since F is order preserving and injective and because of condition (2.3) above.
Hence, every collection of strictly increasing homeomorphisms between $[a_-^J, a_+^J]$ and $[\varphi_F(a_-^J), \varphi_F(a_+^J)]$ – preserving endpoints – extends φ_F to a continuous increasing map $\psi : I \to I$ with $\Delta_\psi = \Delta, C_\psi = C$ and $F_\psi = F$. The set of all collections of such homeomorphisms is easily seen to be in one-to-one correspondence with $\prod_\Gamma \mathrm{Homeo}_+(I)$.

□

2.4. Compositions and Factorizations

We shall now investigate the behaviour of $\mathrm{Rep}_+(I)$ under composition and factorization in view of the description and classification from Proposition 2.13 above. We need to introduce the following notation: For a (continuous) map $\psi : I \to I$, let $\psi_*, (\psi^{-1})^* : \mathfrak{P}_{[]}(I) \to \mathfrak{P}_{[]}(I)$ denote the maps induced on subintervals: $\psi_*(J) = \psi(J) \in \mathfrak{P}_{[]}(I), (\psi^{-1})^*(J) = \psi^{-1}(J)$.

2.4.1. Composition of reparametrizations

The results below follow easily from the definitions of stop-intervals, stop-values and stop-maps:

Lemma 2.14. *Let $\varphi, \psi \in \mathrm{Rep}_+(I)$ denote reparametrizations with associated stop maps $F_\varphi : \Delta_\varphi \to C_\varphi, F_\psi : \Delta_\psi \to C_\psi$. Then*

1. $\Delta_{\varphi\circ\psi} = \{J \in \Delta_\psi \mid F_\psi(J) \notin D_\varphi\} \cup (\psi^{-1})^*(\Delta_\varphi)$,
2. $C_{\varphi\circ\psi} = \varphi(C_\psi) \cup C_\varphi$,
3. $F_{\varphi\circ\psi} : \Delta_{\varphi\circ\psi} \to C_{\varphi\circ\psi}$ *is given by* $F_{\varphi\circ\psi}(J) = \begin{cases} \varphi(F_\psi(J)), & J \in \Delta_\psi \\ F_\varphi(\psi_*(J)), & J \in (\psi^{-1})^*(\Delta_\varphi). \end{cases}$

□

Corollary 2.15. *Let* $\varphi, \psi \in \mathrm{Rep}_+(I)$ *as in Lemma 2.14. If* $\psi \in \mathrm{Homeo}_+(I)$, *resp.* $\varphi \in \mathrm{Homeo}_+(I)$, *then*

1. $\Delta_{\varphi\circ\psi} = (\psi^{-1})^*(\Delta_\varphi)$, *resp.* $\Delta_{\varphi\circ\psi} = \Delta_\psi$,
2. $C_{\varphi\circ\psi} = C_\varphi$, *resp.* $C_{\varphi\circ\psi} = \varphi(C_\psi)$,
3. $F_{\varphi\circ\psi} = F_\varphi \circ \psi_* : \psi_*^{-1}(\Delta_\varphi) \to C_\varphi$, *resp.* $F_{\varphi\circ\psi} = \varphi \circ F_\psi : \Delta_\psi \to \varphi(C_\psi)$.

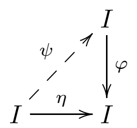
□

2.4.2. Factorizations of reparametrizations

Also factorizations can be studied effectively using stop-data of the reparametrizations involved. It turns out that the following result on factorizations on the *right* will be an essential tool in Section 3:

Proposition 2.16. *Let* $\eta, \varphi \in \mathrm{Rep}_+(I)$ *denote reparametrizations.*

1. *There exists a lift* $\psi \in \mathrm{Rep}_+(I)$ *in the diagram*



(2.6)

   if and only if $C_\varphi \subseteq C_\eta$.
2. *If* $C_\varphi \subseteq C_\eta$ *and* C *is any (at most) countable set with* $\varphi^{-1}(C_\eta \setminus C_\varphi) \subseteq C \subseteq \varphi^{-1}(C_\eta \setminus C_\varphi) \cup D_\varphi$, *then there exists such a lift* $\psi \in \mathrm{Rep}_+(I)$ *with* $C_\psi = C$. *In particular, if* $C_\varphi = C_\eta$, *there exist a lift* $\psi \in \mathrm{Homeo}_+(I)$.
3. *Assume* $C_\varphi \subseteq C_\eta$ *and let* $\Delta_1 := F_\eta^{-1}(C_\varphi) \subseteq \Delta_\eta$. *Then the space of all lifts* $\{\psi \mid \eta = \varphi\circ\psi\} \subseteq \mathrm{Rep}_+(I)$ *is in one-to-one-correspondence with* $\prod_{L\in\Delta_1} \mathrm{Rep}_+(L) = \prod_{\Delta_1} \mathrm{Rep}_+(I)$.

Proof. The "only if" part of 1. follows immediately from Lemma 2.14.2. For the "if" part, we analyse first the set-theoretic requirements to a lift ψ on relevant subintervals. To this end, decompose $\Delta_\eta =: \Delta_1 \sqcup \Delta_2$ with $\Delta_1 := F_\eta^{-1}(C_\varphi)$ and $\Delta_2 = F_\eta^{-1}(C_\eta \setminus C_\varphi)$, and $D_\eta = D_1 \sqcup D_2$ with $D_1 := \bigcup_{J\in\Delta_1}$ and $D_2 = \bigcup_{J\in\Delta_2}$. We construct a lift $\psi : I = D_1 \cup D_2 \cup O_\eta \to I$ by considering each of these three subsets of I:

Remark that Δ_2 necessarily has to be a subset of Δ_ψ and that for $J \in \Delta_2$, $F_\psi(J)$ has to be the unique element of $\varphi^{-1}(F_\eta(J))$.

On any move interval $K \in \Gamma_\eta$ (cf. Definition 2.1.4), the restriction $\eta\,|_K : K \to \eta(K)$ is an increasing homeomorphism; in particular, $\eta(K) \cap C_\eta$ consists at most of the two boundary points. Hence, the restriction $\varphi\,|_{\varphi^{-1}\eta(K)} : \varphi^{-1}\eta(K) \to \eta(K)$

is also an increasing homeomorphism, since $\eta(K)\cap C_\varphi \subseteq \eta(K)\cap C_\eta$ again consists at most of the two boundary points. The restriction of ψ to K has to be defined as $\psi\,|_K = (\varphi\,|_{\varphi^{-1}\eta K})^{-1}\circ \eta\,|_K$; it is onto $\varphi^{-1}\eta K$.

On any interval $L \in \Delta_1$, the restriction of ψ to L can be defined as *any* increasing continuous map $\psi\,|_L : L \to F_\varphi^{-1}(F_\eta(L)) \in \Delta_\varphi$ respecting the boundary points.

The map $\psi : \vec{I} \to \vec{I}$ thus defined altogether is by definition a lift, it is increasing and surjective. Lemma 2.7.2 settles 1.

The only freedom in the construction of a lift ψ is the choice of increasing continuous maps on the intervals $L \in \Delta_1$ (with given end points). As in Lemma 2.10, we can construct the set of stop values (on D_1) to be any countable subset of D_φ. To these one has of course to add the pre-images of the stop values in $C_\eta \setminus C_\varphi$. This settles 2. and 3. □

We will also make use of the following result on factorizations on the *left*.

Proposition 2.17. *Let $\eta, \varphi \in \mathrm{Rep}_+(I)$ denote reparametrizations.*

1. *There exists a factorization with $\psi \in \mathrm{Rep}_+(I)$ in the diagram*

$$\begin{array}{ccc} I & \xrightarrow{\ \eta\ } & I \\ {\scriptstyle\varphi}\downarrow & \nearrow_{\psi} & \\ I & & \end{array} \tag{2.7}$$

   if and only if there exists a map $i_{\varphi\eta} : \Delta_\varphi \to \Delta_\eta$ such that $J \subseteq i_{\varphi\eta}(J)$ for every $J \in \Delta_\varphi$. (Δ_φ is a refinement *of Δ_η).*

2. *If it exists, the factor $\psi \in \mathrm{Rep}_+(I)$ is uniquely determined and satisfies*
   - $C_\psi = C_\eta \setminus \{F_\eta(J) \mid J \in \Delta_\eta \cap \Delta_\varphi\}$,
   - $\Delta_\psi = \{\varphi(K) \mid K \in \Delta_\eta \setminus \Delta_\varphi\}$,
   - *if $K \in \Delta_\psi$, then $F_\psi(K)$ is the unique element of $\eta(\varphi^{-1}(K))$, $K \in \Delta_\psi$.*

Proof. A lift ψ as in (2.7) has to satisfy $\psi(x) := \eta(\varphi^{-1}(x))$. It is well-defined (and then unique and increasing) if and only if the condition of Proposition 2.17 is satisfied. Since η is onto, ψ is onto as well, and thus continuous by Lemma 2.7.2. The description of the invariants of ψ follows by inspection. □

2.5. The algebra of reparametrizations up to homeomorphisms

Consider the group action $\mathrm{Rep}_+(I)\times \mathrm{Homeo}_+(I) \to \mathrm{Rep}_+(I)$ given by composition on the right. An element in the quotient space $\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$ preserves the set of stop values, whereas the exact distribution of stop intervals over the interval is factored out. Using the factorization tools from Section 2.4 above, this intuition will be made more formal in Propositions 2.18 and 2.22 below.

Consider the preorder on $\mathrm{Rep}_+(I)$ (different from the one considered in Section 2.2) given by $\varphi \leqslant \psi \Leftrightarrow \exists\ \eta \in \mathrm{Rep}_+(I) : \psi = \varphi\circ\eta$ ($\Leftrightarrow C_\varphi \subseteq C_\psi$ by Proposition 2.16.1). This preorder factors to yield a partial order on the quotient $\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$ since $\psi = \varphi\circ\eta = (\varphi\circ\rho)\circ(\rho^{-1}\circ\eta)$ for $\rho \in \mathrm{Homeo}_+(I)$.

Moreover, let us consider the set $\mathfrak{P}_c(I)$ of *countable* subsets of I with the partial order given by inclusion.

Proposition 2.18. *The map* $C : \mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)} \to \mathfrak{P}_c(I)$ *given by* $C(\varphi) = C_\varphi$ *is an order-preserving bijection.*

Proof. By Corollary 2.15.2, the map C is well-defined; by Lemma 2.14, it is order-preserving, and by Lemma 2.10, it is surjective. Given two reparametrizations with the same set of stop values, Proposition 2.16 shows that one can construct a lift ψ from one into the other that is a homeomorphism ($C_\psi = \emptyset$); as a consequence, C is also injective. □

Proposition 2.19. *For every* $\varphi_1, \varphi_2 \in \mathrm{Rep}_+(I)$*, there exist* $\psi_1, \psi_2 \in \mathrm{Rep}_+(I)$ *completing the diagram*

$$\begin{array}{ccc} I & \overset{\psi_1}{\dashrightarrow} & I \\ {\scriptstyle \psi_2}\,\vdots & & \downarrow{\scriptstyle \varphi_1} \\ I & \underset{\varphi_2}{\longrightarrow} & I \end{array}$$

with $C_{\varphi_1 \circ \psi_1} = C_{\varphi_2 \circ \psi_2} = C_{\varphi_1} \cup C_{\varphi_2}$.

Proof. Using Lemma 2.10, construct $\psi_1 \in \mathrm{Rep}_+(I)$ with $C_{\psi_1} = \varphi_1^{-1}(C_{\varphi_2} \setminus C_{\varphi_1})$ and hence $C_{\varphi_1 \circ \psi_1} = C_{\varphi_1} \cup C_{\varphi_2}$ (cf. Lemma 2.14.2). Using Proposition 2.16, construct a lift ψ_2 in the diagram

$$\begin{array}{ccc} & & I \\ & {\scriptstyle \psi_2}\nearrow & \downarrow{\scriptstyle \varphi_2} \\ I & \underset{\varphi_1 \circ \psi_1}{\longrightarrow} & I \end{array}$$

with $C_{\psi_2} = \varphi_2^{-1}(C_{\varphi_1} \setminus C_{\varphi_2})$, i.e., without introducing superfluous extra stop values. □

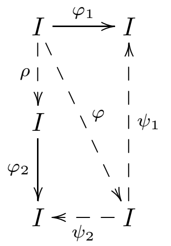
Remark 2.20. In general, it is not possible to complete the dual diagram

$$\begin{array}{ccc} I & \overset{\varphi_1}{\longrightarrow} & I \\ {\scriptstyle \varphi_2}\downarrow & & \vdots\,{\scriptstyle \psi_1} \\ I & \underset{\psi_2}{\dashrightarrow} & I \end{array}$$

since $D_{\psi_1 \circ \varphi_1} = D_{\psi_2 \circ \varphi_2} \supseteq D_{\varphi_1} \cup D_{\varphi_2}$. The latter set might be the entire interval I which is impossible for a reparametrization.

Is there a natural way to construct from two reparametrizations a third one (a common factor) with a set of stop values that is just the *intersection* of the sets of stop values of the given ones? In order to have the stop intervals appear in proper order, it is necessary to modify one of the reparametrizations by a homeomorphism first (which is not a problem if one works in the quotient $\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$!)

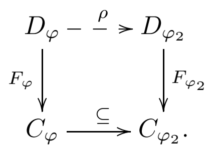
Proposition 2.21. *For every $\varphi_1, \varphi_2 \in \mathrm{Rep}_+(I)$, there exist $\rho \in \mathrm{Homeo}_+(I)$, $\psi_1, \psi_2, \varphi \in \mathrm{Rep}_+(I)$ completing the diagram*



and such that $C_\varphi = C_{\varphi_1} \cap C_{\varphi_2}$.

Proof. Define $\Delta_\varphi := F_{\varphi_1}^{-1}(C_{\varphi_1} \cap C_{\varphi_2}) \subseteq \Delta_{\varphi_1}, C_\varphi := C_{\varphi_1} \cap C_{\varphi_2}$ and define $F_\varphi : \Delta_\varphi \to C_\varphi$ as the restriction of F_{φ_1}. By Proposition 2.13, there exists $\varphi \in \mathrm{Rep}_+(I)$ with F_φ as its stop map. By Proposition 2.17, there exists a lift $\psi_1 \in \mathrm{Rep}_+(I)$ in the right triangle of the diagram above.

In general, the reparametrization φ constructed above does not factor over φ_2 immediately, cf. Proposition 2.17. We need a "correction" homeomorphism $\rho \in \mathrm{Homeo}_+(I)$ whose restriction to D_φ fits into



On intervals $J \in \Delta_\varphi$, ρ can be chosen as the increasing linear map sending J onto $F_{\varphi_2}^{-1}(F_\varphi(J))$ and then extended from D_φ to I as a homeomorphism as in the proof of Proposition 2.13. The condition of Proposition 2.17 is now satisfied to guarantee a lift of φ over $\varphi_2 \circ \rho$ in the left triangle of the diagram above. $\square$

Using the bijection from Proposition 2.18, one may introduce binary operations on the quotient $\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$ in a purely algebraic manner, i.e., one may pull back the operations given by set union and intersection on $\mathfrak{P}_c(I)$. The results of this section allow us to give these operations an intrinsic meaning in terms of reparametrizations. Using the notation from Proposition 2.19, an operation $\vee$ ("least common multiple") is defined by $[\varphi_1] \vee [\varphi_2] := [\varphi_1 \circ \psi_1] = [\varphi_2 \circ \psi_2]$. Likewise, $[\varphi_1] \wedge [\varphi_2]$ can be represented by the reparametrization φ from Proposition 2.21. Altogether we obtain:

Proposition 2.22. *The operations*

$$\vee, \wedge : \mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)} \times \mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)} \to \mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$$

turn $\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}$ into a distributive lattice with the class represented by $\mathrm{Homeo}_+(I)$ as a global minimum. The map

$$C : (\mathrm{Rep}_+(I)/_{\mathrm{Homeo}_+(I)}, \vee, \wedge) \to (\mathfrak{P}_c(I), \cup, \cap)$$

from Proposition 2.18 is then an isomorphism of distributive lattices. $\square$

3. Spaces of traces

3.1. The compact-open topology on path spaces

We now turn to spaces of equivalence classes of paths in a Hausdorff space X. To start with, we give a handy characterization of the compact-open topology of the space of all paths $P(X) = X^I$ in X. By definition, is has as a basis the sets $P(X)(\mathcal{C}, \mathcal{V}) := \{p \in P(X) \mid p(C_i) \subset V_i\}$, indexed by collections $\mathcal{C} = \{C_1, \ldots, C_n\}$ of compact subsets $C_i \subseteq I$, resp. $\mathcal{V} = \{V_1, \ldots, V_n\}$ of open subsets $V_i \subseteq X$, $n \in \mathbf{N}$.

A partition $A = \{0 = a_0 < a_1 < \cdots < a_{n-1} < a_n = 1\}$ of the unit interval I gives rise to the collection $\mathcal{A} = \{[a_{j-1}, a_j]\}$ of closed intervals. We write $P(X)(A, \mathcal{U}) = P(X)(\mathcal{A}, \mathcal{U}) \subseteq P(X)$.

Lemma 3.1. *The sets $P(X)(A, \mathcal{U}) \subset P(X)$ form a basis for the compact-open topology of $P(X)$.*

Proof. Compare [**4**, Ch. XII] for the first part of this proof. Every path q in a basis set $P(X)(\mathcal{C}, \mathcal{V})$ satisfies $C_i \subset q^{-1}(V_i)$; hence C_i is covered by the connected components of $q^{-1}(V_i)$, a set of open disjoint intervals. Finitely many of those, say the intervals J_{ij}, cover C_i; for each of those let $I_{ij} = [\min(J_{ij} \cap C_i), \max(J_{ij} \cap C_i)]$. Then $I_{ij} \subset J_{ij} \subseteq q^{-1}(V_i)$ whence $q(I_{ij}) \subseteq V_i$; moreover, $C_i \subseteq \bigcup_j I_{ij}$.

Let $A' = \bigcup_{i,j} \partial I_{ij} \subset I$ denote the finite subset of interval boundary points. Then, for two successive points $a, b \in A'$ we have: $q([a, b]) \subseteq \bigcap_{[a,b] \subseteq I_{ij}} V_i$; this intersection is to be interpreted as the set X if the index set is empty. We slightly alter the partition A' and arrive at a new partition A if one or several boundary points are both the upper and the lower boundary of some of the intervals I_{ij}: If a is such an upper and a lower boundary of some of the intervals I_{ij} and hence $q(a) \in \bigcap_{a \in I_{ij}} V_i$, we replace a with a pair a_-, a_+ satisfying $a' < a_- < a < a_+ < a''$ for all $a' < a < a''$ in A' and such that $q([a_-, a_+]) \subseteq \bigcap_{a \in I_{ij}} V_i$.

Let $0 = a_0 < a_1 < \cdots < a_{n-1} < a_n = 1$ denote the elements of A in the induced order, and let $\mathcal{U}$ denote the collection of open sets $U_j := \bigcap_{q([a_{j-1}, a_j]) \subseteq V_i} V_i$. The open set $P(X)(\mathcal{C}, \mathcal{V})$ from above then contains the open neighbourhood $P(X)(A, \mathcal{U})$ of q. □

3.2. Regular traces versus traces

In this section we compare several spaces of paths in a Hausdorff space X *up to reparametrization.* Extending Definition 1.1, we get

Definition 3.2. 1. A path $p : I \to X$ in a topological space X is said to be *regular* if $\Delta_p = \emptyset$ or if $\Delta_p = \{I\}$.

2. The set of regular paths in X is denoted $R(X)$ and regarded as a subspace of $P(X) = X^I$.
3. The spaces of (regular) paths p starting in $x \in X$ and ending in $y \in X$ ($p(0) = x$, $p(1) = y$) are denoted by $R(X)(x, y) \subset P(X)(x, y)$; they are equipped with the induced topologies.

Composition on the right yields a group action of the topological group $\mathrm{Homeo}_+(I)$ on $R(X)$ and a monoid action of the topological monoid $\mathrm{Rep}_+(I)$ on $P(X)$. These actions respect the decompositions in subspaces $R(X)(x, y)$, resp. $P(X)(x, y)$.

In Definition 1.2, we called paths $p, q \in \mathrm{Rep}_+(I)$ reparametrization equivalent if there exist reparametrizations $\varphi, \psi \in \mathrm{Rep}_+(I)$ such that $p \circ \varphi = q \circ \psi$.

Corollary 3.3. *1. Reparametrization equivalence of paths is an equivalence relation.*

2. Two reparametrization equivalent paths are thinly homotopic.

For a definition of thin homotopy see [**11**]; essentially, a homotopy $H : I \times I \to X$ fixing the endpoints is thin if it factors through a *tree* (the geometric realisation of an acyclic one-dimensional simplicial set), i.e. if $H : I \times I \to J \to X$ for a tree J. Remark that reparametrization equivalent paths have the same image: $p(I) = q(I) \subseteq X$. This is not necessarily true for thinly homotopic paths; e.g., the cancellation homotopy [**16**, p. 48] between the concatenation of a path with its inverse and the constant path is thin.

Proof. 1. Reparametrization equivalence is clearly a reflexive and symmetric relation. For transitivity, let $p, q, r \in P(X)$ denote three paths and assume that $p \circ \varphi = q \circ \psi$ and $q \circ \varphi' = r \circ \psi'$ for reparametrizations $\varphi, \varphi', \psi, \psi' \in \mathrm{Rep}_+(I)$. By Proposition 2.19, there are $\eta, \eta' \in \mathrm{Rep}_+(I)$ such that $\psi \circ \eta = \varphi' \circ \eta'$; hence $p \circ \varphi \circ \eta = r \circ \psi' \circ \eta'$.

2. It is enough to show that p and $p \circ \varphi$ are thinly homotopic for every $\varphi \in \mathrm{Rep}_+(I)$; consider the homotopy $H : I \times I \to I \xrightarrow{p} X,\ H(s,t) = p((1-s)t + s\varphi(t))$, that even factors over I.

□

Factoring out the respective equivalence relations given by the actions above, we arrive at quotient spaces $T_R(X) = R(X)/_{\mathrm{Homeo}_+(I)}$, resp. $T(X) = P(X)/_{\mathrm{Rep}_+(I)}$ with subspaces $T_R(X)(x,y) = R(X)(x,y)/_{\mathrm{Homeo}_+(I)}$, resp. $T(X)(x,y) = P(X)(x,y)/_{\mathrm{Rep}_+(I)}$ for $x, y \in X$. They are considered as spaces of (regular) *traces* of paths in X and should be compared to the notions of curves or regular curves in elementary differential geometry.

These spaces can be organised in a topological category $T(X)$ (and likewise $T_R(X)$) with the elements of X as objects, with the topological spaces $T(X)(x,y)$ as morphism from x to y and with a composition $T(X)(x,y) \times T(X)(y,z) \to T(X)(x,z)$ induced by concatenation. Remark that one does *not* obtain a category structure on $P(X)$ since concatenation is not associative "on the nose". The categories $T(X)$ and their directed relatives are used as important tools in [**15**].

Lemma 3.4. *Let x, y be elements of a Hausdorff space X and let $p \in R(X)(x,y)$, $\varphi \in \mathrm{Rep}_+(I)$. If $p \circ \varphi = p$, then $\varphi = \mathrm{id}_I$ or p is constant. In particular, for $x \neq y$, the action of* $\mathrm{Homeo}_+(I)$ *on $R(X)(x,y)$ is* free.

Proof. If $\varphi \neq \mathrm{id}_I$, there exists an interval $J = [a,b] \subseteq I$ with $\varphi(a) = a, \varphi(b) = b$ and, without loss of generality, $\varphi(t) < t$ for all $a < t < b$. For all these t we conclude that $p(t) = p(\varphi(t)) = p(\varphi^n(t))$ for all $n > 0$, and hence that $p(t) = p(\lim_{n\to\infty}(\varphi^n(t))) = p(a)$. In particular, there is a non-trivial interval on which p is constant; this is not allowed for a regular path unless p is constant on the entire unit interval I (and thus $x = y$). □

Corollary 3.5. *Let $x \neq y$ be elements of a topological space X. The quotient map $R(X)(x,y) \to T_R(X)(x,y)$ is a weak homotopy equivalence.*

Proof. The free group action yields a fibration with contractible fiber $\mathrm{Homeo}_+(I)$. □

It is not clear to the authors whether one can sort out conditions under which the quotient map is a genuine homotopy equivalence.

3.3. Spaces of (regular) traces

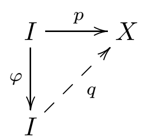
For $x, y \in X$, the inclusion map $R(X)(x,y) \hookrightarrow P(X)(x,y)$ induces a natural map $i : T_R(X)(x,y) \to T(X)(x,y)$ between the corresponding quotient trace spaces. The main aim of this section is a proof of

Theorem 3.6. *For every two elements $x, y \in X$ of a Hausdorff space X, the map $i : T_R(X)(x,y) \to T(X)(x,y)$ is a* homeomorphism.

In particular, every trace can be represented by a regular trace (cf. Proposition 3.7 below). It turns out that many of the results on reparametrizations from the preceding section will be used in the proof. In a first step, we show that the map i is surjective:

Proposition 3.7. *For every path $p \in P(X)$, there exists a regular path $q \in R(X)$ and a reparametrization $\varphi \in \mathrm{Rep}_+(I)$ such that $p = q \circ \varphi$.*

Proof. For every interval $J \subseteq I$ let $m(J)$ denote its midpoint. Let $m : \Delta_p \to I$ denote the map $J \mapsto m(J)$ with image $C := m(\Delta_p) \subseteq I$. In order to arrive at a reparametrization with stop map m, we check the conditions from Proposition 2.13 for the order-preserving bijection $m : \Delta_p \to C$: Let $x_n = \max(J_n) \uparrow x \in I$ denote a strictly monotone sequence. Then the midpoints converge as well: $m(J_n) \uparrow x$. Likewise for a decreasing sequence of lower boundaries and corresponding midpoints. From Proposition 2.13 we conclude that there exists a reparametrization $\varphi \in \mathrm{Rep}_+(I)$ with $\Delta_\varphi = \Delta_p$. Hence, there is a set-theoretic factorization



through a *regular* map $q : I \to X$.

To check that q is continuous, choose an open set $U \subset X$ and note that $q^{-1}(U) = \varphi p^{-1}(U)$. Combining Lemma 2.4 and Lemma 2.8.3, we conclude that $p^{-1}(U)$ is a union of maximal disjoint open intervals that are all mapped onto open intervals under φ. Hence $q^{-1}(U)$ is open as well. □

Proposition 3.8. *Two regular paths $p, q \in R(X)(x,y)$ that are reparametrization equivalent are in fact strictly reparametrization equivalent, i.e., there exists a homeomorphism $\eta \in \mathrm{Homeo}_+(I)$ such that $q = p \circ \eta$.*

The proof of Proposition 3.8 (and also that of Theorem 3.6) makes use of

Lemma 3.9. *Let $p \in R(X)(x,y)$, $p' \in P(X)(x,y)$, $\varphi, \varphi' \in \mathrm{Rep}_+(I)$ with $p \circ \varphi = p' \circ \varphi'$. Then there exists $\eta \in \mathrm{Rep}_+(I)$ with $p \circ \eta = p'$. Unless p is constant, η is unique.*

Proof. We apply Proposition 2.17 to prove the *existence* of such a reparametrization η: For every interval $J' \in \Delta_{\varphi'}$, there is a unique $J \in \Delta_{p'\circ\varphi'} = \Delta_{p\circ\varphi} = \Delta_\varphi$ such that $J' \subseteq J$. Hence, there exists a unique $\eta \in \mathrm{Rep}_+(I)$ with $\varphi = \eta \circ \varphi'$, whence $p' \circ \varphi' = p \circ \varphi = (p \circ \eta) \circ \varphi'$. Since φ' is onto, we conclude that $p' = p \circ \eta$.

To prove *uniqueness* of the factor η, suppose that $\eta_1, \eta_2 \in \mathrm{Rep}_+(I)$ with $p \circ \eta_1 = p \circ \eta_2$. If $\eta_1 \neq \eta_2$, one can choose an interval $J = [a,b]$ such that $\eta_1(a) = \eta_2(a), \eta_1(b) = \eta_2(b)$ and $\eta_1(t) < \eta_2(t)$ for $a < t < b$ (or vice versa). Given $a < t_0 = t'_0 < b$, choose an increasing sequence t_i and a decreasing sequence t'_i such that $\eta_1(t_{i+1}) = \eta_2(t_i)$, resp. $\eta_2(t'_{i+1}) = \eta_1(t_i)$. Both sequences converge to, say, $a \leqslant T' < T \leqslant b$. Since $\eta_1(T) = \eta_2(T)$ and $\eta_1(T') = \eta_2(T')$, we must have $T' = a, T = b$.

On the other hand, $p(\eta_2(t_i)) = p(\eta_1(t_{i+1})) = p(\eta_2(t_{i+1}))$ and hence $p(t_0) = p(b)$, and, by a similar argument: $p(t_0) = p(a)$. Since $t_0 \in]a,b[$ was chosen arbitrarily, p has to be constant on the interval $J = [a,b]$; this is impossible for a regular path p unless it is constant. □

Proof. (of Proposition 3.8) Let $p, q \in R(X)(x,y)$ be reparametrization equivalent, i.e., there exist reparametrizations $\varphi, \psi \in \mathrm{Rep}_+(I)$ such that $p \circ \varphi = q \circ \psi$. If one of p, q is constant, the other is as well. Assume that neither of them is constant. By Lemma 3.9, there is a reparametrization $\rho \in \mathrm{Rep}_+(I)$ with $q = p \circ \rho$. Since q is regular and non-constant, ρ has to be injective and thus a homeomorphism; in particular, p and q represent the same element in $T_R(X)(x,y)$. □

Proof. (of Theorem 3.6) From Propositions 3.7 and 3.8 we conclude that the map i from Theorem 3.6 is a continuous bijection. To see that it is also open, the following diagram will be useful:

$$\begin{array}{ccc} R(X)(x,y) & \xrightarrow{\subseteq} & P(X)(x,y) \\ {\scriptstyle Q_R}\downarrow & & \downarrow{\scriptstyle Q} \\ T_R(X)(x,y) & \xrightarrow{i} & T(X)(x,y) \end{array} \tag{3.1}$$

By Lemma 3.1, a basis for the topology of $R(X)(x,y)$ is given by the sets

$$R(X)(A,\mathcal{U};x,y) = \{p \in R(X)(x,y) \mid p([a_{j-1},a_j]) \subseteq U_j\}$$

indexed by all partitions $A = \{0 = a_0 < \cdots < a_n = 1\}$ of I and collections of n open sets $\mathcal{U} = \{U_1, \ldots, U_n\}$, $n \in \mathbf{N}$.

The sets $Q_R(R(X)(A,\mathcal{U};x,y))$ form a basis for the quotient topology on $T_R(X)(x,y)$ since

$$\begin{aligned} Q_R^{-1}(Q_R(R(X)(A,\mathcal{U};x,y))) &= \bigcup\nolimits_{\rho \in \mathrm{Homeo}_+(I)} \rho \cdot R(X)(A,\mathcal{U};x,y)) \\ &= \bigcup\nolimits_{\rho \in \mathrm{Homeo}_+(I)} R(X)(\rho^{-1}(A),\mathcal{U};x,y)), \end{aligned}$$

and the latter set is open in $R(X)(x,y)$.

We need to show that the sets $i(Q_R(R(X)(A,\mathcal{U};x,y))) = Q(R(X)(A,\mathcal{U};x,y))$ are open in $T(X)(x,y)$, i.e., that the sets $Q^{-1}(Q(R(X)(A,\mathcal{U};x,y)))$ are open in $P(X)(x,y)$. To this end, we note first that $Q^{-1}(Q(R(X)(A,\mathcal{U};x,y))) = \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$: The inclusion $\supseteq$ is obvious. If $p \in R(X)(A,\mathcal{U};x,y)$ and $p' \in P(X)(x,y)$, $\varphi,\varphi' \in \mathrm{Rep}_+(I)$ are such that $p\circ\varphi = p'\circ\varphi'$, then Lemma 3.9 yields $\eta \in \mathrm{Rep}_+(I)$ such that $p' = p\circ\eta$, hence $p' \in \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$.

In the next two steps we show that every element $q \in \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$ has an open neighbourhood in that set. For elements in $R(X)(A,\mathcal{U};x,y)$ it suffices to check that $P(X)(A,\mathcal{U};x,y) \subseteq \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$: According to Proposition 3.7, a path $p \in P(X)(A,\mathcal{U};x,y)$ can be factored in the form $p = q\circ\varphi$ with $q \in R(X)(x,y), \varphi \in \mathrm{Rep}_+(I)$; the factors then have to satisfy $\varphi([a_{i-1},a_i]) \subseteq q^{-1}(U_i)$. By Lemma 2.9, there is an approximation $\rho \in \mathrm{Homeo}_+(I)$ to φ satisfying $\rho([a_{i-1},a_i]) \subseteq q^{-1}(U_i)$ for all i. We can thus factor $p = (q\circ\rho)\circ(\rho^{-1}\circ\varphi)$ with $q\circ\rho \in R(X)(A,\mathcal{U};x,y)$, and hence $p \in \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$.

Now let $q \in R(X)(A,\mathcal{U};x,y)$ and $\varphi \in \mathrm{Rep}_+(I)$. Let $B = \{b_0 = 0 < \cdots < b_n = 1\}$ denote any partition of I with $\varphi(B) = A$. Then $q\circ\varphi \in P(X)(B,\mathcal{U};x,y)$, which is open in $P(X)(x,y)$. Hence it will suffice to show that $P(X)(B,\mathcal{U};x,y) \subseteq \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$. To this end, choose $\rho \in \mathrm{Homeo}_+(I)$ such that $\rho(b_i) = a_i$. Let $p \in P(X)(B,\mathcal{U};x,y)$ and write $p = p_1\circ\rho$. Then $p_1 \in P(X)(A,\mathcal{U};x,y) \subseteq \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$ by the previous step, and hence also $p \in \mathrm{Rep}_+(I)\cdot R(X)(A,\mathcal{U};x,y)$. □

Remark 3.10. All four spaces in (3.1) are (at least) weakly homotopy equivalent: If x,y are not in the same path component, they are all empty. Otherwise, $P(X)(x,y)$ is homotopy equivalent to the loop space $\Omega(X)(x)$ based at x; likewise, $R(X)(x,y)$ is homotopy equivalent to the space of regular loops $\Omega_R(X)(x)$ based at x. Both loop spaces are fibres in fibrations $P_R(X)\downarrow X$, resp. $P(X)\downarrow X$ over X with contractible total spaces $P_R(X;x)$, resp. $P(X;x)$ of (regular) paths starting at x. The Five Lemma shows that the inclusion map $\Omega_R(X)(x) \hookrightarrow \Omega(X)(x)$ is a weak homotopy equivalence.

Note that Lemma 3.9 allows us to give the following "backwards" characterization of reparametrization equivalence:

Proposition 3.11. *Paths $p,q \in P(X)$ are reparametrization equivalent if and only if there exists $r \in P(X)$ and $\varphi,\psi \in \mathrm{Rep}_+(I)$ such that $p = r\circ\varphi$ and $q = r\circ\psi$.*

Proof. For the "only if" part, suppose $p\circ\varphi_1 = q\circ\psi_1, \varphi_1,\psi_1 \in \mathrm{Rep}_+(I)$. We use Proposition 3.7 to write $p = r_1\circ\eta_1$ and $q = r_2\circ\eta_2$, with $r_1, r_2 \in R(X)$ and $\eta_1,\eta_2 \in \mathrm{Rep}_+(I)$. Then $r_1\circ\eta_1\circ\varphi_1 = r_2\circ\eta_2\circ\psi_1$; by Lemma 3.9, there exists $\eta \in \mathrm{Rep}_+(I)$ such that $r_2 = r_1\circ\eta$, whence $q = r_1\circ\eta\circ\eta_2$.

The reverse implication is clear by Proposition 2.19; if $\varphi_2,\psi_2 \in \mathrm{Rep}_+(I)$ are such that $\varphi\circ\varphi_2 = \psi\circ\psi_2$, then $p\circ\varphi_2 = r\circ\varphi\circ\varphi_2 = r\circ\psi\circ\psi_2 = q\circ\psi_2$. □

Let us finally note the following consequence of Proposition 3.7:

Definition 3.12. A path $p : I \to X$ is called *loop-free* if $p(s) = p(t)$ for any $s < t \in I$ implies that the restriction $p|_{[s,t]}$ is the constant path.

Note that a loop-free *regular* path is either constant or injective.

Corollary 3.13. *A loop-free path $p : I \to X$ in a Hausdorff space X has an image $p(I) \subseteq X$ that is either a point or homeomorphic to I.*

Proof. By Proposition 3.7, there is a factorization $p = q \circ \varphi$ with q a loop-free regular and thus either constant or injective path. In the second case, q is a continuous bijection from the compact space I to its image $q(I) = p(I) \subseteq X$. The claim follows since X is Hausdorff. □

4. Directed traces

Originally motivated by models arising in concurrency theory in theoretical computer science, an investigation of topological spaces with "preferred directions" has been launched under the title "Directed Algebraic Topology". Various frameworks (local po-spaces [**6**], d-spaces [**9**], flows [**7**] and others) have been suggested, all modifying in various ways concepts from elementary algebraic topology, in particular replacing relevant groups or groupoids by categories. The d-spaces introduced by M. Grandis [**9**] have turned out to give rise to a particularly successful way to combine homotopy theoretical and categorical methods for the study of spaces with preferred directions.

The question whether one can neglect reparametrizations in a homotopy theoretical study of spaces of directed paths was one of the original motivations for this article. This is – under a certain natural condition – affirmed in Corollary 4.5, which is used as a starting point for the homotopy theoretical study of "directed spaces" in [**15**].

Definition 4.1 ([**9**])**.** A *d-space* is a topological space X together with a set $\vec{P}(X) \subseteq P(X)$ of continuous paths $I \to X$ such that

1. $\vec{P}(X)$ contains all constant paths;
2. $p \circ \varphi \in \vec{P}(X)$ for any $p \in \vec{P}(X)$ and any continuous increasing (not necessarily surjective)s map $\varphi : I \to I$;
3. for all $p, q \in \vec{P}(X)$ such that $p(1) = q(0)$, their concatenation $p * q \in \vec{P}(X)$, cf. (2.1).

Elements of $\vec{P}(X)$ are called d-*paths*. $\vec{P}(X) \subseteq P(X)$ is given the subspace topology (of the compact-open topology).

Definition 4.2. A *d-map* between d-spaces X, Y is a continuous mapping $f : X \to Y$ satisfying $p \in \vec{P}(X) \Rightarrow f \circ p \in \vec{P}(Y)$. Isomorphisms in the category of d-spaces are called *d-homeomorphisms*.

The d-interval $\vec{I}$ is given the standard d-structure $\vec{P}(I) = \text{Rep}_+(I)$.

In general d-spaces, it may occur that a non-d-path becomes directed after reparametrization. To exclude this possibility, we add

Definition 4.3. A d-space X is called *saturated* if it has the following additional property:

4. If $p \in P(X), \varphi \in \mathrm{Rep}_+(I)$ and $p \circ \varphi \in \vec{P}(X)$, then $p \in \vec{P}(X)$.

In words: If a path becomes a d-path after reparametrization, then it has to be a d-path itself already. Remark that for a saturated d-space, two trace equivalent paths are either both d-paths, or neither of them is. The d-space $\vec{I}$ is saturated. It is easy to turn a given d-space X into a saturated one: one just adds all paths $p \in P(X)$ for which there is a reparametrization $\varphi \in \mathrm{Rep}_+(I)$ with $p \circ \varphi \in \vec{P}(X)$ to the d-paths in a new structure $S\vec{P}(X)$, which is easily seen to satisfy the properties of a saturated d-space. So there is no harm in assuming that a d-space is saturated right away.

Among the d-paths in X, we pay particular attention to the *regular* d-paths, cf. Definition 1.1.3; the set of all those will be denoted $\vec{R}(X) = R(X) \cap \vec{P}(X)$ and equipped with the subspace topology. Again, $\mathrm{Homeo}_+(I)$ acts (essentially freely) on $\vec{R}(X)$, and $\mathrm{Rep}_+(I)$ acts on $\vec{P}(X)$. We can now speak of spaces of (regular) traces in a saturated d-space X:

Definition 4.4.
- $\vec{T}_R(X)(x, y) := \vec{R}(X)(x, y)/_{\mathrm{Homeo}_+(I)} \subseteq T_R(X)(x, y)$
- $\vec{T}(X)(x, y) := \vec{P}(X)(x, y)/_{\mathrm{Rep}_+(I)} \subseteq T(X)(x, y)$

These form the morphisms of categories $\vec{T}_R(X)$, resp. $\vec{T}(X)$ that are investigated from a homotopy theory point of view in [**15**]. The following consequence of Theorem 3.6 tells us that it makes no difference in topology which of the two (quotient) trace spaces is chosen:

Corollary 4.5. *Let X denote a saturated d-space and let $x, y \in X$. The map $\vec{i} : \vec{T}_R(X)(x, y) \to \vec{T}(X)(x, y)$ induced by inclusion $\vec{R}(X)(x, y) \hookrightarrow \vec{P}(X)(x, y)$ is a homeomorphism.*

Remark 4.6. It is no longer clear whether the inclusion map $\vec{R}(X)(x, y) \hookrightarrow \vec{P}(X)(x, y)$ is a weak homotopy equivalence. The (weak) homotopy types of both spaces depend on the choice of x and y and it is therefore not possible to argue using loop spaces as in 3.10. From the diagram

$$\begin{array}{ccc} \vec{R}(X)(x, y) & \xrightarrow{\subseteq} & \vec{P}(X)(x, y) \\ {\scriptstyle Q_R}\downarrow \simeq & & \downarrow {\scriptstyle Q} \\ \vec{T}_R(X)(x, y) & \xrightarrow[\cong]{i} & \vec{T}(X)(x, y), \end{array}$$

obtained from (3.1) by restricting to d-paths, we can only deduce that the inclusion maps $\vec{R}(X)(x, y) \hookrightarrow \vec{P}(X)(x, y)$ induce injections and that the quotient maps $Q : \vec{P}(X)(x, y) \to \vec{T}(X)(x, y)$ induce surjections on all homotopy groups.

Definition 4.7.
1. [**9**] A *d-homotopy* from a d-path $p \in \vec{P}(X)$ to a d-path $q \in \vec{P}(X)$ is a d-map $H : \vec{I} \times \vec{I} \to X$ for which $H(0, \cdot) = p$, $H(1, \cdot) = q$, and $H(\cdot, 0)$, $H(\cdot, 1)$ are constant.
2. A d-homotopy is said to be *thin* if it factors through the d-interval $\vec{I}$, i.e. if there are d-maps $\Phi : \vec{I} \times \vec{I} \to \vec{I}$, $r : \vec{I} \to X$ such that $H = r \circ \Phi$.

3. Two d-paths $p, q \in \vec{P}(X)$ are said to be *d-homotopic*, respectively *thinly d-homotopic*, if there exists a sequence $H_1, \dots, H_{2n+1}$ of d-homotopies, respectively thin d-homotopies, such that $H_1(0, \cdot) = p$, $H_{2n+1}(1, \cdot) = q$, $H_{2i-1}(1, \cdot) = H_{2i}(1, \cdot)$, and $H_{2i}(0, \cdot) = H_{2i+1}(0, \cdot)$.

Remark 4.8.

- The directed structure on a product of d-spaces X and Y is given by
$$P(X \times Y) \supseteq \vec{P}(X \times Y) = \vec{P}(X) \times \vec{P}(Y) \subseteq P(X) \times P(Y)$$
under the natural identification $P(X \times Y) \cong P(X) \times P(Y)$. In particular, $\vec{P}(\vec{I} \times \vec{I})$ consists of all paths $p : I \to I \times I$ that are (weakly) increasing in both coordinates.
- The relations $\preccurlyeq$, $\preccurlyeq_T$ on d-paths given by existence of (thin) d-homotopies are preorders on $\vec{P}(X)$. The relations $\simeq$, $\simeq_T$ on d-paths given by being (thinly) d-homotopic are equivalence relations on $\vec{P}(X)$; they are the symmetric, transitive closures of $\preccurlyeq$ respectively $\preccurlyeq_T$.
- In the differentiable setting, thin homotopies of differentiable loops and paths were defined (using homotopies of rank at most 1) and a resulting smooth fundamental groupoid used for holonomy considerations in [**2, 14**].

Proposition 4.9. *Two d-paths p, q in a saturated d-space are reparametrization equivalent if and only if they are thinly d-homotopic.*

This result is needed in the study [**5**] of directed squares. Note that the notion of d-homotopy factors over $\vec{T}(X)$ (and $\vec{T}_R(X)$).

Proof. We use Proposition 3.11. For the forward implication, write $p = r \circ \varphi$, $q = r \circ \psi$, and let $\eta = \max(\varphi, \psi)$. Define $\Phi, \Psi : \vec{I} \times \vec{I} \to \vec{I}$ by
$$\Phi(s,t) = (1-s)\varphi(t) + s\eta(t)$$
$$\Psi(s,t) = (1-s)\psi(t) + s\eta(t)$$
then $r \circ \Phi$, $r \circ \Psi$ are thin d-homotopies connecting p and q.

To show the back implication, it is enough to consider the case where p and q are connected by *one* thin d-homotopy H with $H(0, \cdot) = p$ and $H(1, \cdot) = q$. Write $H = r \circ \Phi : \vec{I} \times \vec{I} \to \vec{I} \to X$; by Corollary 4.5, we can assume r to be regular. Also, by reparametrizing if necessary, we can assume that $\Phi(0,0) = 0$ and $\Phi(1,1) = 1$. Then
$$r(\Phi(s,0)) = H(s,0) = H(0,0) = r(0)$$
$$r(\Phi(s,1)) = H(s,1) = H(1,1) = r(1)$$
hence by regularity of r and following the logic of the final step in the proof of Lemma 3.4, $\Phi(s,0) = 0$ and $\Phi(s,1) = 1$.

Now define $\varphi, \psi : \vec{I} \to \vec{I}$ by $\varphi(t) = \Phi(0,t)$, $\psi(t) = \Phi(1,t)$, then $\varphi, \psi \in \mathrm{Rep}_+(I)$ and $p = r \circ \varphi$, $q = r \circ \psi$. □

Finally, we modify the results from Section 3 about loop-free paths (Definition 3.12) to the d-space environment:

Definition 4.10. A d-space X is said to be *locally loop-free* provided that every point has a neighbourhood in which all non-constant d-paths are loop-free.

Note that d-spaces arising from a space with a locally partial order [**6**] are locally loop-free. The following result applies such to such spaces, in particular. For po-spaces, a result similar to the following had previously been obtained in [**12**, Thm. 5].

Corollary 4.11. *If $p \in \vec{P}(X)$ is a loop-free d-path in a locally loop-free saturated Hausdorff d-space X, then its image $p(\vec{I})$ is either a point or d-homeomorphic to $\vec{I}$.*

Proof. The statement is trivial for a constant d-path. Otherwise, Corollary 4.5 provides us with a *regular* loop-free d-path $q : \vec{I} \to X$ with $p = q \circ \varphi$, $\varphi \in \mathrm{Rep}_+(I)$ which by Corollary 3.13 yields the homeomorphism $q : I \to p(\vec{I})$. All we need to show is that its inverse $q^{-1} : p(\vec{I}) \to \vec{I}$ is a d-map.

Let $r : \vec{I} \to p(\vec{I})$ be a d-path; we need to show that $q^{-1} \circ r \in \vec{P}(I) = \mathrm{Rep}_+(I)$. Let $t_1 < t_2 \in I$ and suppose that $q^{-1}(r(t_1)) > q^{-1}(r(t_2))$.

Restricting to a smaller interval, if necessary, will ensure that $r([t_1, t_2]) \subseteq X$ is contained in a loop-free neighbourhood $U \subseteq X$. The concatenation

$$r\,|_{[t_1,t_2]} * q\,|_{[q^{-1}(r(t_2)),q^{-1}(r(t_1))]}$$

is a d-path and a loop in U and hence constant. Then $q\,|_{[q^{-1}(r(t_2)),q^{-1}(r(t_1))]}$ is constant, in contradiction to being a regular and non-constant path. Hence $q^{-1}(r(t_1)) \leqslant q^{-1}(r(t_2))$ whence q^{-1} is a d-map. ☐

Remark 4.12. It seems plausible that many of the methods and of the results from this article allow generalizations to maps $p : I^n \to X$, resp. $p : \vec{I}^n \to X$, from (directed) cubes to (d)-spaces. The relevant reparametrizations to investigate are the d-maps $\varphi : \vec{I}^n \to \vec{I}^n$ (monotone in every coordinate) that preserve boundaries in the following sense:

$$\varphi(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \text{ and } x_i = 0, \text{ resp. } 1 \Rightarrow y_i = 0, \text{ resp. } 1.$$

In the differentiable setting, such reparametrizations for so-called n-loops and resulting smooth homotopy groups have been defined and studied in [**3, 14**].

References

[**1**] J. Baez and U. Schreiber. Higher Gauge Theory. *arXiv:math.DG/0511710*, 2005.

[**2**] A. Caetano and R. Picken. An axiomatic definition of holonomy. *Internat. J. Math.*, 5(6):835–848, 1994.

[**3**] A. Caetano and R. Picken. On a family of topological invariants similar to homotopy groups. *Rend. Istit. Mat. Univ. Trieste*, 30(1-2):81–90, 1998.

[**4**] J. Dugundji. *Topology.* Allyn and Bacon, 1966.

[**5**] U. Fahrenberg. Homotopy of squares in a directed Hausdorff space. In preparation.

[6] L. Fajstrup, E. Goubault, and M. Raussen. Algebraic topology and concurrency. *Theoret. Comput. Sci.*, 357:241–278, 2006.

[7] P. Gaucher. A model category for the homotopy theory of concurrency. *Homology, Homotopy Appl.*, 5(1):549–599, 2003.

[8] P. Goerss, J.F. Jardine. *Simplicial Homotopy Theory.* Progress in Mathematics 174, Birkhäuser

[9] M. Grandis. Directed homotopy theory I. *Cah. Topol. Géom. Différ. Catég.*, 44:281–316, 2003.

[10] M. Grandis. Absolute lax 2-categories. *Appl. Categ. Structures*, 14:191–214, 2006.

[11] K.A. Hardie, K.H. Kamps, and R.W. Kieboom. A homotopy 2-groupoid of a Hausdorff space. *Appl. Categ. Structures*, 8:209–234, 2000.

[12] E. Haucourt. *Topologie algébrique dirigée et concurrence.* PhD thesis, Université Paris 7, UFR Informatique, 2005.

[13] M. Loève. *Probability Theory.* Univers. Series in Higher Math. van Norstrand, Princeton, NJ, 3 edition, 1963.

[14] M. Mackaay and R. Picken. Holonomy and parallel transport for abelian gerbes. *Adv. Math.*, 170:287–339, 2002.

[15] M. Raussen. Invariants of directed spaces. *Appl. Categ. Structures*, 2007. to appear.

[16] E.H. Spanier. *Algebraic Topology.* Mc Graw-Hill, 1966.

Ulrich Fahrenberg `uli@cs.aau.dk` `www.cs.aau.dk/~uli`

Department of Computer Science
Aalborg University
Denmark
Fredrik Bajersvej 7B
DK-9220 Aalborg Øst

Martin Raussen `raussen@math.aau.dk` `www.math.aau.dk/~raussen`

Department of Mathematical Sciences
Aalborg University
Denmark
Fredrik Bajersvej 7G
DK-9220 Aalborg Øst

Journal of Homotopy and Related Structures, vol. 4(1), 2009, pp.1–5

REPARAMETRIZATIONS WITH GIVEN STOP DATA

MARTIN RAUSSEN

(*communicated by Ronnald Brown*)

Abstract

In [1], we performed a systematic investigation of reparametrizations of continuous paths in a Hausdorff space that relies crucially on a proper understanding of stop data of a (weakly increasing) reparametrization of the unit interval. I am grateful to Marco Grandis (Genova) for pointing out to me that the proof of Proposition 3.7 in [1] is wrong. Fortunately, the statement of that Proposition and the results depending on it stay correct. It is the purpose of this note to provide correct proofs.

1. Reparametrizations with given stop maps

To make this note self-contained, we need to include some of the basic definitions from [1]. The set of all (nondegenerate) closed subintervals of the unit interval $I = [0, 1]$ will be denoted by $\mathcal{P}_{[\,]}(I) = \{[a, b] \mid 0 \leqslant a < b \leqslant 1\}$.

Definition 1.1.

- A *reparametrization* of the unit interval I is a weakly increasing continuous self-map $\varphi : I \to I$ preserving the end points.
- A *non-degenerate* interval $J \subset I$ is a *φ-stop interval* if there exists a value $t \in I$ such that $\varphi^{-1}(t) = J$. The value $t = \varphi(J) \in I$ is called a *φ-stop value*.
- The set of all φ-stop intervals will be denoted as $\Delta_\varphi \subseteq \mathcal{P}_{[\,]}(I)$. Remark that the intervals in Δ_φ are disjoint and that Δ_φ carries a natural total order. We let $D_\varphi := \bigcup_{J \in \Delta_\varphi} J \subset I$ denote the *stop set* of φ; and $C_\varphi \subset I$ the set of all stop values.
- The *φ-stop map* $F_\varphi : \Delta_\varphi \to C_\varphi$ corresponding to a reparametrization φ is given by $F_\varphi(J) = \varphi(J)$.

It is shown in [1] that F_φ is an *order-preserving bijection* between (at most) *countable sets*. It is natural to ask (and important for some of the results in [1]) which order-preserving bijections between such sets arise from some reparametrization:

To this end, let

- $\Delta \subseteq \mathcal{P}_{[\,]}(I)$ denote a subset of *disjoint closed* sub-intervals – equipped with the natural total order;

Received August 19, 2008, revised September 2, 2008; published on February 11, 2009.
2000 Mathematics Subject Classification: 55,68
Key words and phrases: path, regular path, reparametrization, stop map.

- $C \subseteq I$ denote a subset with the same cardinality as Δ;
- $F : \Delta \to C$ denote an order-preserving bijection.

I am grateful to the referee for pointing out the following lemma and its proof:

Lemma 1.2. *A subset $\Delta \subseteq \mathcal{P}_{[\,]}(I)$ of disjoint closed intervals is countable.*

Proof. Given a set Δ of disjoint non-degenerate closed sub-intervals of the unit interval I, each will contain rational numbers by density. By the axiom of choice, choose for each disjoint sub-interval a specific rational number contained in that sub-interval. The chosen set $\Delta' \subset \mathbf{Q}$ of rationals is countable as a subset of $\mathbf{Q}$. Combining an enumeration of Δ' with the bijection between Δ' and Δ mapping each interval to its chosen rational yields an enumeration of Δ. □

Proposition 1.3. *There exists a reparametrization φ with $F_\varphi = F$ if and only if conditions* (1) - (8) *below are satisfied for intervals contained in Δ and for all $0 < z < 1$:*

1. $\min J = \sup_{J'<J} \max J' \Rightarrow F(J) = \sup_{J'<J} F(J')$;
2. $\max J = \inf_{J<J'} \min J' \Rightarrow F(J) = \inf_{J<J'} F(J')$;
3. $\sup_{J'<z} \max J' = \inf_{z<J''} \min J'' \Rightarrow \sup_{J'<z} F(J') = \inf_{z<J''} F(J'')$;
4. $\sup_{J'<z} \max J' < \inf_{z<J''} \min J'' \Rightarrow \sup_{J'<z} F(J') < \inf_{z<J''} F(J'')$;
5. $\inf_{0<J} \min J = 0 \Rightarrow \inf_{0<J} F(J) = 0$;
6. $\inf_{0<J} \min J > 0 \Rightarrow \inf_{0<J} F(J) > 0$;
7. $\sup_{J<1} \max J = 1 \Rightarrow \sup_{J<1} F(J) = 1$;
8. $\sup_{J<1} \max J < 1 \Rightarrow \sup_{J<1} F(J) < 1$.

Proof. Conditions (1) – (3), (5) and (7) are necessary for the stop data of a *continuous* reparametrization φ; (4), (6) and (8) are necessary to avoid further stop intervals.

Given a stop map satisfying conditions (1) – (8), we construct a reparametrization φ_F with $F(\varphi_F) = F$ as follows: For $t \in D = \bigcup_{J\in\Delta} J$, one has to define: $\varphi(t) = F(J)$ with $t \in J$. This defines a weakly increasing function φ_F on D. Conditions (1) and (2) make sure that this function is continuous (on D). Condition (3) makes it possible to extend φ_F uniquely to a weakly increasing continuous function on the closure $\bar{D}$: $\varphi_F(z)$ is defined as $\sup_{J'<z} F(J')$ for $z = \sup_{J'<z} \max J'$ and/or as $\inf_{z<J''} F(J'')$ for $z = \inf_{z<J''} \min J$. By (5) and (7), $\varphi_F(0) = 0$ and $\varphi_F(1) = 1$ if $0, 1 \in \bar{D}$; if not, we have to take these as a definition.

The complement $O = I \setminus \bar{D}$ is an open (possibly empty) subspace of I, hence a union of at most countably many open subintervals $J = [a_-^J, a_+^J]$ with boundary in $\partial D \cup \{0, 1\}$. Condition (4), (6) and (8) make sure, that $\varphi_F(a_-^J) < \varphi_F(a_+^J)$. Hence, every collection of strictly increasing homeomorphisms between $[a_-^J, a_+^J]$ and $[\varphi_F(a_-^J), \varphi_F(a_+^J)]$ – preserving endpoints – extends φ_F to a continuous increasing map $\varphi_F : I \to I$ with $\Delta_{\varphi_F} = \Delta, C_{\varphi_F} = C$ and $F_{\varphi_F} = F$. □

It is natural to ask, whether

- every at most countable subset $C \subset I$ occurs as set of stop values of some reparametrization: This is answered affirmatively in [1], Lemma 2.10;
- every set $\{I\} \neq \Delta \subset \mathcal{P}_{[\,]}(I)$ of closed disjoint intervals arises as set of stop intervals of a reparametrization:

Proposition 1.4. *For every $\{I\} \neq \Delta$ of closed disjoint sub-intervals in the unit interval I, there exists a reparametrization φ with $\Delta_\varphi = \Delta$.*

Proof. We use Lemma 1.2 to provide us with an enumeration j of the totally ordered set Δ (defined either on $\mathbf{N}$ or on a finite integer interval $[1, n]$). Using j, we are going to construct a reparametrization φ with stop value set C_φ included in the set $I[\frac{1}{2}] = \{0 \leqslant \frac{l}{2^k} \leqslant 1\}$ of rational numbers with denominators a power of 2. To this end, we will associate to every number $z \in I[\frac{1}{2}]$ either an interval in Δ or a degenerate one point interval; we end up with an ordered bijection beween $I[\frac{1}{2}]$ and a superset of Δ; all excess intervals will be degenerate one-point sets.

To get started, let I_0 denote either *the* interval in Δ containing 0 or, if no such interval exists, the degenerate interval $[0, 0] = \{0\}$; likewise define I_1. Every number $z \in I[\frac{1}{2}]$ apart from 0 and 1 has a unique representation $z = \frac{l}{2^k}$ with l *odd*, $0 < l < 2^k$. The construction proceeds by induction on k using the enumeration j.

Assume for a given $k \geqslant 1$, I_z and thus the map $I : z \mapsto I_z$ defined for all $z = \frac{l}{2^{k-1}}$, $0 \leqslant l \leqslant 2^{k-1}$ as an ordered map. For $0 < z = \frac{l}{2^k} < 1$ and l odd, both $z_\pm = z \pm \frac{1}{2^k}$ have a representation as fraction with denominator 2^{k-1} and thus $I_{z_-} < I_{z_+}$ are already defined. Let $I_z = j(m)$ with m minimal such that $I_{z_-} < j(m) < I_{z_+}$ if such an m exists; if not, then I_z is defined as the degenerate interval containing the single element $\frac{1}{2}(\max I_{z_-} + \min I_{z_+})$. The map $I : z \mapsto I_z$ thus constructed on $I[\frac{1}{2}]$ is order-preserving. Moreover, this map is onto, since – by an induction over k – $I_{j(k)}$ occurs as I_z with some z of the form $\frac{l}{2^k}$. Hence, there is an order-preserving inverse map $I^{-1} : I_z \mapsto z$.

For $k \geqslant 0$, let φ_k denote the piecewise linear reparametrization that has constant value z on I_z for $z = \frac{l}{2^k}$, $0 \leqslant l \leqslant 2^k$ and that is linear inbetween these intervals. Remark that $\varphi_{k+1} = \varphi_k$ on all I_z with $z = \frac{l}{2^k}$ including all occuring degenerate intervals. As a consequence, $\| \varphi_k - \varphi_{k+1} \| < \frac{1}{2^k}$, and hence for all $l > k$, $\| \varphi_k - \varphi_l \| < \frac{1}{2^{k-1}}$. Hence, the sequence $(\varphi_k)_{k \in \mathbf{N}}$ converges uniformly to a continuous reparametrization φ.

By construction, the resulting reparametrization φ is constant on all intervals in Δ; on every open interval between these stop intervals, it is linear and strictly increasing. In particular, $\Delta_\varphi = \Delta$. □

Remark 1.5. I was first tempted to prove Proposition 1.4 by taking some integral of the characteristic function of the complement of D and to normalize the resulting function. But in general, this does not work out since, as already remarked in [1], it may well be that $\bar{D} = I$!

2. Concluding remarks

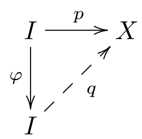
Remark 2.1. 1. Instead of constructing the reparametrization φ in Proposition 1.4, it is also possible to apply the criteria in Proposition 1.3 to the restriction $I_{|\Delta}$ of the map I from the proof above.

2. Proposition 1.3 replaces Proposition 2.13 in [1]. To get sufficiency, requirements (1) and (2) had to be added to those mentioned in [1] in order to make sure that the map φ_F is continuous on D. Moreover, (6) and (8) had to be added to avoid stop intervals containing 0, resp. 1 in case Δ does not contain such intervals.
In particular, the midpoint map m that associates to every interval in Δ its midpoint satisfies the criteria given in [1], Proposition 2.13, but if fails in general to satisfy conditions (1) and (2) in Proposition 1.3 in this note; in particular, the map φ_m will in general not be continuous, as remarked by M. Grandis. The midpoint map m was used in the flawed proof of [1], Proposition 3.7 – stated as Proposition 2.2 below.

The main focus in [1] is on reparametrizations of continuous paths $p : I \to X$ into a Hausdorff space X. A continuous path q is called *regular* if it is constant or if the restriction $q|_J$ to every non-degenerate sub-interval $J \subseteq I$ is *non-constant*.

Proposition 2.2. (Proposition 3.7 in [1])
For every path $p : I \to X$, there exists a regular path q and a reparametrization such that $p = q \circ \varphi$.

Proof. A non-constant path p gives rise to the set of all (closed disjoint) *stop intervals* $\Delta_p \subset \mathcal{P}_{[\,]}(I)$, consisting of the maximal subintervals $J \subset I$ on which p is constant. Proposition 1.4 yields a reparametrization φ with $\Delta_\varphi = \Delta_p$ and thus a set-theoretic factorization



through a map $q : I \to X$ that is not constant on any non-degenerate subinterval $J \subseteq I$. The continuity of q follows as in the remaining lines of the proof in [1]. □

References

[**1**] U. Fahrenberg and M. Raussen *Reparametrizations of continuous paths*, J. Homotopy Relat. Struct. **2** (2007), no.2, 93 – 117.

See also the references in [1].

http://www.emis.de/ZMATH/
http://www.ams.org/mathscinet

This article may be accessed via WWW at http://jhrs.rmi.acnet.ge

Martin Raussen
raussen@math.aau.dk
www.math.aau.dk/~raussen

Department of Mathematical Sciences
Aalborg University
Denmark
Fredrik Bajersvej 7G
DK-9220 Aalborg Øst

Topology and its Applications 156 (2009) 1718–1728

Contents lists available at ScienceDirect

Topology and its Applications

www.elsevier.com/locate/topol

Trace spaces in a pre-cubical complex

Martin Raussen

Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg Øst, Denmark

A R T I C L E I N F O

Article history:
Received 23 October 2008
Received in revised form 6 February 2009
Accepted 6 February 2009

MSC:
54F65
55P10
54E35
54D30
54D65
68Q85

Keywords:
Pre-cubical complex
Trace
Locally compact
Locally contractible
Homotopy type
Higher-dimensional automata

A B S T R A C T

In directed algebraic topology, directed irreversible (d)-paths and spaces consisting of d-paths are studied from a topological and from a categorical point of view. Motivated by models for concurrent computation, we study in this paper spaces of d-paths in a pre-cubical complex. Such paths are equipped with a natural arc length which moreover is shown to be invariant under directed homotopies. D-paths up to reparametrization (called traces) can thus be represented by arc length parametrized d-paths. Under weak additional conditions, it is shown that trace spaces in a pre-cubical complex are separable metric spaces which are locally contractible and locally compact. Moreover, they have the homotopy type of a CW-complex.

1. Introduction

1.1. Background

With motivations arising originally from concurrency theory within computer science, a new field of research, directed algebraic topology, has emerged. Its main characteristic is, that it involves spaces of "directed paths" (or timed paths, executions): these directed paths can be concatenated, but in general *not* reversed; time is not reversable.

A particular model in the investigation of concurrency phenomena leads to Higher-Dimensional Automata (HDA); for a recent report describing and assessing those consult e.g. [25]. The underlying space in these models is then – instead of a directed graph – the geometric realization of a pre-cubical set; defined like a pre-simplicial complex, but with cubes as building blocks; cf. e.g. [4,3]. Every cube carries a natural partial order, and directed paths have to respect the partial orders in their range. In the models, directed paths correspond to executions (calculations); they have two crucial properties:

1. The reverse of a directed path is, in general, *not* directed;
2. Directed paths that are *dihomotopic*, i.e., that can be connected through a one-parameter family of directed paths, are equivalent; this means that HDA – calculations along these "schedules" will always lead to the same result.

E-mail address: raussen@math.aau.dk.

0166-8641/$ – see front matter
doi:10.1016/j.topol.2009.02.003

A nice and flexible framework for directed paths was introduced by Marco Grandis with the notion of d-spaces and, in particular, of d-paths (cf. Definition 2.2) on a topological space X. In [12] and the subsequent paper [11], Grandis developed a framework for directed homotopy theory. The best-studied invariant of such a d-space is its *fundamental category* replacing the fundamental group in ordinary algebraic topology. It is studied in [8] and in [10] and used to decompose the d-space into "components".

A study of other and higher invariants from algebraic topology in the framework of directed topology was initiated in [18]. The general idea is that one ought to study spaces of d-paths with given end points, to organise those in a categorical manner, and, in particular, to find out how the topology of these path spaces changes under variation of the end points.

At present, it is unsatisfactory that only few concrete calculations of algebraic topological invariants of d-spaces (i.e., of their path spaces) are known. For general d-spaces, this is probably a hopeless endeavour; as in ordinary algebraic topology, one needs additional structure. This article studies the topology of path spaces (or spaces homotopy equivalent to those) in pre-cubical complexes, i.e., those spaces that are of interest in the applications. It will be followed up in [20] by an attempt to give the path spaces in this case a combinatorial structure that makes calculations of invariants possible; most of the results from this paper will be needed to get going. A rough outline for this program will be sketched in Section 4.

1.2. Definitions and results from previous work

Spaces of paths *up to reparametrizations* (both undirected and directed) were studied in [6] – with a few corrections in [19]. For the convenience of the reader, we state important definitions and some of the results:

Definition 1.1. Let X denote a Hausdorff space.

1. A reparametrization (of the unit interval I) is a *surjective* weakly increasing self-map $\varphi : I \to I$. The space of all reparametrizations (as subspace of the space of all self-maps I^I with the compact-open topology) is called $\mathrm{Rep}_+(I)$. This space is a monoid under composition.
2. The *strictly* increasing reparametrizations form a sub*group* $\mathrm{Homeo}_+(I) \subset \mathrm{Rep}_+(I)$.
3. The set X^I of all paths in X (with CO-topology) is denoted $P(X)$. For $x_0, x_1 \in X$, $P(X)(x_0, x_1)$ is the subspace of all paths p with $p(0) = x_0$, $p(1) = x_1$.
   Likewise, for $A_0, A_1 \subset X$, $P(X)(A_0, A_1)$ is the subspace of all paths p with $p(0) \in A_0$, $p(1) \in A_1$.
4. Two paths p, q are called *reparametrization equivalent* if and only if there exist reparametrizations $\varphi, \psi \in \mathrm{Rep}_+(I)$ such that $p \circ \varphi = q \circ \psi$.
5. A path $p \in P(X)$ is called *regular* if it is either constant or if there does not exist any non-trivial (stop)-interval $J = [a, b]$, $0 \leqslant a < b \leqslant 1$, such that $p|_J$ is constant.

It is shown in [6, Corollary 3.3], that reparametrization equivalence is in fact an equivalence relation. An equivalence class is called a *trace* in X. To divide out the effect of reparametrizations, we form the quotient space $T_P(X) = P(X)/_{\mathrm{Rep}_+(I)}$ consisting of the traces in X. It is compared in [6] with the quotient $T_R(X)$ of the action of the group $\mathrm{Homeo}_+(I)$ on the space $R(X)$ of regular paths (regular traces). Both trace spaces fibre over $X \times X$ (end points!) with fibres $T_P(X)(x, y)$, resp. $T_R(X)(x, y)$. Using several non-trivial and even surprising lifting results, it is shown in [6, Theorem 3.6]:

Proposition 1.2. *For all $x, y \in X$, the inclusion map $R(X)(x, y) \hookrightarrow P(X)(x, y)$ induces a* homeomorphism $T_R(X)(x, y) \to T_P(X)(x, y)$.

Counterparts of these results for directed paths (and traces) were shown in [6] in the general framework of d-spaces [12], as well. In particular, we will work in this paper with spaces $\vec{T}(X)$ of directed traces and dipointed subspaces $\vec{T}(X)(x, y)$; we take the liberty to represent individual traces by d-paths or by regular d-paths as is suitable.

Trace spaces have several advantages compared to path spaces: First of all, they form a topological category (with pairs of points as objects). Secondly, they are often compact or at least locally compact.

1.3. Summary of results

In this paper, we restrict attention to path and trace spaces in pre-cubical complexes, the geometric realizations of $\square$-sets, cf. [4,3]. We give a definition for the *arc length* of a d-path in such a space that does not extend to general paths. In a similar manner as for smooth paths in differential geometry, one may reparametrize to obtain an arc length (also called natural) parametrization for such a d-path.

This simple idea and its consequences are studied in Section 2: It turns out that length is a *dihomotopy invariant* for d-paths with the same start and end point (unlike for arbitrary paths). Every directed trace (cf. Section 1.2) has a well-defined arc length parametrized representative, and these are easier to work with than abstract traces. This will be exploited several times in Section 3. From a topological point of view, there is no need to distinguish between spaces of d-paths and traces in a pre-cubical complex: we show in Corollary 2.16, that spaces of d-paths, of regular d-paths and of traces (with given end points) are all homotopy equivalent to each other.

In Section 3, we study topological properties of trace spaces in a pre-cubical complex. We show that trace spaces (with elements viewed as naturally parametrized d-paths) are *metrizable* and in fact homeomorphic to *separable* metric spaces. Next, we show that all subspaces of a trace space consisting of traces of a bounded length are *relatively compact*. In particular, a trace space (in a finite pre-cubical complex) is itself *locally compact*, and every connected component (dihomotopy class) in a trace space with given end points is actually *compact*.

Moreover, we show that spaces of d-paths are *locally contractible*; otherwise it would be hopeless to aim for inductive calculations of algebraic topological invariants. The proof uses techniques introduced by John Milnor in [15], that also allow to show that a trace space in a pre-cubical complex has the homotopy type of a CW-complex. This comes in handy in [20] to prove that certain weak homotopy equivalences are actually honest homotopy equivalences.

In the final Section 4, we give a brief outlook on how the results of this paper will be used in [20] to find "condensed" models of path spaces in pre-cubical complexes up to homotopy equivalence.

2. Traces in a pre-cubical complex: Natural parametrization and consequences

2.1. General definitions

Properties of Higher-Dimensional Automata (cf. Section 1.1) are intimately related to the study of directed paths in a pre-cubical set, also called a $\square$-set; this term (cf. [7]) is used in a similar way as a Δ-set – as introduced in [21] – for a simplicial set without degeneracies. We use $\square_n$ as an abbreviation for the n-cube $I^n = [0,1]^n$ with the product topology.

Definition 2.1.

1. A $\square$-set or pre-cubical complex M is a family of disjoint sets $\{M_n \mid n \geqslant 0\}$ with face maps $\partial_i^k : M_n \to M_{n-1}$, $n > 0$, $1 \leqslant i \leqslant n$, $k = 0, 1$, satisfying the pre-cubical relations $\partial_i^k \partial_j^l = \partial_{j-1}^l \partial_i^k$ for $i < j$.
2. The geometric realization $|M|$ of a pre-cubical set M is given as the quotient space $|M| = (\coprod_n M_n \times \square_n)_{/\equiv}$ under the equivalence relation induced from

$$\big(\partial_i^k(x), t\big) \equiv \big(x, \delta_i^k(t)\big), \quad x \in M_{n+1},\ t = (t_1, \dots, t_n) \in \square_n,$$

with

$$\delta_i^k(t) = (t_1, \dots, t_{i-1}, k, t_{i+1}, \dots, t_n).$$

In Section 3, we will make use of particular open sets in $|M|$, the *open stars* of *vertices* in M_0. The open star $St(x, M)$ of $x \in M_0$ consists of the union of the interiors of all cells of which x is a vertex.

Definition 2.2.

1. A continuous path $p = (p_1, \dots, p_k) : I \to \square_k$ is a *d-path* if every component $p_i : I \to I$, $1 \leqslant i \leqslant k$, is (not necessarily strictly) *increasing*.
2. A continuous path $p : I \to |M|$, M a $\square$-set, is a *d-path* if, for every $J \subseteq I$ such that the restriction $p_J : J \to |M|$ of p has a lift $\hat{p}_J^e : J \to e \times \square_k$, $e \in M_k$, that lift $\hat{p}_J^e$ is a d-path; alternatively, if p can be decomposed as a concatenation $p = |p_1| * \cdots * |p_l|$ of *d-paths* p_i in the cells $e_i \times \square_{n_i}$.

   The set of *all* d-paths in $|M|$ will be denoted $\vec{P}(|M|) \subset |M|^I$. It inherits a topology from the CO-topology on $|M|^I$.

Proposition 2.3. *The pair* $(|M|, \vec{P}(|M|))$ *is a d-space* (*in the sense of M. Grandis paper* [12]).

Proof. It is clear that $\vec{P}(|M|)$ contains the constant paths, that it is closed under concatenation and under increasing reparametrizations. $\square$

It is well known, that the topology of the geometric realization of a $\square$-set is metrizable: Choose compatible metrics d_n on $\square_n$ each inducing the standard topology and such that $d_{n+1}((\mathbf{x}, t), (\mathbf{y}, t)) = d_n(\mathbf{x}, \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in \square_n$ and $t \in I$. The inner pseudometric on the identification space $|X|$ given by the infimum of the length of chains in $|X|$ is then in fact a metric (on each connected component; cf. [2,17]).

From now on, without further saying, a pre-cubical complex X will be understood as the geometric realization of a $\square$-set M, with the d-space structure introduced above. For a more general discussion of cubical sets, we refer to [13].

As mentioned in the introduction (Section 1.2) and described in detail in [6], we will mainly work with the quotient space $\vec{T}(X)$ of directed *traces*, i.e., of d-paths *modulo reparametrization equivalence*, cf. Definition 1.1.4.

2.2. L_1-arc length and dihomotopies

[1]We will in the following also consider paths and in particular, generalizing Definition 2.2, d-paths in a pre-cubical complex X, defined on *arbitrary* intervals $[a,b] \subset [0,\infty[$. Moreover, we need to consider *Moore paths* $p : [0,\infty[\to X$ that are constant after some (arbitrary) parameter $T \geqslant 0$: $p(t) = p(T)$ for $t \geqslant T$; $p(T)$ will be called the *end point* for a Moore path. It is clear that any path defined on a finite interval $[0,T]$ can be considered as a Moore path by constant extension from its endpoint $p(T)$. For Moore paths in general compare e.g. [1]. The set $\vec{P}_M(X)$ of *all* Moore d-paths in X inherits a topology from the CO-topology on $X^{[0,\infty[}$.

2.2.1. L_1-arc length

The L_1-distance (aka Manhattan, taxi cab, or city block distance) between two points $\mathbf{x}^0 = (x_1^0,\dots,x_n^0)$ and $\mathbf{x}^1 = (x_1^1,\dots,x_n^1)$ in $\mathbf{R}^n$ is defined as $d_1(\mathbf{x}^0,\mathbf{x}^1) = \sum_1^n |x_i^1 - x_i^0|$. For $\mathbf{x}^0 \leqslant \mathbf{x}^1$, i.e., $x_i^0 \leqslant x_i^1$ for all i, we get: $d_1(\mathbf{x}^0,\mathbf{x}^1) = \sum_1^n x_i^1 - x_i^0$. The L_1-arc length of a *piecewise linear* path $p : [a,b] \to \mathbf{R}^n$ is given by $l_1(p) = \sum_i d_1(p(t_i), p(t_{i+1}))$.

For a piecewise linear *d-path* in $\mathbf{R}^n$, the intermediate terms cancel and $l_1(p) = d_1(p(a), p(b))$. In particular, the L_1-arc length for such d-paths depends *only on the end points*. The L_1-arc length of an arbitrary d-path $p : [a,b] \to \mathbf{R}^n$ is defined as $l_1(p) = \sup_{n\in\mathbf{N}, a=t_0<t_1<\cdots<t_{n-1}<t_n=b} d_1(p(t_i), p(t_{i+1})) = d_1(p(a), p(b))$! As a result, every d-path in $\mathbf{R}^n$ is L_1-rectifiable, and its L_1-arc length is equal to the L_1-distance of its end points – and hence positive unless p is constant. Remark also that the L_1-arc length is additive under concatenation of d-paths.

The definition of L_1-arc length above extends immediately to both piecewise linear paths and for d-paths contained in a cell in a pre-cubical set. Moreover, for $e = \partial_i^k f$, $p \in \vec{P}(|e|)$, the L_1-arc length of p is the same whatever we consider it as a d-path in e or in f. Hence the L_1-arc length of a piecewise linear path $p \in P(X)$ or of an arbitrary d-path $p \in \vec{P}(X)$ in a pre-cubical complex X may be defined as follows:

By Definition 2.2, every d-path $p \in \vec{P}(X)$ (and likewise, any piecewise linear path $p \in P(X)$) can be decomposed as a finite concatenation $p = |p_1| * \cdots * |p_l|$ of d-paths (piecewise linear paths) with p_i contained in *one* cell e_i for every i. The L_1-arc length $l_1(p) = \sum_1^l l_1(p_i)$ is well-defined and additive under concatenation. Moreover, it is invariant under reparametrization (by non-decreasing reparametrizations $\varphi \in \mathrm{Rep}_+(I) \subset I^I$, cf. [6]) and hence, in the directed case, an invariant of *traces* in $\vec{T}(X)$ (cf. Section 1.2).

Example 2.4. Let $\vec{S}^1$ (the oriented circle) denote *the* 1-dimensional cubical complex that realizes a $\square$-set with exactly one 0-cell and one 1-cell. The L_1-arc length of *the shortest* trace t_0 from $[s]$ to $[t]$, $s,t \in I = \square_1$, is $l_1(t_0) = t - s \bmod 1$; i.e., the (non-negative) fractional part of $t - s$. The other traces t_n between the same end points (cycling around n-times) have L_1-arc length $l_1(t_n) = n + l_1(t_0)$. In particular, different d-paths with the same end points may have different L_1-arc lengths; compare with Section 2.2.2.

Remark 2.5. L_1-arc length increases by a factor 2 under cubical barycentric subdivision.

For a pre-cubical complex $X = |M|$, consider furthermore the function $s : X \to S^1$ with $s([e; x_1,\dots,x_n]) = \sum x_i \bmod 1$, $e \in M_n$. The glueing conditions (Definition 2.1) show that s is well-defined and continuous; moreover, it sends d-paths in X into d-paths in $\vec{S}^1$ (cf. Example 2.4) and is thus a d-map [12].

Note that $l_1(s \circ q) \leqslant l_1(q)$ for a piecewise linear (not necessarily directed) path $q \in P(X)$ – essentially since $|\sum y_i| \leqslant \sum |y_i|$; moreover $l_1(s \circ p) = l_1(p)$ for every d-path $p \in \vec{P}(X)$. One may think of $l_1(s \circ p)$ as the length of a lift of $s \circ p$ under the exponential map from the reals onto S^1.

Lemma 2.6. *Let $p \in \vec{P}(X)$ denote a d-path with $l_1(p) \leqslant 0.5$. Then $l_1(p) = d_1(p(0), p(1))$.*

Proof. The condition implies that $l_1(p) = l_1(s \circ p) = d_1((s \circ p)(0), (s \circ p)(1))$. Let $q \in P(X)(p(0), p(1))$ denote any piecewise linear path. Then $l_1(q) \geqslant l_1(s \circ q) \geqslant d_1((s \circ p)(0), (s \circ p)(1)) = l_1(p)$. Taking the infimum over all $q \in P(X)(p(0), p(1))$, it is seen that $d_1(p(0), p(1)) = \inf l_1(q) = l_1(p)$. $\square$

Altogether, for a given pre-cubical complex X, the L_1-arc length provides us with functions $l_1 : \vec{P}(X) \to \mathbf{R}_{\geqslant 0}$ on the space of d-paths and $l_1^M : \vec{P}_M(X) \to \mathbf{R}_{\geqslant 0}$ on the space of Moore d-paths in X – and also on the quotient space of traces $\vec{T}(X)$ (modulo reparametrizations).

Proposition 2.7. *The L_1-arc length functions l_1 and l_1^M are continuous.*

[1] I would like to thank the referee for a question the answer to which allowed me to generalize the results in Section 2.2 compared to the original version.

Proof. For a given d-path $p \in \vec{P}(X)$ and $0 < \varepsilon < 0.25$, there is a decomposition of the unit interval $I = I_1 \cup \cdots \cup I_k = [0, t_1] \cup \cdots \cup [t_{k-1}, t_k]$ such that $l_1(p|I_j) \leqslant 0.5 - 2\varepsilon$. For q with $d_1(q, p) < \varepsilon$ – in the supremum metric – we have $(s \circ q)(I_j) = [(s \circ q)(t_{j-1}), (s \circ q)(t_j)] \subseteq [(s \circ p)(t_{j-1}) - \varepsilon, (s \circ p)(t_j) + \varepsilon]$, an interval on $\vec{S}^1$ of length at most 0.5. Since q is a d-path, $l_1(q|I_j) = l_1((s \circ q)|I_j) \leqslant 0.5$ and thus, by Lemma 2.6, $l_1(q|I_j) = d_1(q(t_{j-1}), q(t_j))$. Hence $|l_1(q|I_j) - l_1(p|I_j)| = |d_1(q(t_{j-1}), q(t_j)) - d_1(p(t_{j-1}), p(t_j))| < 2\varepsilon$, and hence $|l_1(q) - l_1(p)| < 2n\varepsilon$.

In the case of Moore paths, replace the decomposition of I by a decomposition of $\mathbf{R}$ into a finite number of intervals on which p has length bounded by $0.5 - 2\varepsilon$, the last one of type $[t_{k-1}, \infty[$. □

Remark 2.8. It is *not* possible to extend the L_1-arc length function to a *continuous* function on the space of *all* (i.e., not necessarily directed) paths that are L_1-rectifiable on cells; not even to the piecewise linear ones. Consider for instance a sequence of square paths with the same central point (in a square in the Euclidean plane) with side length $\frac{1}{n}$ and winding number n; each path has L_1-arc length 4, but the sequence converges to a constant path of arc length 0 in the compact-open topology.

2.2.2. Dihomotopies preserve the L_1-arc length

A one-parameter family of d-paths in a d-space is called a *dihomotopy* (cf. [9]); formally:

Definition 2.9. Let X denote a d-space.

1. A continuous map $H : \vec{I} \times I \to X$ is called a dihomotopy if every map $H_t : \vec{I} \to X$, $t \in I$, is a d-path.
2. If, moreover, $H(0, t)$ and $H(1, t)$ are *fixed* under the dihomotopy, then H is a dihomotopy rel end points.

Remark 2.10. There are both more general and more special dihomotopies in the literature. D-homotopies $H : X \times I \to Y$ of d-maps (preserving d-paths) between arbitrary d-spaces are investigated in e.g. [12,18]. In order to obtain a van-Kampen theorem for the fundamental category of a d-space, Grandis introduces a d-homotopy between d-paths in which the paths $H(s, t)$ for given s are concatenations of paths that are either d-paths themselves or reverses of d-paths ("zig-zags"). Lisbeth Fajstrup showed in [7] that the classification of d-paths in a (geometric) pre-cubical complex yields the same result whatever one considers classification up to ordinary dihomotopy (as in Definition 2.9) or up to zig-zag d-homotopy (as above).

Corollary 2.11. *For a dihomotopy $H : \vec{I} \times I \to X$ of d-paths in a pre-cubical complex, the L_1-arc length map $l_1(H) : I \to \mathbf{R}_{\geqslant 0}$, $t \mapsto l_1(H_t)$ is continuous.*

Proof. Since the adjoint map $H^* : I \to \vec{P}(X)$, $t \mapsto H_t$, is continuous, this follows from Proposition 2.7. □

In ordinary topology, the (L_1-)arc length of paths can grow arbitrarily in a homotopy class with fixed end points; just concatenate with nulhomotopic zig-zag paths! But for dihomotopies of *d-paths* in precubical complexes this is not at all the case:

Proposition 2.12. *D-paths in a precubical complex X with the same start and end point that are dihomotopic rel end points have the* same *L_1-arc length.*

Proof. Let $p, q \in \vec{P}(X)(x, y)$. Using the map $s : X \to S^1$ introduced after Remark 2.5 above, we get:

$$l_1(p) = l_1(s \circ p) \equiv d_1\big(s(x), s(y)\big) \equiv l_1(s \circ q) = l_1(q),$$

where $\equiv$ is congruence on $\mathbf{R}$ mod $\mathbf{Z}$. By Corollary 2.11, L_1-arc length varies continuously along a dihomotopy; hence it needs to be constant. □

2.3. Natural parametrization of a d-path in a pre-cubical complex

For a d-path $p : [0, l] \to X$ or $p : [0, \infty[\to X$, let $p_t : [0, t] \to X$, $t \leqslant l$, denote the sub-d-path obtained by restriction of the domain. Then the real function $l_p : [0, l] \to [0, l(p)]$, $t \mapsto l_1(p_t)$, is contained in the space $\vec{P}(\mathbf{R}_{\geqslant 0})(0)$ consisting of weakly increasing paths sending 0 into 0 and equipped with the compact-open topology. For a Moore d-path defined on $[0, \infty[$, $l_1(p_t) = l_1(p)$ for $t \geqslant T$ if p is constant on $[T, \infty[$. Furthermore, we define length functionals $l_1 : \vec{P}(X) \to \vec{P}_M(\mathbf{R}_{\geqslant 0})(0)$, resp. $l_1 : \vec{P}_M(X) \to \vec{P}_M(\mathbf{R}_{\geqslant 0})(0)$ by $l_1(p) = l_p$. From Proposition 2.7 we deduce:

Lemma 2.13. *The L_l-arc length functionals l_1 on $\vec{P}(X)$ and on $\vec{P}_M(X)$ are* continuous.

Definition 2.14.

1. A (Moore) d-path p in X is called *regular* [6] if the function l_p is *strictly* increasing. The subspaces of all regular (Moore) d-paths are called $\vec{R}(X) \subset \vec{P}(X)$, resp. $\vec{R}_M(X) \subset \vec{P}_M(X)$.
2. A d-path $p : [0, l] \to X$ is called *natural* if $l_1(p_t) = t$ for all $t \in [0, l]$; in particular, $l = l_1(p)$.
   A Moore d-path $p : [0, \infty[\to X$ is called *natural* if, for some finite T: $l_1(p_t) = t$ for all $t \leqslant T$ and $p(t) = p(T)$ for $t \geqslant T$.
3. The subspace of all natural d-paths in X is denoted $\vec{N}(X) \subset \vec{R}_M(X)$ – with subspaces $\vec{N}(X)(A_0, A_1)$ and $\vec{N}(X)(x_0, x_1)$ of d-paths starting in $A_0 \subseteq X$, resp. $x_0 \in X$, and ending in $A_1 \subseteq X$, resp. $x_1 \in X$.

To compare $\vec{P}(X)$ and $\vec{N}(X)$, we consider the following two maps:

Normalization *norm* $: \vec{P}_M(X) \to \vec{P}(X)$, $norm(p)(t) = p(l_1(p) \cdot t)$.

We shall also make use of the restrictions of the normalization map to the space $\vec{R}_M(X)$ of regular d-paths and to the space $\vec{N}(X)$ of natural d-paths – with image contained in $\vec{R}(X)$.

Naturalization *nat* $: \vec{P}(X) \to \vec{N}(X)$, $nat(q)(s) = q(l_q^{-1}(s))$.

In general, $l_q^{-1}(s)$ may be a non-trivial interval; nevertheless, $q(l_q^{-1}(s))$ is well-defined, and $nat(q)$ is continuous; compare Proposition 3.7 in [6] and Proposition 2.2 in [19].

Proposition 2.15.

1. *The map nat is invariant under the action of the monoid* $\mathrm{Rep}_+(I)$ *of non-decreasing reparametrizations on* $\vec{P}(X)$.
2. *The maps norm and nat are continuous.*

Proof.

1. L_1-arc length is invariant under this action.
2. The continuity of *norm* follows from Proposition 2.7.
   To prove the continuity of *nat* at $p \in \vec{P}(X)$, consider an ε-neighborhood U of p in the metric inherited from the L_1-metric on the individual cells (cf. Section 3.1.1). As in the proof of Proposition 2.7, there is an integer n such that $|l_1(p_t) - l_1(q_t)| < 2n\varepsilon$ for all $t \in I$ and $q \in U$. Given $0 \leqslant s \leqslant l(p)$, choose $t_1, t_2 \in I$ such that $nat(p)(s) = p(t_1)$ and $nat(q)(s) = p(t_2)$; in particular, $l_1(p_{t_1}) = l_1(q_{t_2}) = s$. Then $d_1(nat(p)(s), nat(q)(s)) = d_1(p(t_1), q(t_2)) \leqslant d_1(p(t_1), p(t_2)) + d_1(p(t_2), q(t_2)) \leqslant |l_1(p_{t_2}) - l_1(p_{t_1})| + \varepsilon < (2n+1)\varepsilon$. □

2.4. Homeomorphisms and homotopy equivalences

We use the two maps *norm* and *nat* to show that the spaces of d-paths considered so far (for definitions, we refer to Section 1.2) are all homotopy equivalent. It will be particularly useful (in Section 3) that traces in pre-cubical complexes can be represented by *natural* d-paths up to homeomorphism. Remark that the topology of path spaces only becomes interesting when fixing both end points – the space $\vec{P}(X)$ itself is homotopy equivalent to X!

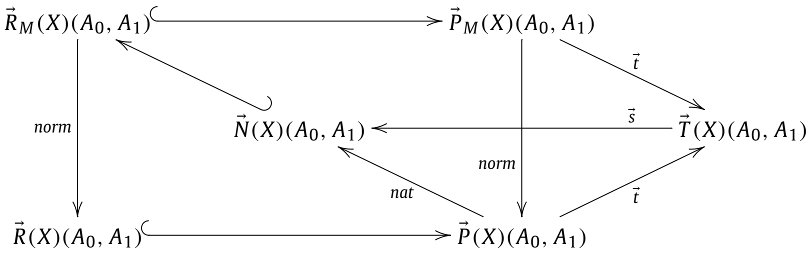
Proposition 2.16. *Let X be a pre-cubical complex with subsets $A_0, A_1 \subset X$.*

1. *The diagrams*

$$\begin{array}{ccc} \vec{P}(X)(A_0, A_1) & \xrightarrow{l_1} & \vec{P}(\mathbf{R})(0) \\ \downarrow{\scriptstyle nat} & & \downarrow{\scriptstyle \text{end point}} \\ \vec{N}(X)(A_0, A_1) & \xrightarrow{l_1} & \mathbf{R}_{\geqslant 0} \end{array} , \qquad \begin{array}{ccc} \vec{R}(X)(A_0, A_1) & \xrightarrow{l_1} & \vec{R}(\mathbf{R})(0) \\ \downarrow{\scriptstyle nat} & & \downarrow{\scriptstyle \text{end point}} \\ \vec{N}(X)(A_0, A_1) & \xrightarrow{l_1} & \mathbf{R}_{\geqslant 0} \end{array}$$

   are pullback diagrams of topological spaces.
2. *The restriction $t_N : \vec{N}(X)(A_0, A_1) \to \vec{T}(X)(A_0, A_1)$ of the trace map $\vec{t} : \vec{P}_M(X)(A_0, A_1) \to \vec{T}(X)(A_0, A_1)$ is a* homeomorphism (*with inverse $\vec{s} : \vec{T}(X)(A_0, A_1) \to \vec{N}(X)(A_0, A_1)$*).

3. *All maps in the diagram*



are homotopy equivalences.

Proof.

1. The inverse map from the fibered product into $\vec{P}(X)(A_0, A_1)$ is given by $(q, \varphi) \mapsto norm(q \circ \varphi)$.
2. We define a section $\vec{s} : \vec{T}(X)(A_0, A_1) \to \vec{N}(X)(A_0, A_1)$ by $\vec{s}(\vec{t}(p)) = nat(p)$ for $p \in \vec{P}(X)(A_0, A_1)$. It is well-defined by Proposition 2.15; it is continuous since *nat* is so, and since $\vec{T}(X)(A_0, A_1)$ has the quotient topology. It is trivial to check that the maps $\vec{t}_N$ and $\vec{s}$ are inverse to each other.
3. By definition, $nat \circ norm$ is the identity on $\vec{N}(X)(A_0, A_1)$.
   Using the pullback diagram in 1. above, the self map $norm \circ nat$ translates to the self map on $\vec{N}(X)(A_0, A_1) \times_{\mathbf{R}_{\geqslant 0}} \vec{P}(\mathbf{R}_{\geqslant 0})(0)$ sending(q, φ) into $(q, l_1(id|_{[0,l_1(q)]}))$. A homotopy to the identity is given by:

$$(q, \varphi; t) \mapsto \big(q, (1-t) \cdot \varphi + t \cdot l_1(id|_{[0,l_1(q)]})\big). \qquad \square$$

Remark 2.17. The only tool used heavily in this section is that of a continuous additive length functional on the space of d-paths. It should be possible to recover the results of this section for d-spaces with such a length function in greater generality.

3. General properties of trace spaces

3.1. Trace spaces as metric spaces

3.1.1. Trace spaces are metrizable

The compact-open topology on $\vec{P}(X)$ is induced from the supremum metric given by $d(p, q) = \max_{t \in I} d(p(t), q(t))$. Likewise, the compact-open topology on $\vec{P}_M(X)$ is inherited from the supremum metric given by $d(p, q) = \max_{s \in [0,\infty[} d(p(s), q(s))$. In the following, we will in particular use that the topology can be seen as inherited from the L_1-metric, cf. Section 2.2.1.

Corollary 3.1. *For a pre-cubical complex X and for $A_0, A_1 \subset X$, the spaces $\vec{N}(X)(A_0, A_1)$, $\vec{T}(X)(A_0, A_1)$, $\vec{R}(X)(A_0, A_1)$ and $\vec{P}(X)(A_0, A_1)$ are all Hausdorff and paracompact.*

Proof. A metric space is Hausdorff (obvious) and paracompact [24,22]. $\square$

3.1.2. Separability

Proposition 3.2. *For a finite pre-cubical complex X, the spaces $\vec{N}(X) \subset \vec{R}(X) \subset \vec{P}(X)$ and $\vec{T}(X)$ are all separable metric spaces. Likewise subspaces such as $\vec{N}(X)(A_0, A_1)$, $\vec{R}(X)(A_0, A_1)$, $\vec{P}(X)(A_0, A_1)$ and $\vec{T}(X)(A_0, A_1)$.*

In the proof, we make use of

Lemma 3.3. *Let $p, q \in \vec{P}(X)$ denote two d-paths in a finite-dimensional pre-cubical complex X sharing start point and carrier sequence. Then $|l_1(p) - l_1(q)| < \dim X$. If they share the same end point, they have equal length.*

Proof. First, we arrive at d-paths p^* and q^* with the same end point by concatenating with linear d-paths to the supremum of the end points within the (same) final cell. Denote the sequence of carriers by $e_1, \ldots, e_k$ and choose $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$ such that $p(t_i) \in e_i$, $1 \leqslant i \leqslant k$. After a reparametrization, we may assume that $q(t_i) \in e_i$, as well (cf. [6]). Since $e_i \subset e_{i+1}$ or $e_i \supset e_{i+1}$, the restrictions of both p^* and of q^* to $[t_i, t_{i+1}]$ are contained in the same cell, and the two paths are therefore dihomotopic by a cellwise linear dihomotopy. By Proposition 2.12, $l_1(p^*) = l_1(q^*)$, and hence $|l_1(p) - l_1(q)| \leqslant d_1(p(1), q(1)) < \dim X$. If p and q share the same end point, then $p = p^*$ and $q = q^*$. $\square$

Proof of Proposition 3.2. At first, we concentrate on the space $\vec{N}(X)$. From Proposition 2.16(2), we then get separability of the space $\vec{T}(X)$ for free. In [7], L. Fajstrup described an explicit method to approximate a given d-path p in a pre-cubical complex X by a (dihomotopic) d-path q on the 1-*skeleton* of that complex; this approximation may be given a natural parametrization, as well. Such an approximation is not well-determined, in general; on the other hand, every such approximation has the property that it shares the carrier sequence with the original d-path. We will now show, that p and q cannot be too far apart in the L_1-metric:

Choose $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$ and $0 \leqslant u_i \leqslant t_i$ such that $p(t_i) \in e_i \cap e_{i+1}$, $1 \leqslant i < k$, and such that $q(u_i)$ is the minimal vertex in $e_i \cap e_{i+1}$. Let p_i, resp. q_i, denote the restrictions of p to $[0, t_i]$, resp. q to $[0, u_i]$ and $q_i^* : [0, t_i] \to X$ the concatenation of q_i with the L_1-arc length parametrized linear path connecting $q(u_i)$ and $p(t_i)$. The d-paths p_i and q_i^* share start and end point and carrier sequence; by Lemma 3.3, $u_i \leqslant t_i = l_1(p_i) = l_1(q_i^*) = u_i + d_1(q(u_i), p(t_i)) < u_i + \dim X$. Hence $d_1(p(t_i), q(t_i)) \leqslant d_1(p(t_i), q(u_i)) + l_1(q|_{[u_i,t_i]}) \leqslant 2 \dim X$. Finally, for $t \in [t_i, t_{i+1}]$, we have: $d_1(p(t), q(t)) \leqslant l(p_{[t_i,t]}) + d_1(p(t_i), q(t_i)) + l_1(q_{[t_i,t]}) < 4 \dim X$.

Next, we apply this argument to the Nth cubical subdivision $X^{(N)}$ of the original complex X. We conclude – with q now on the 1-skeleton of the *subdivided* complex – from Remark 2.5 that $d_1(p(t), q(t)) \leqslant 4 \frac{\dim X}{2^N}$ (with respect to the original distance function *before* subdivision).

For every pair of non-negative integers N and L, there is a *finite* set $\vec{N}_L(X_1^{(N)})$ of natural d-paths of L_1-arc length bounded by $L \in \mathbf{N}$ on the 1-skeleton of the Nth cubical subdivision of a finite pre-cubical complex X; these are in fact determined by the (finitely many) vertices that may occur as values $p(t)$ of such a path p at $t = \frac{k}{2^N}$, $0 \leqslant k \leqslant 2^N L$. The union $\bigcup_{N,L \in \mathbf{N}} \vec{N}_L(X_1^{(N)})$ of these path sets is countable and dense in $\vec{N}(X)$; hence, $\vec{N}(X)$ is separable.

By Proposition 2.16(1), the space $\vec{P}(X)$ can be viewed as a subspace of the product space $\vec{N}(X) \times \vec{P}(\mathbf{R}_{\geqslant 0})(0)$. The second factor is a subspace of the metric space $\mathbf{R}^I$ of continuous functions on the unit interval. This metric space is separable by the Weierstrass approximation theorem – polynomials with rational coefficients form a dense countable subset.

Finally, products of separable metric spaces are separable, and subspaces of separable metric spaces are separable. □

3.2. *Trace spaces are (locally) compact*

A space of d-paths is never compact – unless it only contains constant paths. This is so since the space of reparametrizations $\mathrm{Rep}_+(I) = \vec{P}(I)(0, 1)$ is not compact; it is not even equicontinuous, a necessary condition for compactness by the Arzelà–Ascoli theorem (cf. e.g. [5,16]).

Trace spaces are in general not compact either. If the d-space X contains a non-trivial loop based at $x_0 \in X$, then the closed subspace $\vec{T}(X)(x_0, x_0)$ has d-paths of infinitely many L_1-arc lengths and thus by Proposition 2.12 infinitely many connected components whence it cannot be compact. But compactness results are available if one bounds the L_1-arc lengths of d-paths:

Let $\vec{N}_L(X) \subset \vec{N}(X)$ consist of all natural d-paths of L_1-arc length less than or equal to L introduced in Proposition 3.2. A subset $H \subseteq \vec{T}(X)$ is called *of bounded L_1-arc length* if there exists $L \geqslant 0$ such that $H \subseteq \vec{t}(\vec{N}_L(X))$.

In the following Proposition 3.4 and its corollaries, X will always denote a *finite* – hence compact – pre-cubical complex:

Proposition 3.4. *A subset $H \subseteq \vec{T}(X)$ of bounded L_1-arc length is* relatively compact.

Corollary 3.5. *Trace space $\vec{T}(X)$ is* locally compact.

Corollary 3.6. *For $x_0, x_1 \in X$, every d-homotopy class (connected component) in $\vec{T}(X)(x_0, x_1)$ is compact.*

Proof of Proposition 3.4. Via the homeomorphism $t_{\vec{N}}$, we regard H as a subspace of $\vec{N}_L(X) = \{p = p_1 * p_2: [0, L] \to X \mid p_1 \text{ natural}, p_2 \text{ constant}\}$ and apply the Arzelà–Ascoli theorem (e.g. [5,16]). The conditions are satisfied since X is compact and since $\vec{N}(X)$ – consisting of distance preserving paths – is clearly equicontinuous. □

3.3. *Trace spaces have the homotopy type of a CW complex*

John Milnor investigated in [15] conditions on spaces that ensure that certain mapping spaces have the homotopy type of a CW-complex[2]: We check that these criteria can be applied to spaces of traces in a pre-cubical complex satisfying an extra condition, and we conclude that those spaces have the homotopy type of a CW-complex. This allows us to conclude that a weak homotopy equivalence between trace spaces actually is a (strong) homotopy equivalence; this will be used several times in [20].

[2] I am grateful to W. Lück (Univ. Münster) for drawing my attention to [15].

Definition 3.7. ([15]) A topological space A is called ELCX (equi locally convex) if there exists

1. a neighborhood U of the diagonal $\Delta A \subset A \times A$ and a map $\lambda : U \times I \to A$ satisfying $\lambda(a,b,0) = a$, $\lambda(a,b,1) = b$ for all $(a,b) \in U$, and $\lambda(a,a,t) = a$ for all $a \in A, t \in I$;
2. an open covering of A by sets V_β such that $V_\beta \times V_\beta \subset U$ and $\lambda(V_\beta \times V_\beta \times I) = V_\beta$.

Lemma 3.8. *(A special case of [15, Lemma 4].) Every paracompact ELCX space has the homotopy type of a CW-complex.*

In fact, Milnor shows that a paracompact ELCX space is dominated by a simplicial complex and thus (see e.g. [14, appendix, Proposition A.11]) homotopy equivalent to a CW-complex.

3.3.1. Non-self linked pre-cubical complexes are ELCX

In the proof, we make use of the following additional condition to a pre-cubical complex as the geometric realization of special $\square$-sets:

Definition 3.9. A pre-cubical complex M is called *non-self-linked* (cf. [9]) if for all n, $x \in M_n$ and $0 < i \leqslant n$, the $2^i\binom{n}{i}$ iterated faces $\partial^{k_1}_{l_1} \cdots \partial^{k_i}_{l_i} x \in M_{n-i}$, $k_i = 0, 1$, $1 \leqslant l_1 < \cdots < l_i \leqslant n$, are all different.

In a non-self-linked pre-cubical complex, the map $\square_n \simeq \square_n \times e \to |M|$ is *injective* for every n-cell $e \in M_n$. In particular, every element $m \in |M|$ in the image of this map has *uniquely* determined coordinates in $\square_n$, cf. [9]. Moreover, every element $x \in |M|$ has a unique *carrier* cell $e(x) \in M_n, n \geqslant 0$ such that x comes from an element in the *interior* $\square^o_n$ under the restriction of the quotient map to $\square_n \times e(x)$.

Proposition 3.10. *A non-self linked pre-cubical complex X is ELCX.*

In the proof, we need the following

Lemma 3.11. *There exists a continuous "average" map $m : \{(x,y) \in I^2 \mid |y - x| \neq 1\} \to I$ preserving (partial) orders*[3] *that satisfies for all (x,y) in the domain*:

1. *for $\alpha = 0, 1$, $x = \alpha$ or $y = \alpha$ implies $m(x,y) = \alpha$;*
2. $m(x,x) = x$;
3. $\min(x,y) \leqslant m(x,y) \leqslant \max(x,y)$.

Proof of Lemma 3.11. It is easy to check that the map $m(x,y) = \frac{\min(x,y)}{1-|y-x|}$ (increasing linearly from 0 to 1 on parallels to the diagonal from the lower to the upper boundary of I^2) has properties (1)–(3). To check that it preserves partial orders, use either the heuristic description above or calculate partial derivatives. $\square$

Proof of Proposition 3.10. As in the proof of Lemma 2 in [15], let V_β denote the open star neighborhood (cf. Section 2.1) of a vertex β in the cubical complex and let $U = \bigcup_\beta V_\beta \times V_\beta$. For every $x \in X$, let $e(x)$ denote the carrier cell containing x in its interior. Below, we define an "average" map $\mu : U \to X$ with the property that $\mu(x,y) \in \overline{e(x)} \cap \overline{e(y)}$ for all $(x,y) \in U$ and $\mu(x,x) = x$ for all $x \in X$. Then, as in [15], the path $\lambda(x,y,t), t \in I$, is given as the concatenation of the canonical line parametrizations connecting first x with $\mu(x,y)$ within $e(x)$ and then $\mu(x,y)$ with y within $e(y)$.

For $(x,y) \in V_\beta \times V_\beta \subseteq U$, consider the nonempty cell $\overline{e(x)} \cap \overline{e(y)}$ (containing β) as the closure of an iterated face of both $e(x)$ and $e(y)$. Since X is non-self linked (cf. Definition 3.9, we may assume (after reordering the coordinates) that

- $\overline{e(x)} = I^k \times I^l$, $\overline{e(y)} = I^k \times I^m$,
- $(x_1, \ldots, x_k; x'_1, \ldots, x'_l) \in \overline{e(x)} \cap \overline{e(y)} \subset \overline{e(x)}$ if $x'_i = \alpha_i$, $1 \leqslant i \leqslant l$,
- $(y_1, \ldots, y_k; y'_1, \ldots, y'_m) \in \overline{e(x)} \cap \overline{e(y)} \subset \overline{e(y)}$ if $y'_j = \beta_j$, $1 \leqslant j \leqslant m$,

for certain $\alpha_i, \beta_j \in \{0,1\}$. Using the map m from Lemma 3.11, represent $\mu(x,y)$ by $(m(x_1,y_1), \ldots, m(x_k,y_k); \alpha_1, \ldots, \alpha_l) \in \overline{e(x)} \cap \overline{e(y)} \subset \overline{e(x)}$ – or by $(m(x_1,y_1), \ldots, m(x_k,y_k); \beta_1, \ldots, \beta_m) \in \overline{e(x)} \cap \overline{e(y)} \subset \overline{e(y)}$). Remark that the average $m(x_i,y_i)$ is defined, since x and y both belong to the *open* star of the common vertex β; in particular, $|x_i - y_i| < 1$. Property (1) of m in Lemma 3.11 makes sure that μ factors over the face relations and thus defines a continuous map on U. $\square$

[3] A d-map in the terminology of [12].

Remark 3.12.

1. Since the map m from Lemma 3.11 preserves partial orders, the map μ and therefore also λ in the proof of Proposition 3.10 are *directed* in the following sense: If $x \leqslant x'$ within $\overline{e(x)}$ and $y \leqslant y'$ within $\overline{e(y)}$, then $\mu(x, y) \leqslant \mu(x', y')$ within $\overline{e(x)} \cap \overline{e(y)}$ and thus $\lambda(x, y, t) \leqslant \lambda(x', y', t)$ within $\overline{e(x)}$, resp. $\overline{e(y)}$. This will be essential in the next Section 3.3.2.
2. μ cell-preserving implies: $\lambda(x, y, t) \in \overline{e(x)}$ for $x \leqslant 0.5$ and $\lambda(x, y, t) \in \overline{e(y)}$ for $y \geqslant 0.5$. In particular, both maps preserve open stars of vertices.

Lemma 3.13. *With respect to the metric d_1 on X induced by the L_1-metric d_1 on each cell (cf. Section* 3.1.1*), the maps μ and λ satisfy for $(x, y) \in U$, $t \in I$:*

- $d_1(x, \mu(x, y)), d_1(y, \mu(x, y)) \leqslant d(x, y)$.
- $d_1(x, \lambda(x, y, t)), d_1(y, \lambda(x, y, t)) \leqslant d(x, y)$.

Proof. If x and y are contained in one cell, these properties follow immediately from Lemma 3.11(3). If not, represent $e(x)$ and $e(y)$ as in the proof of Proposition 3.10. Then $d(x, y) = \sum |x_i - y_i| + \sum |x'_j - \alpha_j| + \sum |y'_k - \beta_k| \geqslant \sum |x_i - m(x_i, y_i)| + \sum |x'_j - \alpha_j| = d(x, m(x, y))$ by Lemma 3.11(3); similarly for $d(y, m(x, y))$ and for the distances to $\lambda(x, y, t)$. □

3.3.2. Spaces of d-paths are ECLX

Proposition 3.14. *For every non-self-linked pre-cubical complex X and elements $x_0, x_1 \in X$, the spaces $\vec{P}(X)$ and $\vec{P}(X)(x_0, x_1)$ are ELCX.*

Proof. For a partition $I = I_1 \cup \cdots \cup I_k$ into finitely many closed intervals and a sequence $\beta_1, \ldots, \beta_k$ of vertices in X, let $\vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k)$ denote the subspace of all d-paths p such that $p(I_j)$ is contained in the open star of β_j for $1 \leqslant j \leqslant k$; it is an open subspace in the topology induced on $\vec{P}(X)$ from the compact-open topology. Those subsets play the role of the V_β in Definition 3.7.

Let $U \subset \vec{P}(X) \times \vec{P}(X)$ denote the union of all "squares" of these open subspaces; obviously an open neighborhood of the diagonal in $\vec{P}(X)$. A continuous map $\Lambda : U \times I \to X^I$ is given by $\Lambda((p, q), t)(s) = \lambda(p(s), q(s), t)$. By Remark 3.12(1), the image of Λ is in fact contained in $\vec{P}(X)$. Since λ, by Remark 3.12(2), preserves open stars, Λ maps $\vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k) \times \vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k) \times I$ into $\vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k)$.

By Lemma 3.11(2), the map λ is constant on the diagonal. Hence the map Λ above preserves end points of d-paths along the parameter interval I, whence $\vec{P}(X)(x_0, x_1)$ is ELCX, as well. □

3.3.3. Spaces of d-paths and trace spaces have the homotopy type of a CW-complex

A combination of Corollary 3.1, Proposition 3.14, Lemma 3.8 and Proposition 2.16 yields:

Proposition 3.15. *For every non-self-linked pre-cubical complex X and for all elements $x_0, x_1 \in X$, the spaces*

1. $\vec{P}(X)$ *and* $\vec{P}(X)(x_0, x_1)$;
2. $\vec{T}(X)$ *and* $\vec{T}(X)(x_0, x_1)$

have the homotopy type of a CW-complex.

The interest here is in the di-pointed versions, since, as remarked in Section 2.4, $\vec{P}(X) \simeq \vec{T}(X) \simeq X$.

3.4. Trace spaces are locally contractible

Using similar techniques, we prove local contractibility of path spaces and of trace spaces; this property is necessary and gives hope for inductive calculations of algebraic topological invariants of such spaces. Note that there are two versions of local contractibility:

In a strongly locally contractible space Y, for every $y \in Y$ and every open neighborhood U of y, there exists a neighborhood of $V \subseteq U$ of y that contracts to y in V. In a weakly locally contractible space Z, for every open neighborhood U of z, there exists a neighborhood of $V \subseteq U$ of z that contracts to z in U.

Proposition 3.16. *For every non-self-linked pre-cubical complex X and for all elements $x_0, x_1 \in X$, the spaces $\vec{P}(X)$ and $\vec{P}(X)(x_0, x_1)$, are strongly locally contractible.*

Proof. For every d-path $p \in \vec{P}(X)$, there exist a partition $I_1, \ldots, I_k$ of the unit interval I and an open star sequence $V_{\beta_1}, \ldots, V_{\beta_k}$ such that p is contained in the open set $\vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k) \subset \vec{P}(X)$; see the proof of Proposition 3.14.

Choose $\varepsilon > 0$ such that $U_\varepsilon(p) \subseteq \vec{P}(X)(I_1, \ldots, I_k; \beta_1, \ldots, \beta_k)$ (with respect to the metric induced by the infinity metric on X and thus $\vec{P}(X)$; cf. Section 3.1.1). By Lemma 3.13, the map Λ from the proof of Proposition 3.14 restricts to a contraction $U_\varepsilon(p) \times \{p\} \times I$ to p *within* $U_\varepsilon(p)$.

The same proof applies to the relative case, since Λ preserves end points. □

Proposition 3.17. *For every non-self-linked pre-cubical complex X and for all elements $x_0, x_1 \in X$, the spaces $\vec{T}(X)$ and $\vec{T}(X)(x_0, x_1)$ are weakly locally contractible.*

Proof. As earlier, Proposition 2.16(2) allows us to represent $\vec{T}(X)$ by the homeomorphic space $\vec{N}(X)$ of L_1-arc length parametrized d-paths. The trouble is that the map Λ, defined on a neighborhood of the diagonal in $\vec{N}(X)$, will in general leave $\vec{N}(X)$. We have to replace it by $\Lambda_N : (U \cap \vec{N}(X)) \times I \to \vec{N}(X)$ with $\Lambda_N(p, q, t) = nat(\Lambda(p, q, t))$ with naturalization *nat* as defined in Section 2.3.

With $\varepsilon > 0$ chosen as in the proof of Proposition 3.16, we prove that Λ_N contracts $U_{\frac{\varepsilon}{2}}(p)$ to $p \in \vec{N}(X)$ within $U_\varepsilon(p)$: By Proposition 3.16, the contraction Λ itself preserves $U_{\frac{\varepsilon}{2}}(p)$. Using Lemma 3.13 and the argument comparing L_1-arc length functions of close dihomotopic d-paths from the proof of Proposition 3.2, we conclude for a given $q \in U_{\frac{\varepsilon}{2}}(p)$ for the L_1-arc lengths of all intermediate d-paths $\Lambda((q, p), t)$ at s: $|l_1(\Lambda((q, p), t)(s)) - s| < \frac{\varepsilon}{2}$. Thus, the reparametrization $nat(\Lambda((q, p), t))$ differs only slightly from the original $\Lambda((q, p), t)$: There is a function $s(t)$ with $|s(t) - s| < \frac{\varepsilon}{2}$ for all $t \in I$ such that $nat(\Lambda((q, p), t))(s) = \Lambda((q, p), t)(s(t))$. As a result, $d(p(s), nat(\Lambda((q, p), t))(s)) < d(p(s), p(s(t))) + d(p(s(t)), \Lambda((q, p), t)(s(t))) < \varepsilon$ for all $s, t \in I$. □

Question Are the results from Sections 3.3 and 3.4 still valid for traces in pre-cubical complexes with self-links?

4. Conclusion and further work

In Section 2, we have established that traces in pre-cubical complexes have a nice and useful representation by *natural* d-paths. Making use of this representation, it is shown in Section 3 that trace spaces have nice topological properties: they can be viewed as separable metric spaces, they are locally contractible and locally compact, and they have the homotopy type of a CW-complex.

All these properties are applied in [20] to compare trace spaces with subspaces of particular traces that give rise to additional combinatorial structure; e.g., traces that are *piecewise linear*, i.e., linear on each individual cube in the range. Weak homotopy equivalences between several of these trace spaces will be established using Smale's version [23] of the Vietoris–Begle theorem; the properties of trace spaces granted by the results of this paper are needed as conditions to apply this theorem. In many cases, Proposition 3.15 will allow us to conclude moreover, that the spaces involved are actually homotopy equivalent.

References

[1] D.R. Andersson, F.X. Conolly, H.J. Munkholm, A comparison of continuously controlled and controlled K-theory, Topology Appl. 71 (1) (1996) 9–46.
[2] M.R. Bridson, A. Haefliger, Metric Spaces of Non-Positive Curvature, Grundlehren Math. Wiss., vol. 319, Springer-Verlag, 1999.
[3] R. Brown, P.J. Higgins, Colimit theorems for relative homotopy groups, J. Pure Appl. Algebra 22 (1981) 11–41.
[4] R. Brown, P.J. Higgins, On the algebra of cubes, J. Pure Appl. Algebra 21 (1981) 233–260.
[5] J. Dugundji, Topology, Allyn & Bacon, 1966.
[6] U. Fahrenberg, M. Raussen, Reparametrizations of continuous paths, J. Homotopy Relat. Struct. 2 (2) (2007) 93–117.
[7] L. Fajstrup, Dipaths and dihomotopies in a cubical complex, Adv. in Appl. Math. 35 (2) (2005) 188–206.
[8] L. Fajstrup, É. Goubault, E. Haucourt, M. Raussen, Components of the fundamental category, Appl. Categ. Structures 12 (1) (2004) 81–108.
[9] L. Fajstrup, É. Goubault, M. Raussen, Algebraic topology and concurrency, Theoret. Comput. Sci. 357 (2006) 241–278.
[10] É. Goubault, E. Haucourt, Components of the fundamental category II, Appl. Categ. Structures 15 (2007) 387–414.
[11] M. Grandis, Directed homotopy theory II. Homotopy constructs, Theory Appl. Categ. 10 (14) (2002) 369–391.
[12] M. Grandis, Directed homotopy theory I. The fundamental category, Cahiers Topologie Géom. Différentielle Catégoriques 44 (2003) 281–316.
[13] M. Grandis, L. Mauri, Cubical sets and their site, Theory Appl. Categ. 11 (8) (2003) 185–211.
[14] A. Hatcher, Algebraic Topology, Cambridge Univ. Press, 2002.
[15] J. Milnor, On spaces having the homotopy type of a CW-complex, Trans. Amer. Math. Soc. 90 (2) (1959) 272–280.
[16] J.R. Munkres, Topology: A First Course, Prentice–Hall Inc., 1975.
[17] G.A. Niblo, L.D. Reeves, The geometry of cube complexes and the complexity of their fundamental groups, Topology 37 (3) (1998) 621–633.
[18] M. Raussen, Invariants of directed spaces, Appl. Categ. Structures 15 (2007) 355–386.
[19] M. Raussen, Reparametrizations with given stop data, Tech. Report R-2008-09, Department of Mathematical Sciences, Aalborg University, J. Homotopy Relat. Struct. 4 (1) (2009), in press.
[20] M. Raussen, Combinatorial structures on trace spaces, in preparation.
[21] C.P. Rourke, B.J. Sanderson, Δ-sets I. Homotopy theory, Quart. J. Math. Oxford Ser. (2) 22 (1971) 321–338.
[22] M.E. Rudin, A new proof that metric spaces are paracompact, Proc. Amer. Math. Soc. 20 (2) (1969) 603.
[23] S. Smale, A Vietoris mapping theorem for homotopy, Proc. Amer. Math. Soc. 8 (3) (1957) 604–610.
[24] A.H. Stone, Paracompactness and product spaces, Bull. Amer. Math. Soc. 54 (1948) 977–982.
[25] R.J. van Glabbeek, On the expressiveness of higher dimensional automata, in: Express 2004, Proceedings of the 11th International Workshop on Expressiveness in Concurrency, 2005, pp. 5–34.

Algebraic & Geometric Topology 10 (2010) 1683–1714

Simplicial models of trace spaces

MARTIN RAUSSEN

Directed algebraic topology studies topological spaces in which certain directed paths (d-paths) are singled out; in most cases of interest, the reverse path of a d-path is no longer a d-path. We are mainly concerned with spaces of directed paths between given end points, and how those vary under variation of the end points. The original motivation stems from certain models for concurrent computation. So far, homotopy types of spaces of d-paths and their topological invariants have only been determined in cases that were elementary to overlook.

In this paper, we develop a systematic approach describing spaces of directed paths – up to homotopy equivalence – as finite prodsimplicial complexes, ie with products of simplices as building blocks. This method makes use of a certain poset category of binary matrices related to a given model space. It applies to a class of directed spaces that arise from a certain class of models of computation – still restricted but with a fair amount of generality. In the final section, we outline a generalization to model spaces known as Higher Dimensional Automata.

In particular, we describe algorithms that allow us to determine not only the fundamental category of such a model space, but all homological invariants of spaces of directed paths within it. The prodsimplical complexes and their associated chain complexes are finite, but they will, in general, have a huge number of cells and generators.

55P10, 55P15, 55U10; 68Q55, 68Q85

1 Introduction

1.1 Background

With motivations arising originally from concurrency theory within computer science, a new field of research, directed algebraic topology, has emerged; for a comprehensive overview from a categorical perspective, we refer to the recent book by Grandis [13]. Directed algebraic topology involves spaces of ”directed paths” (or timed paths, executions): these directed paths can be concatenated, but in general *not* reversed; time is not reversible.

Published: 11 August 2010 DOI: 10.2140/agt.2010.10.1683

A particular model for concurrent computation, called Higher Dimensional Automata (HDA), was introduced by V Pratt [23] back in 1991. Mathematically, HDA can be described as (labelled) precubical sets (cf Brown and Higgins [2; 1]) with a preferred set of directed paths respecting the natural partial orders in any of the cubes of the model; (di-)homotopies of such directed paths have to respect the order along a deformation; cf Fajstrup, Goubault and Raussen [6].

Compared to other well-studied concurrency models like labelled transition systems, event structures, Petri nets etc (for a survey on those, see Winskel and Nielsen [30]), it has been shown by R J van Glabbeek [9] that Higher Dimensional Automata have the highest expressivity; on the other hand, they are certainly less studied and less often applied so far.

All concurrency models deal with sets of states and with sets of execution paths – with some further structure. The interest is mainly in the path spaces; typically, it is difficult to get an overview and to infer valuable information about the path space from the state space model.

A general framework for topological spaces with directed paths was defined and investigated as the category of d-spaces (d for *directed*). The objects are topological spaces with a preferred set of d-paths; the morphisms are the continuous maps preserving d-paths; cf in particular Grandis [12; 11; 13]. Grandis investigates first of all the fundamental category of a d-space – generalising the fundamental group of a topological space. Unlike the classical case, spaces of d-paths depend critically on the chosen end points. This makes it interesting to investigate how spaces of d-paths (with given fixed end points) vary under variation of these end points (cf Raussen [26]) and how this gives rise to a suitable decomposition of the state space into "components"; cf Fajstrup, Goubault, Haucourt and Raussen [7; 10].

General topological properties of spaces of d-paths and of *traces* (= d-paths up to monotone reparametrizations; cf Fahrenberg and Raussen [4; 27]) in semicubical complexes were investigated in Raussen [28]. But so far, apart from low-dimensional examples with convincing drawings, there have been very few explicit examples of actual computations of spaces of such traces (for an attempt in dimension two, see Raussen [24]); let alone a general method to perform such computations.

It is the aim of this article to make the homotopy types of trace spaces computable for a restricted class of Higher Dimensional Automata – those arising from the semaphore or PV–models introduced by Dijkstra [3] back in 1968. The state spaces for such models are complements of a number of hyperrectangular "holes" in a partially ordered hypercube $\vec{I}^n$. We describe trace spaces for these models explicitly as finite-dimensional prodsimplicial complexes (cf Kozlov [20]; with products of simplices as their building blocks) with the nerve of a particular poset category as barycentric subdivision.

For applications in concurrency, it is already very important to know the Betti number β_0 and to get hold on the connected components of a trace space: The reason is that traces in each connected component will always lead to the same result in a concurrent computation. Using the prodsimplicial structure makes it possible – at least in principle, the complexes may have lots of cells – to calculate algebraic topological invariants of such trace spaces.

We will finally hint on how to extend our results to general HDA. The overall philosophy reminds of the analysis of the topology of path spaces in CW–complexes in Milnor's article [22]: Also the spaces of d-paths in a precubical complex with given end points are equi locally convex (ELCX) (cf Raussen [28]) and thus locally contractible; for the PV–models analysed here, suitably chosen contractible subsets can be described explicitly, by a blend of order and combinatorics. They and their intersections form a poset category with a geometric realization that is homotopy equivalent to the path space under consideration.

Note that Jardine [16], using a different approach, has recently described a method to calculate the path category $P(X)$ of a simplicial or cubical set X as the path component category of a related explicitly constructed 2–category. It is a challenge to compare the two methods.

1.2 Structure and overview of results

Dijkstra's PV–models [3] are a particular class of models for *linear* concurrent computations with *semaphores*, a particularly simple, but instructive class of Higher Dimensional Automata. These models are introduced in Section 2; the state space X for such a model is embedded in a hypercube $\vec{I}^n$ and inherits a partial order. To get going, we define certain subspaces of the model space and show that the space of d-paths within any of these subspaces (for simplicity, from the bottom $\mathbf{0} \in \vec{I}^n$ to the top $\mathbf{1} \in \vec{I}^n$) is *empty or contractible* by a specific contraction. Moreover we show, that every d-path in the model space is contained in at least one of these subspaces.

This allows us in Section 3 to define a poset category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ indexing the nonempty subspaces of restricted d-paths in X (for simplicity, with paths starting at $\mathbf{0}$ and ending at $\mathbf{1}$) described above and their nonempty intersections. That category is naturally isomorphic to a poset subcategory of a *product* of a number of order categories of nonempty subsets of the positive integers $[1:n]$ less than or equal to n. A topological realization of this subcategory can thus be modelled on products of simplices and gives rise to a *prodsimplicial complex* [20]. Using standard methods (nerve lemma, projection lemma etc; cf Kozlov [20]), we show that the space of d-paths (or rather traces, ie, d-paths modulo monotone reparametrizations; cf Fahrenberg and Raussen [4]) $\vec{T}(X)(\mathbf{0},\mathbf{1})$ in such a model space is in fact homotopy equivalent to an explicit prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ that arises as geometric realization of the poset category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ – with the nerve $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$ of that category as barycentric subdivision.

It is the aim of Section 4 to exploit this theoretical result by achieving an explicit description of the index category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$. To this end, it is necessary to decide, for each of the subspaces mentioned above, whether it is empty or not, ie whether there exists a d-path within it from bottom to top. Every subspace can be described as the complement of a number l of homothetic hyperrectangles (with faces parallel to the coordinate planes) extending the original holes. It turns out that it is enough to find out whether there exist *deadlock* points (the only d-path with a deadlock as source is trivial) in these extended models. A combinatorial search algorithm for deadlocks was earlier described in Fajstrup, Goubault and Raussen [5]. The outcome of a systematic search for deadlocks (in all extended models) is a set $D(X)(\mathbf{0},\mathbf{1})$ of *minimal nonfaces* – all of dimension $n-1$ – of the prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ within the prodsimplicial complex $(\Delta^{n-1})^l$. The *maximal faces* of $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ can then be determined via minimal transversals in an associated hypergraph.

The explicit determination of the complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ thus achieved makes the calculation of algebraic topological invariants of the trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ possible. Even if, for complicated model spaces, the "curse of dimensionality" might prohibit explicit calculations, it will still be interesting and possible to study the change of invariants under change of end points (in rounds of computation; compare Herlihy and Rajsbaum [15] and other sources in distributed computing for this point of view).

For simplicity of the presentation, we restrict in this paper attention to an investigation of the traces from the bottom corner $\mathbf{0}$ to the top corner $\mathbf{1}$ in a state space without holes on the boundary $\partial I^n \subset I^n$. More general situations are important: First of all, most semaphore models allow holes intersecting the boundary ∂I^n. A study of the fundamental category as in Grandis [12] or of associated categories modelling higher invariants as in Raussen [26] of the model space relies on information about the topology of general spaces $\vec{T}(X)(\mathbf{c},\mathbf{d})$ at intermediate points $\mathbf{c},\mathbf{d} \in \vec{I}^n$. A modification of the setup discussed here has been described in Raussen [29]; it will be published elsewhere.

The final Section 5 takes first steps in generalizing the methods described so far. Dijkstra's PV–models can easily be generalized to a state space that is a product of *digraphs* with "hyperrectangular holes" modelling processes that may branch, merge and loop. For these, the topology of the trace space can be determined in two steps: First determine (the components of) the traces in the product of digraphs *without* holes. That space is homotopy equivalent to a product of trace spaces of the 1–dimensional digraphs; it is thus homotopy discrete. For each of the components, one can pull back (or "unloop") to a state space, including holes, of the type previously investigated. It will still have to be investigated how to unloop in a coherent manner in order to reuse calculations (of deadlocks etc) performed during previous steps.

For general HDA (modelled on general precubical sets), it is no longer possible to use the explicit contraction method for specific subspaces yielding local contractability used in this article. Instead, it is probably necessary to use the method described in Raussen [28] with a higher combinatorial complexity still to be sorted out.

2 Models of computation and subspaces

2.1 A simple higher dimensional automaton

To start with, we analyse spaces of directed paths in a simple model space that can be described as follows: A (linear) schedule for each of a number of n individual processors P_j, $1 \leq j \leq n$, is modelled on the directed interval $\vec{I}_j = [0,1]$. On subintervals $I_j^i =]a_j^i, b_j^i[\subseteq I_j$, $1 \leq i \leq l$, there is potential conflict with the schedules of the other processors. Let $\mathbf{a}^i = (a_1^i, \dots, a_n^i), \mathbf{b}^i = (b_1^i, \dots, b_n^i) \in I^n \setminus \partial I^n$ and let $R^i = \{\mathbf{x} \in I^n \mid a_j^i < x_j < b_j^i, 1 \leq j \leq n\}$ denote the "homothetic" hyperrectangle (faces parallel to the coordinate planes) with bottom corner $\mathbf{a}$ and top corner $\mathbf{b}$.

The state space for concurrent executions of n linear processes is then the space $X = \vec{I}^n \setminus F \subset \vec{I}^n$ excluding the *forbidden region* $F = \bigcup_{i=1}^l R^i$. The forbidden region F models conflicts and may not be entered. Consult Figure 1 for a simple 3–dimensional example.

The space X inherits a partial order $\leq$ from the componentwise partial order $\leq$ on $\vec{I}^n$. We study compound schedules (execution paths) in such a state space X: A *d-path* in X is a continuous path $p\colon \vec{I} \to X$ that is continuous and *order-preserving*: each coordinate $\pi_j \circ p\colon \vec{I} \to X \subset \vec{I}^n \to \vec{I}$, $1 \leq j \leq n$, is *weakly increasing*. The set $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ consists of all d-paths in X starting at $\mathbf{c} \in X$ and ending at $\mathbf{d} \in X$; in particular, these d-paths avoid the "forbidden region" $F \subset \vec{I}^n$. Consult eg Gunawardena [14] and Fajstrup, Goubault and Raussen [5] for detailed descriptions.

As a topological space, $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ is given the subspace topology inherited from the space $P(X)(\mathbf{c}, \mathbf{d}) = [(I, 0, 1); (X, \mathbf{c}, \mathbf{d})]$ of *all* paths in X from $\mathbf{c}$ to $\mathbf{d}$ in the compact-open topology (= uniform convergence topology).

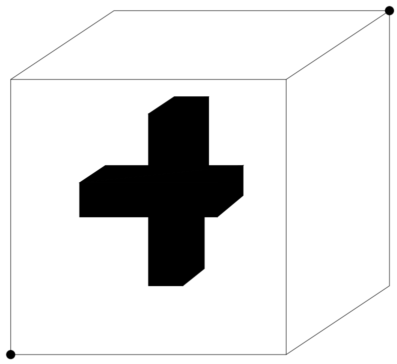


Figure 1: Three-dimensional state space X as complement of a forbidden region F consisting of two hyperrectangles within a unit cube

Reparametrization equivalent d-paths [4] in X have the same directed image (= *trace*) in X. Dividing out the action of the monoid of (weakly-increasing) reparametrizations of the parameter interval $\vec{I}$, we arrive at trace space $\vec{T}(X)(\mathbf{c},\mathbf{d})$ (cf Fahrenberg and Raussen [4; 27]) which is shown in Raussen [28] to be homotopy equivalent to path space $\vec{P}(X)(\mathbf{c},\mathbf{d})$ for a far wider class of directed spaces X. In the latter paper, it is also shown that trace spaces enjoy nice properties: They are metrizable, locally compact, locally contractible, and they have the homotopy type of a CW–complex.

It is the aim of the present paper to describe and analyze a combinatorial/topological model of spaces of traces (and hence also of d-paths) in a model space X up to homotopy equivalence in order to make calculations of their algebraic topological invariants feasible.

2.2 Subspaces of the model space

We will now describe certain subspaces of X and then prove that associated spaces of d-paths *within* these subspaces are *either empty or contractible*.

We use the following notation:

- The set of elements "below" $\mathbf{d} = (d_1, \dots d_n) \in X$ is denoted

$$\downarrow\mathbf{d} := \{\mathbf{x} \in X \mid \mathbf{x} \le \mathbf{d}\} = \{\mathbf{x} \in I^n \mid \mathbf{x} \le \mathbf{d}, \mathbf{x} \notin F\}.$$

  Remark that it is not always possible to reach $\mathbf{d}$ from every $\mathbf{x} \in\downarrow\mathbf{d}$ by a d-path. Likewise $\uparrow\mathbf{c} = \{\mathbf{x} \in X \mid \mathbf{c} \le \mathbf{x}\}$ denotes the set of elements "above" $\mathbf{c}$.
- The upper boundary $\{\mathbf{x} \in\downarrow\mathbf{d} \mid \exists 1 \le i \le n\colon x_i = d_i\}$ of the hyperrectangle with upmost vertex in $\mathbf{d}$ with X will be denoted $\partial_+\downarrow\mathbf{d}$.
- $\mathbf{a}^i = (a^i_1, \dots, a^i_n)$, $\mathbf{b}^i = (b^i_1, \dots, b^i_n)$.

Consider Example 2.3 and Figure 2 for 2–dimensional illustrations of the following definition:

Definition 2.1 (1) For $1 \leq i \leq l, 1 \leq j_i \leq n$, let

$$X_{j_1,\dots,j_l} := \{\mathbf{x} \in X \mid \forall i\colon\ x_{j_i} \leq a^i_{j_i} \text{ or } \exists k\colon x_k \geq b^i_k\}.$$

(2) For nonempty subsets $J_i \subseteq [1:n]$, $1 \leq i \leq l$, let

$$X_{J_1,\dots,J_l} := \{\mathbf{x} \in X \mid \forall i\colon\ x^i_{j_i} \leq a^i_{j_i}, j_i \in J_i, \text{ or } \exists k\colon x_k \geq b^i_k\}.$$

For later use, we note an equivalent formulation of these conditions:

$$\mathbf{x} \in X,\ \forall i : (\forall k\ x_k < b^i_k \Rightarrow x_{j_i} \leq a^i_{j_i}) \tag{2-1}$$

$$\mathbf{x} \in X,\ \forall i : (\forall k\ x_k < b^i_k \Rightarrow x_{j_i} \leq a^i_{j_i} \text{ (for all } j \in I_j)). \tag{2-2}$$

Remark 2.2 An execution path in $X_{j_1,\dots,j_l}$ has the following characterization: Processor j_i is late at R^i: it has not yet reached the "conflict" interval $J^i_{j_i}$ when one of the others, say k_i, has already left the corresponding conflict interval $J^i_{k_i}$.

Example 2.3 Figure 2 shows in each of the two rows an example of a model space $X = \vec{I}^2 \setminus F$ given as the complement of the forbidden region F consisting of two black squares. The grey-shaded areas show, in both cases, the subspaces X_{11}, X_{12}, X_{21}, resp. X_{22}, in that order. Remark that an empty space of d-paths $\vec{P}(X_{ij})(\mathbf{0},\mathbf{1}) = \varnothing$ occurs only in the second row – and only for X_{12}.

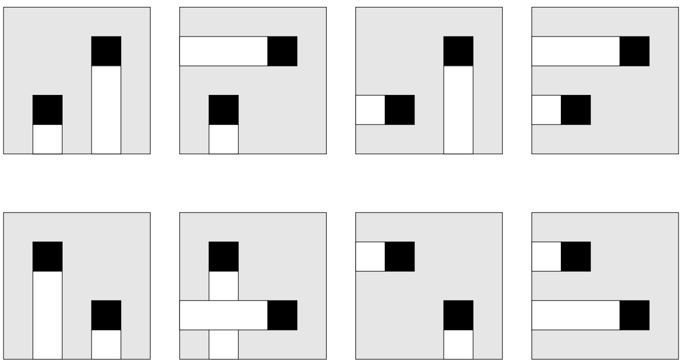


Figure 2: Two examples of a model space X and of subspaces $X_{ij}, 1 \leq i, j \leq 2$, the grey-shaded areas

Example 2.4 In Figure 3, $X = \vec{I}^3 \setminus F, F = \vec{J}^3$ and $\vec{J} \subset \vec{I}$ is an interior open interval. Apart from the forbidden region "black box" $\vec{J}^3$ with upper corner $\mathbf{b}$, you see the shaded areas $X_j \cap \partial_+ \downarrow \mathbf{b}, 1 \leq j \leq 3$. Remark that every pair X_{j_1}, X_{j_2} of these areas intersect, whereas the intersection $X_1 \cap X_2 \cap X_3$ is empty. In particular, $\vec{P}(X_J)(\mathbf{0},\mathbf{1}) = \varnothing$ for $\varnothing \neq J \subseteq [1:3]$ if and only if $J = [1:3]$.

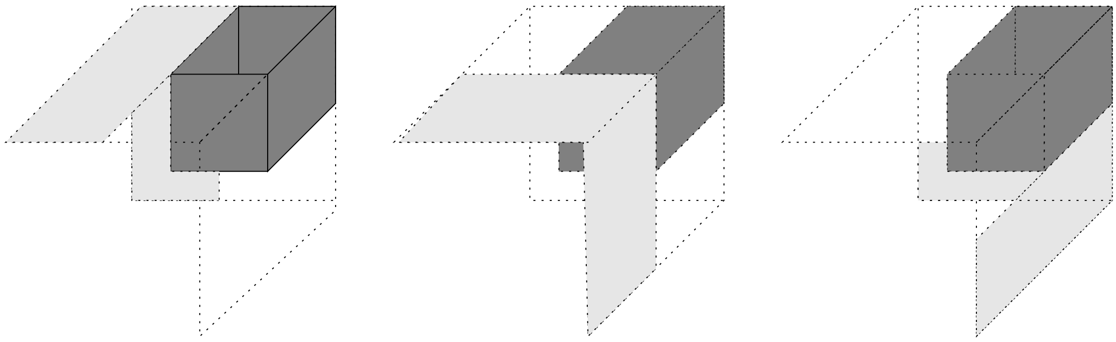


Figure 3: Intersections of X_i with the upper boundary $\partial_+\downarrow\mathbf{b}$ of the box with upper corner $\mathbf{b}$

The subspaces from Definition 2.1 above have the following obvious property:

Lemma 2.5 $X_{J_1,\dots,J_l} = \bigcap_{j_i\in J_i} X_{j_1,\dots,j_l}$. □

2.3 Restricted path spaces are empty or contractible

With $\mathbf{a} = (a_1,\dots,a_n), \mathbf{b} = (b_1,\dots,b_n)$, the binary operation $\vee$ on $\mathbf{R}^n$ (the least upper bound) is given by $\mathbf{a}\vee\mathbf{b} = (\max(a_1,b_1),\dots,\max(a_n,b_n))$. Observe ((1) as a consequence of Definition 2.1):

Lemma 2.6 (1) *$X_{j_1,\dots,j_l}$ is closed under $\vee$ for every choice $j_i\in[1:n], i\in[1:l]$.*

(2) *Intersections of $\vee$–closed sets are $\vee$–closed.*

(3) *$X_{J_1,\dots,J_l}$ is closed under $\vee$ for every collection of* nonempty *subsets $J_i\subseteq[1:n]$, $1\le i\le l$.* □

A similar result holds for the binary operation $\wedge$ (the greatest lower bound) given by $\mathbf{a}\wedge\mathbf{b} = (\min(a_1,b_1),\dots,\min(a_n,b_n))$ after a suitable change of definition for $X_{j_1,\dots,j_l}$.

Note that Lemma 2.6 no longer holds if one of the sets J_i may be empty!

Remark 2.7 It should also be possible to exploit the least upper bound operation $\vee$ for the definition and analysis of (future) components (cf Fajstrup, Goubault, Haucourt and Raussen [7; 10]) as follows: $\mathbf{x},\mathbf{y}\in X$ are elementarily future related if $\mathbf{z}_1\vee\mathbf{z}_2\in X$, for every $\mathbf{z}_1,\mathbf{z}_2\in X$ with $\vec{P}(X)(\mathbf{x},\mathbf{z}_i)\neq\varnothing\neq\vec{P}(X)(\mathbf{z}_i,\mathbf{y})$. Consider the equivalence relation *future equivalent* generated by symmetric and transitive closure. This idea will be pursued elsewhere.

The next observation that is valid for more general binary operations $*$ is essential for our purposes:

Proposition 2.8 *Assume* $A \subseteq X, \mathbf{a}, \mathbf{b} \in A$.

Let $*\colon A \times A \to A$ *denote a commutative continuous map satisfying*

- $\mathbf{x}_i \leq \mathbf{y}_i \Rightarrow \mathbf{x}_1 * \mathbf{x}_2 \leq \mathbf{y}_1 * \mathbf{y}_2$, *ie,* $*$ *is a d-map;*
- $\mathbf{x} \leq \mathbf{y} \in A \Rightarrow \mathbf{x} * \mathbf{y} = \mathbf{y}$.

Then the following holds:

(1) *The trace space* $\vec{T}(A)(\mathbf{c}, \mathbf{d})$ *is either empty or contractible.*

(2) *Let* $\varnothing \neq J_i \subseteq [1:n], 1 \leq i \leq l$ *and* $\mathbf{c}, \mathbf{d} \in X_{J_1,\dots J_l}$. *Then the trace space* $\vec{T}(X_{J_1,\dots J_l})(\mathbf{c}, \mathbf{d})$ *is either empty or contractible.*

Proof (2) follows from (1) and Lemma 2.6 with $* = \vee$.

To prove (1), we show first that $\vec{P}(A)(\mathbf{c}, \mathbf{d})$ is either contractible or empty. If $\vec{P}(A)(\mathbf{c}, \mathbf{d})$ is nonempty, then, for any pair $p, q \in \vec{P}(A)(\mathbf{c}, \mathbf{d})$, define a one-parameter family $H(p, q)\colon \vec{P}(A)(\mathbf{c}, \mathbf{d}) \times I \to \vec{P}(A)(\mathbf{c}, \mathbf{d})$ by

$$H_t(p, q)(s) := q(s) * p(ts),\ t \in I.$$

Remark that $H_0(p, q)(s) = q(s) * \mathbf{c} = q(s)$, $H_t(p, q)(0) = \mathbf{c} * \mathbf{c} = \mathbf{c}$, $H_t(p, q)(1) = \mathbf{d} * p(t) = \mathbf{d}$ and that $H_1(p, q)(s) = q(s) * p(s)$. Thus $H(p, q)$ defines an increasing d-homotopy (cf Grandis [12]) $q \mapsto p * q$ between d-paths in $\vec{P}(A)(\mathbf{c}, \mathbf{d})$. Likewise, $H(q, p)$ is an increasing d-homotopy $p \mapsto q * p = p * q$. Their concatenation $G(q, p) = H(p, q) * H^-(q, p)$ (orientations are reversed for the second d-homotopy) is a "zig-zag" d-homotopy from q to p; in particular a path from q to p within $\vec{P}(\mathbf{c}, \mathbf{d})$. The map $G(-,-)$ defines a continuous section of the "end path map" $\mathrm{ev}_0 \times \mathrm{ev}_1\colon \vec{P}(\mathbf{c}, \mathbf{d})^I \to \vec{P}(\mathbf{c}, \mathbf{d}) \times \vec{P}(\mathbf{c}, \mathbf{d})$ that associates to a pair (q, p) the d-homotopy $G(p, q)$.

Given an arbitrary $p \in \vec{P}(A)(\mathbf{c}, \mathbf{d})$, the map $G(-, p)\colon \vec{P}(\mathbf{c}, \mathbf{d}) \times I \to \vec{P}(\mathbf{c}, \mathbf{d})$ is a contraction of $\vec{P}(\mathbf{c}, \mathbf{d})$ to p. By Raussen [28], Proposition 2.16, the trace space $\vec{T}(\mathbf{c}, \mathbf{d})$ is homotopy equivalent to the space of d-paths $\vec{P}(\mathbf{c}, \mathbf{d})$ and is thus also contractible. □

Remark 2.9 (1) A similar result holds for a map $*$ satisfying $\mathbf{x} \leq \mathbf{y} \Rightarrow \mathbf{x} * \mathbf{y} = \mathbf{x}$.

(2) If $J_i = [1:n]$ for at least one i, then $\vec{T}(X_{J_1,\dots,J_l})(\mathbf{0}, \mathbf{1})$ is always empty; in this case, condition (2) from Definition 2.1 amounts to $\mathbf{x} < \mathbf{b}^i \Rightarrow \mathbf{x} \leq \mathbf{a}^i$. But every d-path from $\mathbf{0}$ to $\mathbf{1}$ needs to pass through the region $\downarrow \mathbf{b}^i \setminus (\downarrow \mathbf{a}^i)$ in between.

The trace spaces considered above *cover* the total trace space: With notation as in Proposition 2.8, we obtain:

Lemma 2.10 $\vec{T}(X)(\mathbf{c}, \mathbf{d}) = \bigcup_{[1:n]^l} \vec{T}(X_{j_1,\dots,j_l})(\mathbf{c}, \mathbf{d})$ *for any* $\mathbf{c}, \mathbf{d} \in X$.

Proof For a given d-path $p=(p_1,\dots,p_n)\in\vec{P}(X)(\mathbf{c},\mathbf{d})\subseteq\vec{P}(\vec{I}^n)(\mathbf{c},\mathbf{d})$ and $1\le i\le l$, choose a minimal t_i such that there exists $k_i\in[1:n]$ with $p_{k_i}(t_i)=\widetilde{b}^i_{k_i}:=\min(b^i_{k_i},d_i)$. If $\widetilde{b}^i_{k_i}\le a^i_{k_i}$, then $j_i\in[1:n]$ can be chosen arbitrarily; otherwise choose $s_i<t_i$ such that $p_{k_i}(]s_i,t_i[)=]a^i_{k_i},\widetilde{b}^i_{k_i}[$. Since $p(t)\not\in R^i$ for every t and $p_j(t)<b^i_j$ for all (j,t) with $t<t_i$, there exists j_i such that $p_{j_i}(t)\le a^i_{j_i}$ for $s_i<t<t_i$ and hence, by monotonicity, for $t<t_i$. In conclusion, $p\in\vec{P}(X_{j_1,\dots,j_l})(\mathbf{c},\mathbf{d})$. □

In the following Section 3, we need to cover a trace space $\vec{T}(X)(\mathbf{c},\mathbf{d})$ by *open* subsets. Therefore, we carefully augment the spaces $X_{J_1,\dots,J_l}$: Choose $\varepsilon>0$ such that all distances $|a^i_j-a^k_j|,|a^i_j-b^k_j|,|b^i_j-a^k_j|>4\varepsilon$ unless they vanish.

Definition 2.11 (1) $Y_{j_1,\dots,j_l}:=\{\mathbf{x}\in X\mid \forall i\colon\ x_{j_i}<a^i_{j_i}+\varepsilon \text{ or } \exists k\colon\ x_k>b^i_k-\varepsilon\}\supset X_{j_1,\dots,j_l}$.

(2) $Y_{J_1,\dots,J_l}=\bigcap_{j_i\in J_i}Y_{j_1,\dots,j_l}\supset X_{J_1,\dots,J_l}$.

Proposition 2.12 *Suppose that, for every $1\le j\le n$, no upper boundary b^i_j is equal to a lower boundary a^k_j; ie, that $\{a^i_j\}_i\cap\{b^i_j\}_i=\varnothing$ for every j.*

(1) *There exists a d-map $\varphi\colon X\to X$ (continuous and order preserving) and a d-homotopy (cf Grandis [12]) $\Phi=(\Phi_t)\colon X\times\vec{I}\to X,\varphi\to\mathrm{id}_X$ keeping $X_{j_1,\dots,j_l}$ pointwise fix that satisfy*

$$\varphi(Y_{J_1,\dots,J_l})\subseteq X_{J_1,\dots,J_l}\quad\text{and}\quad\Phi(Y_{J_1,\dots,J_l}\times I)\subseteq Y_{J_1,\dots,J_l}$$

for all $(J_1,\dots,J_l)\subseteq[1:n]^l$.

(2) *$X_{j_1,\dots,j_l}$ is a deformation retract of $Y_{j_1,\dots,j_l}$.*

(3) *$X_{J_1,\dots,J_l}$ is a deformation retract of $Y_{J_1,\dots,J_l}$.*

Proof Choose weakly increasing reparametrizations $\varphi_j\colon\vec{I}\to\vec{I},\ 1\le j\le n$, of the unit interval I that are piecewise linear, equal to the identity map outside the intervals $]a^i_j,a^i_j+2\varepsilon[$ and $]b^i_j-2\varepsilon,b_j[$ and that map the intervals $[a^i_j,a^i_j+\varepsilon]$ constantly to a^i_j and $[b^i_j-\varepsilon,b^i_j]$ constantly to b^i_j. The product $\varphi=\prod_{j=1}^n\varphi_j\colon\vec{I}^n\to\vec{I}^n$ restricts to a map $\varphi\colon X\to X$ such that $\varphi(Y_{j_1,\dots,j_l})\subseteq X_{j_1,\dots,j_l}$ and hence $\varphi(Y_{J_1,\dots,J_l})\subseteq X_{J_1,\dots,J_l}$.

The linear homotopy Φ that connects φ and the identity map id_{I^n} is a d-homotopy that restricts to d-homotopies on the spaces $Y_{j_1,\dots,j_l}$ and $Y_{J_1,\dots,J_l}$; it induces homotopies between the identity map and the maps induced by the restrictions of φ on associated trace spaces. □

Corollary 2.13 *Let* $\mathbf{c},\mathbf{d}\in X$.

(1) *Every subspace* $Y_{j_1,\dots,j_l}\subset X$ *is* open. *The path spaces* $\vec{T}(Y_{j_1,\dots,j_l})(\mathbf{c},\mathbf{d})$ *form an* open cover *of* $\vec{T}(X)(\mathbf{c},\mathbf{d})$.

(2) $\vec{T}(Y_{J_1,\dots,J_l})(\mathbf{c},\mathbf{d})$ *is contractible, resp. empty, if and only if* $\vec{T}(X_{J_1,\dots,J_l})(\mathbf{c},\mathbf{d})$ *is contractible, resp. empty.*

Proof Immediate from Proposition 2.8 and Proposition 2.12. □

3 (Prod)simplicial models for trace spaces

In this and the next Section 4, we concentrate on an investigation of trace spaces $\vec{T}(X)(\mathbf{0},\mathbf{1})$ from the bottom vertex $\mathbf{0}$ to the top vertex $\mathbf{1}$ of $X\subseteq\vec{I}^n$ under the further simplifying restriction that all forbidden hyperrectangles $R^i\subset F$ are contained in the *interior* of I^n. The necessary modifications arising for more general state spaces and for trace spaces of type $\vec{T}(X)(\mathbf{c},\mathbf{d})$, resp. $\vec{T}(X)(\mathbf{c},\partial_+(\downarrow\mathbf{d}))$ are discussed in Raussen [29]; these results will be published elsewhere.

3.1 The index category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$

3.1.1 A matrix representation of a power poset The index multisets $(J_1,\dots,J_l)$ with $J_i\subseteq[1:n]$ considered in the previous Section 2 may be viewed as elements of $(\mathcal{P}([1:n]))^l\cong\mathcal{P}([1:l]\times[1:n])$. Elements of the latter power set can be encoded by their characteristic functions which can be viewed as binary $l\times n$–matrices:

Let $M_{l,n}=M_{l,n}(\mathbf{Z}/2)$ denote the set of all binary $l\times n$–matrices – with 2^{ln} elements. Componentwise logical *aut* and logical *and* define addition and multiplication of a ring structure on $M_{l,n}$ – that does not concern us here. The total order on $\mathbf{Z}/2$ given by $a\le b$ unless ($a=1$ and $b=0$) extends to a componentwise given partial order $\le$ on $M_{l,n}$. With this partial order defining the morphisms, $M_{l,n}$ will be viewed as a *poset category*.

There is a natural order-preserving bijection between the subsets of $[1:l]\times[1:n]$ (elements of the power set $\mathcal{P}([1:l]\times[1:n])$ with partial order given by inclusion) and elements in $M_{l,n}$ given by

$$J=(J_1,\dots,J_l)\mapsto M^J=(m_{ij}^J),\quad m_{ij}^J=1\Leftrightarrow j\in J_i \tag{3-1}$$

with inverse $M=(M_{ij})\mapsto J^M,\ j\in J_i^M\Leftrightarrow m_{ij}=1$.

Under this bijection, the relevant multisets $J=(J_1,\dots,J_l)$ with $J_i\neq\varnothing, 1\le i\le l$, correspond to matrices in the subset $M_{l,n}^R\subset M_{l,n}$ consisting of the $(2^n-1)^l$ matrices such that *no row vector is a zero vector*. We view $M_{l,n}^R$ as the *full subposet category* within $M_{l,n}$.

3.1.2 Subcategories and pasting functors To ease notation, we will in the following write $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ instead of $\vec{T}(X_{J^M})(\mathbf{0},\mathbf{1})$. The relevant index category to consider here is the full subposet category $\mathcal{C}(X)(\mathbf{0},\mathbf{1}) \subset M_{l,n}^R \subset M_{l,n}$ consisting of all matrices M such that

$$\vec{T}(X_M)(\mathbf{0},\mathbf{1}) \text{ is nonempty.} \tag{3-2}$$

This index category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ gives rise to functors $\mathcal{D}$ and $\mathcal{E}$ into **Top**:

- The functor $\mathcal{D}$: $\mathcal{C}(X)(\mathbf{0},\mathbf{1})^{\mathrm{op}} \to \mathbf{Top}$ associates $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ to the matrix M; the reverse partial order on $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ corresponds to inclusion in **Top**.
- The functor $\mathcal{E}$: $\mathcal{C}(X)(\mathbf{0},\mathbf{1}) \to \mathbf{Top}$ restricts from a functor $\mathcal{E}_n^l$: $M_{l,n}^R \to \mathbf{Top}$; it associates to M^J with $J = (J_1,\dots,J_l)$ – all $J_i \neq \varnothing$! – the standard simplex product $\Delta^{|J_1|-1} \times \cdots \times \Delta^{|J_l|-1} \subset \mathbf{R}^{|J_1|} \times \cdots \times \mathbf{R}^{|J_l|} \subset (\mathbf{R}^n)^l$; $\mathbf{R}^{|J_i|}$ is included in $\mathbf{R}^n$ as the subspace given by the equations $x_j = 0$, $j \notin J_i$. For this functor, the original partial order on $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ corresponds to inclusion in **Top**.

The functor $\mathcal{E}_n^l$ should be considered as a pasting scheme for the product of simplices $(\Delta^{n-1})^l$; the functor $\mathcal{E}$ becomes then a pasting scheme for a subprodsimplicial complex (cf Kozlov [20]) $X_M \subseteq (\Delta^{n-1})^l$ to be explained below.

Remark 3.1 We prefer $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ as indexing category to the nerve of the covering given by the spaces $X_{j_1,\dots,j_l}$, since an intersection X_M – in view of Lemma 2.5 – can arise in many ways as intersection of the basic spaces $X_{j_1,\dots,j_l}$ corresponding to matrices in which every row is a *unit* vector; even as intersection of a varying number of the basic covering sets. The nerve of that latter covering carries redundant information: it does not take care of the product structure that gives rise to automatically commuting morphisms. It is in fact a barycentric subdivision of $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ as will be explained below.

3.2 Trace spaces and prodsimplicial complexes as colimits

Regarding the functors $\mathcal{E}$ and $\mathcal{D}$ as pasting schemes, we consider their colimits:

- $\operatorname{colim}(\mathcal{D}) = \vec{T}(X)(\mathbf{0},\mathbf{1})$ by Lemma 2.10.
- $\operatorname{colim}(\mathcal{E}_n^l) = (\Delta^{n-1})^l$.

- $\mathbf{T}(X)(\mathbf{0},\mathbf{1})\colon = \mathrm{colim}(\mathcal{E}) \subset \mathrm{colim}(\mathcal{E}_n^l) = (\Delta^{n-1})^l$ is a prodsimplicial complex (in the terminology of Kozlov [20]) consisting of all those products of simplices $\Delta^{|J_1|-1} \times \cdots \times \Delta^{|J_l|-1}$ that correspond to tuples $(J_1, \ldots, J_l)$ such that $M^J \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$; in other words, the functor $\mathcal{E}$ is a pasting scheme for a prodsimplicial complex with one simplex product for each $M \in M_{l,n}^R$ giving rise to a nonempty trace space $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$.

Remark 3.2 This prodsimplicial complex is *not* a general complex of morphisms in the sense of Kozlov [20, 9.2.4]. Whether $\vec{T}(X_{J_1,\ldots J_l})(\mathbf{0},\mathbf{1})$ is nonempty cannot be decided by investigating whether all $\vec{T}(X_{j_1,\ldots j_l})(\mathbf{0},\mathbf{1})$, $j_i \in J_i$, are nonempty. The topology of the complex does not only depend on its 1–skeleton [20, Proposition 18.1].

Comparing with $\mathrm{colim}(\mathcal{E}_n^l) = (\Delta^{n-1})^l$, we obtain at once:

Lemma 3.3 *The prodsimplicial complex* $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ *is a subcomplex of* $(\partial\Delta^{n-1})^l \cong (S^{n-2})^l$. *It has at most* n^l *vertices, and* $\dim(\mathbf{T}(X)(\mathbf{0},\mathbf{1})) \le (n-2)l$.

Proof From Remark 2.9, it follows that $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$ as soon as M has a row vector consisting of digits one only; in particular, no product can have a (full) factor Δ^{n-1}. The complex $(\partial\Delta^{n-1})^l$ has the number of vertices and the dimension given in the lemma. □

Example 3.4 Assume that the obstruction hyperrectangles $R^i =]\mathbf{a}^i, \mathbf{b}^i[$ have the property $\mathbf{b}^i < \mathbf{a}^{i+1}$, $1 \le i < l$; ie, the holes are totally ordered with respect to the partial order in $\mathbf{R}^n$. This is the case in the first row of Example 2.3. It is not difficult to see that then $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$ if and only if M has a row in which *all* coefficients are equal to 1. We conclude that $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) = (\partial\Delta^{n-1})^l \cong (S^{n-2})^l$ in this case; compare Example 4.11(4). Hence the bounds given in Lemma 3.3 are sharp!

3.3 Homotopy equivalences

Theorem 3.5 *Assume that, for every* $1 \le i \le n$, *no upper boundary coordinate* b_j^i *is equal to a lower boundary coordinate* a_j^k. *Then trace space* $\vec{T}(X)(\mathbf{0},\mathbf{1})$ *is homotopy equivalent to the prodsimplicial complex* $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) \subset (\partial\Delta^{n-1})^l$ *and to the nerve of the category* $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$; *the latter simplicial complex arises as a barycentric subdivision of* $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$.

Proof First, we determine the *homotopy* colimits of the functors defining the pasting schemes above. We apply the homotopy lemma [20, Theorem 15.12] to the natural transformation $\Psi\colon \mathcal{D} \Rightarrow \mathcal{T}^*$ from $\mathcal{D}$ to the trivial functor $\mathcal{T}^*\colon \mathcal{C}(X)(\mathbf{0},\mathbf{1})^{\mathrm{op}} \to \mathbf{Top}$ which sends every object into the same one-point space. Since the maps corresponding to Ψ are homotopy equivalences at any object M in $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ (from a contractible space $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ – by Proposition 2.8(2) – to a point), the map $\operatorname{hocolim}\mathcal{D} \to \operatorname{hocolim}\mathcal{T}^*$ induced by Ψ is a homotopy equivalence by the homotopy lemma. By definition, $\operatorname{hocolim}\mathcal{T}^*$ is the nerve $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$ of the indexing category.

A similar argument shows that also the trivial natural transformation from $\mathcal{E}$ to $\mathcal{T}\colon \mathcal{C}(X)(\mathbf{0},\mathbf{1}) \to \mathbf{Top}$ induces a homotopy equivalence of homotopy colimits.

Next, we wish to apply the projection lemma [20, Theorem 15.19] – with two twists – to the fiber projection maps $\operatorname{hocolim}\mathcal{D} \to \operatorname{colim}\mathcal{D}$ and $\operatorname{hocolim}\mathcal{E} \to \operatorname{colim}\mathcal{E}$. If applicable, that lemma ensures that these projection maps are homotopy equivalences. Altogether, the maps discussed above fit to yield a homotopy equivalence

$$\begin{array}{ccccc}
\vec{T}(X)(\mathbf{0},\mathbf{1}) = \operatorname{colim}(\mathcal{D}) & \longleftarrow & \operatorname{hocolim}(\mathcal{D}) & \longrightarrow & \operatorname{hocolim}(\mathcal{T}^*) \\
 & & & & \updownarrow \\
\mathbf{T}(X)(\mathbf{0},\mathbf{1}) = \operatorname{colim}(\mathcal{E}) & \longleftarrow & \operatorname{hocolim}(\mathcal{E}) & \longrightarrow & \operatorname{hocolim}(\mathcal{T})
\end{array}$$

since the two opposite categories $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ and $\mathcal{C}(X)(\mathbf{0},\mathbf{1})^{\mathrm{op}}$ have the same classifying space $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$. In particular, $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ is also homotopy equivalent to the nerve $\Delta(\mathcal{C}(X)(\mathbf{0},\mathbf{1}))$ – which is thus a barycentric subdivision of $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$.

The first twist alluded to above consists in using, instead of the nerve diagram of the covering given by the spaces $X_{j_1,\dots,j_l}$, the functors $\mathcal{D}$ and $\mathcal{E}$ with the smaller indexing category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$; cf Remark 3.1. To get to the conclusion in case of the functor $\mathcal{E}$, we use moreover, that $\mathbf{T}(X)(\mathbf{c},\mathbf{d})$ has a prodsimplicial and thus a CW–structure; cf Kozlov [20, Remark 15.20].

As to the functor $\mathcal{D}$, we need to verify the conditions of the projection lemma: It was shown in Raussen [28], that $\vec{T}(X)(\mathbf{0},\mathbf{1})$ is paracompact – even under much weaker assumptions to X. Furthermore, Proposition 2.12 allows us to replace the cover given by the subspaces $\vec{T}(X_{j_1,\dots,j_l})(\mathbf{0},\mathbf{1})$ to that given by the homotopy equivalent *open* subspaces $\vec{T}(Y_{j_1,\dots,j_l})(\mathbf{0},\mathbf{1})$ from Definition 2.11, with the same colimit and a homotopy equivalent homotopy colimit. □

Remark 3.6 A modified version of Theorem 3.5 holds without assuming that the obstruction hyperrectangles are contained in the interior of I^n; also for trace spaces of type $\vec{T}(X)(\mathbf{c},\mathbf{d})$ and $\vec{T}(X)(\mathbf{c},\partial_+\!\downarrow\!\mathbf{d})$ described in Section 2. The only necessary change is a different description of the corresponding index category; this has been worked out in Raussen [29].

When these index categories $\mathcal{C}(X)(\mathbf{c},\mathbf{d})$ are determined (cf Section 4 and Raussen [29, Chapter 5]) one may replace the morphism sets $\vec{T}(X)(\mathbf{c},\mathbf{d})$ in the trace category $\vec{T}(X)$ [26] by the corresponding prodsimplicial complexes $\mathbf{T}(X)(\mathbf{c},\mathbf{d})$ and calculate their homological invariants. In particular, their homology groups in dimension 0 describe the fundamental category of the d-space X.

4 Determination of the index category

To determine, using Theorem 3.5, the prodsimplicial model $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ of trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$, we need to describe the indexing category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ explicitly. We have to determine which of the subspaces $\vec{T}(X_M)(\mathbf{0},\mathbf{1})$ corresponding to matrices $M \in M_{l,n}^R$ are empty and which not; cf (3-2).

It turns out that (non)emptyness can be investigated by a method that was originally designed in Fajstrup, Goubault and Raussen [5] for the detection of deadlocks and associated unsafe regions in models for the simple Higher Dimensional Automata described in Section 2.1.

A *deadlock* in X is an element $\mathbf{x} \in X$ that admits only the constant path as d-path with source $\mathbf{x}$. The *unsafe region* corresponding to the deadlock $\mathbf{x}$ consists of all $\mathbf{y} \in X$ such that *no* d-path in X with source $\mathbf{y}$ can leave the hyperrectangle spanned by $\mathbf{y}$ and $\mathbf{x}$.

It will be shown that $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$ is equivalent to the existence of a deadlock ($\neq \mathbf{1}$) within X_M. This in turn depends on whether a certain set of *inequalities* – determined by M – between coordinates of the obstruction hyperrectangles R^i holds.

A simple-minded version of the procedure worked out below was described earlier in Raussen [24]; it was restricted entirely to dimension $n = 2$.

4.1 Empty path spaces and deadlocks

Remember the notation convention: $1 \le i \le l$ enumerates the obstruction hyperrectangles R^i; $1 \le j \le n$ enumerates the n coordinate directions in $\mathbf{R}^n$.

We begin with a "dual" look at the spaces $X_{J_1,\dots,J_l}$ from Definition 2.1, resp. X_M from (3-1) in Section 3.1. For each of the original forbidden hyperrectangles $R^i = \prod_{j=1}^n I_j^i$ (cf Section 2), we define n *extended* hyperrectangles

$$R_j^i = \prod_{k=1}^{j-1} \tilde{I}_k^i \times I_j^i \times \prod_{k=j+1}^{n} \tilde{I}_k^i, \quad 1 \le i \le l,\ 1 \le j \le n, \tag{4-1}$$

with $\widetilde{I}^i_k = [0, a^i_k] \cup I^i_k = [0, b^i_k[$; an interval with 0 as its lower boundary. For illustrations, consult Figure 2 – the black rectangles are extended in several ways by white rectangles – and Figure 4 showing an extension of the forbidden black cube to a box containing also the grey-shaded part.

Remark 4.1 Each of the hyperrectangles R^i_j has a lowest vertex for which all apart from one of the coordinates are 0.

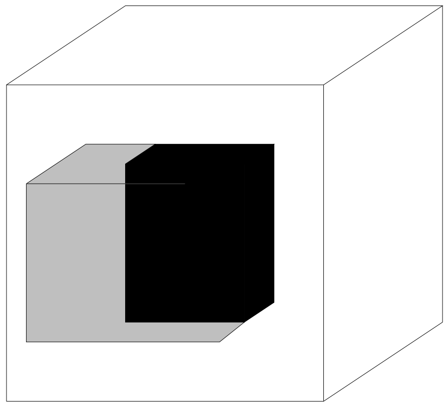


Figure 4: Extension of a three-dimensional hyperrectangle

By negating (2-2) from Definition 2.1, one obtains immediately for every matrix $M = (m_{ij}) \in M_{l,n}$:

Lemma 4.2 $X_M = \vec{I}^n \setminus \bigcup_{m_{ij}=1} R^i_j$. □

The following result shows that (non)-emptyness of the relevant trace spaces can be established by checking a bunch of inequalities. These inequalities arise via the detection of *deadlocks* in the subspaces X_M by identifying *nonempty intersections* of n *extended* hyperrectangles among the R^i_j, $m_{ij} = 1$, and their associated unsafe regions; cf Fajstrup, Goubault and Raussen [5, Theorem 2.1].

Proposition 4.3 *For a matrix $M \in M^R_{l,n}$, the following are equivalent:*

(1) *M is* not *an object in $\mathcal{C}(X)(\mathbf{0}, \mathbf{1})$.*

(2) $\vec{T}(X_M)(\mathbf{0}, \mathbf{1}) = \varnothing$.

(3) *There is a map $i\colon [1:n] \to [1:l]$ such that*

$$m_{i(j),j} = 1 \text{ for all } 1 \le j \le n \quad \text{and} \quad \bigcap_{1 \le j \le n} R^{i(j)}_j \neq \varnothing.$$

(4) *There is a map* $i\colon [1:n] \to [1:l]$ *with*

$$a_j^{i(j)} < b_j^{i(k)} \text{ for all } j, k \in [1:n].$$

Proof The equivalence of (1) and (2) follows from the definition of $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$; cf (3-2). To establish the equivalence of (3) and (4), note that an intersection of (homothetic) hyperrectangles is nonempty if and only if each coordinate of the bottom vertex of one of the participating hyperrectangles is smaller than the corresponding coordinates of the top vertices of all the other participating hyperrectangles. From Remark 4.1 we know that all but one of the lower coordinates of the $R_j^{i(j)}$ are *zero*; the requirement has thus only to be checked for $a_j^{i(j)}$; exactly what is required in (4).

Assuming (2), ie $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$, the bottom vertex $\mathbf{0}$ must be contained in the unsafe region associated to a deadlock ($\neq \mathbf{1}$) for some configuration of n forbidden hyperrectangles chosen among the R_j^i, $j \in J_i$. (If a deadlock making $\mathbf{0}$ unsafe arises by a configuration containing one or several of the *original* hyperrectangles R^i, then extending R^i to some $R_j^i, m_{ij} = 1$, will enlarge the compound obstruction and certainly again give rise to a configuration with the same property. Hence, we may restrict attention to configurations consisting of *extended* hyperrectangles only. It is important that the matrix $M \in M_{l,n}^R$ has no zero row vector for this argument to hold; cf also Remark 4.4 below.) The existence of a deadlock in X_M is equivalent to the existence of a nonempty intersection $\bigcap_{1\le j\le n} R_j^{i(j)}$ (see Fajstrup, Goubault and Raussen [5, Theorem 2.1]), ie of a map as given in (3).

On the other hand, granted (3), if $\bigcap_{1\le j\le n} R_j^{i(j)} \neq \varnothing$, the intersection gives rise to a deadlock $\mathbf{e} = (e_1, \dots, e_n) \neq \mathbf{1}$ in X_M; in fact the coordinates e_j of $\mathbf{e}$ are *maximal* among the j–th coordinates of the $R_j^{i(j)}$; in our case $e_j = a_j^{i(j)}$, cf [5, Theorem 2.1]. The associated unsafe region has as its *bottom* vertex the point in X the n coordinates of which are *next to* maximal among these lower coordinates of the $R_j^{i(j)}$, $1 \le j \le n$ [5, Theorem 2.2].

Now we exploit that the extended hyperrectangles are special (cf Remark 4.1), in the sense that all these coordinates (next to maximal among the lower coordinates) are 0! Hence $\mathbf{0}$ is automatically in the unsafe region associated to (any) deadlock $\mathbf{e} \neq \mathbf{1}$ in X_M. In particular, there is no d-path with source $\mathbf{0}$ leaving $\downarrow\mathbf{e}$. This proves (2): $\vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$. □

Remark 4.4 In proving (2) implies (3) above, it is crucial that all index sets J_i are nonempty. Otherwise, a number of extended hyperrectangles R_j^i might, together with some of the original R^i, generate a deadlock with $\mathbf{0}$ in the unsafe region that does *not* arise from a nonempty intersection of *extended* hyperrectangles. Figure 5 gives an illustration for that phenomenon in dimension two.

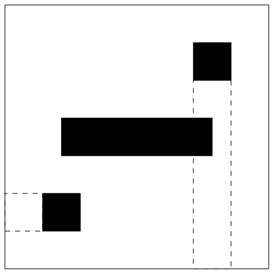


Figure 5: Deadlock arising from a combination of extended and nonextended rectangles

4.2 Algorithmic determination of $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$

4.2.1 The map Ψ and its properties We will also consider the following subset of the set of binary matrices in $M_{l,n}$:

- $M^C_{l,n} \subset M_{l,n}$ consists of the matrices such that *every column vector is a unit vector* – they form a subset with l^n elements. Every such matrix M represents the characteristic function of the graph of some map $i\colon [1:n] \to [1:l]$; cf Proposition 4.3(3-4).

To $i\colon [1:n] \to [1:l]$, we associate the matrix $M(i) \in M^C_{l,n}$ whose only nonzero coefficients are given as $m_{i(j),j}$.

Define the map $\Psi\colon M_{l,n} \to \mathbf{Z}/2$ by $\Psi(M) = 1 \Leftrightarrow \vec{T}(X_M)(\mathbf{0},\mathbf{1}) = \varnothing$; equivalently, $\Psi(M) = 0 \Leftrightarrow M \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$ for matrices $M \in M^R_{l,n}$.

Proposition 4.5 (1) *Ψ is order-preserving.*

(2) *$\Psi(M) = 0$ if $M \in M_{l,n}$ has a zero vector among its column vectors.*

(3) *$\Psi(M) = 1 \Leftrightarrow$ there exists $N \in M^C_{l,n}$ with $\Psi(N) = 1$ and $N \leq M \in M_{l,n}$.*

Proof (1) If $M \leq M' \in M_{l,n}$, then $\vec{T}(X_{M'})(\mathbf{0},\mathbf{1}) \subseteq \vec{T}(X_M)(\mathbf{0},\mathbf{1})$. If the latter set is empty, the first set needs to be empty, as well.

(2) Assume that the j–th column in M is the zero vector. Then no obstruction hyperrectangle R^i is extended in direction j. Hence, all j–th lower coordinates chosen from the *extended* hyperrectangles corresponding to M are *strictly positive*. In particular, $\mathbf{0}$ is *not* contained in the unsafe region of any deadlock occurring in X_M; in particular, there exists a d-path from $\mathbf{0}$ to $\mathbf{1}$ since neither the lower face $x_j = 0$ in I^n nor the upper boundary $\partial_+ \downarrow \mathbf{1}$ intersect any of the extended hyperrectangles.

(3) One implication is an immediate consequence of (1).

For the other implication, we may assume in view of (2) that a matrix M with $\Psi(M) = 1$ has no zero vector among its column vectors. By Proposition 4.3(3), there is a map $i\colon [1:n] \to [1:l]$ giving rise to a deadlock with $\mathbf{0}$ in the associated unsafe region. In particular, we conclude for the associated matrix $N = M(i) \in M_{l,n}^C$ with $M(i) \le M$: $\vec{T}(X_{M(i)})(\mathbf{0},\mathbf{1}) = \varnothing$ and hence $\Psi(M(i)) = 1$. □

The determination of Ψ can thus be performed in two steps. First, we determine the restriction of Ψ to the subset $M_{l,n}^C$ consisting of matrices $M(i)$ that correspond to maps $i\colon [1:n] \to [1:l]$. In particular, we determine the set of matrices (D for "dead")

$$D(X)(\mathbf{0},\mathbf{1}) := \{M \in M_{l,n}^C \mid \Psi(M) = 1\}. \tag{4-2}$$

Using this set $D(X)(\mathbf{0},\mathbf{1})$, we will then apply Proposition 4.5(3) to determine the set of matrices

$$\mathcal{C}(X)(\mathbf{0},\mathbf{1}) := \{M \in M_{l,n}^R \mid \Psi(M) = 0\}. \tag{4-3}$$

describing the objects of the relevant index category.

4.2.2 Determination of $D(X)(\mathbf{0},\mathbf{1})$ To a matrix $M \in M_{l,n}^C$, we associate its *row set* $R(M) := \{1 \le i \le l \mid \mathbf{m}^i \ne \mathbf{0}\} \subseteq [1:l]$ – indexing the *nonzero rows* $\mathbf{m}^i$ of M. The row set $R(M(i))$ is equal to the image $i([1:n]) \subseteq [1:l]$.

The condition from Proposition 4.3(4) leads us to consider the *same* upper bounds $b_j^{i(k)}$ for matrices M with the *same* row set $R(M) = B \subseteq [1:l]$: To each of the

$$\sum_{k=1}^{\min(n,l)} \binom{n}{k} \le 2^l - 1$$

nonempty subsets $B \subseteq [1:l]$ of cardinality at most $\min(n,l)$ corresponds an upper bound $\mathbf{b}^B = (b_1^{r_1}, \dots, b_n^{r_n}) \in [0,1]^n$ with $b_j^{r_j} = \min_{i \in B} b_j^i$.

Ordering the j–th coordinates a_j^i, resp. b_j^i of subinterval boundaries for $\vec{I}_j^i \subset \vec{I}_j$ (eg by a quicksort algorithm) gives rise to $2n$ (not necessarily well-determined) permutations $\pi_j^0, \pi_j^1 \in \Sigma_l$ such that

$$a_j^{\pi_j^0(1)} \le \dots \le a_j^{\pi_j^0(n)} \quad \text{and} \quad b_j^{\pi_j^1(1)} \le \dots \le b_j^{\pi_j^1(n)}.$$

A comparison of these two ordered lists leads to maps $C_j\colon [1:l] \to [1:l]$ given by

$$C_j(k) := \max\{r \mid 1 \le r \le l, a_j^{\pi_j^0(r)} < b_j^{\pi_j^1(k)}\}, \quad 1 \le j \le n.$$

Note that $C_j(k) \ge k$ for all j and that C_j is monotone. Only the relative order of the a_j^i, b_j^i matters!

Every n–tuple $(r_1,\dots,r_n)$ in $[1:l]^n$ corresponds to a "mixed vertex"

$$(a_1^{\pi_1^0(r_1)},\dots,a_n^{\pi_n^0(r_n)});$$

likewise $(s_1,\dots s_n)$ corresponds to

$$(b_1^{\pi_1^1(s_1)},\dots,b_n^{\pi_n^1(s_n)}).$$

The upper bound $\mathbf{b}^B$ corresponding to a nonempty (row) subset $B\subseteq[1:l]$ is determined by the (componentwise) minimum $(s_1^B,\dots,s_n^B)$ of the $(s_1^i,\dots,s_n^i)$ corresponding to the $\mathbf{b}^i, i\in B$; these upper bounds can be calculated recursively.

For such a nonempty row set $B\subseteq[1:l]$, let

$$\widetilde{R}_j(B):=\{i\in[1:l]\mid a_j^i<b_j^{\pi^1(s_j^B)}\}=\pi_j^0([1:C_j(s_j^B)])$$

and $R_j(B):=\widetilde{R}_j(B)\cap B$. From condition (4) in Proposition 4.3, we conclude:

Lemma 4.6 *A map $i\colon[1:n]\to[1:l]$ gives rise to a matrix $M=M(i)\in D(X)(\mathbf{0},\mathbf{1})$ if and only if*

$$i(j)\in R_j(i([1:n]))\text{ for every }1\le j\le n. \qquad\square$$

What is left is to describe a method that determines the sets $R_j(B)$ for every nonempty subset $B\subseteq[1:l]$ of cardinality at most n. For the determination of the sets $R_j(B)$, the following properties – in particular (4) – are helpful:

Lemma 4.7 (1) *If $B=\{i\}$ is a one-element set, then $R_j(B)=B$ for $j\in[1:n]$. Hence $\Psi(M)=1$ for each of the l matrices $M\in M_{l,n}^C$ with a one-element row set $R(M)$.*

(2) $\emptyset\neq B\subseteq C\subseteq[1:l]\Rightarrow\widetilde{R}_j(B)\supseteq\widetilde{R}_j(C),\ 1\le j\le n.$

(3) $R_j(B\cup C)=(R_j(B)\cap\widetilde{R}_j(C))\cup(\widetilde{R}_j(B)\cap R_j(C)).$

(4) *For $i\notin B$, let $\mathbf{b}^B$ correspond to $(s_1^B,\dots,s_n^B)$, $\mathbf{b}^i$ to $(s_1^i,\dots s_n^i)$ and $\mathbf{a}^i$ to $(r_1^i,\dots r_n^i)$. Then*

$$R_j(B\cup\{i\})=\begin{cases}R_j(B), & C_j(s_j^B)<r_j^i,\\ R_j(B)\cup\{i\}, & r_j^i<C_j(s_j^B), s_j^B<s_j^i,\\ (\pi_j^0([1:C_j(s_j^i)])\cap R_j(B))\cup\{i\}, & s_j^i<s_j^B.\end{cases}$$

Proof (1) follows from (the proof of) Lemma 3.3; (2) is obvious. For (3), note that $R_j(B\cup C)=(\widetilde{R}_j(B)\cap\widetilde{R}_j(C))\cap(B\cup C)$ and use distributivity. (4) is an easy consequence. $\square$

4.2.3 Determination of $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ From Proposition 4.5 and Lemma 4.6, we can conclude immediately:

Proposition 4.8 *Let $M \in M_{l,n}^R$.*

(1) $\Psi(M) = 1$ *if and only if there is a matrix $N \in D(X)(\mathbf{0},\mathbf{1})$ (cf Lemma 4.6) such that $n_{ij} \le m_{ij}$ for all $1 \le i \le l, 1 \le j \le n$.*

(2) $\Psi(M) = 0 \Leftrightarrow M \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$ *if and only if, for every matrix $N \in D(X)(\mathbf{0},\mathbf{1})$, there is a pair $(i,j) \in [1:l]\times[1:n]$ such that $m_{ij} = 0, n_{ij} = 1$.* □

Matrices that are *maximal* with respect to the partial order $\le$ on binary matrices within $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ constitute – by definition – the subset $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1}) \subseteq \mathcal{C}(X)(\mathbf{0},\mathbf{1})$. By Proposition 4.5(1) they determine the index category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$. They correspond to maximal simplex products in the prodsimplicial space $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$.

To determine the matrices contained in these two sets, we consider (choice) subsets $C \subseteq [1:l]\times[1:n]$ characterized by the property:

- For every matrix $N \in D(X)(\mathbf{0},\mathbf{1})$ there exists $(i,j) \in C$ with $n_{ij} = 1$.

Remark that one index (i,j) can count for several matrices N.

Functions $m_C = 1-\chi(C)$ for such choices are then exactly the characteristic functions for matrices $M_C = (m_{ij}) \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$.

A choice C is *minimal*, if for every $C' \subset C$ there is a matrix $N \in D(X)(\mathbf{0},\mathbf{1})$ with $n_{ij} = 0$ for each $(i,j) \in C'$. The function $m_C = 1-\chi(C)$ for a *minimal* choice function is then the characteristic function of a *maximal* matrix $M_C \in \mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})$.

We describe a simple-minded algorithm constructing $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})$ step by step given $D(X)(\mathbf{0},\mathbf{1}) = \{D_1,\dots,D_p\}$ starting with $A^0_{\max}(X)(\mathbf{0},\mathbf{1})$ with the matrix $\mathbf{1}$ consisting of only 1s as the only element. Assume $A^{h-1}_{\max}(X)(\mathbf{0},\mathbf{1}) = \{M_1,\dots,M_{q_{h-1}}\}$ to consist of the maximal binary matrices M such that $N_k \not\le M$ for $1 \le k \le h-1 < p$.

Compare the matrices $M_l \in A^{h-1}_{\max}(X)(\mathbf{0},\mathbf{1})$ to N_h. If $N_h \not\le M_l$, then keep M_l unchanged as an element of $A^h_{\max}(X)(\mathbf{0},\mathbf{1})$; if $M(i) = N_h \le M_l$ (cf Section 4.2.1), then replace M_l by the n matrices $M_l^1,\dots M_l^n \in A^h_{\max}(X)(\mathbf{0},\mathbf{1})$ determined as follows: M_l^j arises from M_l by replacing $m^l_{i(j),j} = 1$ in M_l by $m^{l,j}_{i(j),j} = 0$ in M_l^j.

Assessing whether $N \le M$ is easy, given N: form the binary product $\bigwedge_{n_{pq}=1} m_{pq}$; this product is always over n entries for $N \in M^C_{l,n}$.

The maximal matrices in $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})$ correspond to the *maximal simplex products* that are patched together in $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) \subset (\Delta^{n-1})^l$ while the matrices in $D(X)(\mathbf{0},\mathbf{1})$

correspond to *minimal nonfaces* in $(\Delta^{n-1})^l$. The construction above reminds of a similar construction of a simplicial complex $K(\mathcal{F})$ associated to a set system $\mathcal{F}$ used for topological investigations of colouring problems; cf eg Matoušek and Ziegler [21]; in our case, we have the product structure in the underlying category $M_{l,n}$ as an additional feature.

Corollary 4.9 *Let $C \subset [1:l] \times [1:n]$ denote a choice subset.*

(1) *The simplex product in $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ corresponding to $M_c \in \mathcal{C}(X)(\mathbf{0},\mathbf{1})$ considered as an object in $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ (cf Section 3.2) has dimension $(n-1)l - |C|$.*

(2) $\dim \mathbf{T}(X)(\mathbf{0},\mathbf{1}) = (n-1)l - \min |C|$.

Proof The simplex product corresponding to M_C has type $\prod_{1 \le i \le l} \Delta^{n-1-c_i}$ with $c_i = |\{j \mid (i,j) \in C\}|$. □

Corollary 4.10 *The Lusternik–Schnirelmann category of trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ satisfies the inequality* $\mathrm{cat}(\vec{T}(X)(\mathbf{0},\mathbf{1})) \le |\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})|$.

Proof The prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ homotopy-equivalent to $\vec{T}(X)(\mathbf{0},\mathbf{1})$ is covered by maximal products of simplices; there are $|\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})|$ of those. As products of simplices, they are contractible; they are deformation retracts of contractible open neighbourhoods in $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$. □

4.2.4 Examples

Example 4.11 (1) X is a square with two square holes as in the first row of Figure 2, Example 2.3: $D(X)(\mathbf{0},\mathbf{1})$ consists then of the two matrices

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

since a deadlock arises only if a rectangle is extended in both directions from the *same* obstruction. There are then four minimal choices $C \subset [1:2] \times [1:2]$; each of them has two elements. Hence $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1}) = \mathcal{C}(X)(\mathbf{0},\mathbf{1})$ consists of the four matrices

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Extending each of the rectangles in one direction according to the recipe encoded in one of these matrices yields the subspaces shown in the upper row of Figure 2. Each of them allows exactly one d-homotopy class around the (extended) holes.

The corresponding prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ of type $(\partial\Delta^1)^2$ consists of four points.

(2) X is a square with two square holes as in the second row of Figure 2, Example 2.3: This time, $D(X)(\mathbf{0},\mathbf{1})$ consists of the three matrices

$$\begin{bmatrix}1&1\\0&0\end{bmatrix},\quad \begin{bmatrix}0&0\\1&1\end{bmatrix},\quad \begin{bmatrix}0&1\\1&0\end{bmatrix}:$$

one additional deadlock configuration corresponding to X_{12} comes up with one extension for every hole. This time, there are only three minimal choices C, each of cardinality 2, corresponding to the matrices

$$\begin{bmatrix}0&1\\0&1\end{bmatrix},\quad \begin{bmatrix}1&0\\0&1\end{bmatrix},\quad \begin{bmatrix}1&0\\1&0\end{bmatrix}$$

in $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})=\mathcal{C}(X)(\mathbf{0},\mathbf{1})$. $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ consists thus of three points.

(3) X a square with three holes as in Figure 5: In this case,

$$D(X)(\mathbf{0},\mathbf{1})=\left\{\begin{bmatrix}1&1\\0&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\1&0\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&0\\1&1\end{bmatrix}\right\}.$$

There are four minimal choices giving rise to the matrices in

$$\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})=\left\{\begin{bmatrix}1&0\\1&0\\1&0\end{bmatrix},\begin{bmatrix}1&0\\1&0\\0&1\end{bmatrix},\begin{bmatrix}1&0\\0&1\\0&1\end{bmatrix},\begin{bmatrix}0&1\\0&1\\0&1\end{bmatrix}\right\}.$$

The complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ consists of four points.

(4) $X=\vec{I}^n\setminus\vec{J}^n$ as in Example 2.4: In this case, $D(X)(\mathbf{0},\mathbf{1})$ has only one element, the 1–row matrix $[1,1,\dots,1]$. The n minimal choices correspond to the one row matrices with exactly one entry 0 in $\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1})$. These matrices correspond to the maximal simplices in $\mathbf{T}(X)(\mathbf{0},\mathbf{1})=\partial\Delta^{n-1}$. Products of spheres arise likewise in the case considered in Example 3.4.

We conclude from Theorem 3.5:

Corollary 4.12 $\vec{T}(\vec{I}^n\setminus\vec{J}^n)(\mathbf{0},\mathbf{1})\simeq\partial\Delta^{n-1}$. □

Previous attempts to prove Corollary 4.12 directly were far more complicated.

Example 4.13 (1) For the state space X from Figure 1, $D(X) = M^C_{2,3}$: Since the two forbidden boxes intersect, every set of extensions *in all three directions* will produce a deadlock and hence an empty trace space. As a consequence,

$$\mathcal{C}_{\max}(X)(\mathbf{0},\mathbf{1}) = \left\{ \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \right\}.$$

Hence, $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ is the union of the three diagonal squares in a flat 2–dimensional torus $(\partial\Delta^2)^2$ covered by three times three squares. In particular, $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to a circle S^1.

This example can be generalized as follows: Let $X = I^n \setminus F$ with the forbidden region $F = \bigcup_{i=1}^l R^i$ consisting of $l \le n$ hyperrectangles R^i with nonempty intersection $\bigcap_{i=1}^l R^i$. Then, $\mathbf{T}(X)(\mathbf{0},\mathbf{1}) = \bigcup_f (\partial\Delta^{n-1} \setminus f)^l$ with f ranging over the n faces of $\partial\Delta^{n-1}$. That latter trace space is homotopy equivalent to S^{n-2} – and the homotopy type is thus independent of the number l of contributing hyperrectangles – by the following argument: Consider the diagonal inclusion

$$i\colon S^{n-2} \simeq \partial\Delta^{n-1} = \bigcup_f \partial\Delta^{n-1} \setminus f \hookrightarrow \bigcup_f (\partial\Delta^{n-1} \setminus f)^l = \mathbf{T}(X)(\mathbf{0},\mathbf{1}).$$

Both spaces are colimits of contractible spaces over the same poset index category given by the nonempty proper subsets of $[1:n]$ (subsets of faces), and the inclusion respects this filtration. The result follows thus from the homotopy lemma [20, Theorem 15.12].

More generally, it seems promising to use the intersection pattern among the contributing hyperrectangles as an input simplifying the determination of the index categories.

(2) The space X in Figure 6 below shows a cube from which two wedges, each of them composed of two intersecting rectangular boxes are removed. Remark that the two wedges do not touch each other. The trace in that drawing from bottom to top is homotopic but *not dihomotopic* (homotopic through a 1–parameter deformation of d-paths) to a trace on the boundary of the cube. A simple-minded analysis of this model in Raussen [25] showed by a quite intricate argument that the trace space for this d-space (from bottom to top) is not connected.

The general method described in this article yields a model for trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ as a subspace of the 4-torus $(\partial\Delta^2)^4 \cong (S^1)^4$. It turns out by inspection (similar to (1) above) that one can handle each of the two wedges as one obstruction in this case. It turns out that the trace space can then be seen as the union of five squares and a disjoint extra ("corner") point in the two-torus $(\partial\Delta^2)^2 = (S^1)^2$ on the right hand side of Figure 6 above. This subspace is of course homotopy equivalent to the disjoint union of a wedge of circles and of an extra point: $\vec{T}(X)(\mathbf{0},\mathbf{1}) \simeq (S^1 \vee S^1) \sqcup *$.

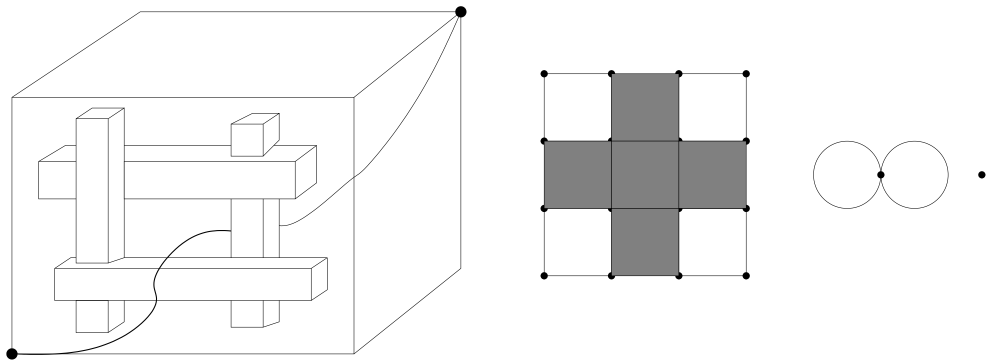


Figure 6: State space X and associated models for trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$

It would be interesting to find more general methods for dimension reduction as the one described above.

4.2.5 A reformulation: Minimal transversals in hypergraphs The search for minimal choices in $D(X)(\mathbf{0},\mathbf{1})$ can be translated into a well-known and well-investigated problem in combinatorics.[1] The set $D(X)(\mathbf{0},\mathbf{1})$ may be considered as a *hypergraph* (with hyperedges = simplices connecting a number of vertices; every matrix in $D(X)(\mathbf{0},\mathbf{1})$ defines a hyperedge) on the vertex set $[1:l]\times[1:n]$. A minimal choice – that allows to find the maximal elements in $\mathcal{C}_{\max}(\mathbf{0},\mathbf{1})$; cf Section 4.2.3 – is then a *minimal transversal* (or *hitting set*) of that hypergraph: This means that it has nonempty intersection with every hyperedge and it is minimal with this property. Computing minimal transversals has many applications (eg, machine learning, indexing of databases, data mining and optimization). There are several articles about algorithms for finding minimal transversals and their complexity in the literature; cf eg Khachiyan, Boros, Elbassioni and Gurvich [19].

The hypergraph given by the matrices in $D(X)(\mathbf{0},\mathbf{1})$ has special properties: All hyperedges have the same cardinality n; even more so, they are graphs of functions from $[1:n]$ to $[1:l]$. This ought to simplify the setting.

4.3 Homology of the trace space

By Theorem 3.5, the homology of the trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$ may be calculated as the homology of the associated prodsimplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$. Given the poset category $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$, this is the homology of a particular chain complex $C(X)(\mathbf{0},\mathbf{1})$ with one generator for every product of simplices.

[1] I would like to thank my colleague Leif Kjær Jørgensen, Aalborg University, for mentioning hypergraphs and their transversals to me.

More precisely, let $C_k(X)(\mathbf{0},\mathbf{1})$ denote the free R–module generated by all matrices in $\mathcal{C}(X)(\mathbf{0},\mathbf{1})$ with $(k+l)$ entries 1; R denotes the chosen coefficient ring. For a matrix $M \in M_{l,n}$ with $m_{pq} = 1$, let M_{pq} be given by

$$(m_{pq})_{ij} = \begin{cases} m_{ij} & (i,j) \neq (p,q), \\ 0 & (i,j) = (p,q). \end{cases}$$

The boundary operator ∂ on $C(X)(\mathbf{0},\mathbf{1})$ is then given by

$$\partial(M) = \sum_{m_{pq}=1} (-1)^{|(p,q)|} M_{pq}$$

with alternating sign: the integer $|(p,q)| = \sum_{i=1}^{p-1} \sum_{j=1}^{n} m_{ij} + \sum_{j=1}^{q} m_{pj} - 1$ takes account of the ones in M preceding $m_{pq} = 1$.

It should be interesting to perform actual homology calculations in "real life" examples that give rise to huge chain complexes. The algorithms for the calculation of homology in Kaczynski, Mrozek and Slusarek [18] by reduction of chain complexes (with field coefficients) might be helpful. Likewise a modification of the algorithms in Kaczynski, Mischaikow and Mrozek [17] for the homology of cubical complexes.

5 Models for more general trace spaces

5.1 Trace spaces corresponding to concurrent nonlinear programs

So far, we have only looked at model spaces corresponding to concurrent *linear* programs; without branchings, mergings and loops. More realistic models can be investigated using the same tools – but with a twist: Let $\Gamma = \prod_{j=1}^{n} \Gamma_j$ denote a product of *directed* graphs (branchings, mergings and loops allowed); each Γ_j represents a program run by a single processor. The graphs Γ_j are regarded as d-spaces (realizations of precubical sets of dimension one), and Γ is given the product structure: as an n–dimensional precubical complex with d-space structure (cf Grandis [12]).

A *directed interval* J_j from a_j to b_j in the geometric realization of a component Γ_j is *uniquely* given by the image $p_j(I)$ of a trace $p_j \in \vec{T}(\Gamma_j)(a_j, b_j)$ arising from an *injective* d-path; no loops allowed here! Remark on the other hand, that such a directed interval is not always determined uniquely by its end points.

A *(generalized) hyperrectangle* in Γ is a product $R = \prod_j J_j \subseteq \prod_j \Gamma_j = \Gamma$ of such directed intervals. A forbidden region $F = \bigcup_i R^i$ is a union of such generalized hyperrectangles, and the state space $X = \Gamma \setminus F$ is its complement.

The aim is to analyse the space of d-paths $\vec{P}(X)(\mathbf{x},\mathbf{y}) \subseteq \vec{P}(\Gamma)(\mathbf{x},\mathbf{y})$ (or the space of traces $\vec{T}(X)(\mathbf{x},\mathbf{y})$ homotopy equivalent to it) between two points $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ in the space X. First, we have a look at the (bigger) space $\vec{T}(\Gamma)(\mathbf{x},\mathbf{y})$ and then, we will use the map induced on traces by the inclusion map $i_X\colon X \hookrightarrow \Gamma$.

For a directed graph – no cubes of higher dimension supporting homotopies are available – dihomotopy of d-paths (with fixed end points) is equivalent to reparametrization equivalence; cf Fahrenberg and Raussen [4]. In particular, each factor $\vec{T}(\Gamma_j)(x_j, y_j)$ is *discrete*; every component is represented by a (reparametrization equivalence class of) a particular directed path from x_j to y_j. The product structure of Γ yields:

Lemma 5.1 $\vec{T}(\Gamma)(\mathbf{x},\mathbf{y}) \simeq \prod \vec{T}(\Gamma_j)(x_j, y_j)$. *In particular,* $\vec{T}(\Gamma)(\mathbf{x},\mathbf{y})$ *is* homotopy discrete.

Proof The d-space structure and the dihomotopy relations factor:

$$\vec{T}(\Gamma) \simeq \vec{P}(\Gamma) \cong \prod \vec{P}(\Gamma_j) \simeq \prod \vec{T}(\Gamma_j). \qquad \square$$

Remark 5.2 To enumerate the components (=traces) of the space of d-paths in a directed graph Γ, one should first reduce Γ to normal form $N(\Gamma)$: Vertices with valency $(1, -1)$ – exactly one ingoing and one outgoing arrow, different from each other – are suppressed; the two arrows are concatenated to one. The normal form $N(\Gamma)$ does no longer have such vertices.

Attach a unique label to each arrow in a directed graph Γ *in normal form* and form words in these labels along concatenable arrows. Then $\vec{T}(\Gamma)$ corresponds to the discrete set of such words; $\vec{T}(\Gamma)(x, y)$ to the words starting and ending with one or several specific labels, depending on whether x, y correspond to vertices or to points on a directed edge. There is no need to distinguish between points on the interior of the same edge.

Each component $C \in \vec{T}(\Gamma)(\mathbf{x},\mathbf{y})$ can thus be represented by an n–tuple of (traces of) specific d-paths $c_j \in \vec{P}(\Gamma_j)(x_j, y_j)$. As representatives, we choose $c_j \in \vec{R}(\Gamma_j)(x_j, y_j) \subset \vec{P}(\Gamma_j)(x_j, y_j)$ to be *regular* (ie, locally injective; cf Fahrenberg and Raussen [4])); every other d-path in $\vec{P}(\Gamma_j)(x_j, y_j)$ dihomotopic to c_j is then a reparametrization $c_j \circ \varphi_j$ of c_j with $\varphi_j \in \vec{P}(\vec{I})(0,1)$ an (increasing) d-path in the standard ordered unit interval $\vec{I}$ [4, Theorem 3.6 and Proposition 3.8]. The d-paths c_j altogether define a d-map (a specific "delooping") $c\colon \vec{I}^n \to \Gamma$ given as $c(t_1, \dots, t_n) = (c_1(t_1), \dots, c_n(t_n))$, and:

Lemma 5.3 *The d-map $c\colon \vec{I}^n \to \Gamma$ induces a homeomorphism*

$$c\circ\colon \vec{T}(\vec{I}^n)(\mathbf{0},\mathbf{1}) \to C \subset \vec{T}(\Gamma)(\mathbf{x},\mathbf{y}),\ p \mapsto c \circ p. \qquad \square$$

Comparing d-paths in Γ and in $X = \Gamma\backslash F$ and given such a component $C \in \pi_0(\vec{T}(\Gamma)) \cong \prod \pi_0(\vec{T}(\Gamma_j))$, the following two questions arise naturally:

(1) Does C lift to X at all (ie, can it be represented by an – interleaving – d-path in X rather than in Γ)?

(2) Determine the topology of $i_X^{-1}(C)$, ie, of the space of all d-paths in X whose projections to the Γ_j are (reparametrizations of) these specified execution paths ("interleavings").

Every directed interval $J =]a_j, b_j[\subset \Gamma_j$ (in the sense above) pulls back to the standard interval $c_j^{-1}(]a_j, b_j[) \subset I$ – which is an open subinterval of I in the subspace topology, possibly empty. To each generalized hyperrectangle $R^i = \prod J_j^i \subset \Gamma$ corresponds thus an (honest) hyperrectangle $\widetilde{R}^i = c^{-1}(R^i) = \prod c_j^{-1}(J_j^i)$, possibly empty. The forbidden region $F \subset \Gamma$ corresponds to a forbidden region $\widetilde{F} = c^{-1}(F) = \bigcup_i \widetilde{R}^i \subset I^n$, leaving $\widetilde{X} = I^n \setminus \widetilde{F} \subset I^n$ as state space (with the d-structure inherited from $\vec{I}^n$). By restricting the homeomorphism $c\circ$ from Lemma 5.3, we obtain

Corollary 5.4 *The d-map $c\colon \widetilde{X} \to X$ induces a homeomorphism*

$$c\circ\colon \vec{T}(\widetilde{X})(\mathbf{0},\mathbf{1}) \to i_X^{-1}(C) \subset \vec{T}(X)(\mathbf{x},\mathbf{y}). \qquad \square$$

Example 5.5 Let $\Gamma_1 = \vec{S}^1$ denote a circular digraph and $\Gamma_2 = \vec{I}$ a linear one. Let $X = (\Gamma_1 \times \Gamma_2) \setminus J^2$ with $J^2 \subset (\Gamma_1 \times \Gamma_2)$ an open rectangular hole; cf Figure 7. The component C_r in $\vec{T}(\Gamma_1 \times \Gamma_2)(\mathbf{0},\mathbf{1})$ corresponding to $r + 1/2$ spiral tours leads to a state space $\widetilde{X}_r$ with r rectangular holes with an exponential covering map back to X:

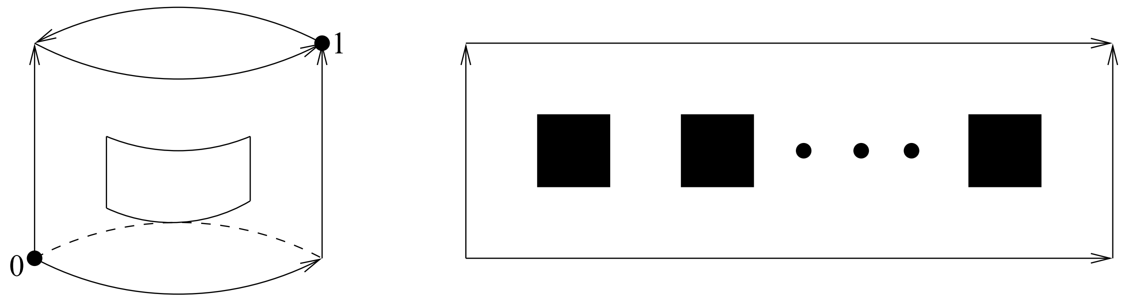


Figure 7: State spaces X (directed cylinder with hole) and $\widetilde{X}_r$

The contractible(!) components of $\vec{T}(X)(\mathbf{0},\mathbf{1})$ in this case correspond to pairs (r,s) of integers $r \geq s \geq 0$. The integer r counts the number of rounds (the delooping); the integer s describes where a d-path in the component passes from the lower to the upper annulus.

Corollary 5.4 allows us to attack the questions asked above:

(1) This question is equivalent to: Is $\vec{T}(\widetilde{X})(\mathbf{0},\mathbf{1})$ nonempty? This is the case if $\mathbf{0}$ is *not* contained in the unsafe region corresponding to any deadlock in $\widetilde{X}$ – this can be settled using the techniques described in Fajstrup, Goubault and Raussen [5]; compare also Fajstrup and Sokolowski [8].

(2) The topology of a nonempty space $i_X^{-1}(C) \cong \vec{T}(\widetilde{X})(\mathbf{0},\mathbf{1})$ can be analysed as that of the prodsimplicial complex $\mathbf{T}(\widetilde{X})(\mathbf{0},\mathbf{1})$ as described in Section 4.

Remark 5.6 The components $C \subset \vec{T}(\Gamma)(-,-)$ form the morphisms of the fundamental category $\vec{\pi}_1(\Gamma)$ (composition induced by concatenation of paths) with the elements of Γ as objects. In particular, loop components act on (the left and on the right) on components with matching end points. In [8], Fajstrup and Sokolowski have shown that unsafe areas corresponding to a specific deadlock point can look quite different for components ("deloopings") with the same end points. It should be interesting to investigate how the topology of the spaces $i_X^{-1}(C)$ behaves under composition with loops. For applications, it is essential to find out whether there is an algorithm determining them in a recursive fashion.

5.2 Simplicial models for trace spaces in precubical complexes

The methods used in this paper can certainly be applied more generally. General Higher Dimensional Automata can be described as labelled precubical complexes with a compatible d-structure defined originally for every single cube; cf Fajstrup, Goubault and Raussen [5]. In Raussen [28], we investigated spaces of d-paths in a non-self-linked precubical complex X and showed that, for all $x_0, x_1 \in X$, the path spaces $\vec{P}(X)(x_0,x_1)$ are ELCX (equi locally convex) in the sense of Milnor [22]; hence that they are locally contractible and possess the homotopy type of a CW–complex.

The main ingredient in the proof is the construction of a locally defined average map $\mu\colon U \to X$ defined on a neighbourhood $U = \bigcup_\beta V_\beta \times V_\beta$ of the diagonal with V_β the open star neighbourhood of a vertex β in X. This average map plays a role very similar to that of the least upper bound $\vee$ in Section 2.

In particular, a directed sequence of adjacent vertices and their open star neighbourhoods gives rise to a *contractible* space of d-paths (or traces) progressing consecutively through that sequence of neighbourhoods. The spaces of d-paths in all possible such sequences give rise to a covering of the space of *all* d-paths (with given end points) by contractible subspaces [28, Proposition 3.16]. Using the same method (and restrictions of the map μ above), it can be shown that intersections of such subspaces (through intersections of open stars of certain vertices) are empty or also contractible.

The nerve lemma [20, Theorem 15.21] shows then that spaces of d-paths (and thus of traces) in a precubical complex (with given end points) are homotopy equivalent to the *nerve* of the covering described above. In particular, $\vec{T}(X)(x_0, x_1)$ has an explicit structure of a simplicial complex. To describe it explicitly, one needs to know which sets of sequences of adjacent vertices give rise to open star neighbourhoods with a (common) d-path contained in each of them.

Remark 5.7 More abstractly, one may describe a category of *contractible cube paths in semicubical complexes* (with contractible trace spaces and such that all subcube paths are contractible, as well) and then consider the induced category *over* X (objects = semicubical maps from a contractible cube path into X respecting given end points). The nerve of that category (or of any subcategory that covers all d-paths in X with given end points) is then homotopy equivalent to $\vec{T}(X)(x_0, x_1)$.

References

[1] **R Brown**, **P J Higgins**, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Algebra 22 (1981) 11–41 MR621285

[2] **R Brown**, **P J Higgins**, *On the algebra of cubes*, J. Pure Appl. Algebra 21 (1981) 233–260 MR617135

[3] **E W Dijkstra**, *Co-operating sequential processes*, from: "Programming Languages", (F Genuys, editor), Academic Press, New York (1968) 43–110

[4] **U Fahrenberg**, **M Raussen**, *Reparametrizations of continuous paths*, J. Homotopy Relat. Struct. 2 (2007) 93–117 MR2369163

[5] **L Fajstrup**, **E Goubault**, **M Raussen**, *Detecting deadlocks in concurrent systems*, from: "CONCUR'98: concurrency theory (Nice)", (D Sangiorgi, R de Simone, editors), Lecture Notes in Comput. Sci. 1466, Springer, Berlin (1998) 332–347 MR1683333

[6] **L Fajstrup**, **M Raussen**, **E Goubault**, *Algebraic topology and concurrency*, Theoret. Comput. Sci. 357 (2006) 241–278 MR2242768

[7] **L Fajstrup**, **M Raussen**, **E Goubault**, **E Haucourt**, *Components of the fundamental category. Homotopy theory*, Appl. Categ. Structures 12 (2004) 81–108 MR2057412

[8] **L Fajstrup**, **S Sokolowski**, *Infinitely running concurrents processes with loops from a geometric viewpoint*, Elec. Notes Theor. Comput. Sci. 39 (2000)

[9] **R J van Glabbeek**, *On the expressiveness of higher dimensional automata*, Theoret. Comput. Sci. 356 (2006) 265–290 MR2223695

[10] **E Goubault**, **E Haucourt**, *Components of the fundamental category. II*, Appl. Categ. Structures 15 (2007) 387–414 MR2350213

[11] **M Grandis**, *Directed homotopy theory. II. Homotopy constructs*, Theory Appl. Categ. 10 (2002) No. 14, 369–391 MR1921751

[12] **M Grandis**, *Directed homotopy theory. I*, Cah. Topol. Géom. Différ. Catég. 44 (2003) 281–316 MR2030049

[13] **M Grandis**, *Directed algebraic topology. Models of non-reversible worlds*, New Math. Monogr. 13, Cambridge Univ. Press (2009) MR2562859

[14] **J Gunawardena**, *Homotopy and concurrency*, Bull. EATCS 54 (1994) 184–193

[15] **M Herlihy**, **S Rajsbaum**, *Algebraic topology and distributed computing—a primer*, from: "Computer science today", (J van Leeuwen, editor), Lecture Notes in Comput. Sci. 1000, Springer, Berlin (1995) 203–217 MR1389577

[16] **J F Jardine**, *Path categories and resolutions*, Preprint (2009) Available at `http://www.math.uwo.ca/~jardine/papers/preprints/pathcat-paper3.pdf`

[17] **T Kaczynski**, **K Mischaikow**, **M Mrozek**, *Computational homology*, Applied Math. Sciences 157, Springer, New York (2004) MR2028588

[18] **T Kaczynski**, **M Mrozek**, **M Ślusarek**, *Homology computation by reduction of chain complexes*, Comput. Math. Appl. 35 (1998) 59–70 MR1604791

[19] **L Khachiyan**, **E Boros**, **K Elbassioni**, **V Gurvich**, *A new algorithm for the hypergraph transversal problem*, from: "Computing and combinatorics", (L Wang, editor), Lecture Notes in Comput. Sci. 3595, Springer, Berlin (2005) 767–776 MR2190900

[20] **D Kozlov**, *Combinatorial algebraic topology*, Algorithms and Computation in Math. 21, Springer, Berlin (2008) MR2361455

[21] **J Matoušek**, **G M Ziegler**, *Topological lower bounds for the chromatic number: a hierarchy*, Jahresber. Deutsch. Math.-Verein. 106 (2004) 71–90 MR2073516

[22] **J Milnor**, *On spaces having the homotopy type of a* CW*–complex*, Trans. Amer. Math. Soc. 90 (1959) 272–280 MR0100267

[23] **V Pratt**, *Modeling concurrency with geometry*, from: "POPL '91: Proceedings of the 18th ACM SIGPLAN-SIGACT symposium on Principles of programming languages", ACM, New York (1991) 311–322

[24] **M Raussen**, *On the classification of dipaths in geometric models for concurrency. Geometry and concurrency*, Math. Structures Comput. Sci. 10 (2000) 427–457 MR1786470

[25] **M Raussen**, *Deadlocks and dihomotopy in mutual exclusion models*, Theoret. Comput. Sci. 365 (2006) 247–257 MR2269456

[26] **M Raussen**, *Invariants of directed spaces*, Appl. Categ. Structures 15 (2007) 355–386 MR2350212

[27] **M Raussen**, *Reparametrizations with given stop data*, J. Homotopy Relat. Struct. 4 (2009) 1–5 MR2481616

[28] **M Raussen**, *Trace spaces in a pre-cubical complex*, Topology Appl. 156 (2009) 1718–1728 MR2521708

[29] **M Raussen**, *Simplicial models for trace spaces*, Tech. Report R-2010-02, Dept. of Math. Sciences, Aalborg University (2010) Available at `http://www.math.aau.dk/fileadmin/user_upload/www.math.aau.dk/Forskning/Rappor% tserien/R-2010-02.pdf`

[30] **G Winskel**, **M Nielsen**, *Models for concurrency*, from: "Handbook of logic in computer science, Vol. 4", (S Abramsky, D M Gabbay, T S E Maibaum, editors), Handb. Log. Comput. Sci. 4, Oxford Univ. Press, New York (1995) 1–148 MR1365754

Department of Mathematical Sciences, Aalborg University
Fredrik Bajersvej 7G, DK-9220 Aalborg Øst, Denmark

`raussen@math.aau.dk`

`http://people.math.aau.dk/~raussen`

Received: 30 March 2010

Trace Spaces: An Efficient New Technique for State-Space Reduction

Lisbeth Fajstrup[1], Éric Goubault[2], Emmanuel Haucourt[2],
Samuel Mimram[2], and Martin Raussen[1]

[1] Department of Mathematical Sciences, Aalborg University
[2] CEA, LIST*

Abstract. State-space reduction techniques, used primarily in model-checkers, all rely on the idea that some actions are independent, hence could be taken in any (respective) order while put in parallel, without changing the semantics. It is thus not necessary to consider all execution paths in the interleaving semantics of a concurrent program, but rather some equivalence classes. The purpose of this paper is to describe a new algorithm to compute such equivalence classes, and a representative per class, which is based on ideas originating in algebraic topology. We introduce a geometric semantics of concurrent languages, where programs are interpreted as directed topological spaces, and study its properties in order to devise an algorithm for computing dihomotopy classes of execution paths. In particular, our algorithm is able to compute a control-flow graph for concurrent programs, possibly containing loops, which is "as reduced as possible" in the sense that it generates traces modulo equivalence. A preliminary implementation was achieved, showing promising results towards efficient methods to analyze concurrent programs, with very promising results compared to partial-order reduction techniques.

Introduction

Formal verification of concurrent programs is traditionally considered as a difficult problem because it might involve checking all their possible schedulings, in order to verify all the behaviors the programs may exhibit. This is particularly the case for checking for liveness or reachability properties, or in the case of verification methods that imply traversal of some important parts of the graph of execution, such as model-checking [4] and abstract testing [6]. Fortunately, many of the possible executions are equivalent (we say *dihomotopic*) in the sense that one can be obtained from the other by permuting independent instructions, therefore giving rise to the same results. In order to analyze a program, it is thus enough (and much faster) to analyze one representative in each dihomotopy class of execution traces.

We introduce in this paper a new algorithm to reduce the state-space explosion during the analysis of concurrent systems. It is based on former work of some of the authors, most notably [24] where the notion of trace space is introduced and studied, and also builds up considerably on the geometric semantics approach to concurrent systems,

* This work has been supported by the PANDA ("Parallel and Distributed Analysis", ANR-09-BLAN-0169) French ANR project and by ESF project ACAT.

H. Seidl (Ed.): ESOP 2012, LNCS 7211, pp. 274–294, 2012.

as developed in [13]. Some fundamentals of the mathematics involved can be found in [19]. The main contributions of this article are the following: we develop and improve the algorithms for computing trace spaces of [24] by reformulating them in order to devise an efficient implementation for them, we generalize this algorithm to programs which may contain loops and thus exhibit an infinite number of behaviors, we apply these algorithms to a toy shared-memory language whose semantics is given in the style of [12], but in this paper, formulated in terms of d-spaces [19], and we report on the implementation and experimentation of our algorithms on trace spaces – an industrial case-study using those methods is also detailed in [3].

Stubborn sets [25], sleep sets and persistent sets [15] are among the most popular methods used for diminishing the complexity of model-checking using transition systems; they are in particular used in SPIN [1], with which we compare our work experimentally in Section 2.5. They are based on semantic observations using Petri nets in the first case and Mazurkiewicz trace theory in the other one. We believe that these are special forms of dihomotopy-based reduction as developed in this paper when cast in our geometric framework, using the adjunctions of [18]. Of course, the trace spaces we are computing have some acquaintance with traces as found in trace theory [7]: basically, traces in trace theory are points of trace spaces, and composition of traces modulo dihomotopy is concatenation in trace theory. Trace spaces are more general in that they consider general directed topological spaces and not just partially commutative monoids; they also include all information related to higher-dimensional (di-)homotopy categories, and not just the fundamental category, as in trace theory. Trace spaces are also linked with component categories, introduced by some of the authors [14,17], and connected components of trace spaces can also be computed using the algorithm introduced in [16].

Contents of the paper. We first define formally the programming language we are considering (Section 1.1) as well as an associated geometric semantics, (Section 1.2). We then introduce an algorithm for computing an effective combinatorial representation of trace spaces as well as an efficient implementation of it (Section 2), and extend this algorithm in order to handle program containing loops (Section 3). Finally, we discuss various applications, in particular to static analysis (Section 3.5) and possible extensions of the algorithm and conclude.

1 Geometric Semantics of Concurrent Processes

1.1 A Toy Shared-Memory Concurrent Language

In this paper, we consider a toy imperative shared-memory concurrent language as grounds for experimentation. In this formalism, a program can be constituted of multiple subprograms which are run in parallel. The environment provides a set of resources $\mathcal{R}$, where each resource $a \in \mathcal{R}$ can be used by at most κ_a subprograms at the same time, the integer $\kappa_a \in \mathbb{N}$ being called the *capacity* of the resource a. In particular, a *mutex* is a resource of capacity 1.

Whenever a program wants to access a resource a, it should acquire a lock by performing the action P_a which allows access to a, if the lock is granted. Once it does not need the resource anymore, the program can release the lock by performing the action V_a, following again the notation set up by Dijkstra [8]. If a subprogram tries to acquire a lock on a resource a when the resource has already been locked κ_a times, the subprogram is stuck until the resource is released by an other subprogram. In order to be realistic even though simple, the language considered here also comprises a sequential composition operator ., a non-deterministic choice operator $+$ and a loop construct $(-)^*$, with similar semantics as in regular languages (it should be thought as a `while` construct), as well as a parallel composition operator $|$ to launch two subprograms in parallel.

Programs p are defined by the following grammar:

$$p \quad ::= \quad \mathbf{1} \quad | \quad P_a \quad | \quad V_a \quad | \quad p.p \quad | \quad p|p \quad | \quad p+p \quad | \quad p^*$$

Programs are considered modulo a *structural congruence* $\equiv$ which imposes that operators ., $+$ and $|$ are associative and admit $\mathbf{1}$ as neutral element. A *thread* is a program which does not contain the parallel composition operator $|$.

1.2 Geometric Semantics

We introduce here a semantics based on (directed) topological spaces. The geometric semantics will allow a different representation of n pairwise independent actions (as the surface of an n-cube) and n truly concurrent actions as the full n-cube.

We denote by $I = [0,1] \subseteq \mathbb{R}$ the standard euclidean interval. A *path* p in a topological space X is a continuous map $p : I \to X$, and the points $p(0)$ and $p(1)$ are respectively called the *source* and *target* of the path. Given two paths p and q such that $p(1) = q(0)$, we define their *concatenation* as the path $p \cdot q$ defined by

$$(p \cdot q)(t) = \begin{cases} p(2t) & \text{if } 0 \leqslant t \leqslant 1/2 \\ q(2t-1) & \text{if } 1/2 \leqslant t \leqslant 1 \end{cases}$$

A topological space can be equipped with a notion of “direction” as follows [19]:

Definition 1. *A* directed topological space *(or* d-space *for short)* $X = (X, dX)$ *consists of a topological space* X *together with a set* dX *of paths in* X *(the* directed paths*) such that*

1. constant paths*: every constant path is directed,*
2. reparametrization*:* dX *is closed under precomposition with (non necessarily surjective) increasing maps* $I \to I$*, which are called* reparametrizations*,*
3. concatenation*:* dX *is closed under concatenation.*

A morphism of d-spaces $f : X \to Y$*, a* directed map*, is a continuous function* $f : X \to Y$ *which preserves directed paths, in the sense that* $f(dX) \subseteq dY$*.*

The category of d-spaces is complete and cocomplete [19]. This allows us to abstractly define some constructions on d-spaces, which extend usual constructions on topological spaces, that we detail here explicitly by describing the associated directed paths.

– The *terminal d-space* $\star$ is the space reduced to one point.
– The *cartesian product* $X \times Y$ of two d-spaces X and Y has $d(X \times Y) = dX \times dY$.
– The *disjoint union* $X \uplus Y$ of two d-spaces X and Y is such that $d(X \uplus Y) = dX \uplus dY$.

– The *amalgamation* $X[x = y]$ of two points x and y in a d-space X is the d-space X where x and y have been identified, together with the expected set of directed paths.
– Given a d-space X and a topological space $Y \subseteq X$, the *subspace* Y can be canonically equipped with a structure of d-space by $dY = \{p \in dX \mid p(I) \subseteq Y\}$.

The geometric semantics of a program is defined using those constructions as follows:

Definition 2. *To every program p, we associate a d-space G_p together with a pair of points $b_p, e_p \in G_p$, respectively called* beginning *and* end*, and a* resource function *$r_p : \mathcal{R} \times G_p \to \mathbb{Z}$ which indicates the number of locks the program holds at a given point. The definition of these is done by induction on the structure of p as follows:*

$G_{\mathbf{1}} = \star, \quad b_{\mathbf{1}} = *, \quad e_{\mathbf{1}} = *, \quad r_{\mathbf{1}}(a, x) = 0$	
$G_{P_a} = \vec{I}, \quad b_{P_a} = 0, \quad e_{V_a} = 1,$ $r_{P_a}(b, x) = \begin{cases} -1 & \text{if } b = a \text{ and } x > 0 \\ 0 & \text{if } b \neq a \text{ or } x = 0 \end{cases}$	$G_{V_a} = \vec{I}, \quad b_{V_a} = 0, \quad e_{V_a} = 1,$ $r_{V_a}(b, x) = \begin{cases} 1 & \text{if } b = a \text{ and } x = 1 \\ 0 & \text{if } b \neq a \text{ or } x < 1 \end{cases}$
$G_{p.q} = (G_p \uplus \vec{I} \uplus G_q)[e_p = 0, 1 = b_q],$ $b_{p.q} = b_p, \quad e_{p.q} = e_q,$ $r_{p.q}(a, x) = \begin{cases} r_p(a, x) & \text{if } x \in G_p \\ r_p(a, e_p) + r_q(a, x) & \text{if } x \in G_q \end{cases}$	$G_{p+q} = (G_p \uplus G_q)[b_p = b_q, e_p = e_q],$ $b_{p+q} = b_p, \quad e_{p+q} = e_q,$ $r_{p+q}(a, x) = \begin{cases} r_p(a, x) & \text{if } x \in G_p \\ r_q(a, x) & \text{if } x \in G_q \end{cases}$
$G_{p\|q} = G_p \times G_q,$ $b_{p\|q} = (b_p, b_q), \quad e_{p\|q} = (e_p, e_q),$ $r_{p\|q}(a, (x, y)) = r_p(a, x) + r_q(a, y)$	$G_{p^*} = G_p[b_p = e_p],$ $b_{p^*} = b_p, \quad e_{p^*} = b_p,$ $r_{p^*}(a, x) = r_p(a, x)$

Given a program p, the forbidden region *is the d-space $F_p \subseteq G_p$ defined by*

$$F_p = \{x \in G_p \mid \exists a \in \mathcal{R},\ \kappa_a + r_p(a, x) < 0 \ \text{ or } \ r_p(a, x) > 0\}$$

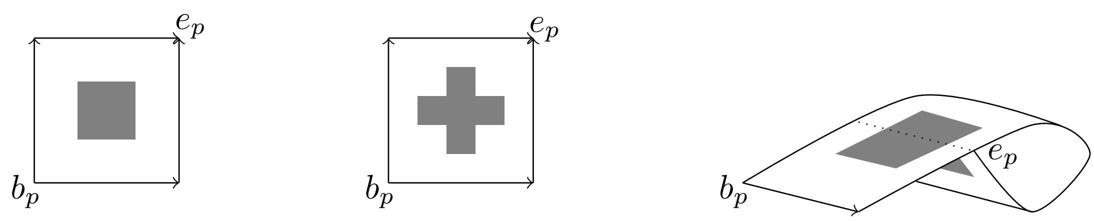
The geometric realization *of a process p, is defined as the d-space $H_p = G_p \setminus F_p$.*

We sometimes write 0 and ∞ for the beginning and the end points respectively of a geometric realization, and say that a path $p : \vec{I} \to G_p$ is *total* when it has 0 as source and ∞ as target. It is easy to show that the geometric semantics of a program is well-defined in the sense that two structurally congruent programs give rise to isomorphic geometric realizations.

Example 1. The processes

$$P_a.V_a | P_a.V_a \qquad\qquad P_a.P_b.V_b.V_a | P_b.P_a.V_a.V_b \qquad\qquad P_a.(V_a.P_a)^* | P_a.V_a$$

respectively have the following geometric realizations, which all consist of a space with some "holes", drawn in gray, induced by the forbidden region:



The space in the middle is sometimes called the "Swiss flag" because of its form and is interesting because it exhibits both a deadlock and an unreachable region [13].

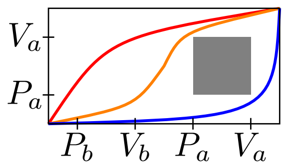
2 Computing Trace Spaces

2.1 Trace Spaces

In topology, two paths p and q are often considered as equivalent when q can be obtained by deforming continuously p (or vice versa), this equivalence relation being called *homotopy*. The corresponding variant of this relation in the case of directed topological spaces is called *dihomotopy* and is formally defined as follows. In the category of d-spaces, the object $\vec{I}$ is *exponentiable*, which means that for every d-space Y, one can associate a d-space $Y^{\vec{I}}$ such that there is a natural bijection between morphisms $X \times \vec{I} \to Y$ and morphisms $X \to Y^{\vec{I}}$. The underlying space of $Y^{\vec{I}}$ is the set of functions $\vec{I} \to Y$ with the compact-open topology (also called uniform convergence topology), and the directed paths $h : \vec{I} \to Y^{\vec{I}}$ are the functions such that $t \mapsto h(t)(u)$ is increasing for every $u \in \vec{I}$. Finally, two paths are said to be dihomotopic when one can be continuously deformed into the other:

Definition 3. *The* dihomotopy *is defined as the smallest equivalence relation on paths such that two directed paths $p, q : \vec{I} \to X$ are dihomotopic when there exists a directed path $h : \vec{I} \to X^{\vec{I}}$ with p as source and q as target.*

Example 2. In the geometric semantics of the program $P_b.V_b.P_a.V_a \mid P_a.V_a$, the two paths above the hole are dihomotopic, whereas the path below is not dihomotopic to the two others:



The intuition underlying the geometric semantics is that two dihomotopic paths correspond to execution traces differing by inessential commutations of instructions, thus giving rise to the same result.

Given two points x and y of a d-space X, we write $X(x, y)$ for the subset of $X^{\vec{I}}$ consisting of dipaths from x to y. A *trace* is the equivalence class of a path modulo surjective reparametrization, and a *scheduling* is the equivalence class of a trace modulo dihomotopy. We write $\vec{T}(X)(x, y)$ for the *trace space* obtained from $X(x, y)$ by identifying paths equivalent up to reparametrization, and simply $\vec{T}(X)$ for $\vec{T}(X)(0, \infty)$. In particular, we have $\vec{T}(X)(x, y) \neq \emptyset$ if and only if there exists a directed path in X going from x to y.

In this section, we reformulate the algorithm for computing the trace space $\vec{T}(X)$ up to dihomotopy equivalence, originally introduced in [24], in order to achieve an efficient implementation of it. For simplicity, we restrict here to spaces which are geometric realizations of programs of the form

$$p \quad = \quad p_0 \mid p_1 \mid \ldots \mid p_{n-1} \tag{1}$$

where the p_i are built up only from $\mathbf{1}$, concatenation, resource locking and resource unlocking (extending the algorithm to programs which may contain loops requires significant generalizations which are described in Section 3). In this case, the geometric realization is of the form

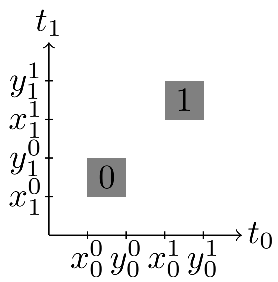
$$G_p \quad = \quad \vec{I}^n \setminus \bigcup_{i=0}^{l-1} R^i$$

where $\vec{I}^n$ denotes the cartesian product of n copies of $\vec{I}$, and each $R^i = \prod_{j=0}^{n-1} \vec{I}_j^i$ is a rectangle. We suppose here that each R^i is homothetic to the n-dimensional open rectangle, i.e. each directed interval $\vec{I}_j^i$ is of the form $\vec{I}_j^i =]x_j^i, y_j^i[$, and generalize this at the end of the section. The restrictions on the form of the programs are introduced here only to simplify our exposition: programs with choice can be handled by computing the trace spaces on each branch and program with loops can be handled by suitably unfolding the loops so that all the possible behaviors are exhibited (a detailed presentation of this is given in Section 3, which will enable to handle the full language). We suppose fixed a program with n threads and l forbidden open rectangles, and consistently use the notations above.

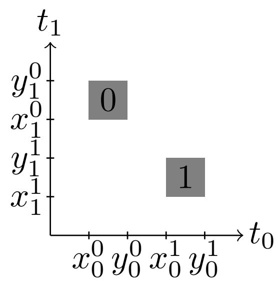
Example 3. The geometric realization of the programs

$$P_a.V_a.P_b.V_b | P_a.V_a.P_b.V_b \qquad \text{and} \qquad P_a.V_a.P_b.V_b | P_b.V_b.P_a.V_a$$

are respectively



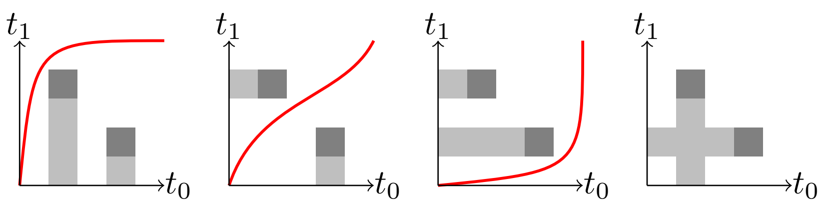
and



2.2 The Index Poset

Let us come back to the second program of Example 3. We will determine the different traces, and their relationships in the trace space, by combinatorially looking at the way they can turn around holes. To see this in that example, we extend each hole in parallel to the axes, below or leftwards from the holes, until they reach the boundary of the state space. These new obstructions impose traces to go the other way around each hole: the existence of deadlocks, given these new constraints in the trace space allows us to determine whether traces going one way or the other around each hole exist. In fact, this combinatorial information precisely computes all of the trace space [24].

In the second program of Example 3, there are four possibilities to extend once each of the two holes:



(2)

Notice that there exists a total path in the first three spaces (as depicted above), whereas there is none in the last one.

A simple way to encode the combinatorial information about the extension of holes is through boolean matrices. We write $\mathcal{M}_{l,n}$ for the poset of $l \times n$ matrices, with l rows (the number of holes R^i) and n columns (the dimension of the space, i.e. the number of threads in the program), with coefficients in $\mathbb{Z}/2\mathbb{Z}$, with the pointwise ordering such that $0 \leqslant 1$: we have $M \leqslant N$ whenever

$$\forall(i,j) \in [0:l[\times[0:n[, \qquad M(i,j) \leqslant N(i,j) \tag{3}$$

where $[m:n[$ denotes the set $\{m,\ldots,n-1\}$ of integers and $M(i,j)$ denotes the (i,j)-th coefficient of M. We also write $\mathcal{M}^R_{l,n}$ for the subposet of $\mathcal{M}_{l,n}$ consisting of matrices whose row vectors are all different from the zero vector, and $\mathcal{M}^C_{l,n}$ for the subposet of $\mathcal{M}_{l,n}$ consisting of matrices whose column vectors are all unit vectors (containing exactly one coefficient 1).

Given a matrix $M \in \mathcal{M}_{l,n}$, we define X_M as the subspace of X obtained by extending downwards each forbidden rectangle R^i in every direction j' different from j for every j such that $M(i,j) = 1$. Formally,

$$X_M \quad = \quad \vec{I}^n \setminus \bigcup_{M(i,j)=1} \tilde{R}^i_j$$

where $\tilde{R}^i_j = \prod_{j'=0}^{j-1}[0, y^i_{j'}[\times]x^i_j, y^i_j[\times \prod_{j'=j+1}^{n-1}[0, y^i_{j'}[$, see (2) and Example 4 below.

In order to study whether there is a total path in the space associated to a matrix, we define a map $\Psi : \mathcal{M}_{l,n} \to \mathbb{Z}/2\mathbb{Z}$ by $\Psi(M) = 1$ iff $\vec{T}(X_M) = \emptyset$, i.e. there is no total path in X_M. A matrix M is *dead* when $\Psi(M) = 1$ and *alive* otherwise. The map Ψ can easily be shown to be order preserving.

Definition 4. *We write*

$$\mathcal{D}(X) \quad = \quad \{M \in \mathcal{M}^C_{l,n} \;/\; \Psi(M) = 1\}$$

for the set of (column) dead matrices and

$$\mathcal{C}(X) \quad = \quad \{M \in \mathcal{M}^R_{l,n} \;/\; \Psi(M) = 0\}$$

for the set of alive matrices (with non-empty rows), which is called the index poset *– it is implicitly ordered by the relation* (3).

Example 4. In the example above, the three extensions of holes (2) are respectively encoded by the following matrices:

$$\begin{pmatrix}1&0\\1&0\end{pmatrix} \qquad \begin{pmatrix}0&1\\1&0\end{pmatrix} \qquad \begin{pmatrix}0&1\\0&1\end{pmatrix} \qquad \begin{pmatrix}1&0\\0&1\end{pmatrix}$$

The last matrix is dead and the three others are alive. The last matrix being dead indicates that there is no way a trace can pass left of the upper left hole and carry on passing below the lower right hole.

A reason why the matrices in the index poset are convenient objects to study the schedulings is that they are topologically very simple [24]:

Proposition 1. *For any matrix* $M \in \mathcal{M}^R_{l,n}$*, the space* $X_M(x, y)$ *is either empty or contractible: any two paths with the same source* x *and target* y *are dihomotopic. In particular, for any matrix* $M \in \mathcal{C}(X)$*, the space* $X_M(0, \infty)$ *is always contractible.*

Our main interest in the index poset is that it enables us to compute the schedulings (i.e. maximal paths modulo dihomotopy) of the space: these schedulings are in bijection with alive matrices in $\mathcal{C}(X)$ modulo an equivalence relation called *connexity*, which is defined as follows. Given two matrices $M, N \in \mathcal{M}_{l,n}$, their *intersection* $M \wedge N$ is defined as the matrix $M \wedge N$ such that $(M \wedge N)(i, j) = \min(M(i, j), N(i, j))$.

Definition 5. *Two matrices* M *and* N *are* connected *when their intersection does not contain any row filled with* 0.

The dihomotopy classes of total paths in X can finally be computed thanks to the following property:

Proposition 2. *The connected components of* $\mathcal{C}(X)$ *are in bijection with schedulings in* X.

Example 5. Consider the program $p = q|q|q$ where $q = P_a.V_a$. The associated trace space X_p is a cube minus a cube (as shown in Example 8). The matrices in $\mathcal{C}(X_p)$ are

$$\begin{pmatrix}1 & 0 & 0\end{pmatrix} \qquad \begin{pmatrix}0 & 1 & 0\end{pmatrix} \qquad \begin{pmatrix}0 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix}0 & 1 & 1\end{pmatrix} \qquad \begin{pmatrix}1 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix}1 & 1 & 0\end{pmatrix}$$

and they are all (transitively) connected. For instance, $\begin{pmatrix}0 & 1 & 1\end{pmatrix} \wedge \begin{pmatrix}1 & 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0 & 1\end{pmatrix}$. The program p thus has exactly one total scheduling, as expected.

Intuitively, alive matrices describe sets of dihomotopic total paths (Proposition 1) and the fact that two matrices have non-zero rows in their intersection means that there are paths which satisfy the constraints imposed by both matrices, i.e. the two matrices describe the same dihomotopy class of total paths.

2.3 Computing Dihomotopy Classes

The computation of the dihomotopy classes of total paths in the geometric semantics X of a given program will be performed in three steps:

1. we compute the set $\mathcal{D}(X)$ of dead matrices,
2. we use $\mathcal{D}(X)$ to compute the index poset $\mathcal{C}(X)$,
3. we deduce the homotopy classes of total paths by quotienting $\mathcal{C}(X)$ by the connexity relation.

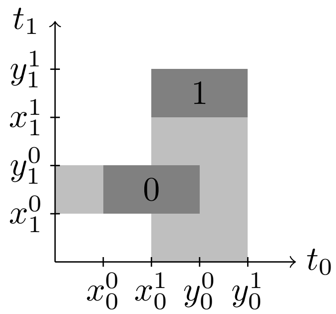
These steps are detailed below.

Given a subset I of $[0 : l[$ and an index $j \in [0 : n[$, we write $y^I_j = \min\{y^i_j \mathbin{/} i \in I\}$ (by convention $y^{\emptyset}_j = \infty$). Given a matrix $M \in \mathcal{M}_{l,n}$, we define the set of *non-zero rows* of M by $R(M) = \{i \in [0 : l[\mathbin{/} \exists j \in [0 : n[,\ M(i, j) \neq 0\}$. It can be shown that a matrix M is dead if and only if the space X_M contains a deadlock. From the characterization of deadlocks in geometric semantics given in [11], the following characterization of dead matrices can therefore be deduced:

Proposition 3. *A matrix $M \in \mathcal{M}^C_{l,n}$ is in $\mathcal{D}(X)$ iff it satisfies*

$$\forall (i,j) \in [0:l[\times[0:n[, \quad M(i,j) = 1 \Rightarrow x^i_j < y^{R(M)}_j \tag{4}$$

Example 6. In the example below with $l = 2$ and $n = 2$, the matrix $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is dead (we suppose that $x^i_j = 1 + i(j+1)$ and $y^i_j = 3 + i(j+1) - j$):



$$x^0_1 = 1 < 2 = y^{\{0,1\}}_1$$
$$x^1_0 = 2 < 3 = y^{\{0,1\}}_0$$

The above proposition enables us to compute the set of dead matrices, for instance by enumerating all matrices and checking whether they satisfy condition 4 (a more efficient method is described in Section 2.4). From this set, the index poset $\mathcal{C}(X)$ can be determined using the following property:

Proposition 4. *A matrix $M \in \mathcal{M}_{l,n}$ is not in $\mathcal{C}(X)$ iff there exists a matrix $N \in \mathcal{D}(X)$ such that $N \leqslant M$. In other words, $M \in \mathcal{C}(X)$ iff for every matrix $N \in \mathcal{D}(X)$ there exists indexes $i \in [0:l[$ and $j \in [0:n[$ such that $M(i,j) = 0$ and $N(i,j) = 1$.*

Notice that the poset $\mathcal{C}(X)$ is downward closed (because Ψ is order preserving) and one is naturally interested in the subset $\mathcal{C}_{\max}(X)$ of *maximal* matrices in order to describe it. Proposition 4 provides a simple-minded algorithm for computing (maximal) matrices in $\mathcal{C}(X)$. We write $\mathcal{D}(X) = \{D_0, \ldots, D_{p-1}\}$. We then compute the sets C_k of maximal matrices M such that for every $i \in [0:k[$ we have $D_i \not\leqslant M$. We start from the set $C_0 = \{\mathbf{1}\}$ where $\mathbf{1}$ is the matrix containing only 1 as coefficients. Given a matrix M, we write $M^{\neg(i,j)}$ for the matrix obtained from M by replacing the (i,j)-th coefficient by $1 - M(i,j)$. The set C_{k+1} is then computed from C_k by doing the following for all matrices $M \in C_k$ such that $D_k \leqslant M$:

1. remove M from C_k,
2. for every (i,j) such that $D_k(i,j) = 1$,
   – remove every matrix $N \in C_k$ such that $N \leqslant M^{\neg(i,j)}$,
   – if there exists no matrix $N \in C_k$ such that $M^{\neg(i,j)} \leqslant N$, add $M^{\neg(i,j)}$ to C_k.

The set $\mathcal{C}_{\max}(X)$ is obtained as C_p. If we remove the second point and replace it by

2'. for every (i,j) such that $D_k(i,j) = 1$ and $M^{\neg(i,j)} \in \mathcal{M}^R_{l,n}$, add $M^{\neg(i,j)}$ to C_k.

we compute a set C_p such that $\mathcal{C}_{\max}(X) \subseteq C_p \subseteq \mathcal{C}(X)$, which is enough to compute connected components and has proved faster to compute in practice.

Example 7. Consider again Example 3. The algorithm starts with

$$C_0 \quad = \quad \left\{ M_0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$$

For C_1, we must have $D_0 \nleqslant M_0$ so we swap any of the two ones in the first row:

$$C_1 \quad = \quad \left\{ M_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\}$$

Similarly for C_2, we have to swap the bits on the second row so that $D_1 \nleqslant M_i$:

$$C_2 = \left\{ M_3 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, M_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, M_5 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M_6 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

Finally, we have $D_2 \nleqslant M_i$, excepting $D_2 \leqslant M_5$, so we swap the bits in position $(1, 1)$ and in position $(2, 2)$:

$$M_5' = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \leqslant M_3 \qquad\qquad M_5'' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \leqslant M_6$$

Since we are only interested in maximal matrices, we end up with $C_3 = \{M_6, M_4, M_3\}$. The trace spaces corresponding to those matrices are the three first depicted in (2). None of those matrices being connected, the trace space up to dihomotopy consists of exactly 3 distinct points.

Other implementations of the algorithm can be obtained by reformulating the computation of $\mathcal{C}_{\max}(X)$ as finding a minimal transversal in a hypergraph, for which efficient algorithms have been proposed [21].

We have supposed up to now that the forbidden region was a union of rectangles R^i, each such rectangle being a product of open intervals $\vec{I}^i_j =]x^i_j, y^i_j[$. The algorithm given above can easily be generalized to the case where the rectangles R^i can "touch the boundary" in some dimensions, i.e. the intervals $\vec{I}^i_j$ are either of the form $]x^i_j, y^i_j[$ or $[0, y^i_j[$ or $]x^i_j, \infty]$ or $[0, \infty]$. For example, the process $P_a.V_a|P_a.V_a|P_a.V_a$, with $\kappa_a = 1$, generates such a forbidden region. We write $B \in \mathcal{M}_{l,n}$ for the *boundary matrix*, which is the matrix such that $B(i, j) = 0$ whenever $x^i_j = 0$ (i.e. the i-th interval touches the lowest boundary in dimension j) and $B(i, j) = 1$ otherwise. The matrices of $\mathcal{D}(X)$ are the matrices $M \in \mathcal{M}_{n,l}$ of the form $M = N \wedge B$, for some matrix $N \in M^C_{n,l}$, which satisfy (4) and such that

$$\forall j \in C(M), \qquad y_j^{R(M)} = \infty \tag{5}$$

where $C(M)$ is the set of indexes of null columns of M.

2.4 An Efficient Implementation

In order to compute the set $\mathcal{D}(X)$ of dead matrices, the general idea is to enumerate all the matrices $M \in \mathcal{M}^C_{l,n}$ and check whether they satisfy the condition (4). Of course, a direct implementation of this idea would be highly inefficient since there are l^n matrices in $\mathcal{M}^C_{l,n}$. In order to improve this, we try to detect "as soon as possible" when a matrix

```
let rec compute_dead j m rows yrows =
  if j = n then dead := m :: !dead else
    for i = 0 to l − 1 do
      try
        let changed_rows = not (Set.mem i rows) in
        let rows = Set.add i rows in
        let m = Array.copy m in
        if bounds(i,j) = 1 then m.(j) ←None else m.(j) ←Some i;
        (match m.(j) with
            | Some i →if x_j^i ⩾ yrows.(j) then raise Exit
            | None → if yrows.(j) ≠∞ then raise Exit);
        let yrows =
          let j′ = j in
          if not changed_rows then yrows else
            Array.mapi (fun j yrj →
                if yrj ⩽ y_j^i then yrj else
                  match m.(j) with
                    | None →
                        if j ⩽ j′ && y_j^i≠∞ then raise Exit; y_j^i
                    | Some i →
                        if x_j^i ⩾ y_j^i then raise Exit; y_j^i
                        ) yrows
        in
        compute_dead (j+1) m rows yrows
      with Exit → ()
    done
```

Fig. 1. Algorithm for computing dead matrices

does not satisfy the condition: we first fix the coefficient in the first column of M and check whether it is possible for a matrix with this first column to be dead, then we fix the second column and so on. In fact, we have to check that every coefficient (i, j) such that $M(i, j) = 1$ satisfies $x_j^i < y_j^{R(M)}$. Now, suppose that we know some of the coefficients (i, j) for which $M(i, j) = 1$. We therefore know a subset $I \subseteq R(M)$ of the non-zero rows. If for one of these coefficients we have $x_j^i \geqslant y_j^I$, we know that the matrix cannot satisfy the condition (4) because $x_j^i \geqslant y_j^I \geqslant y_j^{R(M)}$. A similar reasoning can be held for condition (5).

The actual function computing the dead matrices is presented in Figure 1, in pseudo-OCaml code. This recursive function fills j-th column of the matrix M (whose columns with index below j are supposed to be already fixed) and performs the check: it tries to set the i-th coefficient to 1 (and all the others to 0) for every $i \in [0 : l[$. If a matrix beginning as M (up to the j-th column) cannot be dead, the computation is aborted by raising the Exit exception. When all the columns have been computed the matrix is added to the list *dead* of dead matrices. Since a matrix $M \in \mathcal{M}_{l,n}^C$ has at most one non-null coefficient in a given column, it will be coded as an array of length n whose j-th element is either None when all the elements of the j-th column are null, or Some i

when the i-th coefficient of the j-th column is 1 and the others are 0. The argument $rows$ is the set of indexes of known non-null rows of M and $yrows$ is an array of length n such that $yrows.(j)= y_j^{rows}$. The matrix $bounds$ is the matrix previously noted B used to perform the check (5). Notice that the algorithm takes advantage of the fact that when the coefficient i chosen for the j-th column is already in $rows$ (i.e. when the variable $changed\_rows$ is false) then many computations can be spared because the coefficients y_j^{rows} are not changed.

Once the set of dead matrices computed, the set $\mathcal{C}(X)$ of alive matrices is then computed using the naive algorithm of Section 2.3, exemplified in Example 7. We have also implemented a simple hypergraph transversal algorithm [2] but it did not bring significant improvements, more elaborate algorithms might give better results though. Finally, the representatives of traces are computed as the connected components (in the sense of Proposition 2) of $\mathcal{C}(X)$, in a straightforward way. An explicit sequence of instructions corresponding to every representative M can easily be computed: it corresponds to the sequence of instructions crossed by any increasing total path in the d-space X_M.

2.5 An Example: The n Dining Philosophers

In order to illustrate the performances of our algorithm, we present below the computation times for the well-known n dining philosophers program [9] whose schedulings are in $O(2^n)$, hence is pushing any algorithm that would determine the essential schedules to its (exponential) limits. It is constituted of n processes p_k in parallel, using n mutexes a_i, defined by $p_k = P_{a_k}.P_{a_{k+1}}.V_{a_k}.V_{a_{k+1}}$, where the indexes on mutexes a_i are taken modulo n. Such a program generates $2^n - 2$ distinct schedulings, which our program finds correctly. The table below summarizes the execution time and memory consumption for our tool ALCOOL (programmed in OCaml), as well as for the model checker SPIN [1] implementing partial order reduction techniques. Whereas SPIN is not significantly slower, it consumes much more memory and starts to use swap from $n = 12$ (thus failing to give an answer in a reasonable time for $n > 12$). Notice that the implementation of SPIN is finely tuned and also benefits from gcc optimizations, whereas there is room for many improvements in ALCOOL. In particular, most of the time is spent in computing dead matrices and the algorithm of Section 2.4 could be improved by finding a heuristic to suitably sort holes so that failures to satisfy condition (4) are detected earlier. The present algorithm is also significantly faster than some of the author's previous contribution [16]: for instance, it was unable to generate these maximal dipaths because of memory requirements, for n philosophers with $n > 8$ (in the benchmarks of [16], it was taking already 13739s, on a 1GHz laptop computer though, to generate just the component category for 9 philosophers).

n	sched.	ALC. (s)	ALC. (MB)	SP. (s)	SP. (MB)
10	1022	5	4	8	179
11	2046	32	9	42	816
12	4094	227	26	313	3508
13	8190	1681	58	∞	∞
14	16382	13105	143	∞	∞

Since the size of the output is generally exponential in the size of the input, there is no hope to find an algorithm which has less than an exponential worst-case complexity (which our algorithm clearly has). However, since our goal is to program actual tools to very concurrent programs, practical improvements in the execution time or memory consumption are really interesting from this point of view. We have of course tried our tool on many more examples, which confirm the improvement trend, and shall be presented in a longer version of the article.

3 Programs with Loops

3.1 Paths in Deloopings

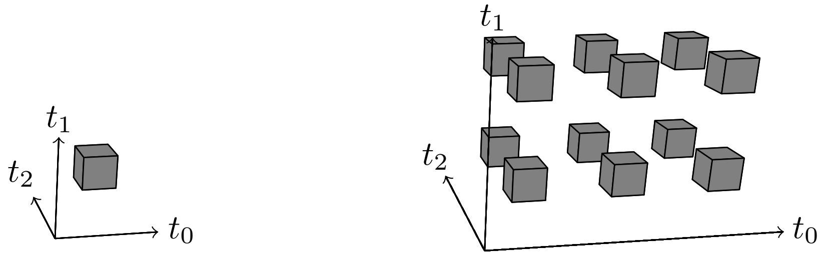
One of the most challenging part of verifying concurrent programs consists in verifying programs with loops since those contain a priori an infinite number of possible execution traces. We extend here the previous methodology and, given a program containing loops, we compute a (finite!) automaton whose accepted paths describe the schedulings of the program: this automaton, can thus be considered as a control flow graph of the concurrent program. Of course, we are then able to use the traditional methods in static analysis, such as abstract interpretation, to study the program (this is briefly presented in Section 3.5). This section builds on some ideas being currently developed by Fajstrup [10], however most of the properties presented in this section are entirely new. To the best of our knowledge, this is the first works in which geometric methods are used in order devise a practical algorithm to handle programs containing loops. A particularly interesting feature of our method lies in the fact that it consider the broad "geometry of holes" and can thus associate a small control flow graph to a given program, see Section 3.4.

In the following, we suppose fixed a program of the form $p = p_0|p_1|\ldots|p_{n-1}$ as in (1), with n threads. We write

$$p^* \quad = \quad p_0^* \mid p_1^* \mid \ldots \mid p_{n-1}^*$$

for the associated "looping program". Our goal in this section is to describe the schedulings of such a program p^* (the restriction on the form of the programs considered here was only done to simplify our presentation and the methodology can be extended to handle all well-bracketed programs generated by the grammar, without any essential technical difficulty added). Following Section 1.2, its geometrical semantics consists of an n-dimensional torus with rectangular holes. As previously, for simplicity, we suppose that these holes do not intersect the boundaries, i.e. that p satisfies the hypothesis of Section 2.1. Given an n-dimensional vector $v = (v_0, \ldots, v_{n-1})$ with coefficients in $\mathbb{N}$, the *v-delooping* of p, written p^v, is the program $p_0^{v_0}|p_1^{v_1}|\ldots|p_{n-1}^{v_{n-1}}$, where $p_j^{v_j}$ denotes the concatenation of v_j copies of p_j. A *scheduling* in p is a scheduling in the previous sense (i.e. a total path modulo homotopy) in p^v for some vector v.

Example 8. Consider the program $p = q|q|q$ of Example 5, where $q = P_a.V_a$. Its geometric realization X_p is pictured on the left, and its $(3,2,2)$-delooping $X_{p^{(3,2,2)}}$ is pictured on the right.



Given two spaces X and Y which are hypercubes with holes (which is the case for the geometric realizations of the programs we are considering here), we write $X \oplus_j Y$ for the space obtained by identifying the j-th target face of the hypercube X with the j-th source face of the hypercube Y, and call it the *j-gluing* of X and Y. Formally, this can be defined as in Section 1.2 as $X \oplus_j Y = X \uplus Y/\sim$, where the relation $\sim$ identifies points $x \in X$ and $y \in Y$ such that $x_j = \infty$, $y_j = 0$ and $x_{j'} = y_{j'}$ for every dimension $j' \neq j$, and directed paths are defined in a similar fashion. Notice that, by definition, there is a canonical embedding of X (resp. Y) into $X \oplus_j Y$, which will allow us to implicitly consider X (resp. Y) as a subspace of $X \oplus_j Y$ in the following.

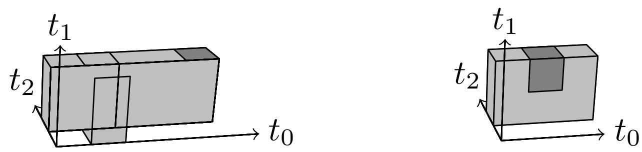
Example 9. The $(3, 2, 2)$-delooping of Example 8 is

$$X_{p^{(3,2,2)}} \quad = \quad (Y \oplus_1 Y) \oplus_2 (Y \oplus_1 Y) \qquad \text{with} \qquad Y = X_p \oplus_0 X_p \oplus_0 X_p$$

More generally, any v-delooping p^v of a program p of the form (1) can be obtained by gluing copies X_p^w of X_p, indexed by a vector w such that for every dimension i with $0 \leqslant i < n$, we have $0 \leqslant w_i < v_i$ (what we will simply write $0 \leqslant w < v$).

Given two scheduling matrices M and N encoding extensions of holes of such a program p (cf. Section 2.2), we reuse the notation and write $M \oplus_j N$ for the obvious matrix coding extension of holes in the space $X_p \oplus_j X_p$. At this point, it is crucial to notice that the holes described by N in the second copy of X_p can have an effect on the first copy of X_p (when they are extended to 0 in the direction j), what we call the *j-shadow of N*, and write $X_{N|_j}$.

Example 10. With the program p of Example 8, consider the matrices $M = (1\ 0\ 0)$ and $N = (0\ 0\ 1)$. We have $M \oplus_0 N = \begin{pmatrix} 1\ 0\ 0 \\ 0\ 0\ 1 \end{pmatrix}$, the space $X_{M \oplus_0 N}$ is pictured on the left, and the 0-shadow $X_{N|_0}$ of N is pictured on the right:



The above example makes clear that the space corresponding to a scheduling $M \oplus_j N$ is of the form $X_{M \oplus_j N} = (X_M \cap X_{N|_j}) \otimes_j X_N$, i.e. the holes in the first copy come either from M or from shadows of N. Moreover, the holes in the space $X_{N|_j}$ are hypercubes which are products of intervals of the form $\prod_{0 \leqslant j < n} \vec{I_j}$, where each interval $\vec{I_j}$ is of the form $]x_j^i, y_j^i[$ or $[0, y_j^i[$ or $[0, \infty]$, with $0 \leqslant i < l$. The shadows can therefore be coded as matrices (using a slightly different coding from the one used up to now, the precise way they are coded being quite irrelevant) and we write $N|_j$ for the matrix coding the j-shadow of n, which can easily be computed from N and j. A scheduling matrix M can obviously be seen as a particular "shadow", enabling us to use the same notation for both, and we write $M \cup N$ for the union of two shadows M and N, so that $X_{M \cup N} = X_M \cap X_N$. Finally, given a shadow M, the algorithm described in Section 2.3 can easily be adapted to the new coding in order to determine whether the space X_M is alive.

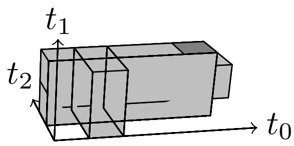
3.2 The Shadow Automaton

The trace space of a program p^* is not finite in the general case. We show here that it can however be described as the set of paths of an automaton that we call the *shadow automaton*: this automaton provides us with a *finite presentation* of the set of schedulings.

Consider the v-delooping p^v of a program p. The space X_{p^v} consists of the gluing of copies of X_p indexed by vectors w such that $0 \leqslant w < v$ and similarly, a scheduling M of X_{p^v} consists of the gluing of matrices M^w. Clearly, if some submatrix M^w is dead then the whole matrix M is dead:

Lemma 1. *If a matrix M is alive then all its submatrices M^w are alive.*

However, the converse is not true because a scheduling M^w might create a deadlock with the shadows coming from matrices above it. For instance in Example 8, the matrix $M = (1\ 0\ 0) \oplus_0 (0\ 1\ 1)$ is not alive because the space $X_{M^{(0,0,0)}}$ induced by the submatrix $M^{(0,0,0)}$ is contained in the space X_N, where $N = (1\ 1\ 1)$ is a dead matrix:



In order to generate all the possible schedulings M^w visited by a total path in X_{p^v}, we therefore have to take in account the shadows dropped by scheduling of copies of X_p in its future. We will construct an automaton which will consider the visited schedulings of the path, starting from the end, and maintains the shadow they produce on the next state in a given direction j, so that we can compute the possible previous matrices in direction j such that the whole matrix is not dead. Formally,

Definition 6. *The* shadow automaton *of a program p is a non-deterministic automaton whose*

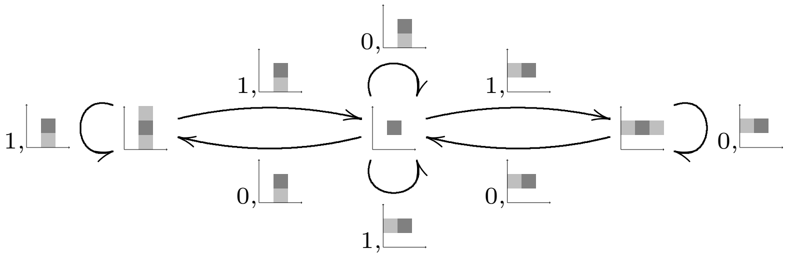
- *states are shadows*
- *transitions $N \xrightarrow{j,M} N'$ are labeled by a direction j (with $0 \leqslant j < n$) and a scheduling M*

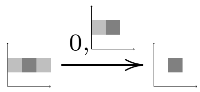
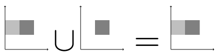
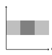
defined as the smallest automaton

- *containing the empty scheduling* $\emptyset$
- *and such that for every state* N'*, for every direction* j *and for every scheduling* M *such that the scheduling* $M \cup N'$ *is alive, and* M *is maximal with this property, there is a transition* $N \xrightarrow{j,M} N'$ *with* $N = (M \cup N')|_j$.

All the states of the automaton are both initial and final.

Example 11. Consider the program $p = q|q$ with $q = P_a.V_a$ whose geometric semantics is a square with a square hole. The associated shadow automaton is



For instance the transition  is computed as follows: we take the shadow $M =$  and compute its shadow in direction 0, i.e. on the left, to compute the source of the transition. This shadow is , namely: .

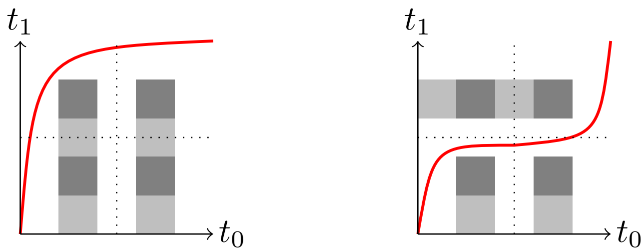
The interest of the automaton lies in the fact that fully describes the possible schedulings crossed by a total path in a scheduling of a delooping X_{p^v}:

Theorem 1. *Suppose that* M *is a scheduling of* X_{p^v}*, obtained by gluing schedulings* M^w *of* X_p*. Then there exists a total path in* X_M *going through the subspaces* $X_{M^{w_0}}, X_{M^{w_1}}, \ldots, X_{M^{w_m}}$ *in this order, such that* w_k *and* w_{k+1} *only differ by one coordinate* j_k *(i.e. the path exits from* $X_{M^{w_k}}$ *through its* j_k*-th face), if and only if there exists a path labeled as follows in the shadow automaton:*

$$N_0 \xrightarrow{j,M^{w_0}} N_1 \xrightarrow{j_0,M^{w_1}} N_2 \qquad \cdots \qquad N_m \xrightarrow{j_{m-1},M^{w_m}} N_{m+1}$$

for some states N_i *and dimension* j.

Example 12. With the program p of Example 11, the following paths in the $(2, 2)$-delooping



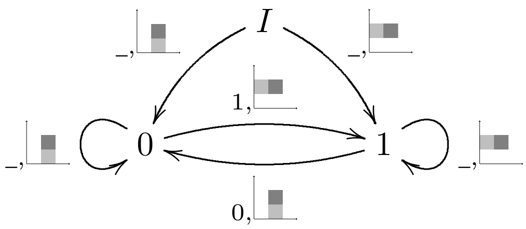
are respectively witnessed by the following paths of the shadow automaton:

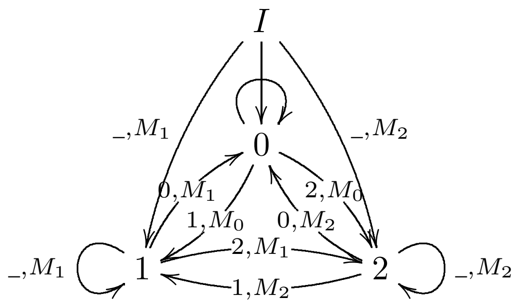
3.3 Reducing the Size of the Shadow Automaton

The size of the shadow automaton grows very quickly when the complexity of the trace space grows. For instance, for the program p of Example 8, the shadow automaton has already 19 states and 80 transitions. We describe here some ways to reduce the automaton while preserving Theorem 1. Namely, we should remark that the automaton is not minimal in the following sense. By Proposition 1, given a scheduling M two total paths X_M are necessarily homotopic: an alive scheduling thus describes an homotopy class of total paths. By Theorem 1, the schedulings "visited" by a total path in X_{p^v} are described by a path in the shadow automaton, therefore every homotopy class of total paths in X_{p^v} is described by at least one path in the scheduling automaton. The shadow automaton is not minimal in the sense that generally, an homotopy class is described by more than one path in the scheduling automaton.

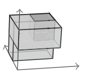
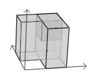
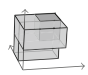
Determinization. First, our non-deterministic automaton can be determinized using classical algorithms of automata theory, which in practice greatly reduce their size: the determinized automaton for the program of Example 8 has only 4 states and 24 transitions.

Example 13. The determinized automata for Examples 11 and 8 are respectively:



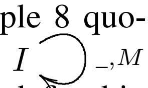


with

$M_0 =$  $M_1 =$  $M_2 =$ 

where "_" means any direction j. The state I is initial and all the states are final.

Quotient under connexity. A way to further reduce the automaton consists in quotienting the scheduling matrices labeling the arrows of the automaton under the connexity relation of Definition 5 before determinizing the automaton, which is formally justified by Proposition 2.

Example 14. The shadow automaton corresponding to the program Example 8 quotiented under connexity, determinized and minimized is simply the automaton 

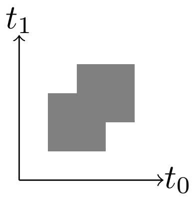
where $M = M_1 = M_2 = M_3$ up to connexity (the matrices M_i are those defined in Example 13).

We are currently investigating further conditions in order to construct the minimal automaton describing the trace space associated to a looping program, but the conditions mentioned above are already providing us with promisingly small automata.

3.4 Preliminary Implementation and Benchmark

A preliminary implementation of the computation of the shadow automaton was done. The algorithm implemented is currently quite simple, but we plan to generalize the algorithm of Section 2.4 soon, which is not complicated from a theoretical point of view but much more involved technically, in order to achieve better performances. Most experiments lead so far are already promising and make it clear that taking in account the geometry of the state-space enables us to reduce, sometimes drastically, the size of the control flow graph corresponding to the program to be analyzed.

Example 15. The *two-phase locking protocol* is a simple discipline for distributed databases, in which the processes first lock all the mutexes for the resources they are going to use and free all of them in the end [20]. This can be modeled as a program $q_{n,l}$ consisting of n copies of the process $p = P_{a_1} \ldots P_{a_l}.V_{a_1} \ldots V_{a_l}$ in parallel (each of these process is using l resources). For instance, the geometric semantics of $q_{2,2} = p|p$ is shown below. Notice that this state space is equivalent to a space with only one hole up to dihomotopy. More generally, given $l \geqslant 1$, it can be shown that the geometric semantics of $q_{n,l}$ is equivalent to $q_{n,1}$, which our algorithm is able to take into account! Namely, the size of the shadow automaton associated to $q^*_{n,l}$ only depends on n whereas the number of states of the automaton produced by SPIN is exponential in l (with n fixed). Below are presented the size (states, transitions) of the non-deterministic automaton (s, t), determinized automaton (s',t') and SPIN's automaton (s_{SPIN}, t_{SPIN}) for the two-phase locking process described in Example 15, for some values of n and l.



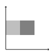
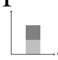
n	l	s	t	s'	t'	s_{SPIN}	t_{SPIN}
2	1	3	8	3	10	58	65
2	2	3	8	3	10	112	129
2	3	3	8	3	10	180	209
3	1	19	90	4	24	171	218
3	2	19	90	4	24	441	602
3	3	19	90	4	24	817	1128

3.5 An Application to Static Analysis

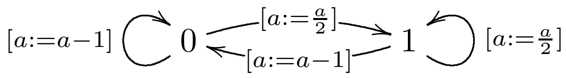
Now that we have the reduced shadow automaton, we can explain how one can perform static analysis by *abstract interpretation* [5] on concurrent systems, in an economic way. The systematic design and proof of correctness of such abstract analysis is left for a future article, the aim of this section is to give an intuition why the computations of Section 3 are relevant to static analysis by abstract interpretation. The idea is to associate, to each node n of the shadow automaton, a set of values A_n that program variables can take if computation follows a transition path whose last vertex is n. Among the actions the program can take along this scheduling, we consider only the *greedy* ones, that is the ones which execute all possible actions permitted by the dihomotopy class of schedulings ending by n.

Suppose that we want to analyze the program

$$p^* \quad = \quad \Big(P_a.\,(a := a-1)\,.V_a\Big)^* \Big| \Big(P_a.\Big(a := \frac{a}{2}\Big).V_a\Big)^* \tag{6}$$

What are the possible sets of values reached, for a, starting with $a \in [0,1]$? The associated shadow automaton S_p has been determined in Example 13 (this automaton is reduced) together with relations, that we will not be using in this article, yet. In many ways, this reduced shadow automaton plays the role of a compact *control flow graph* for the program we are analyzing. Calling $M_0 =$  and $M_1 =$ , X_{M_0} has the effect on environment: $a := a/2$ and X_{M_1} has as effect: $a := a - 1$.

We are now in a position to interpret the arrows of the shadow automaton as simple *abstract transfer functions* and produce a system of equations for which we want to determine a least-fixed point, to get the invariant of the program at the (multi-)control point which is the pair of the heads of the loops of each process. The interpretation on the shadow automaton now gives (ignoring the initial state I in that picture, for simplicity's sake) can be graphically pictured as:



Given the abstract transfer functions on each edge of the shadow automaton, we produce as customary the abstract semantic equations, one per node, by joining all transfer functions correspond to ingoing edges to that node:

$$\begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = F\begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = \begin{pmatrix} I \cup (A_0 - 1) \cup (A_1 - 1) \\ I \cup \frac{A_1}{2} \cup \frac{A_0}{2} \end{pmatrix} \tag{7}$$

This set of semantic equations can be seen as a least-fixed point equation, that we can solve using any of our favorite tool, for instance Kleene iteration and widening/narrowing, on any abstract domain, such as the domain of intervals as in the example below. The least-fixed point formulation that we are looking for is thus $A^\infty = \bigvee_{[0,1]} F$, where F is the function defined in (7) and $I = [0,1]$. A Kleene iteration on this monotonic function F on the lattice of intervals over $\mathbb{R}$ reveals that $A_0^\infty = A_1^\infty =]-\infty, 1]$.

We have presented this example in order to show how the reduced shadow automaton can be used in order to use usual static analysis methods on concurrent programs, avoiding state-space explosion as much as possible. It has the advantage of being short, however it does not really show the main interest of our technique: the scheduling automaton allows us to take in account properties which tightly depend on the way the synchronizations constraint the executions of the programs.

4 Conclusion and Future Work

We have presented an algorithm in order to compute a finite presentation of the trace space of concurrent programs, which may contain loops. An application to abstract interpretation has also described but remains to be implemented. In order to give a simple presentation of the algorithm, we have restricted ourselves here to programs of a simple form (in particular, we have omitted non-determinism). We shall extend our algorithm to more realistic programming languages in a subsequent article. Our approach can also be applied to languages with other synchronization primitives (monitors, send/recv, etc.), for which there are simple geometric semantics available. There are also many possible general improvements of the algorithm; the most appealing one would perhaps be to find a way to have a more modular way of computing the total schedulings by combining locally computed schedulings in $\vec{T}(X)(x, y)$ with varying endpoints x and y. In a near future, the schedulings provided by the algorithm will be used by our tool ALCOOL to analyze concurrent programs using abstract interpretation, thus providing one of the first tools able to do such a static analysis on concurrent programs without forgetting most of the possible synchronizations during their execution.

On the theoretical side, we envisage to study in details and use the structure of the index poset $\mathcal{C}(X)$ which contains much more information than only the schedulings of the program. Namely, it can be equipped with a structure of *prodsimplicial set* [22] (a structure similar to simplicial sets but whose elements are products of simplexes), whose geometric realization provides a topological space which is homotopy equivalent to the trace space $\vec{T}(X)$ [24]. This essentially means that $\mathcal{C}(X)$ contains all the geometry of the trace space and we plan to try to benefit from all the information it provides about the possible computations of a program. Our ALCOOL prototype actually implements this computation – using a combinatorial presentation of the prodsimplicial sets known as *simploidal sets* [23] – which will be reported elsewhere.

References

1. The SPIN Model-Checker, `http://spinroot.com/`
2. Berge, C.: Hypergraphs, vol. 445. North Holland Mathematical Library (1989)
3. Bonichon, R., Canet, G., Correnson, L., Goubault, E., Haucourt, E., Hirschowitz, M., Labbé, S., Mimram, S.: Rigorous Evidence of Freedom from Concurrency Faults in Industrial Control Software. In: Flammini, F., Bologna, S., Vittorini, V. (eds.) SAFECOMP 2011. LNCS, vol. 6894, pp. 85–98. Springer, Heidelberg (2011)
4. Clarke Jr., E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press (1999)
5. Cousot, P., Cousot, R.: Systematic design of program analysis frameworks. In: Proceedings of Principles Of Programming Languages, pp. 269–282. ACM Press (1979)
6. Cousot, P., Cousot, R.: Abstract interpretation based program testing. In: Proc. of the SSGRR 2000 Computer & eBusiness International Conference (2000)
7. Diekert, V., Rozenberg, G.: The Book of Traces. World Scientific (1995)
8. Dijkstra, E.: The Structure of the THE Operating System. Com. of the ACM 11(15) (1968)
9. Dijkstra, E.W.: Hierarchical ordering of sequential processes. Acta Informatica 1(2) (1971)
10. Fajstrup, L.: Trace spaces of directed tori with rectangular holes. Technical Report R-2011-08, Aalborg Univ. (2001)
11. Fajstrup, L., Goubault, É., Raußen, M.: Detecting Deadlocks in Concurrent Systems. In: Sangiorgi, D., de Simone, R. (eds.) CONCUR 1998. LNCS, vol. 1466, pp. 332–347. Springer, Heidelberg (1998)
12. Fajstrup, L., Sokolowski, S.: Infinitely running concurrent processes with loops from a geometric viewpoint. ENTCS 39(2) (2000)
13. Fajstrup, L., Raußen, M., Goubault, E.: Algebraic topology and concurrency. Theor. Comput. Sci. 357(1-3), 241–278 (2006)
14. Fajstrup, L., Raußen, M., Goubault, E., Haucourt, E.: Components of the fundamental category. Appl. Cat. Struct. 12(1), 81–108 (2004)
15. Godefroid, P., Wolper, P.: Using Partial Orders for the Efficient Verification of Deadlock Freedom and Safety Properties. In: Larsen, K.G., Skou, A. (eds.) CAV 1991. LNCS, vol. 575, pp. 332–342. Springer, Heidelberg (1992)

16. Goubault, É., Haucourt, E.: A Practical Application of Geometric Semantics to Static Analysis of Concurrent Programs. In: Abadi, M., de Alfaro, L. (eds.) CONCUR 2005. LNCS, vol. 3653, pp. 503–517. Springer, Heidelberg (2005)
17. Goubault, E., Haucourt, E.: Components of the fundamental category II. Applied Categorical Structures 15(4), 387–414 (2007)
18. Goubault, E., Mimram, S.: Formal relationships between geometrical and classical models for concurrency. CoRR, abs/1004.2818 (2010)
19. Grandis, M.: Directed Algebraic Topology, Models of Non-Reversible Worlds. New Mathematical Monographs, vol. 13. Cambridge University Press (2009)
20. Gunawardena, J.: Homotopy and concurrency. Bulletin of the EATCS 54, 184–193 (1994)
21. Kavvadias, D.J., Stavropoulos, E.C.: Evaluation of an Algorithm for the Transversal Hypergraph Problem. In: Vitter, J.S., Zaroliagis, C.D. (eds.) WAE 1999. LNCS, vol. 1668, pp. 72–84. Springer, Heidelberg (1999)
22. Kozlov, D.: Combinatorial Algebraic Topology. Springer, Heidelberg (2007)
23. Peltier, S., Fuchs, L., Lienhardt, P.: Simploidals sets: Definitions, Operations and Comparison with Simplicial Sets. Discrete Applied Math. 157, 542–557 (2009)
24. Raussen, M.: Simplicial models of trace spaces. Alg. & Geom. Top. 10, 1683–1714 (2010)
25. Valmari, A.: A Stubborn Attack on State Explosion. In: Clarke, E., Kurshan, R.P. (eds.) CAV 1990. LNCS, vol. 531, pp. 156–165. Springer, Heidelberg (1991)

AAECC (2012) 23:59–84
DOI 10.1007/s00200-012-0165-9

ORIGINAL PAPER

Execution spaces for simple higher dimensional automata

Martin Raussen

Received: 5 November 2010 / Revised: 15 August 2011 / Accepted: 25 August 2011 /
Published online: 28 March 2012

Abstract Higher dimensional automata (HDA) are highly expressive models for concurrency in Computer Science, cf van Glabbeek (Theor Comput Sci 368(1–2): 168–194, 2006). For a topologist, they are attractive since they can be modeled as cubical complexes—with an inbuilt restriction for directions of allowable (d-)paths. In Raussen (Algebr Geom Topol 10:1683–1714, 2010), we developed a new method describing, for a certain subclass of HDA, the homotopy type of the space of execution paths (d-paths) as a finite simplicial complex. Several restrictions that were made to ease the presentation in that latter paper will be removed in the present article in order to make the results applicable in greater generality. Furthermore, we take a close look at semaphore models with semaphores all of arity one. It turns out that execution spaces for these are always homotopy discrete with components representing sets of "compatible" permutations. Finally, we describe a model for the complement of the execution space seen as a subspace of a product of spheres—with the aim to make the calculation of topological invariants easier and faster.

Keywords Higher dimensional automata · Execution path · Poset category · Prodsimplicial complex · Duality

Mathematics Subject Classification 55P10 · 55P15 · 55U10 · 68Q55 · 68Q85

M. Raussen (✉)
Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, 9220 Aalborg Øst, Denmark
e-mail: raussen@math.aau.dk

1 Introduction

1.1 Background

A particular model for concurrent computation in Computer Science, called *Higher dimensional automata* (HDA), was introduced by Pratt [22] back in 1991. Mathematically, HDA can be described as (labelled) pre-cubical sets (with n-dimensional cubes instead of simplices as building blocks; cf Brown and Higgins [3,4]) with a preferred set of *directed paths* (respecting the natural partial orders) in any of the cubes of the model.

Compared to other well-studied concurrency models like labelled transition systems, event sturctures, Petri nets etc. (for a survey on those cf Winskel and Nielsen [31]), it has been shown by van Glabbeek [30] that Higher Dimensional Automata have the highest expressivity; on the other hand, they are certainly less studied and less often applied so far.

All concurrency models deal with sets of states and with associated sets of execution paths (with some further structure). The interest is mainly in the structure of the spaces of *execution paths*; typically, it is difficult to extract valuable information about the path space from the state space model. We use topological models for both state space and the execution (=path) space consisting of the directed paths in state space. It is particularly important to know whether the path space is *path-connected*; and, if not, to get an overview over its path components: Executions in the same path component yield the *same result* (decision) in a concurrent computation; different components *may* lead to different results. From a topological perspective, the ultimate aim is to determine the *homotopy type* of these path spaces.

Higher Dimensional Automata are prototypes of *directed* topological spaces, cf Grandis [15,16]; a directed topological space consists of an (ordinary) topological space X together with a subspace $\vec{P}(X) \subseteq X^I = [I; X]$ of "directed" d-paths satisfying several natural requirements: $\vec{P}(X)$ is

- closed under concatenation
- closed under weakly increasing reparametrizations (order-preserving self maps of the unit interval I)
- contains the constant paths.

General topological properties of spaces of d-paths and of *traces* (=d-paths up to monotone reparametrizations; cf Fahrenberg and Raussen [7,25]) in pre-cubical complexes were investigated in Raussen [26]. But so far, apart from low-dimensional examples with convincing drawings, there have been very few explicit examples of actual computations of spaces of such traces (for an attempt in dimension two, cf Raussen [23]); let alone a general method to perform such computations.

The paper Raussen [27] describes an algorithmic method to determine the homotopy types of trace spaces for Higher Dimensional Automata (and thus in particular to calculate and describe their components) through explicitly constructed *finite simplicial complexes*, but only under several restrictions for the HDA under consideration:

1. We had to stick to *semaphore*—or PV—models as described by Dijkstra [5]—an important but restricted class of HDA. Loosely speaking, a PV-model space is a

hypercube I^n—with I the unit interval [0, 1]—from which a number of n-dimensional isothetic hyperrectangles has been removed; cf Sect. 2.1 for details.
2. In order to make matters mathematically "clean", we restricted attention to models in which the forbidden hyperrectangles do not intersect the boundary of I^n. For most natural models, intersections with the boundary *will occur*.
3. Once again, in order to get a mathematically easy description, we described only path spaces from the bottom vertex $\mathbf{0}$ to the top vertex $\mathbf{1}$ in I^n. It is important also to collect information about "intermediate" path spaces between arbitrary points in the model—for example for investigations of its fundamental category (Grandis [15], Goubault et al. [10,14]) but also for inductive reasoning and calculations. In this case, it is typically necessary to consider obstruction hyperrectangles intersecting the boundary (of a smaller hypercube).

In this paper, we will elaborate how to get rid of the last two restrictions; we will still only deal with PV-models. The general idea how a path space associated to the model can be represented in simplicial terms is the same as in Raussen [27]. We will explain it in Sect. 2 and moreover sketch a mathematical proof for the fact that the simplicial model is in fact homotopy equivalent to the space of directed paths (executions). This proof follows the ideas already described in Raussen [27].

As already explained in Raussen [27, Section 5.1], it is not difficult to include cases in which individual processes are allowed to branch, merge and loop. Each individual program is then modelled by a digraph; the state space is a product of digraphs from which a number of "hyperrectangles" has been removed. For some first ideas on how to achieve simplicial models for execution spaces on general HDA (not necessarily semaphores), we refer to Raussen [27, Section 5.2] and, in particular to Raussen [28].

1.2 Structure and overview of results

In Sect. 2, we introduce model spaces for semaphore models in concurrency and review and modify methods from Raussen [27] that allow to determine combinatorial/topological models of the associated spaces of directed paths (in these spaces). The key idea is to decompose such a model space into pieces that are *contractible*, i.e., homotopy equivalent to a point; even more important, the spaces of *d-paths within every such subspace* are *either contractible or empty*.

Fix a model space X and a pair $(\mathbf{c}, \mathbf{d})$ of start and end point within X. We wish to derive a finite description of the space $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ of directed paths joining $\mathbf{c}$ to $\mathbf{d}$, or rather of the homotopy equivalent trace space $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ consisting of *traces*, ie equivalence class of directed paths up to weakly increasing reparametrizations, cf Fahrenberg and Raussen [7,25].

We associate to X and to $(\mathbf{c}, \mathbf{d})$ a *poset category* $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$. That category is naturally isomorphic to a subcategory of a *product* of a number of poset categories consisting of non-empty subsets of the positive integers $[1 : n]$ less than or equal to n. A topological realization of this subcategory can thus be modelled on *products of simplices* and gives rise to an explicit *prodsimplicial complex*, cf Kozlov [20], called $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$. Using standard methods from algebraic topology already explained in Raussen [27]—modifications of the nerve lemma, cf Kozlov [20, Theorem 15.21]—

we show that the space of directed paths $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ in X from $\mathbf{c}$ to $\mathbf{d}$ or, equivalently, the trace space $\vec{T}(X)(\mathbf{c}, \mathbf{d})$, is homotopy equivalent to that finite complex $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$. The latter in turn has the nerve $\Delta(\mathscr{C}(X)(\mathbf{c}, \mathbf{d}))$ of the poset category as a barycentric subdivision.

A similar technique works also for spaces of directed paths starting at a given point and ending on the *upper boundary of a hypercube*. This is interesting both for inductive reasoning but also for the investigation of the decision power of distributed concurrent processes of which some may die (compare Herlihy and Rajsbaum [18] and more recent papers in distributed computing).

For calculations, it is essential to determine the category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ explicitly. It can be described as a subcategory of the order category of *binary* $l \times n$-matrices $M_{l,n} \cong (\mathbf{Z}/2)^{ln}$ (with l the number of obstructions and n the dimension of the model space) with the componentwise partial order. In Sect. 3, we describe how to achieve this: One needs to decide, for every of the contractible subspaces mentioned above, whether there exists a directed path *within* that subspace from $\mathbf{c}$ to $\mathbf{d}$. It turns out that it is enough to find out whether there exist *deadlock* points (the only d-path with a deadlock as source is trivial) in certain *associated* models and to apply a combinatorial search algorithm for deadlocks described in Fajstrup, Goubault and Raussen [11]. The outcome of a systematic search for deadlocks (in all associated models) is a set $D(X)(\mathbf{c}, \mathbf{d})$ of *minimal non-faces*—all of the same dimension $n-1$—describing the prodsimplicial complex $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$ within the prodsimplicial complex $(\Delta^{n-1})^l$. The *maximal faces* of $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$ can then be determined via minimal transversals in an associated hypergraph, as already described in Raussen [27, Section 4.2].

In Sect. 4, we indicate how the topology of execution spaces $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ *alters* under variation of end points; and in particular, when it does not change! This is important for inductive reasoning and for obtaining a complete overview over the trace category; cf Raussen [24]. The trace category determines the fundamental category $\vec{\pi}_1(X)$ of the model space by taking the connected components of the morphism spaces; this is the information needed to classify the possible outcome of partial executions.

Section 5 is devoted to an application of the results to a specific case: these are sempaphore models in which all semaphores are of *arity one*, ie they allow only *one* process to proceed at any given time. In that case, the space of executions is shown to be *homotopy discrete*; all homotopy information is contained in the fundamental category—with *finite* sets of morphisms. These morphism sets can be described as subsets of *compatible* permutations within $(\Sigma_n)^k$ where k is the number of semaphores involved.

The final Sect. 6 deals with a computational issue from a theoretical perspective: The prodsimplicial complex $\mathbf{T}(X)(-,-)$ modelling execution spaces embeds naturally in a product $(\partial\Delta^{n-2})^l \cong (S^{n-2})^l$ of l spheres. It seems to be algorithmically easier and quicker to determine the complement $\mathbf{U}(X)(-,-) := (\partial\Delta^{n-2})^l \setminus \mathbf{T}(X)(-,-)$ of the trace complex by giving it a prodsimplicial structure. Poincaré–Alexander–Lefschetz duality can then be applied to infer information from the complement to trace space itself.

1.3 Implementation issues

Some first steps towards implementation, in particular the determination of the category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$, have been taken in collaboration with a team at CEA/LIST in Saclay, France; first results and benchmarks can be found in Fajstrup et al. [9].

The outcome of this algorithm has been combined with fast algorithms for the calculation of the homology of big chain complexes as eg in Kaczynski, Mrozek and Slusarek [19]; this is a still ongoing project. Moreover, we suggest a systematic investigation of how to use and implement these new methods to improve applications of geometric semantics to the static analysis of concurrent programs, cf Goubault and Haucourt [13].

2 Prodsimplicial models for execution spaces

2.1 A model space and contractible subspaces

2.1.1 Geometric semaphore models

To start with, we analyse spaces of directed paths in a simple model space that can be described as follows: A (linear) schedule for each of a number of n individual processors P_j, $1 \leq j \leq n$, is modelled on the directed interval $\vec{I}_j = [0, 1]$. On subintervals $I^i_j =]a^i_j, b^i_j[\subseteq I_j$, $1 \leq i \leq l$, there is potential conflict with the schedules of the other processors. Let $\mathbf{a}^i = (a^i_1, \ldots, a^i_n)$, $\mathbf{b}^i = (b^i_1, \ldots, b^i_n) \in I^n \setminus \partial I^n$ and let $R^i = \{\mathbf{x} \in I^n | a^i_j < x_j < b^i_j, 1 \leq j \leq n\}$ denote the "homothetic" hyperrectangle (with faces parallel to the coordinate planes) with bottom corner $\mathbf{a}^i$ and top corner $\mathbf{b}^i$.

The state space for concurrent executions of these n linear processes is the space $X = \vec{I}^n \setminus F \subset \vec{I}^n$ with the *forbidden region* $F = \bigcup_{i=1}^{l} R^i$. The forbidden region F models conflicts and cannot be entered due to guarding semaphores (Dijkstra's PV-models [5]; an interval $]a^i_j, b^i_j[$ corresponds to a call $PcVc$ to a semaphore). See Fig. 2 for an example of a forbidden region. The space X inherits a partial order $\leq$ from the componentwise partial order $\leq$ on $\vec{I}^n$.

We study compound schedules (execution paths) in such a state space X: A *d-path* in X is a continuous path $p\colon \vec{I} \to X$ that is continuous and *order-preserving*: each coordinate $\pi_j \circ p\colon \vec{I} \to X \subset \vec{I}^n \to \vec{I}$, $1 \leq j \leq n$, is *weakly increasing*. The set $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ consists of all d-paths in X starting at $\mathbf{c} \in X$ and ending at $\mathbf{d} \in X$; in particular, these d-paths avoid the "forbidden region" $F \subset \vec{I}^n$. Consult eg Gunawardena [17] and Fajstrup, Goubault and Raussen [11] for detailed descriptions.

As a topological space, $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ is given the subspace topology inherited from the space $P(X)(\mathbf{c}, \mathbf{d}) = [(I, 0, 1); (X, \mathbf{c}, \mathbf{d})]$ of *all* paths in X from $\mathbf{c}$ to $\mathbf{d}$ in the compact-open topology (= uniform convergence topology).

Reparametrization equivalent d-paths, cf Fahrenberg and Raussen [7] in X have the same directed image (= *trace*) in X. Dividing out the action of the monoid of (weakly-increasing) reparametrizations of the parameter interval $\vec{I}$, we arrive at trace space

$\vec{T}(X)(\mathbf{c},\mathbf{d})$ (cf Fahrenberg and Raussen [7,25] which is shown in Raussen [26] to be homotopy equivalent to path space $\vec{P}(X)(\mathbf{c},\mathbf{d})$ for a far wider class of directed spaces X; in the latter paper, it is also shown that trace spaces enjoy nice properties: They are metrizable, locally compact, locally contractible, and they have the homotopy type of a CW-complex.

It is possible to generalise these models to incorporate concurrent executions of non-linear processes that are allowed to branch, to merge and to loop, still governed by semaphores; cf Raussen [27, Section 5.2]. Simple semaphore models with loops constructed from spaces of the form $X = T^n \setminus F$ with $T^n = (S^1)^n$ a directed n-torus and F a collection of forbidden rectangles have been investigated in considerable depth in Fajstrup [8] and Fajstrup et al. [9].

2.1.2 *Contractible subspaces*

We will now describe certain subspaces of X and then prove that associated spaces of d-paths *within* these subspaces are *either empty or contractible*. We need some notation:

- The set of elements "below" $\mathbf{d} \in X$ is denoted $\downarrow\mathbf{d} := \{\mathbf{x} \in X \mid \mathbf{x} \le \mathbf{d}\} = \{\mathbf{x} \in I^n \mid \mathbf{x} \le \mathbf{d}, \mathbf{x} \notin F\}$.
  Remark that it is not always possible to reach $\mathbf{d}$ from every $\mathbf{x} \in\downarrow\mathbf{d}$ by a d-path. Likewise $\uparrow\mathbf{c} = \{\mathbf{x} \in X \mid \mathbf{c} \le \mathbf{x}\}$ denotes the set of elements above $\mathbf{c}$.
- The upper boundary $\{\mathbf{x} \in\downarrow\mathbf{d} \mid \exists 1 \le i \le n\colon x_i = d_i\}$ of the hyperrectangle with corners in $\mathbf{0}$ and $\mathbf{d}$ within X will be denoted $\partial_+\downarrow\mathbf{d}$.
- $\mathbf{a}^i = (a^i_1, \dots, a^i_n)$, $\mathbf{b}^i = (b^i_1, \dots, b^i_n)$.

Definition 1 1. For $1 \le i \le l$, $1 \le j_i \le n$, let

$$X_{j_1,\dots,j_l} := \{\mathbf{x} \in X \mid \forall i\colon\ x_{j_i} \le a^i_{j_i} \text{ or } \exists k\colon\ x_k \ge b^i_k\} \tag{1}$$

$$= \{\mathbf{x} \in X \mid \forall i\colon\ (\forall k\ x_k < b^i_k \Rightarrow x_{j_i} \le a^i_{j_i})\}. \tag{2}$$

2. For non-empty subsets $J_i \subseteq [1:n]$, $1 \le i \le l$, let

$$\begin{aligned} X_{J_1,\cdots,J_l} &:= \{\mathbf{x} \in X \mid \forall i\colon\ x^i_{j_i} \le a^i_{j_i},\ j_i \in J_i, \quad \text{or} \quad \exists k\colon\ x_k \ge b^i_k\} \\ &= \{\mathbf{x} \in X \mid \forall i\colon\ (\forall k\ x_k < b^i_k \Rightarrow x_{j_i} \le a^i_{j_i} \text{ for all } j \in I_j)\} \\ &= \bigcap_{j_i \in J_i} X_{j_1,\dots,j_l}. \end{aligned}$$

A graphical illustration for an example of several such subspaces in two and three dimensions can be found in Fig. 1.

Proposition 1 *For X as above and $(\mathbf{c},\mathbf{d}) \in X \times X$ we have:*

1. $\vec{T}(X)(\mathbf{c},\mathbf{d}) = \bigcup_{[1:n]^l} \vec{T}(X_{j_1,\dots,j_l})(\mathbf{c},\mathbf{d})$*; and*
$\vec{T}(X)(\mathbf{c},\partial_+\downarrow\mathbf{d}) = \bigcup_{[1:n]^l} \vec{T}(X_{j_1,\cdots,j_l})(\mathbf{c},\partial_+\downarrow\mathbf{d})$.

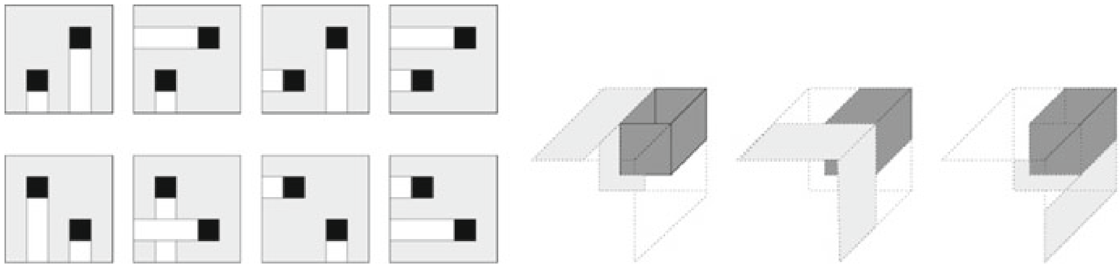


Fig. 1 From Raussen [27]: In both *rows* on the *left hand side*, the space X is the complement of the *black boxes* within a *square*. The subspaces $X_{ij},\ 1 \le i, j \le 2$ are given by the *shaded areas*. On the *right hand side*, the space X is the complement of the *black box* within a unit *cube*. The *shaded areas* show the intersections of the subspaces $X_i,\ 1 \le i \le 3$ with the upper boundary of the subcube extended to the lower boundary

2. *Let* $\emptyset \neq J_i \subseteq [1 : n], 1 \le i \le l, \mathbf{c}, \mathbf{d} \in X_{J_1,\dots J_l}$. *Then the trace spaces* $\vec{T}(X_{J_1,\dots J_l})(\mathbf{c}, \mathbf{d})$, *resp.* $\vec{T}(X_{J_1,\dots J_l})(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$ *are either empty or contractible.*

Proof For trace spaces with fixed end points as in (1), this was proved in Raussen [27, Lemma 2.10, Proposition 2.8(2)]. We repeat the essential arguments to make this article self-contained:

1. We can represent a trace in $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ by a parametrized d-path $p : I \to X$. For an obstruction R^i, choose a maximal $t_i \in I$ such that there exists j_i with $p_{j_i}(t_i) = c_{j_i}$; and hence $t \le t_i \Rightarrow p_{j_i}(t) \le c_i$. Then $p_j(t) > c_j$ for all $t > t_i$ and all j; there has to exist $1 \le k \le n$ such that $p_k(t) \ge d_k$. We conclude that $p(t) \in X_{j_1,\dots,j_l}$ for all t.
2. It is easy to see that all spaces $X_{j_1,\dots,j_l}$ are closed under the least upper bound operation $\vee$ ($\mathbf{a} \vee \mathbf{b} = (\max(a_1, b_1), \dots, \max(a_n, b_n))$. Since intersections of $\vee$-closed subsets are $\vee$-closed, also all spaces $X_{J_1,\dots,J_l}$ are $\vee$-closed. In particular, given two d-paths p and q in $\vec{P}(X_{J_1,\dots J_l})(\mathbf{c}, \mathbf{d})$, the d-path $p \vee q$ is also contained in $\vec{P}(X_{J_1,\dots J_l})(\mathbf{c}, \mathbf{d})$; likewise also for the spaces of d-paths ending in $\partial_+ \downarrow \mathbf{d}$, since this target space is closed under $\vee$ itself.
   Let us write $\vec{P}(X_J)$ as short form for $\vec{P}(X_{J_1,\dots J_l})(\mathbf{c}, \mathbf{d})$ or $\vec{P}(X_{J_1,\dots J_l})(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$. Consider the map $H : \vec{P}(X_J) \times \vec{P}(X_J) \times I \to \vec{P}(X_J)$ given by $H(p, q, t)(s) = q(s) \vee p(ts)$. Its concatenation with $H(q, p, t)$ defines a d-homotopy from q to p (via $p \vee q = q \vee p$). If the second coordinate is restricted to a particular d-path $p \in \vec{P}(X_J)$, we get a contraction of $\vec{P}(X_J)$ to p as a result. □

2.2 Index categories, matrix representations and homotopy equivalences

2.2.1 A matrix representation of a power poset

The index multisets $(J_1, \dots, J_l)$ with $J_i \subseteq [1 : n]$ considered in the previous Sect. 2.1.2 may be viewed as elements of $(\mathcal{P}([1 : n]))^l \cong \mathcal{P}([1 : l] \times [1 : n])$. Elements of the latter power set can be encoded by their characteristic functions which can be viewed as *binary* $l \times n$-matrices:

Let $M_{l,n} = M_{l,n}(\mathbf{Z}/2)$ denote the set of all binary $l \times n$-matrices—with 2^{ln} elements. The total order on $\mathbf{Z}/2$ given by $a \le b$ unless ($a = 1$ and $b = 0$) extends to a componentwise given *partial order* $\le$ on $M_{l,n}$. With this partial order defining the morphisms, $M_{l,n}$ will be viewed as a *poset category*.

There is a natural order-preserving bijection between the subsets of $[1:l] \times [1:n]$ (elements of the power set $\mathscr{P}([1:l] \times [1:n])$ with partial order given by inclusion) and elements in $M_{l,n}$ given by

$$J = (J_1, \ldots, J_l) \mapsto M^J = (m^J_{ij}), \quad m^J_{ij} = 1 \Leftrightarrow j \in J_i \tag{3}$$

with inverse $M = (M_{ij}) \mapsto J^M$, $j \in J^M_i \Leftrightarrow m_{ij} = 1$. Under this bijection, the relevant multisets $J = (J_1, \ldots, J_l)$ with $J_i \neq \emptyset, 1 \le i \le l$, correspond to matrices in the subset $M^R_{l,n} \subset M_{l,n}$ consisting of the $(2^n - 1)^l$ matrices such that *no row vector is a zero vector*. We view $M^R_{l,n}$ as the *full subposet category* within $M_{l,n}$.

To ease notation, we will in the following write X_M instead of X_{J^M}. The relevant index category to consider here is the full subposet category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d}) \subset M^R_{l,n} \subset M_{l,n}$ consisting of all binary matrices M such that

$$\vec{T}(X_M)(\mathbf{c}, \mathbf{d}) \text{ is non-empty.} \tag{4}$$

Likewise, $\mathscr{C}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d}) \subset M^R_{l,n} \subset M_{l,n}$ consists of all binary matrices M such that

$$\vec{T}(X_M)(\mathbf{c}, \partial_+ \downarrow \mathbf{d}) \text{ is non-empty.} \tag{5}$$

Departing from the index category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ we construct a *prodsimplicial complex* (in the terminology of Kozlov [20]) $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$ as follows:

To every matrix $M \in \mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ the product of simplices $\Delta(M) = \prod_{i=1}^{l} \Delta_i(M) \subset (\Delta^{n-1})^l \subset \mathbf{R}^{nl}$ with

$$\Delta_i(M) := \left\{ (t_1, \ldots, t_n) |\ 0 \le t_j \le 1, \sum_{j=1}^{n} t_j = 1, m_{ij} = 0 \Rightarrow t_j = 0 \right\} \subset \Delta^{n-1}.$$

Remark that $M \le N \Rightarrow \Delta(M) \subset \Delta(N)$.
Using the category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ as a pasting scheme we define the colimits

$$\mathbf{T}(X)(\mathbf{c}, \mathbf{d}) := \bigcup_{M \in \mathscr{C}(X)(\mathbf{c}, \mathbf{d})} \Delta(M), \tag{6}$$

and in a completely analogous way

$$\mathbf{T}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d}) := \bigcup_{M \in \mathscr{C}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d})} \Delta(M). \tag{7}$$

2.2.2 Homotopy equivalences between trace spaces and finite prodsimplicial complexes

For the following result, we need a technical, but natural and generic assumption about the placement of the hyperrectangles R^i making up the forbidden region F: For every $1 \leq j \leq n$, no upper boundary coordinate b^i_j is equal to a lower boundary coordinate a^k_j. Under this assumption we get a homotopy equivalence between the infinite dimensional trace space and a finite prodsimplicial model:

Theorem 1 *1. Trace space* $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ *is homotopy equivalent to the prodsimplicial complex* $\mathbf{T}(X)(\mathbf{c}, \mathbf{d}) \subset (\partial\Delta^{n-1})^l$ *and to the nerve* $\Delta(\mathscr{C}(X)(\mathbf{c}, \mathbf{d}))$ *of the category* $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$.

2. Trace space $\vec{T}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$ *is homotopy equivalent to the prodsimplicial complex* $\mathbf{T}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d}) \subset (\partial\Delta^{n-1})^l$ *and to the nerve* $\Delta(\mathscr{C}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d}))$ *of the category* $\mathscr{C}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$.

Proof The proof is completely analogous to that of Raussen [27, Theorem 3.5] in the case $\mathbf{c} = \mathbf{0}$ and $\mathbf{d} = \mathbf{1}$: The homotopy colimit of the functor associating the contractible spaces $\mathbf{T}(X_M)(\mathbf{c}, \mathbf{d})$, resp. Δ_M to a matrix $M \in \mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ is homotopy equivalent to the functor associating the same point to every such M, ie to the nerve of the category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$. Homotopy colimit and colimit of the first two functors are homotopy equivalent to each other by the projection lemma (cf Kozlov [20, Theorem 15.19]). By Proposition 1, this colimit is the entire trace space; resp. the prodsimplicial complex $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$, by definition. □

Remark 1 If a matrix $M \in M_{l,n}$ contains a row $\mathbf{m}_i = \mathbf{1} = (1, \ldots, 1)$, $1 \leq i \leq l$, corresponding to $J_i = [1 : n]$, then it is easy to see from Definition 1 that $\vec{T}(X_M)(\mathbf{c}, \mathbf{d})$ is empty. This is why trace space can be embedded in $(\partial\Delta^{n-1})^l \subset (\Delta^{n-1})^l$. This observation will be exploited in Sect. 6 in which we view $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ as a subset of the space of matrices $\tilde{M}^R_{l,n}$ for which every row vector contains at least one 0 and at least one 1.

3 State spaces with forbidden region intersecting the boundary

3.1 Introduction

In Raussen [27], we had assumed that all obstruction hyperrectangles R^i are contained in the *interior* of I^n and obtained a method to enumerate the index category $\mathscr{C}(X)(\mathbf{0}, \mathbf{1})$. It is the purpose of Sect. 3 to modify that method and to obtain descriptions of the index category $\mathscr{C}(X)(\mathbf{0}, \mathbf{1})$ in the general case and, by more or less the same method, to generalize to the categories derived in Theorem 1 and described below. We shall use the notation for subspaces $X_M \subseteq X$ introduced in Sect. 2.2.1, we will describe the index categories (with morphisms given by the partial order)

– $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ with set of objects $\{M \in M^R_{l,n} \mid \vec{T}(X_M)(\mathbf{c}, \mathbf{d}) \neq \emptyset\}$ corresponding to trace space $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ with traces starting at $\mathbf{c}$ and ending at $\mathbf{d}$; and

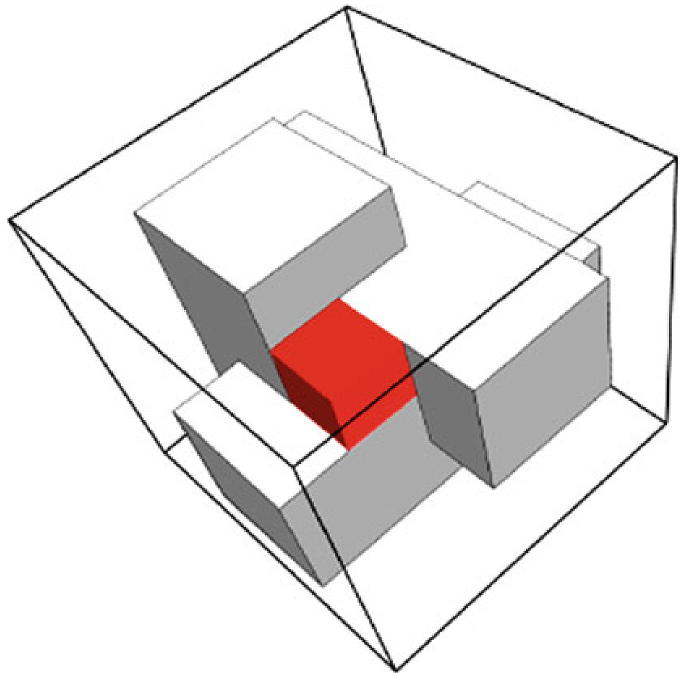


Fig. 2 Cubical complex describing the 3-philosophers problem within the 3-cube I^3 with forbidden region F intersecting its boundary ∂I^3; the small *red interior cube* represents the unsafe region and is not a part of the model. Figure courtesy to E. Haucourt and A. Lang. (Color figure online)

– $\mathscr{C}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$ with set of objects $\{M \in M_{l,n}^R \mid \vec{T}(X_M)(\mathbf{c}, \partial_+\downarrow\mathbf{d}) \neq \emptyset\}$ corresponding to traces ending on the *upper boundary* of the box with corners at $\mathbf{c}$ and $\mathbf{d}$ within $X \subset I^n$.

Remark 2 1. PV-models for higher dimensional automata have often obstructions intersecting the boundary ∂I^n; those arise as soon as semaphores of an arity $r < n-1$ (at most r processors can proceed concurrently) are involved. This is for example the case for the cubical model describing the dining philosophers problem, cf Dijkstra [6] and Fig. 2 below.

2. The trace space $\vec{T}(X)(\mathbf{0}, \partial_+\downarrow\mathbf{1})$ is interesting in the analysis of algorithms for wait-free protocols (cf eg Herlihy and Rajsbaum [18]) in which all processors with at least one exception are allowed to “die”, ie cease to communicate. Such an algorithm may end in a state where only one processor P_i ends in the final state $x_i = 1$; for all others, we only know that $0 \leq x_j \leq 1$. In this case, the accepting states correspond therefore to the points contained in $\partial_+\downarrow\mathbf{1}$.

We have seen in Sect. 2.2.1 that the matrix poset categories $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ as well as $\mathscr{C}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$ serve as pasting schemes that give rise to prodsimplicial complexes $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$—cf (6); and $\mathbf{T}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$—cf (7). Under the general conditions of Raussen [27, Theorem 3.5], but allowing obstruction hyperrectangles to intersect the boundary of $[\mathbf{c}, \mathbf{d}]$, we obtain using Raussen [27, Proposition 2.8] as an analogue to that theorem with essentially the same proof:

Theorem 2 *1. Trace space $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ is homotopy equivalent to the prodsimplicial complex $\mathbf{T}(X)(\mathbf{c}, \mathbf{d})$ and to the nerve of the category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$.*

2. Trace space $\vec{T}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$ is homotopy equivalent to the prodsimplicial complex $\mathbf{T}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$ and to the nerve of the category $\mathscr{C}(X)(\mathbf{c}, \partial_+\downarrow\mathbf{d})$.

For an algorithmic determination of these index categories (as in Raussen [27, Section 4]), we need to describe several modifications of the subsets $D(X)(-,-)$ of "dead" matrices (cf Sect. 3.2 below) and of "alive" matrices $\mathscr{C}(X)(-,-)$ with respective boundaries in parentheses.

3.2 Which trace spaces are (non-)empty?

3.2.1 Extended obstructions

As in Raussen [27, Section 4.2], we define the map $\Psi : M_{l,n} \to \mathbf{Z}/2$ by
$\Psi(M) = 1 \Leftrightarrow \overrightarrow{T}(X_M)(-,-) = \emptyset$ (with relevant boundaries).
We wish to determine the matrices $M \in \mathscr{C}(X)(-,-) \subset M^R_{l,n} \subset M_{l,n}$ with $\Psi(M) = 0$:

Again, one first determines "generating" matrices with $\Psi(M) = 1$ arising from a deadlock condition; it is here that several modifications become necessary as compared to Raussen [27, Section 4]. First of all, we may have fewer matrices to consider:

In both cases, hyperrectangles $R^i \subseteq [\mathbf{0}, \mathbf{1}]$ that do not intersect the box $[\mathbf{c}, \mathbf{d}]$ between $\mathbf{c}$ and $\mathbf{d}$ become irrelevant. This can be handled by reducing the number of rows in the matrices representing the index categories: We separate $[1 : l] = [1 : l]^{in} \sqcup [1 : l]^{out}$ with $i \in [1 : l]^{in} \Leftrightarrow (1 \le j \le n \Rightarrow a^i_j < d_j, c_j < b^i_j)$ and let $l' := |[1 : l]^{in}|$.

Remark 3 Comparing trace spaces with varying end points, it may be necessary to take intor account these irrelevant rectangles nevertheless. On the prodsimplicial side this will result in taking a product with one or several simplices Δ^{n-1}; cf. Sect. 4.

Lemma 1 *Suppose $a^i_j \le c_j$. Then $M \notin \mathscr{C}(X)(\mathbf{c}, -)$ for every matrix $M \in M^R_{l',n}$ with $m_{ij} = 1$.*

Proof Suppose $\mathbf{x} \in X_M$ with $m_{ij} = 1$. Then $\mathbf{x} \le \mathbf{b}^i$ implies $x_j < a^i_j \le c_j$, i.e., $\mathbf{x} \notin [\mathbf{c}, \mathbf{1}]$; in particular, $\overrightarrow{T}(X_M)(\mathbf{c}, -) = \emptyset$. □

Under these circumstances, we will thus only have to investigate matrices

$$M \in M^R_{l',n}(\mathbf{c}) \text{ with the additional property: } a^i_j \le c_j \Rightarrow m_{ij} = 0. \tag{8}$$

As in Raussen [27, Section 4], we will deal with *extensions* $R^i_j, 0 \le i \le l'$, $1 \le j \le n$, of the (relevant) obstruction hyperrectangles; these are given as

$$R^i_j = \prod_{k=1}^{j-1}[0, b^i_k[\times]a^i_j, b^i_j[\times \prod_{k=j+1}^{n}[0, b^i_k[, \quad 1 \le i \le l, 1 \le j \le n, \text{ and} \tag{9}$$

$$R^0_j = [0, 1]^{j-1} \times [d_j, 1] \times [0, 1]^{n-j}, \quad i = 0, 1 \le j \le n. \tag{10}$$

For an illustration of the extensions from (9), we refer to Raussen [27, Figure 4]. The hyperrectangles from (10) are new compared to Raussen [27]; they intersect the box $[\mathbf{c}; \mathbf{d}]$ only on an upper boundary facet and may generate deadlocks on that facet.

3.2.2 *Combinatorial descriptions of index categories*

We will deal with the easier case of the index category $\mathcal{C}(X)(\mathbf{c}, \partial_+(\downarrow \mathbf{d}))$ first. The result Raussen [27, Proposition 4.3] has the following immediate modification:

Proposition 2 *For $M \in M^R_{l',n}(\mathbf{c})$, the following are equivalent:*

1. *M is not an object in $\mathcal{C}(X)(\mathbf{c}, \partial_+(\downarrow \mathbf{d}))$.*
2. *$\vec{T}(X_M)(\mathbf{c}, \partial_+(\downarrow \mathbf{d})) = \emptyset$.*
3. *There is a map $i : [1:n] \to [1:l']$ such that $m_{i(j),j} = 1$ and $\bigcap_{1\le j\le n} R_j^{i(j)} \neq \emptyset$.*
4. *There is a map $i : [1:n] \to [1:l']$ with $a_j^{i(j)} < b_j^{i(k)}$ for all $j, k \in [1:n]$.*

Proof The proof is an easy modification of the one given for Raussen [27, Proposition 4.3]: trace space is empty if every trace is bound to end in a deadlock arising from an n-tuple of extended hyperrectangles. Note that a deadlock on the boundary $\partial_+(\downarrow \mathbf{d})$ is irrelevant for paths/traces in $\vec{T}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$. □

For the analysis of $\mathcal{C}(X)(\mathbf{c}, \mathbf{d})$ in general, we have to deal with obstruction hyperrectangles *intersecting* $\partial_+ \downarrow \mathbf{d}$ as well; this may also happen in the case $\mathbf{d} = \mathbf{1}$ for hyperrectangles R^i intersecting the boundary of I^n. For every such "intersection direction" $1 \le j \le n$ with a hyperrectangle intersecting the jth facet $x_j = d_j$ of the upper boundary of $[\mathbf{0}; \mathbf{d}]$, we apply the new obstruction hyperrectangles R_j^0 introduced above.

The result from Raussen [27, Proposition 4.3] can then be modified as follows:

Proposition 3 *For $M \in M^R_{l',n}(\mathbf{c})$, the following are equivalent:*

1. *M is not an object in $\mathcal{C}(X)(\mathbf{c}, \mathbf{d})$.*
2. *$\vec{T}(X_M)(\mathbf{c}, \mathbf{d}) = \emptyset$.*
3. *There is a map $i : [1:n] \to [0:l']$ such that $i(j) \neq 0 \Rightarrow m_{i(j),j} = 1$ and such that $\bigcap_{1\le j\le n} R_j^{i(j)} \neq \emptyset$.*
4. *There is a map $i : [1:n] \to [0:l']$ such that $i(j) \neq 0 \Rightarrow m_{i(j),j} = 1$ and such that*
$$\begin{cases} a_j^{i(j)} < b_j^{i(k)} & \text{for } j,k \in [1:n], i(j) > 0; \text{or } i(j) = 0, a_j^0 = d_j < 1; \\ b_j^{i(k)} = 1 & \text{for } j,k \in [1:n], i(j) = 0, a_j^0 = d_j = 1. \end{cases}$$

Compared to the result Raussen [27, Proposition 4.3], remark that further intersections involving hyperrectangles R_j^0—but only those corresponding to intersection directions—need to be considered.

3.3 Modified algorithms

3.3.1 *D versus $\mathcal{C}$*

The matrix sets and categories from Raussen [27, Section 4.2.1] need a few modifications: First of all, in both cases, only obstructions intersecting $[\mathbf{c}, \mathbf{d}]$ need to be taken care of, and this may reduce the number of rows from l to l' in the matrices to be considered. For the category $\mathscr{C}(X)(\mathbf{c}, \partial_+ \downarrow \mathbf{d})$, one may then proceed as in Raussen [27, Section 4.2.1]—with the simplification that only matrices in $M^R_{l',n}(\mathbf{c})$—cf (8)—need to be considered.

As in Raussen [27, Section 4.2], we determine the index category $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$ in two steps: First, we calculate the restriction of the map Ψ to the subset $M^C_{l',n}(\mathbf{c})$ consisting of matrices $M(i)$ corresponding to maps $i : [1 : n] \to [0, l']$—cf Sect. 3.3.2 below—as in Proposition 3. As in Raussen [27, Proposition 4.5], one obtains $\Psi(M)$ (cf Sect. 3.2) for a matrix $M \in M^R_{l',n}(\mathbf{c})$:

Proposition 4 *A matrix $M \in M^R_{l',n}(\mathbf{c})$ satisfies $\Psi(M) = 1$ if and only if there exists a matrix $N \in M^C_{l',n}(\mathbf{c})$ with $\Psi(N) = 1$ and $N \leq M$.*

In particular, we determine the set of matrices (D for "dead")

$$D(X)(-, -) := \{M \in M^C_{l',n}(\mathbf{c}) |\ \Psi(M) = 1\}. \tag{11}$$

Using this set $D(X)(-, -)$—upward closed under $\leq$—we will then apply (4) to determine the set of matrices

$$\mathscr{C}(X)(-, -) := \{M \in M^R_{l',n}(\mathbf{c}) |\ \Psi(M) = 0\}. \tag{12}$$

describing the objects of the relevant index category; the latter is downward closed under $\leq$.

3.3.2 Determination of $D(X)(\mathbf{c}, \mathbf{d})$

For a category of type $\mathscr{C}(X)(\mathbf{c}, \mathbf{d})$, we replace the matrix set $M^C_{l,n}$ by the set $M^C_{l',n}(\mathbf{c})$ consisting of matrices M has the following properties:

- $a^i_j < c_j \Rightarrow m_{ij} = 0$;
- every column vector $\mathbf{m}_j$ is either a unit vector or the zero vector $\mathbf{0}$;
- if $\mathbf{m}_j = \mathbf{0}$, then j is an intersection direction.

A matrix $M \in M^C_{l',n}(\mathbf{c})$ codes the map

$$i_M : [1 : n] \to [0 : l'], i_M(j) = \begin{cases} i(j) & \mathbf{m}_j = \mathbf{e}_{i(j)} \\ 0 & \mathbf{m}_j = \mathbf{0} \end{cases}.$$

Vice versa, a (relevant) map $i : [1 : n] \to [0 : l']$ comes with a characteristic matrix $M(i) = (m_{ij}) \in M^C_{l',n}(\mathbf{c}), m_{ij} = 1 \Leftrightarrow i(j) = i \neq 0$.

In order to determine $D(X)(\mathbf{c}, \mathbf{d})$—cf (11), one has to consider both the row set $R(M) \subset [1 : l']$ (cf Raussen [27, Section 4.2.2]) and moreover the *column set* $C(M) \subset [1 : n]$ indexing the *non-zero* rows, resp. columns of M.

The method described in Raussen [27, Lemma 4.6/4.7] has to be extended as follows: for a given non-empty (row) subset $B \subset [1 : l']$, one determines first the bound $\mathbf{b}^B = [b^B_1, \ldots, b^B_n]$, $b_j = \min_{i \in B} b^i_j$; the greatest lower bound of the maxima of the hyperrectangles indexed by B. Given $\mathbf{b}^B$, we can determine a *maximal* column set $C(B) := \{j \in [1:n] |\ i(j) = 0 \Rightarrow d_j < b_j^{r_j} \text{ or } d_j = b_j^{r_j} = 1\} \subset [1:n]$; this requires checking n (in)equalities.

For every subset of $C \subseteq C(B)$, the sets $R_j(B;C) := \{i \in B,\ j \notin C \Rightarrow a_j^i < b_j^B\}$ have then to be determined—decrementally—as in Raussen [27, Lemma 4.7]. We end up determining the set of matrices $M \in D(X)(\mathbf{c},\mathbf{d}) := \{M \in M_{l',n}^C(\mathbf{c}) |\ \Psi(M) = 1\}$—as in Raussen [27, Lemma 4.6]—as follows:

Lemma 2 *A map* $i : [1:n] \to [0:l']$ *gives rise to a matrix* $M = M(i) \in D(X)(\mathbf{c},\mathbf{d})$ *if and only if*

1. $i(j) = 0 \Rightarrow j \in C(i([1:n])$ *and*
2. $i(j) > 0 \Rightarrow i(j) \in R_j(i([1:n]); i^{-1}(0))$.

Having found $D(X)(\mathbf{c},\mathbf{d})$, we can now determine the matrices in $\mathscr{C}(X)(\mathbf{c},\mathbf{d}) := \{M \in M_{l',n}^R(\mathbf{c}) |\ \Psi(M) = 0\}$ as described in Raussen [27, Proposition 4.8]; again, only matrices in $M_{l',n}^R(\mathbf{c})$ need to be checked. Alternatively, one can determine the complement of $\mathscr{C}(X)(\mathbf{c},\mathbf{d})$ as described in Sect. 6.

4 The trace category: varying end points

4.1 Induced maps

By concatenation, traces $\sigma \in \vec{T}(X)(\mathbf{d},\mathbf{d}')$, $\tau \in \vec{T}(X)(\mathbf{c}',\mathbf{c})$ induce continuous maps $\sigma_\sharp : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{c},\mathbf{d}')$ and $\tau^\sharp : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{c}',\mathbf{d})$.

In order to find out what happens "between" $\mathbf{d}$ and $\mathbf{d}'$, one has to study the effect of these induced maps. d-homotopic d-paths yield induced maps that are homotopic to each other and hence it suffices to study the effect of $[\sigma] \in \vec{\pi}_1(X)(\mathbf{d},\mathbf{d}')$ when comparing the homotopy types of $\vec{T}(X)(\mathbf{c},\mathbf{d})$ and of $\vec{T}(X)(\mathbf{c},\mathbf{d}')$.

In categorical terms, cf. Raussen [24], one studies the category $D(X)$ with objects pairs of points $(\mathbf{x},\mathbf{x}')$ such that $\mathbf{T}(X)(\mathbf{x},\mathbf{x}') \neq \emptyset$ and morphisms $\vec{T}(X)(\mathbf{y},\mathbf{x}) \times \vec{T}(X)(\mathbf{x}',\mathbf{y}')$ from $(\mathbf{x},\mathbf{x}')$ to $(\mathbf{y},\mathbf{y}')$ and the functor $\vec{T}^X : \vec{D}(X) \to Ho-Top$. Combining with the homotopy equivalences from Theorem, we wish thus to understand the induced map on the right hand side of the diagram

$$
\begin{array}{ccccc}
(\mathbf{x},\mathbf{x}') & \xrightarrow{\vec{T}^X} & \vec{T}(X)(\mathbf{x},\mathbf{x}') & \xrightarrow{\simeq} & \mathbf{T}(X)(\mathbf{x},\mathbf{x}') \\
\downarrow{\scriptstyle(\sigma,\tau)} & & \downarrow{\scriptstyle(\sigma_\sharp,\tau^\sharp)} & & \downarrow \\
(\mathbf{y},\mathbf{y}') & \xrightarrow{\vec{T}^X} & \mathbf{T}(X)(\mathbf{y},\mathbf{y}') & \xrightarrow{\simeq} & \mathbf{T}(X)(\mathbf{y},\mathbf{y}')
\end{array}
$$

as a prod-simplicial map. Similarly, one may analyse what happens between $\partial_+ \downarrow \mathbf{d}$ and $\partial_+ \downarrow \mathbf{d}'$.

For brevity, we restrict to an anlysis of the concatenation map $\sigma_\sharp : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{c},\mathbf{d}')$. For a discussion, or better, a determination of so-called components in

X (cf Fajstrup et al. [10,14,24]), one would particularly like to know which traces σ induce homotopy equivalences (or at least bijections on sets of path components).

Assume that l of the hyperrectangles intersect $[\mathbf{c},\mathbf{d}]$ whereas $l' \geq l$ intersect $[\mathbf{c},\mathbf{d}']$. In order to compare, we will use the *same* larger index set $[1:l']$ in both cases. A hyperrectangle R^i not intersecting $[\mathbf{c},\mathbf{d}]$ does not pose any conditions to the question $\vec{T}(X_M)(\mathbf{c},\mathbf{d}) \neq \emptyset$ defining index categories; the corresponding ith row in the matrix M is irrelevant. As a result, the index category $\mathscr{C}(X)(\mathbf{c},\mathbf{d}) \subset M^R_{l,n}$ will be replaced by the pullback

$$\begin{array}{ccccc}
\mathscr{C}(X)(\mathbf{c},\mathbf{d}) \times M^R_{l'-l,n} & \cong & \tilde{\mathscr{C}}(X)(\mathbf{c},\mathbf{d}) & \xrightarrow{\pi} & \mathscr{C}(X)(\mathbf{c},\mathbf{d}) \\
 & & \downarrow \subseteq & & \downarrow \subseteq \\
M^R_{l,n} \times M^R_{l'-l,n} & \cong & M^R_{l',n} & \xrightarrow{\pi} & M^R_{l,n}
\end{array}$$

with $\pi : M^R_{l',n} \to M^R_{l,n}$ leaving out superfluous rows.

The pasting scheme corresponding to $\tilde{\mathscr{C}}(X)(\mathbf{c},\mathbf{d})$ gives rise to the prodsimplicial complex $\tilde{\mathbf{T}}(X)(\mathbf{c},\mathbf{d}) = \mathbf{T}(X)(\mathbf{c},\mathbf{d}) \times (\Delta^{n-1})^{l'-l}$ homotopy equivalent to $\mathbf{T}(X)(\mathbf{c},\mathbf{d})$.

The index category $\mathscr{C}(X)(\mathbf{c},\mathbf{d}')$ becomes then a *sub*category of $\tilde{\mathscr{C}}(X)(\mathbf{c},\mathbf{d})$ with certain matrices eliminated; one needs to analyse the effect of the associated *inclusion* of prodsimplicial complexes $\mathbf{T}(X)(\mathbf{c},\mathbf{d}') \hookrightarrow \tilde{\mathbf{T}}(X)(\mathbf{c},\mathbf{d})$.

It is also relevant to ask what happens if one digs an additional forbidden hyperrectangle R out of the state space $X \subset I^n$ to get $X' = X \setminus R$; this is interesting in particular for an inductive determination of index categories and associated prodsimplicial models of trace spaces. Again, the associated map between prodsimplicial models is a combination of a homotopy equivalence (taking the product with a simplex) and an inclusion map reflecting the additional obstruction. The effect of this map (and the maps induced by it on homology etc) have still to be investigated more closely.

4.2 Homotopy equivalences

In the following, we give sufficient conditions making sure that the induced maps become homotopy equivalences. This should be useful for the investigation of component categories, cf Fajstrup, Goubault, Haucourt and Raussen [10,14]:

Proposition 5 *1. Assume that $\vec{T}(X_M)(\mathbf{c},\mathbf{d}) = \emptyset$ implies $\vec{T}(X_M)(\mathbf{c},\mathbf{d}') = \emptyset$ for all matrices $M \in M^R_{l,n}$. Then every trace $\sigma \in \vec{T}(X)(\mathbf{d},\mathbf{d}')$ induces a homotopy equivalence $\sigma_\sharp : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{c},\mathbf{d}')$.*

2. Assume that $\vec{T}(X_M)(\mathbf{c}',\mathbf{d}) = \emptyset \Rightarrow \vec{T}(X_M)(\mathbf{c},\mathbf{d}) = \emptyset$ for all $M \in M^R_{l,n}$. Then every trace $\tau \in \vec{T}(X)(\mathbf{c},\mathbf{c}')$ induces a homotopy equivalence $\tau^\sharp : \vec{T}(X)(\mathbf{c}',\mathbf{d}) \to \vec{T}(X)(\mathbf{c},\mathbf{d})$.

Proof The maps $\sigma_\sharp$, resp. $\tau^\sharp$ induce always inclusions of subcategories $\mathscr{C}(X)(\mathbf{c},\mathbf{d}) \hookrightarrow \mathscr{C}(X)(\mathbf{c},\mathbf{d}')$, resp. $\mathscr{C}(X)(\mathbf{c}',\mathbf{d}) \hookrightarrow \mathscr{C}(X)(\mathbf{c},\mathbf{d})$ within $M^R_{l,n}$. The conditions in Proposition 5 ensure that these are in fact equalities. As a consequence, the prodsimplicial models agree: $\mathbf{T}(X)(\mathbf{c},\mathbf{d}) = \mathbf{T}(X)(\mathbf{c},\mathbf{d}')$, resp. $\mathbf{T}(X)(\mathbf{c}',\mathbf{d}) = \mathbf{T}(X)(\mathbf{c},\mathbf{d})$. The results follow from Theorem 2. □

To check the conditions, one may apply the method from Sect. 3.3.2 to see whether $D(X)(-,-)$ does (not) change under variation of end points.

A different induced map arises, under certain conditions, from taking the least upper bound with some element $\mathbf{e} \in X$: Suppose that

$$\mathbf{e} \vee \mathbf{y} \in X \text{ for all } \mathbf{y} \text{ satisfying } \vec{T}(X)(\mathbf{c},\mathbf{y}) \neq \emptyset \neq \vec{T}(X)(\mathbf{y},\mathbf{d}); \tag{13}$$

ie $\mathbf{y}$ is on a trace connecting $\mathbf{c}$ and $\mathbf{d}$. Then there is an induced map

$$\mathbf{e}\vee : \vec{P}(X)(\mathbf{c},\mathbf{d}) \to \vec{P}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d}),\ (\mathbf{e}\vee p)(t) = \mathbf{e}\vee p(t) \tag{14}$$

inducing the map $\mathbf{e}\vee : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d})$.

Proposition 6 *Assume that* $\mathbf{e} \in \bigcap_{M\in\mathscr{C}(X)(\mathbf{c},\mathbf{d})} X_M$ *and that* $\vec{T}(X_M)(\mathbf{c},\mathbf{d}) = \emptyset \Rightarrow \vec{T}(X_M)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d}) = \emptyset$ *for all* $M \in M^R_{l,n}$.
Then the map $\mathbf{e}\vee : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d})$ *is a homotopy equivalence.*

The proof is analogous to that of Proposition 5; remark that $\mathbf{e}\vee$ has restrictions on the respective trace spaces in X_M, $M \in \mathscr{C}(X)(\mathbf{c},\mathbf{d})$ since these spaces X_M are closed under $\vee$, cf Raussen [27, Lemma 2.6]. Again, the methods from Sect. 3.3.2 may be applied to verify the conditions.

Corollary 1 *Assume that*

1. $\mathbf{c} \leq \mathbf{e} \in \bigcap_{M\in\mathscr{C}(X)(\mathbf{c},\mathbf{d})} X_M$;
2. $\vec{T}(X_M)(\mathbf{c},\mathbf{e}) \neq \emptyset$ *for* $M \in \mathscr{C}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d})$;
3. $\vec{T}(X_M)(\mathbf{c},\mathbf{d}) = \emptyset \Rightarrow \vec{T}(X_M)(\mathbf{c},\mathbf{e}\vee\mathbf{d}) = \emptyset$ *for all* $M \in M^R_{l,n}$.

Then the map $\mathbf{e}\vee : \vec{T}(X)(\mathbf{c},\mathbf{d}) \to \vec{T}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d})$ *is a homotopy equivalence.*

Proof For $M \in \mathscr{C}(X)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d})$ choose $\sigma \in \vec{T}(X_M)(\mathbf{c},\mathbf{e})$. Concatenation $\sigma_\sharp$ with σ shows: $\vec{T}(X_M)(\mathbf{e}\vee\mathbf{c},\mathbf{e}\vee\mathbf{d}) \neq \emptyset \Rightarrow \vec{T}(X_M)(\mathbf{c},\mathbf{e}\vee\mathbf{d}) \neq \emptyset$. Condition (3) above implies that $\vec{T}(X_M)(\mathbf{c},\mathbf{d}) \neq \emptyset$, as well. Apply Proposition 6. □

4.3 Future work

Remark 4 Discussing components following Fajstrup et al. [10,14,24], one would typically apply Proposition 6 to investigate all $\mathbf{c} \leq \mathbf{e}$ for which the map $\mathbf{e}\vee$ from (14) is a homotopy equivalence; and then establish a maximal region such that all such maps $\mathbf{e}\vee$ within it yield homotopy equivalences. Details have still to be worked out.

Remark 5 Connections to topological complexity as discussed by Farber, cf the book [12] and its list of references, should also be interesting to investigate. Taking directions into account makes matters more complicated since the end point map ev_{01} : $\vec{P}(X) \to X \times X$ is no longer a fibration; it is not even surjective. Nevertheless, one may ask for coverings of $\{(\mathbf{x}, \mathbf{y}) \in X \times X | \; \vec{T}(X)(\mathbf{x}, \mathbf{y}) \neq \emptyset\}$ by subsets on which there is a continuous section of the restricted map ev_{01}. It is not difficult to produce such a section on sets of type $X_{j_1,\dots,j_l} \times X_{j_1,\dots,j_l} \subset X \times X$—in the notation of Sect. 2.1.2.

5 A particular case: semaphores of arity one

5.1 Trace spaces are homotopy discrete

Matters become more specific and combinatorial in nature for a semaphore or PV-model (cf Sect. 2.1.1) in which *every* semaphore allows only a *single* process to proceed. In this case, the forbidden region F is the union of hyperrectangles of a particular type:

Whenever two processes $1 \le j_1 < j_2 \le n$ call the same semaphore h on intervals $I_{j_1} =]a^{h,m_1}_{j_1}, b^{h,m_1}_{j_1}[$ and $I_{j_2} =]a^{h,m_2}_{j_2}, b^{h,m_2}_{j_2}[$, a hyperrectangle $R_h(j_1, m_1; j_2, m_2) := I_{j_1} \times I_{j_2} \times I^{n-2}$—with I_{j_i} inserted as factor j_i—is added to the forbidden region; remark that all but two subintervals correspond to the full interval I.

Let us assume that semaphore h, $1 \le h \le k$, is called upon by the processes labelled by the subset $J_h \subseteq [1 : n]$. For every $j \in J_h$, there is a number r_{hj} of calls Ph by process j, on an interval $]a^{h,m}_j, b^{h,m}_j[$, $1 \le m \le r_{hj}$. A hyperrectangle $R_h(j_1, m_1; j_2, m_2)$ arises for every semaphore h, every pair $j_1 < j_2$, $j_i \in J_h$ and every pair of calls corresponding to the two processes j_1, j_2. The total number l of forbidden hyperrectangles is thus

$$l = \sum_{h=1}^{k} \sum_{j_1 < j_2 \in J_h} r_{hj_1} r_{hj_2}. \tag{15}$$

Example 1 For $k \ge 2$ dining philosophers, cf Dijkstra [6] and Fig. 2 below, every semaphore (=fork) is called upon once by exactly two processes (philosophers). Hence $l = k$, and F is the union of k hyperrectangles R^i.

In the following, we will stick to endpoints $\mathbf{c} = \mathbf{0}$, $\mathbf{d} = \mathbf{1}$—both for simplicity, and because this is the most interesting case. Apart from the sets of binary matrices introduced in Sect. 2, we also need the set $M^1_{l,n} \subset M^R_{l,n}$ consisting of matrices in which *every row* vector is a *unit* vector. Remark that every contributing hyperrectangles R^i, $1 \le i \le l$ is of type $R_h(j_1(i), k_1; j_2(i), k_2)$ described above.

Proposition 7 *Let* $M \in M^R_{l,n}$.

1. *If* $\vec{T}(X_M)(\mathbf{0}, \mathbf{1})$ *is non-empty, then* $M \in M^1_{l,n}$.

2. $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ *is homotopy equivalent to a finite discrete space; its (contractible) connected components are the non-empty ones among the spaces* $\vec{T}(X_M)(\mathbf{0}, \mathbf{1})$ *with* $M \in M^1_{l,n}$.

Proof 1. Suppose $M \notin M^1_{l,n}$, ie $m_{i,j_1} = m_{i,j_2} = 1$ for some $1 \leq i \leq l$; $1 \leq j_1 < j_2 \leq n$.
If, say, $j_1 \notin \{j_1(i), j_2(i)\}$, then $a^i_{j_1} = 0$ and hence trace space is empty by Raussen [27, Proposition 4.3(1)] or by Lemma 1 from this paper.
If $\{j_1, j_2\} = \{j_1(i), j_2(i)\}$, we define a map $i : [1 : n] \to [0 : l]$ with $i(j_1) = i(j_2) = i$ and $i(j) = 0$ for all other j. We check that condition (4) in Proposition 3 is satisfied: $a^i_{j_1} < b^i_{j_1}, a^i_{j_2} < b^i_{j_2}$, and $b^i_j = 1$ for $j_1 \neq j \neq j_2$; hence trace space is empty also in this case.
2. It follows from (1), that the subposet category $\mathscr{C}(X)(\mathbf{0}, \mathbf{1})$ has no non-trivial morphisms, and hence that the prodsimplicial complex $\mathbf{T}(X)(\mathbf{0}, \mathbf{1})$ has dimension zero. □

It remains thus to determine for *which* $M \in M^1_{l,n}$ the spaces $\vec{T}(X_M)(\mathbf{0}, \mathbf{1})$ are *non-empty*. Remark that there are 2^l such spaces with l as in (15)—for every i corresponding to a pair of calls, one may choose either $j_1(i)$ or $j_2(i)$.

5.2 A single call to a semaphore of arity one

Let us first consider just a *single* concurrent call to a semaphore of arity one. Without restriction of generality, we assume that all n processes call to it. If only $m < n$ processes call the semaphore, then the forbidden region has type $F = F_m \times I^{n-m}$. Hence the state space is $X = X_m \times I^{n-m}$ with $X_m = I^m \setminus F^m$, and $\vec{T}(X)(\mathbf{0}, \mathbf{1}) \simeq \vec{T}(X_m)(\mathbf{0}, \mathbf{1})$.

Assume that the semaphore calls are given by intervals $]a_j, b_j[\subset [0, 1], 1 \leq j \leq n$. The associated forbidden region is the union $F = \bigcup_{1\leq j_1<j_2\leq n} R(j_1, j_2)$ of the $\binom{n}{2}$ hyperrectangles $R(j_1, j_2) = \{\mathbf{x} \in I^n |\ a_{j_i} < x_{j_i} < b_{j_i}, i = 1, 2\}$. As usual, let $X = I^n \setminus F$.

In the proof of the next result, we will also need the extended hyperrectangles $R_{j_1}(j_1, j_2) = \{\mathbf{x} \in I^n |\ 0 \leq x_{j_2} < b_{j_2}, a_{j_1} < x_{j_1} < b_{j_1}\}$ and likewise $R_{j_2}(j_1, j_2)$; moreover, as in Sect. 3.2, the degenerate hyperrectangles $R^0_j = [0, 1]^{j-1} \times \{1\} \times [0, 1]^{n-j}$ for $1 \leq j \leq n$.

Proposition 8 *Trace space* $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ *is homotopy equivalent to the discrete space whose underlying set is the symmetric group* Σ_n.
A homotopy equivalence $\vec{x} : \Sigma_n \to \vec{T}(X)(\mathbf{0}, \mathbf{1})$ *is given by*

$$\vec{x}(\pi)(t) = (x_1(t), \ldots, x_n(t)) \text{ with } x_{\pi(k)}(t) = \begin{cases} 0 & t \leq \frac{k-1}{n} \\ (nt - (k-1)) & \frac{k-1}{n} \leq t \leq \frac{k}{n} \\ 1 & \frac{k}{n} \leq t \leq n \end{cases}.$$

Proof Note that every $\vec{x}(\pi)$ describes a d-path on the 1-skeleton of $\vec{I}^n$ that does not intersect the forbidden region F; these are in fact *all* d-paths on the 1-skeleton up to trace equivalence.

Let $\mathscr{P}_2(n)$ denote the set of all 2-element subsets of $[1:n]$ (with $l = \frac{n(n-1)}{2}$ elements indexing the obstruction hyperrectangles), and let $c : \mathscr{P}_2(n) \to [1:n]$ denote a *choice* function with the property $c(\{j_1, j_2\}) \in \{j_1, j_2\}$. For such a choice function c—determining in which order to pass the obstructions $R(j_1, j_2)$—let $F_c = \bigcup_{1\leq j_1<j_2\leq n} R_{c(j_1,j_2)}(j_1, j_2)$ and $X_c = I^n \setminus F_c$. By Theorem 2 and Proposition 7, the (contractible) components of $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ correspond to those choice functions c giving rise to *non-empty* trace spaces $\vec{T}(X_c)(\mathbf{0}, \mathbf{1})$; more precisely, c corresponds to the matrix $M_c \in M^1_{\binom{n}{2},n}$—every row a unit vector—with $m_{(j_1,j_2),c(j_1,j_2)} = 1$.

A choice function c gives rise to a relation on the set $[1:n]$ defined by $j_1 \leq_c j_2$ if $c(j_1, j_2) = j_1$ and its reflexive and transitive closure $\preceq_c$. If $\preceq_c$ defines a *total order* on $[1:n]$, then this total order is given by a permutation $\pi \in \Sigma_n$: $\pi(1) \preceq_c \pi(2) \preceq_c \cdots \preceq_c \pi(n)$. On the other hand, every permutation $\pi \in \Sigma_n$ orders the elements of $[1:n]$ and gives thus rise to a choice function: $c(j_1, j_2) = 1 \Leftrightarrow \pi^{-1}(j_1) < \pi^{-1}(j_2)$. We claim: $\vec{T}(X_c)(\mathbf{0}, \mathbf{1}) \neq \emptyset$ if and only if $\preceq_c$ is a total order.

If $\preceq_c$ is *not* a total order, then there is a chain $j_1 \preceq_c \cdots \preceq_c j_k \preceq_c j_1$ with $k < n$; let $j_{k+1}, \ldots, j_n$ denote the remaining elements of $[1:n]$. The extended, resp. degenerate hyperrectangles $R_{j_1}(j_1, j_2), \ldots, R_{j_k}(j_k, j_1), R^0_{j_{k+1}}, \ldots, R^0_{j_n}$ have a non-empty intersection $\{\mathbf{x} \in I^n \mid a_{j_i} < x_{j_i} < b_{j_i}, i \leq k; x_{j_i} = 1, i > k\}$ giving rise to a deadlock $\mathbf{x} = [x_1, \ldots, x_n]$ with $x_{j_i} = a_{j_i}, i \leq k; x_{j_i} = 1, i > k$. Hence $\vec{T}(X_c)(\mathbf{0}, \mathbf{1}) = \emptyset$.

Suppose now that $\preceq_c$ *is* a total order. No non-degenerate hyperrectangle contributes with an index corresponding to the direction that is maximal under $\preceq_c$. We claim that for every choice of n—one for every direction $j \in [1:n]$—among the extended and (at least one) degenerate hyperrectangles $R_{j_p}(j_p, j_q)$, $p \preceq_c q$ and $R^0_{j_r}$, their intersection has to be empty: Since there is no loop with respect to $\preceq_c$, the union of all $\{j_p, j_q\}$ corresponding to extended hyperrectangles has at least one element j in common with the set $\{j_r\}$ corresponding to degenerate hyperrectangles.

An element $\mathbf{x} \in \bigcap R_{j_p}(j_p, j_q) \cap \bigcap R^0_{j_r}$ would have to satisfy both $x_j < b_j$ and $x_j = 1$. Hence, all these intersections are empty, there are no deadlocks in X_c, and trace space $\vec{T}(X_c)(\mathbf{0}, \mathbf{1})$ is non-empty. □

To fix notation, we let to every permutation $\pi \in \Sigma_n$ correspond

- the $\binom{n}{2}$ extended hyperrectangles $R_\pi(j_1, j_2) = R_j(j_1, j_2)$ with $j = j_1$ if $\pi(j_1) < \pi(j_2)$ and $j = j_2$ else;
- the forbidden region $F_\pi = \bigcup_{1\leq j_1<j_2\leq n} R_\pi(j_1, j_2)$; and
- the state space $X_\pi = I^n \setminus F_\pi$.

Proposition 8 can be reformulated as follows:

Corollary 2 *The trace space is a disjoint union* $\vec{T}(X)(\mathbf{0}, \mathbf{1}) = \bigsqcup_{\pi\in\Sigma_n} \vec{T}(X_\pi)(\mathbf{0}, \mathbf{1})$. *All components* $\vec{T}(X_\pi)(\mathbf{0}, \mathbf{1})$, $\pi \in \Sigma_n$, *are contractible.*

Intuitively, a component $\vec{T}(X_\pi)(\mathbf{0}, \mathbf{1})$, $\pi \in \Sigma_n$ consists of all interleaving d-paths that access the semaphore in the order given by the permutation π.

5.3 Several calls to semaphores of arity one

Let us now consider a state space corresponding to a collection of k semaphores of arity one. Every semaphore h is called upon by a subset $J_h \subseteq [1:n]$ of processes, and this inclusion induces an inclusion $\Sigma_{J_h} \subseteq \Sigma_n$ of permutation groups: Σ_{J_h} is the stabilizer of $[1:n] \setminus J_h$.

Suppose that semaphore h is locked by process $j \in J_h$ at intervals $]a_j^{m_j(h)}, b_j^{m_j(h)}[$ with $1 \leq m_j(h) \leq r_j(h)$. A concurrent call c consists of the choice of a semaphore $h = h(c)$ with $1 \leq h \leq k$ and an unordered $|J_h|$-tuple $(m_j(h))_{j \in J_h}$, $1 \leq m_j(h) \leq r_j(h)$. It gives rise to a forbidden region $F(c) = F(h; m_1(h), \ldots, m_{|J_h|}(h))$ determined as in Sect. 5.2. That forbidden region has extensions $F_\pi(c)$, one for every permutation $\pi \in \Sigma_{J_h}$.

Let C denote the set of calls, ie of tuples of the form $(h; m_1(h), \ldots, m_{|J_h|}(h))$. The total forbidden region is given by

$$F = \bigcup_{c \in C} F(c) = \bigcup_{1 \leq h \leq k} \bigcup_{j \in J_h; 1 \leq m_j(h) \leq r_j(h)} F(h; m_1(h), \ldots, m_{|J_h|}(h)),$$

and the state space is $X = I^n \setminus F$. We need to consider one permutation per call, i.e., elements $\boldsymbol{\pi} = (\pi(c))_{c \in C}$ in the product $\Sigma = \prod_{c \in C} \Sigma_{J_h}$. Every such permutation determines in which order the processes pass semaphore h under call c.

Proposition 9 *Trace space is a disjoint union* $\vec{T}(X)(\mathbf{0}, \mathbf{1}) = \bigsqcup_{\boldsymbol{\pi} \in \Sigma} \vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0}, \mathbf{1})$. *Each subspace* $\vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0}, \mathbf{1})$ *is either empty or contractible.*

Proof According to Proposition 7, $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ is homotopy equivalent to a disjoint union of spaces of the form $\vec{T}(X_M)(\mathbf{0}, \mathbf{1})$, $M \in M_{l,n}^R$; each of those is either empty or contractible. By Proposition 8 and Corollary 2, only matrices $M \in M_{l,n}^R$ arising from collections of *permutations* can give rise to non-empty trace spaces. □

It remains to study, for *which* collections of permutations $\boldsymbol{\pi} = (\pi_c)_{c \in C} \in \Sigma = \prod_{c \in C} \Sigma_{J_{h(c)}}$ the space $\vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0}, \mathbf{1})$ is non-empty: Consider the set of boundary coordinates $a_j^i, b_j^i \in I$, $1 \leq j \leq n$, $i \in \bigcup_{j \in J_h} \{m_j(h)\}$, of *all* calls to semaphores. For every collection $\boldsymbol{\pi} = (\pi_c)_{c \in C} \in \Sigma = \prod_{c \in C} \Sigma_{J_{h(c)}}$, we consider several order relations on subsets of these real numbers:

- The natural order $\leq$, inherited from the reals, on numbers a_j^i, b_j^i with the *same* subscript (direction) j;
- $b_{\pi_c(j)}^{m_{\pi_c(j)}(h)} \preceq a_{\pi_c(j')}^{m_{\pi_c(j')}(h)}$ for $c \in C$, $j < j' \in J_{h(c)}$ for the *same* call $c = (h; m_1(h), \ldots, m_{|J_h|}(h)) \in C$.

We call the collection $\boldsymbol{\pi}$ *compatible* if the transitive hull $\sqsubseteq_{\boldsymbol{\pi}}$ of these relations is a *partial order*.

Proposition 10 *Let* $X = I^n \setminus F$ *denote the state space corresponding to a collection of* k *semaphores of arity one. Then* $\vec{T}(X)(\mathbf{0}, \mathbf{1})$ *is homotopy equivalent to the discrete space*

$\{\boldsymbol{\pi} = (\pi_c)_{c\in C} \in \prod_{c\in C} \Sigma_{J_{h(c)}} \mid \boldsymbol{\pi}\ compatible\} \subseteq \prod_{c\in C} \Sigma_{J_{h(c)}} \subseteq (\Sigma_n)^l.$

Proof We need to show: $\vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0},\mathbf{1}) \neq \emptyset$ if and only if $\boldsymbol{\pi}$ is compatible.

Assume first that $\vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0},\mathbf{1}) \neq \emptyset$. Any d-path $p : \vec{I} \to X_{\boldsymbol{\pi}}$ from $\mathbf{0}$ to $\mathbf{1}$ leads to a *total order* of the boundary coordinates $c^i_j = a^i_j, b^i_j$ given by

$$t_1 \le t_2 \quad \text{and} \quad x_{j_1}(t_1) = c^{i_1}_{j_1}, x_{j_2}(t_2) = d^{i_2}_{j_2} \Rightarrow c^{i_1}_{j_1} \le d^{i_2}_{j_2}$$

compatible with the relations defining $\sqsubseteq_{\boldsymbol{\pi}}$. In particular, $\sqsubseteq_{\boldsymbol{\pi}}$ is a partial order.

Now assume that $\vec{T}(X_{\boldsymbol{\pi}})(\mathbf{0},\mathbf{1}) = \emptyset$, i.e., the forbidden hyperrectangles give rise to a deadlock. A deadlock arises as lower corner of a non-empty intersection of $m \le n$ hyperrectangles among the $R_{\pi(c)}(j_1, j_2)$, and $n-m$ among the degenerate hyperrectangles R^0_j, $1 \le j \le n$; it is enough to consider the case of a non-empty intersection of n extended hyperrectangles:

In fact, we consider n hyperrectangles of type $R_j :=]a^{s(j)}_j, b^{s(j)}_j[\times[0, c^{s(j)}_{r(j)}[\times[0,1]^{n-2}$ where the first named interval is on the jth coordinate axis and the second named on an axis in position $r(j) \neq j$; the superscript $s(j)$ refers to one of the original hyperrectangles. The map $r : [1:n] \to [1:n]$ need not be onto; it has no fixed point: $r(j) \neq j$ for all $1 \le j \le n$. Hence, there is a cycle $k, r(k), \dots, r^{q-1}(k), r^q(k) = k$ for some $1 \le k \le n$; of length $2 \le q \le n$. We consider only indices $j = r^i(k)$ within such a cycle, and on these, r is a bijection.

By definition, we have that $c^{s(j)}_{r(j)} \preceq a^{s(j)}_j$. Moreover, since $(a^{s(1)}_1, \dots, a^{s(n)}_n)$ is a deadlock, $a^{s(j)}_j \le c^{s(j')}_j$ with $r(j') = j$. Hence, with subscripts restricted to the cycle under r, every a-coordinate is less than ($\le$) at least one of the c-coordinates, and every c-coordinate precedes ($\preceq$) one of the a-coordinates. This gives rise to a cycle under the relation $\sqsubseteq_{\boldsymbol{\pi}}$ which therefore cannot be a partial order. □

Remark 6 1. For an illustration of the last step in the proof above, consider Fig. 1, row two, case two, leading to a deadlock of extended rectangles of the form $R_1 =]a^1_1, b^1_1[\times[0, c^1_2[$, $R_2 = [0, c^2_1[\times]a^2_2, b^2_2[$ and to a cycle $c^1_2 \preceq a^1_1 \le c^2_1 \preceq a^2_2 \le c^1_2$.

2. The relation $\sqsubseteq_{\boldsymbol{\pi}}$ generated by $\le$ and by $\preceq$ defines a digraph $G_{\boldsymbol{\pi}}$ with the boundary coordinates a^i_j, b^i_j as vertices; the k-tuple $\boldsymbol{\pi}$ is compatible if and only if $G_{\boldsymbol{\pi}}$ does not contain a directed cycle. This can be checked algorithmically by a depth-first search as in Tarjan's strongly connected component algorithm [29]. In general, there are exponentially many collections of permutations to check; the existence of a particular cycle may decide for several collections at a time.

Example 2 Let $X_k \subset I^k$ denote the state space corresponding to the PV-model describing k dining philosophers (each protocol of type $PaPbVaVb$; cf Example 1, Fig. 2 and Dijkstra [6]). Then only two out of 2^k permutations (those in the "diagonal"—all elements the identity or all the nontrivial transposition) in $\Sigma^k_2 \subset (\Sigma^k_k)$ lead to a relation with a non-trivial cycle under the relation $\sqsubseteq_{\boldsymbol{\pi}}$; all others give rise to partial orders. As a consequence, $\vec{T}(X_k)(\mathbf{0},\mathbf{1})$ consists of $2^k - 2$ contractible components: There are $2^k - 2$ essentially different interleavings of the d-paths corresponding to each individual protocol—indicating who of the two neighbouring philosophers uses a fork first. The number $2^k - 2$ of schedules is, for $k > 3$, considerably smaller than the number $k!$ of ordered k-tuples of philosophers. This is due to the fact that several philosophers can serve themselves concurrently for $k > 3$.

6 The complement of a trace space

The aim of this last section is to describe a combinatorial method that yields a prodsimplicial model of the *complement* of $\mathbf{T}(X)(-, -)$ in a product $(S^{n-2})^l$ of spheres. Duality allows to obtain information about $\mathbf{T}(X)(-, -)$ and thus $\vec{T}(X)(-, -)$ itself; cf Remark 7. The advantage of this method is, that it is far easier to determine the poset category describing the homotopy type of the complement that that of trace space itself—by upward completion of the set $D(X)(-, -)$ in the category underlying the product of spheres. It is hoped that this will also make implementations easier and faster.

6.1 A combinatorial description of the complement of trace space

For simplicity, we restrict to traces from $\mathbf{0}$ to $\mathbf{1}$. First some notation regarding matrices and associated poset categories: Let $\tilde{M}^R_{l,n} \subset M^R_{l,n} \subset M_{l,n}$ consist of the binary matrices such that *every row vector contains at least one 0 and at least one 1*. This subset is a sublattice of the Boolean lattice $M_{l,n}$ (with least upper bound, greatest lower bound, and coordinatewise involution I switching 0s and 1s). It has the matrices in which every row vector contains *exactly one* 1, resp *exactly one* 0 as minimal, resp. maximal elements. All elements can be written as least upper bounds of minimal matrices and (hence!) as greatest lower bounds of maximal matrices.

In Raussen [27, Section 4.2], we introduced the subset $D(X)(\mathbf{0}, \mathbf{1}) := \{M \in M^C_{l,n} | \Psi(M) = 1\} \subset M_{l,n}$. Remark that, by Raussen [27, Lemma 3.3], matrices in $D(X)(\mathbf{0}, \mathbf{1})$ containing a row with only 1s can and will be neglected right away; we let $\tilde{D}(X)(\mathbf{0}, \mathbf{1}) \subseteq M^R_{l,n}$ consist of those matrices in $D(X)(\mathbf{0}, \mathbf{1})$ without a row vector consisting of 1s only.

We define an upward completion of the matrix set $\tilde{D}(X)(\mathbf{0}, \mathbf{1})$ within $\tilde{M}^R_{l,n}$ as follows:

$$\bar{D}(X)(\mathbf{0}, \mathbf{1}) := \{M \in \tilde{M}^R_{l,n} | \; \exists N \in \tilde{D}(X)(\mathbf{0}, \mathbf{1}) : \; N \leq M\}. \tag{16}$$

Obviously, this completed matrix set $\bar{D}(X)(\mathbf{0}, \mathbf{1})$ forms an *upward* closed subcategory (with respect to $\leq$) of the poset category $\tilde{M}^R_{l,n}$. Reversing the arrows (using $\geq$ instead of $\leq$ as partial order) yields a *downward* closed subcategory $\bar{D}(X)(\mathbf{0}, \mathbf{1})^{op} \subset (M^R_{l,n})^{op}$.

By Raussen [27, Lemma 3.3], the prodsimplicial complex $\mathbf{T}(X)(\mathbf{0}, \mathbf{1})$ is contained in the complex $(\partial\Delta^{n-1})^l$—a product of $(n-1)$-spheres. Let $U(X)(\mathbf{0}, \mathbf{1}) := (\partial\Delta^{n-1})^l \setminus \mathbf{T}(X)(\mathbf{0}, \mathbf{1})$ denote its *complement* within the latter; this is an open set, which does not have a (prod-)simplicial structure right away.

But it turns out that $U(X)(\mathbf{0},\mathbf{1})$ is homotopy equivalent to a prodsimplicial complex with a pasting scheme construction analogous to (6): To every matrix $M \in \bar{D}(X)(\mathbf{0},\mathbf{1})$ we associate the product of simplices $\bar{\Delta}(M) = \prod_{i=1}^{l} \bar{\Delta}_i(M) \subset (\Delta^{n-1})^l \subset \mathbf{R}^{nl}$ with

$$\bar{\Delta}_i(M) := \left\{ (t_1,\ldots,t_n) |\ 0 \le t_j \le 1, \sum_{j=1}^{n} t_j = 1, m_{ij} = 1 \Rightarrow t_j = 0 \right\} \subset \Delta^{n-1}.$$

This time, $M \ge N$ implies $\Delta(M) \subseteq \Delta(N)$. Now, we can define a prod-simplicial complex

$$\mathbf{U}(X)(\mathbf{0},\mathbf{1}) := \bigcup_{M \in \bar{D}(X)(\mathbf{0},\mathbf{1})} \bar{\Delta}(M). \tag{17}$$

Remark 7 As was pointed out in Raussen [27] for trace space, this is a colimit construction: $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$ is the colimit of a functor $\mathscr{F}_n^l : \bar{D}(X)(\mathbf{0},\mathbf{1})^{op} \to \mathbf{Top}$ where

$$\mathscr{F}_n^l : \bar{D}(X)(\mathbf{0},\mathbf{1})^{op} \hookrightarrow (\tilde{M}_{l,n}^R)^{op} \xrightarrow{I} \tilde{M}_{l,n}^R \hookrightarrow M_{l,n}^R \xrightarrow{\mathscr{E}_n^l} \mathbf{Top}$$

with the involution I introduced above and the pasting functor $\mathscr{E}_n^l$ from Raussen [27, Section 3.1.2].

For a binary matrix $M \in M_{l,n}$, let $o(M)$ denote the number of rows that are zero vectors. Let $r(M) = \sum_{1\le i\le l, 1\le j\le n} m_{ij} + o(M)$. Remark that $r(M) \ge l$ for $M \in M_{l,n}^R$.

Lemma 3 *The complex* $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$ *has dimension* $\dim \mathbf{U}(X)(\mathbf{0},\mathbf{1}) = l(n-1) - \min_{M \in \tilde{D}(X)(\mathbf{0},\mathbf{1})} r(M) \le l(n-2)$.

Proof A vertex corresponds to a matrix with $l(n-1)$ ones. A k-face corresponds to a matrix with k additional zeros; zero rows are not considered in $\tilde{M}_{l,n}^R$. □

6.2 A homotopy equivalence with several consequences

The proof of the following result is to be found at the end of this article:

Theorem 3 *The complement* $U(X)(\mathbf{0},\mathbf{1})$ *is homotopy equivalent to the prodsimplicial complex* $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$.

Remark 8 If one knows the homotopy type of $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$ or has calculations of its homology, respectively cohomology at hand, one may apply Poincaré–Alexander–Lefschetz duality (cf eg Bredon [2, Theorem VI.8.3]) in order to obtain information about $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ and hence about $\vec{T}(X)(\mathbf{0},\mathbf{1})$:

$$H^p(\mathbf{T}(X)(\mathbf{0},\mathbf{1})) \cong H_{(n-2)l-p}((S^{n-2})^l, \mathbf{U}(X)(\mathbf{0},\mathbf{1}));$$

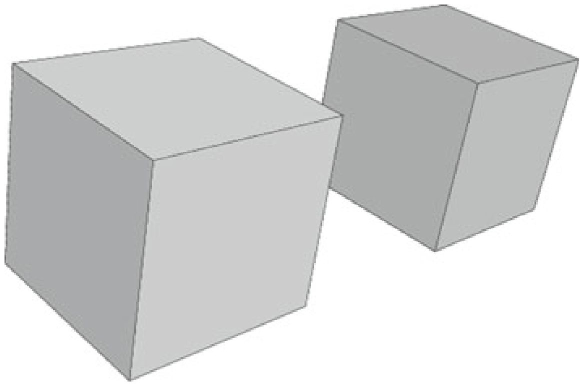


Fig. 3 Forbidden region within I^3 consisting of two adjacent cubes; figure courtesy to Lang [21]

and, using a compact deformation retract of $U(X)(\mathbf{0},\mathbf{1})$,

$$H^p(\mathbf{U}(X)(\mathbf{0},\mathbf{1})) \cong H_{(n-2)l-p}((S^{n-2})^l, \mathbf{T}(X)(\mathbf{0},\mathbf{1})).$$

From the exact homology sequence of the pair $((S^{n-2})^l, \mathbf{U}(X)(\mathbf{0},\mathbf{1}))$ we conclude in particular:

$$\tilde{H}^0(\overrightarrow{T}(X)(\mathbf{0},\mathbf{1})) \cong H_{(n-2)l-1}(\mathbf{U}(X)(\mathbf{0},\mathbf{1}))$$

since, for $\overrightarrow{T}(X)(\mathbf{0},\mathbf{1}) \neq \emptyset$, the inclusion $\mathbf{U}(X)(\mathbf{0},\mathbf{1}) \hookrightarrow (S^{n-2})^l$ is not onto; and

$$H^p(\overrightarrow{T}(X)(\mathbf{0},\mathbf{1})) \cong H_{(n-2)l-p-1}(\mathbf{U}(X)(\mathbf{0},\mathbf{1})) \quad \text{for } p \not\equiv 0, 1 \bmod n-2.$$

In the remaining cases, one has to study the maps induced on $(n-2)$-dimensional homology by the components $i_k : U \to S^{n-2}$ of inclusion $U \hookrightarrow (S^{n-2})^l$. The maps i_k are geometric realizations corresponding to the functors $\bar{D}(X)(\mathbf{0},\mathbf{1}) \hookrightarrow \tilde{M}^R_{l,n} \downarrow \tilde{M}^R_{1,n}$; the last map projects a matrix to its kth row.

For $n = 2$, both $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ and $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$ are (discrete) complements within $(S^0)^l$; in particular:

$$|\mathbf{T}(X)(\mathbf{0},\mathbf{1})| + |\mathbf{U}(X)(\mathbf{0},\mathbf{1})| = 2^l.$$

Example 3 Let X denote the complement of two "adjacent" cubes in I^3; cf Fig. 3.

Using Raussen [27, Lemma 4.6], it is not difficult to see that in this case

$$\tilde{D}(X)(\mathbf{0},\mathbf{1}) = \left\{ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \right\}$$

with completion given by

$$\bar{D}(X)(\mathbf{0},\mathbf{1}) = \tilde{D}(X)(\mathbf{0},\mathbf{1}) \cup \left\{ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \right\}.$$

The three matrices in $\tilde{D}(X)(\mathbf{0},\mathbf{1})$ correspond to the *edges*, the last four matrices to the *vertices* of the (prod)simplicial complex $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$. In this case, that $1=(4-3)$-dimensional complex is just an interval (with four vertices and three edges joining them), which is clearly contractible.

In this case, the complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$—which is homotopy equivalent to trace space $\vec{T}(X)(\mathbf{0},\mathbf{1})$—is thus homotopy equivalent to the complement of a contractible set within the 2-dimensional torus $(\partial\Delta^2)^2$, i.e., $\vec{T}(X)(\mathbf{0},\mathbf{1})\simeq\mathbf{T}(X)(\mathbf{0},\mathbf{1})\simeq S^1\vee S^1$.

Proof (of Theorem 3:) It is folklore that the complement of a subcomplex of $\partial\Delta^{n-1}$ can be given a dual simplicial structure; cf Björner and Tancer [1] for a combinatorial proof. We outline a proof using the nerve lemma for subcomplexes of products of simplicial spheres. To this end, we define a cover of the space $U(X)(\mathbf{0},\mathbf{1})$ by open contractible subspaces with contractible or empty intersections and compare the space with the nerve of that cover.

The sphere $\partial\Delta^{n-1}$ can be covered by contractible open neighbourhoods $U(f_j)$ of its $(n-2)$-facets f_j; it is not difficult to write down contractions to the barycenter of the facet that respect the *complements* of subsimplices. The product $(\partial\Delta^{n-1})^l$ is then covered by the contractible open sets $\prod_{i=1}^{l}U(f_{j_i}^i)$ with contractions respecting complements of products of subsimplices. Intersections of neighbourhoods are neighbourhoods of lower dimensional faces, also contractible to their barycenters with contractions respecting complements of subsimplices and their products.

This has consequences for the space $U(X)(\mathbf{0},\mathbf{1})$, the complement of the simplicial complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ in $(\partial\Delta^{n-1})^l$: First of all, it is covered by the open sets $\prod_{i=1}^{l}U(f_{j_i}^i)$ corresponding to those matrices $M\in\bar{D}(X)(\mathbf{0},\mathbf{1})$ with exactly one 0 in each row corresponding to the vertices of $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$. Moreover, for such a matrix M with $m_{i,j_i}=0$, the set $\prod_{i=1}^{l}U(f_{j_i}^i)\cap U(X)(\mathbf{0},\mathbf{1})$ is (non-empty and) contractible, since the contraction respects the complement of the simplicial complex $T(X)(\mathbf{0},\mathbf{1})$. The same argument holds also for the non-empty intersections with $U(X)(\mathbf{0},\mathbf{1})$ corresponding to matrices in $\bar{D}(X)(\mathbf{0},\mathbf{1})$ with additional zeroes.

Finally, one can argue as in the proof of Raussen [27, Theorem 3.5]: Both spaces $U(X)(\mathbf{0},\mathbf{1})$ and $\mathbf{U}(X)(\mathbf{0},\mathbf{1})$ are colimits of functors over $\bar{D}(X)(\mathbf{0},\mathbf{1})^{op}$ homotopy equivalent to the homotopy colimits of these functors; since the functor takes contractible values everywhere, those are in turn homotopy equivalent to the nerve of the category $\bar{D}(X)(\mathbf{0},\mathbf{1})^{op}$. □

Remark 9 Björner and Tancer [1] have given an entirely combinatorial proof for an analogue of Theorem 3 in the case of the complement of a simplicial complex *within a sphere*. I believe it should be possible to give a smilar argument for the complement in a product of spheres.

Acknowledgments I wish to thank the referee for several suggestions improving the presentation and for a pointer to Tarjan's algorithm, cf Remark 6.

References

1. Björner, A., Tancer, M.: Note: combinatorial Alexander duality—a short and elementary proof. Discret. Comput. Geom. **42**(4), 586–593 (2009)

2. Bredon, G.: Topology and Geometry, GTM, vol. 139. Springer, New York (1993)
3. Brown, R., Higgins, P.: Colimit theorems for relative homotopy groups. J. Pure Appl. Algebra **22**, 11–41 (1981)
4. Brown, R., Higgins, P.: On the algebra of cubes. J. Pure Appl. Algebra **21**, 233–260 (1981)
5. Dijkstra, E.: Co-operating sequential processes. In: Genuys, F. (ed.) Programming Languages, pp. 43–110. Academic Press, New York (1968)
6. Dijkstra, E.: Hierarchical ordering of sequential processes. Acta Informatica **1**(2), 115–138 (1971)
7. Fahrenberg, U., Raussen, M.: Reparametrizations of continuous paths. J. Homotopy Relat. Struct. **2**(2), 93–117 (2007)
8. Fajstrup, L.: Trace spaces of directed tori with rectangular holes. Tech. Rep. R-2011-08, Department of Mathematical Sciences, Aalborg University (2011)
9. Fajstrup, L., Goubault, E., Haucourt, E., Mimram, S., Raussen, M.: Trace Spaces: an efficient new technique for State-Space Reduction. In: Seidl, H. (ed.) Programming Languages and Systems, Lecture Notes in Computer Science, vol. 7211. 21st European Symposium on Programming, ESOP 2012, Springer-Verlag (2012)
10. Fajstrup, L., Goubault, E., Haucourt, E., Raussen, M.: Components of the fundamental category. Appl. Categ. Struct. **12**(1), 81–108 (2004)
11. Fajstrup, L., Goubault, E., Raussen, M.: Detecting Deadlocks in Concurrent Systems. In: Sangiorgi, D., de Simone, R. (eds.) Proceedings of 9th International Conference on Concurrency Theory (CONCUR '98). Lecture Notes in Computer Science, vol. 1466, pp. 332–347. Springer, Nice, France (1998)
12. Farber, M.: Invitation to Topological Robotics. Zurich Lectures in Advanced Mathematics. EMS Publishing House (2008)
13. Goubault, E., Haucourt, E.: A Practical Application of Geometric Semantics to Static Analysis of Concurrent Programs. In: Abadi, M., de Alfaro, L. (eds.), CONCUR, Lecture Notes in Computer Science, vol. 3653, pp. 503–517. Springer, New York (2005)
14. Goubault, E., Haucourt, E.: Components of the fundamental category II. Appl. Categ. Struct. **15**, 387–414 (2007)
15. Grandis, M.: Directed Homotopy Theory I. The Fundamental Category. Cah. Topol. Géom. Différ. Catég. **44**, 281–316 (2003)
16. Grandis, M.: Directed Algebraic Topology. New Mathematical Monographs. Cambridge University Press, Cambridge (2009)
17. Gunawardena, J.: Homotopy and concurrency. Bull. EATCS **54**, 184–193 (1994)
18. Herlihy, M., Rajsbaum, S.: Algebraic topology and distributed computing—a primer. In: Computer Science Today, Lecture Notes in Comput. Sci., vol. 1000, pp. 203–217, Springer, New York (1995)
19. Kaczynski, T., Mrozek, M., Slusarek, M.: Homology computation by reduction of chain complexes. Comput. Math. Appl. **35**(4), 59–70 (1998)
20. Kozlov, D.: Combinatorial Algebraic Topology, Algorithms and Computation in Mathematics, vol. 21, Springer, New York (2008)
21. Lang, A.: Paths in directed spaces. Tech. rep., Ecole Polytechique Paris, CEA LIST/LMeASI (2010)
22. Pratt, V.: Modelling concurrency with geometry. In: Proceedings of the 18th ACM Symposium on Principles of Programming Languages, pp. 311–322 (1991)
23. Raussen, M.: On the classification of dipaths in geometric models for concurrency. Math. Struct. Comput. Sci. **10**(4), 427–457 (2000)
24. Raussen, M.: Invariants of directed spaces. Appl. Categ. Struct. **15**, 355–386 (2007)
25. Raussen, M.: Reparametrizations with given stop data. J. Homotopy Relat. Struct. **4**(1), 1–5 (2009)
26. Raussen, M.: Trace spaces in a pre-cubical complex. Topol. Appl. **156**(9), 1718–1728 (2009)
27. Raussen, M.: Simplicial models for trace spaces. Algebr. Geom. Topol. **10**, 1683–1714 (2010)
28. Raussen, M.: Simplicial models for trace spaces II: General higher-dimensional automata. Tech. Rep. R-2011-11, Department of Mathematical Sciences, Aalborg University (submitted) (2011)
29. Tarjan, R.: Depth-first search and linear algorithms. SIAM J. Comput. **1**, 146–160 (1972)
30. van Glabbeek, R.: On the expressiveness of higher dimensional automata. Theor. Comput. Sci. **368**(1–2), 168–194 (2006)
31. Winskel, G., Nielsen, M.: Models for concurrency. In: Handbook of Logic in Computer Science, pp. 1–148. Oxford University Press, Oxford (1995)

Simplicial models for trace spaces II: General higher dimensional automata

MARTIN RAUSSEN

Higher Dimensional Automata (HDA) are topological models for the study of concurrency phenomena. The state space for an HDA is given as a pre-cubical complex in which a set of directed paths (d-paths) is singled out. The aim of this paper is to describe a general method that determines the space of directed paths with given end points in a pre-cubical complex as the nerve of a particular category.

The paper generalizes the results from Raussen [18; 19] in which we had to assume that the HDA in question arises from a semaphore model. In particular, important for applications, it allows for models in which directed loops occur in the processes involved.

55P10, 55P15, 55U10; 68Q85, 68Q55

1 Introduction

1.1 Background

A particular model for concurrent computation in computer science, called *Higher Dimensional Automata* (HDA), was introduced by Pratt [15] back in 1991. Mathematically, HDA can be described as (labelled) pre-cubical sets (with n–dimensional cubes instead of simplices as building blocks; cf Brown and Higgins [2; 1]) with a preferred set of *directed paths* (respecting the natural partial orders) in any of the cubes of the model.

Compared to other well-studied concurrency models like labelled transition systems, event structures, Petri nets, etc (for a survey on those, see Winskel and Nielsen [23]), R J van Glabbeek [22] showed that Higher Dimensional Automata have the highest expressivity; on the other hand, they are certainly less studied and less often applied so far.

All concurrency models deal with sets of *states* and with associated sets of *execution paths* (with some further structure). The interest is mainly in the structure of the *spaces of execution paths*; typically, it is difficult to extract valuable information about the path

Published: 4 August 2012 DOI: 10.2140/agt.2012.12.1741

space from the state space model. We use topological models for both state space and the execution (=path) space consisting of the *directed paths* (called d-paths) in state space. It is particularly important to know whether the path space is *path-connected*; and, if not, to get an overview over its path components: Executions in the same path component yield the *same result* (decision) in a concurrent computation; different components *may* lead to different results. From a topological perspective, the ultimate aim is to determine the *homotopy type* of these path spaces.

Higher Dimensional Automata are prototypes of *directed* topological spaces; see Grandis [12; 13]. General topological properties of spaces of d-paths and of *traces* (=d-paths up to monotone reparametrizations; see Fahrenberg and Raussen [5; 16]) in pre-cubical complexes were investigated in Raussen [17]. Raussen [18; 19] describes an algorithmic method to determine the homotopy types of trace spaces for Higher Dimensional Automata (and thus in particular to calculate and describe their components) through explicitly constructed *finite simplicial complexes* for a *restricted class* of model spaces:

(1) We had to stick to *semaphore* – or PV – models as described by Dijkstra [4] – an important but restricted class of HDA. Loosely speaking, a PV–model space is a hypercube I^n – with I the unit interval $[0,1]$ – from which a number of n–dimensional hyperrectangles has been removed; see Raussen [18].
(2) We only considered model spaces *without nontrivial directed loops*.

For these restricted class of models, the resulting algorithm has meanwhile been implemented with encouraging results; see Fajstrup et al [8].

In the present paper, we propose an algorithm extending the framework to full generality yielding (generalized simplicial) models for spaces of traces in general pre-cubical complexes; hence we cover models for general (unlabeled) HDA. For these, the homotopy type of trace spaces between given end points is identified with an explicitly constructed complex (a generalization of a simplicial complex); all components of that complex are finite. Using this complex, topological invariants (eg homology) can be calculated.

A price has to be paid: the algorithm determining this complex is, at least in general, more intricate than in the semaphore case. Data structures can be much more complicated, and we have no experience with running times yet.

1.2 Structure and overview of results

Section 2 introduces pre-cubical complexes as HDA; we abstract away from labels. We introduce a signed L_1–arc length on *general* paths in a pre-cubical complex with positive or negative values extending the definition from Raussen [17] for d-paths. It is shown that this signed L_1–arc length is invariant under homotopy *with fixed end points* for *all* paths and that the range of the L_1–arc length map is *discrete* given a pair of end points.

We introduce the class of *nonbranching* (and nonlooping) pre-cubical complexes in Section 3. We show, that the space of traces between two points in such a complex is always *either empty or contractible*.

In the central Section 4, we consider traces in pre-cubical complexes *with branch points* but without nontrivial directed loops. We decompose such a complex into subcomplexes *without* branch points and such that the associated trace spaces cover the trace space corresponding to the entire complex. This decomposition can be quite complicated in the presence of higher order branch points. The nerve of the poset category associated to this cover is homotopy equivalent to trace space. Moreover, we construct a complex (with cones of products of simplices as building blocks) homotopy equivalent to trace space and "more economical" than this nerve.

In Section 5, we show that trace spaces for a pre-cubical complex *with nontrivial directed loops* can be analysed through trace spaces in an associated covering space in which lifts of paths depend on their L_1–arc length – and in which (nontrivial) d-loops lift to nonloops.

In the final Section 6, we give a few hints about a possible implementation that, with a pre-cubical set as input and using an associated directed graph, allows to determine the poset category describing the associated trace space.

2 Pre-cubical complexes and length maps

2.1 Directed paths and traces in a pre-cubical complex

Properties of Higher Dimensional Automata (cf Section 1.1) are intimately related to the study of directed paths in a pre-cubical set, also called a $\square$–set; this term (cf [6]) is used in a similar way as a Δ–set – as introduced in [20] – for a simplicial set without degeneracies. We use $\square_n$ as an abbreviation for the n–cube $I^n = [0,1]^n$ with the product topology.

Definition 2.1 (1) A $\square$–set or pre-cubical complex X is a family of disjoint sets $\{X_n \mid n \geq 0\}$ with face maps $\partial_i^k\colon X_n \to X_{n-1}$, $1 \leq i \leq n$, $k = 0, 1$, satisfying the pre-cubical relations $\partial_i^k \partial_j^l = \partial_{j-1}^l \partial_i^k$ for $i < j$.

(2) The geometric realization $|X|$ of a pre-cubical set X is given as the quotient space $|X| = (\coprod_n X_n \times \square_n)/\equiv$ under the equivalence relation induced from

$$(\partial_i^k(x), t) \equiv (x, \delta_i^k(t)), \quad x \in X_{n+1}, \quad t = (t_1, \dots, t_n) \in \square_n,$$

with $$\delta_i^k(t) = (t_1, \dots, t_{i-1}, k, t_{i+1}, \dots, t_n).$$

(3) A pre-cubical complex M is called *non-self-linked* (cf [9; 17]) if, for all n, $x \in M_n$ and $0 < i \le n$, the $2^i \binom{n}{i}$ iterated faces

$$\partial_{l_1}^{k_1} \cdots \partial_{l_i}^{k_i} x \in M_{n-i}, \quad k_i = 0, 1, \quad 1 \le l_1 < \cdots < l_i \le n,$$

are all different.

In the future, we will not distinguish between a pre-cubical complex X and its geometric realization and just write X for both. We will tacitly assume that all pre-cubical complexes are non-self-linked; if necessary, after a barycentric subdivision.

We are interested in *directed* paths in X. A continuous path within a cube $\square_n$ is a d-path, if all n component functions are (not necessarily strictly) *increasing*. A path in X is a d-path if it is the concatenation of d-paths within cubes; see Raussen [17, Definition 2.2] for details. The set of all d-paths in X will be denoted $\vec{P}(X) \subset X^I$ with subspaces $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ consisting of paths with $p(0) = \mathbf{c}$ and $p(1) = \mathbf{d}$. These spaces inherit a topology from the CO–topology on X^I (the uniform convergence topology).

Reparametrization equivalent d-paths [5] in X have the same directed image ($=$*trace*) in X. Dividing out the action of the monoid of (weakly increasing) reparametrizations of the parameter interval $\vec{I}$, we arrive at trace space $\vec{T}(X)(\mathbf{c}, \mathbf{d})$ (see Fahrenberg and Raussen [5; 16]); it is shown in Raussen [17] to be homotopy equivalent to path space $\vec{P}(X)(\mathbf{c}, \mathbf{d})$ for a far wider class of directed spaces X; in the latter paper, it is also shown that trace spaces enjoy nice properties; eg, they are metrizable, locally compact, locally contractible, and they have the homotopy type of a CW–complex.

Notation Within X and for $x \in X$, we let $\downarrow x := \{y \in X \mid \vec{P}(X)(y, x) \neq \emptyset\}$ denote the *past* of x.

2.2 Length maps

The L_1–arc length of a d-path in a pre-cubical complex was introduced and studied in Raussen [17]. The definition and important properties can be extended to general nondirected paths; for these the (signed) L_1–arc length may be negative. This goes roughly as follows:

The *signed* L_1–length $l_1^{\pm}(p)$ of a path $p\colon I \to \square_n$ within a cube $\square_n$ is defined as $l_1^{\pm}(p) = \sum_{j=1}^{n} p_j(1) - p_j(0)$. For any path p, that is the concatenation of finitely many paths each of which is contained in a single cube, the signed L_1–length is defined as the sum of the lengths of the pieces; the result is independent of the choice of decomposition – and of the parametrization! Moreover, it is nonnegative for every d-path and positive for every nonconstant d-path.

This construction can be phrased more elegantly using differential one-forms on a cubical complex (a special case of the PL differential forms introduced by D Sullivan [21] in his approach to rational homotopy theory, or of the closed one-forms on topological spaces by Farber [10; 11]): On an n–cube $e \simeq \square_n$, consider the particular 1–form $\omega_e = dx_1 + \cdots + dx_n \in \Omega^1(\square_n)$. It is obvious that $\omega_{\partial_i^k e} = (i_i^k)^* \omega_e$ with $i_i^k\colon |\partial_i^k e| \hookrightarrow |e|$ denoting inclusion. Pasting together, one arrives at a particular (closed!) 1–form ω_X on every pre-cubical set X – the one-form that reduces to ω_e on every cell e in X.

The signed length of a (piecewise differentiable) path γ on X can then be defined as $l_1^{\pm}(\gamma) = \int_0^1 \gamma^* \omega_X$ and extended to continuous paths using uniformly converging sequences of such piecewise differentiable paths. This length is invariant under orientation preserving reparametrization; it changes sign under orientation reversing reparametrization; it is additive under concatenation and nonnegative for d-paths. It yields a *continuous* map $l_1^{\pm}\colon P(X)(x_0, x_1) \to \mathbf{R}$. An application of Stokes' theorem shows:

Proposition 2.2 *Two paths $p_0, p_1 \in P(X)(x_0, x_1)$ that are homotopic rel end points have the same signed length: $l_1^{\pm}(p_0) = l_1^{\pm}(p_1)$.* □

A more direct proof can be given along the lines of Raussen [17] using the continuous map $s\colon X \to S^1 = \mathbf{R}/\mathbf{Z}$ given by $s(e; x_1, \dots, x_n) = \sum x_i \bmod 1$. It is clear from the construction, that $l_1^{\pm}(p) \equiv s(p(1)) - s(p(0)) \bmod 1$. As a consequence, $l_1^{\pm}(P(X)(x_0, x_1))$ is constant mod 1 and, in particular, a discrete subset of the reals. Hence, $l_1^{\pm}$ is constant on a connected component, ie, a homotopy class of paths in $P(X)(x_0, x_1)$.

Remark 2.3 As remarked in Raussen [17, Remark 2.8] it is *not* possible to extend *nonnegative* L_1–arc length continuously to nondirected paths.

3 Trace spaces for nonbranching pre-cubical complexes

In the following two sections, we will only consider *nonlooping* pre-cubical complexes. In such a complex X, the only *directed loops* are trivial, ie, *constant*.

A (finite) such pre-cubical complex X will be called *nonbranching* if it satisfies the following additional property:

(NB) Every vertex $v \in X_0$ is the lower corner vertex of a *unique maximal* cube c_v in X. This maximal cube c_v contains thus all cubes with lower corner vertex v as a (possibly iterated) lower face.

On a nonbranching cubical complex, there is a *privileged directed* flow

$$F^X \colon X \times \mathbf{R}_{\geq 0} \to X :$$

Every element $x \in X$ is contained in the interior or the lower boundary of a uniquely determined maximal cube, ie, the maximal cube c_v of its lowest vertex v. On the interior and the lower faces of such a cube c_v, this flow is locally given by the *diagonal* flow:

$$F_c^X(c;(x_1,\dots,x_n);t) = (c;x_1+t,\dots,x_n+t) \quad \text{for } 0 \leq t \leq 1 - \max_{1\leq i\leq n} x_i. \tag{3-1}$$

On a maximal vertex v_1 with $c = c_{v_1} = v_1$ (a deadlock), F_c^X is defined to be constant in t for $0 \leq t$.

On cubes, that are not lower boundaries of others, these flow lines are the gradient lines of the 1–form ω_X from Section 2.2; this is *not* true on such lower boundaries. Piecing together these local flows so that they satisfy the flow semigroup property yields a piecewise-linear (hence Lipschitz continuous) global flow all of whose flow lines are d-paths; note from the construction that this flow can only have equilibria at deadlocks.

Remark 3.1 At a branch point v_0 in a general (branching) pre-cubical complex X (see Definition 4.3), it is *not* possible to construct such a flow. Diagonal flows on several maximal cubes do not fit together on their intersections.

Lemma 3.2 *A finite nonbranching* connected *pre-cubical complex X has a* unique *maximal vertex v_1.*

Proof First of all, there is at least one maximal vertex. Otherwise, one would have d-paths of arbitrary length in X; hence X – without nontrivial loops – could not be finite.

Suppose $v_1, v_2, \dots, v_k \in X, k > 1$, is a list of *all* maximal vertices. Consider the maximal vertices in the common past subcomplexes $\downarrow v_i \cap \downarrow v_j, i \neq j$, and choose among those the *maximal* ones (that cannot reach any of the others). Pick such a maximal vertex v and consider the associated maximal cube c_v.

There is at least one edge in c_v with v as lower boundary from which one can reach v_i and not v_j; likewise another edge from which one can reach v_j and not v_i. From the top edge of c_v, at least one of the v_r in the list is reachable. As a consequence, from at least one of the two edges mentioned before, *two* maximal vertices can be reached. Contradiction to maximality! □

The key Proposition 2.8 from Raussen [18] generalizes as follows:

Proposition 3.3 *For every pair of elements $x_0, x_1 \in X$ in the geometric realization X of a pre-cubical* nonbranching *complex X, trace space $\vec{T}(X)(x_0, x_1)$ is either* empty *or* contractible.

Proof We assume that $\vec{T}(X)(x_0, x_1) \neq \varnothing$ and, without restriction, that x_1 is *the* maximal vertex in X; in general, just replace X by $\downarrow x_1 \subset X$, still a nonbranching complex; without deadlocks and unsafe regions.

The directed flow line corresponding to the flow F^X (cf (3-1)) starting at $x \in X$ and ending at x_1 (after linear renormalization so that its domain becomes the unit interval I) will be called $p_x \in \vec{P}(X)(x, x_1)$.

A contraction $H\colon \vec{P}(X)(x_0, x_1) \times I \to \vec{P}(X)(x_0, x_1)$ to the flow path $H_0 = p_{x_0}$ is constructed as follows: For $p \in \vec{P}(X)(x_0, x_1)$, let

$$H(p,t)(s) = \begin{cases} p(s) & t \leq s, \\ p_{p(t)}(\frac{s-t}{1-t}) & s \leq t. \end{cases}$$

Remark that $H_1 = p$ and that an intermediate d-path H_t follows p until $p(t)$ and then it follows the flow line starting at $p(t)$ automatically ending at x_1.

Finally, the quotient map $\vec{P}(X)(x_0, x_1) \to \vec{T}(X)(x_0, x_1)$ is a homotopy equivalence [17]. □

This proof, using the diagonal flow F^X, is different from the one given in Raussen [18, Proposition 2.8] for the special case of cubical complexes arising from semaphore models; but it is certainly similar in spirit.

4 Trace spaces for nonlooping pre-cubical complexes

In this section, we study traces in a more general finite pre-cubical complex X; still *without nontrivial loops*, but allowing for branch points: How to find subcomplexes $Y \subseteq X$ satisfying (NB)? Investigating the space of d-paths between $\mathbf{x}_0$ and $\mathbf{x}_1$ in X, we assume that $X = [\mathbf{x}_0, \mathbf{x}_1] = \uparrow\mathbf{x}_0 \cap \downarrow\mathbf{x}_1$. In particular, X contains neither unsafe nor unreachable regions. We start with an abstract description:

4.1 An abstract simplicial model

The subcomplex given by the *carrier sequence* corresponding to any directed path (see Fajstrup [6]), the sequence of cubes containing segments of that path, is obviously a subcomplex satisfying (NB).

One may order NB subcomplexes of X by inclusion – chains are of bounded length since there are only finitely many cubes – and focus on the *maximal* nonbranching subcomplexes. Every d-path with a given start point is contained in a maximal NB subcomplex, that is in general not uniquely determined. Traces contained in maximal NB subcomplexes cover thus the space of all d-paths (with given start and end point).

Lemma 4.1 *An intersection $S = \bigcap X_i$ of subcomplexes X_i each satisfying* (NB) *satisfies* (NB) *as well. Hence the trace space $\vec{T}(S)(\mathbf{x}_0, \mathbf{x}_1)$ is either contractible or empty.*

Proof The intersection of maximal cubes at every vertex will be the maximal cube in the intersection and hence unique. For contractibility, use Proposition 3.3. Empty path spaces may arise when S is not connected. □

The subcomplexes $X_i \subset X, i \in I$, satisfying (NB) that are *maximal* with respect to inclusion give thus rise to a covering $\vec{T}(X_i)(\mathbf{x}_0, \mathbf{x}_1)$ of trace space $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ by contractible sets; in fact:

Theorem 4.2 *For a finite pre-cubical complex X, the trace space $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ is homotopy equivalent to the nerve of the covering given by the subspaces $\vec{T}(X_i)(\mathbf{x}_0, \mathbf{x}_1)$.*

Proof The theorem is an almost immediate consequence of the nerve lemma; see Kozlov [14, Theorem 15.21]. The subspaces X_i are in general not open, and the subspaces $\vec{T}(X_i)(\mathbf{x}_0, \mathbf{x}_1)$ will in general not give rise to an open covering. According to Raussen [17, Proposition 3.15], the trace space $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ has the homotopy type of a CW–complex; the subspaces $\vec{T}(X_i)(\mathbf{x}_0, \mathbf{x}_1)$ correspond to subcomplexes. The projection lemma [14, Theorem 15.19 and Remark 15.20] comparing homotopy colimits with colimits holds also in this case. □

4.2 An index category

4.2.1 (Higher order) branch points The maximal subcomplexes X_i from Theorem 4.2 may be very difficult to identify for a complex X with many cells. In the following, we describe an algorithmic method that determines an index category $\mathcal{C}(X)(\mathbf{x}_0, \mathbf{x}_1)$

that can be represented by a complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ which is homotopy equivalent to trace space $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$. The building blocks of the complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ are products of simplices and of cones of such spaces. The construction is similar in spirit to that in Raussen [18], albeit, in the details slightly more complicated.

We investigate vertices in the 0–skeleton $X_0 \subset X$ at which condition (NB) from Section 4 are violated:

Definition 4.3 A vertex $v \in X_0$ is called a *branch point* if there are *more than one* maximal cube c having v as lower vertex (ie, an iterate of ∂_*^0 yields v). The set of all such maximal cells with lower vertex v is called the *branch set* B_v with $|B_v| > 1$.

Remark 4.4 Let $v \in X_0$ be a branch point with several maximal cubes $c_1,\dots,c_r$ with lower vertex v. Obviously, at most one of the cubes c_j can be contained in every of the (NB) subcomplexes X_i from Section 4.1.

For a cell c in X, we denote by c^- the geometric realization of c and of all (iterated) *lower* boundaries – not including mixed or upper boundaries. Hence, $|c^-| \cong [0,1[^n$ for an n–cell c. For a fixed branch point v and a branch cell $c_j \in B_v$, let

$$X_j^v := \downarrow c_j^- \cup \mathcal{C}\Big(\downarrow \bigcup_{c_i \in B_v} c_i^-\Big)$$

consist of all points that, as far as they can reach any branch in B_v, they have to stay in the past of the particular branch c_j; $\mathcal{C}$ denotes the complement within X. Clearly, $X = \bigcup_{c_j \in B_v} X_j^v$.

Lemma 4.5 $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1) = \bigcup_{c_j \in B_v} \vec{T}(X_j^v)(\mathbf{x}_0,\mathbf{x}_1)$ *for every branch point* v.

Proof We need to show that $\vec{T}(\bigcup_{c_j \in B_v} X_j^v)(\mathbf{x}_0,\mathbf{x}_1) = \bigcup_{c_j \in B_v} \vec{T}(X_j^v)(\mathbf{x}_0,\mathbf{x}_1)$: Every d-path from $\mathbf{x}_0$ to $\mathbf{x}_1$ starts in the (past closed) set $\downarrow \bigcup_{c_j \in B_v} c_j^-$ and then leaves it for its (future closed) complement. The sets $\downarrow c_j^-$ are all past closed; a d-path p that has left one of these sets will never get back to it. In particular, there is at least one (last) set $X_j^v, c_j \in B_v$, containing p. □

Contrary to the special situation of pre-cubical complexes arising from PV–protocols discussed in [18], it is not enough to consider only (1. order) branch points as the following example (cf Figure 1) shows:

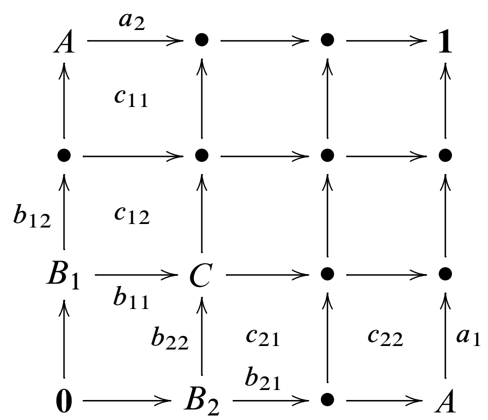


Figure 1: Branch points in a 2D complex

Example 4.6 The complex X to be discussed arises from the 9 planar 2–cubes in Figure 1 by identifying the two vertices denoted A. Remark the two special “horizontal” and “vertical” d-paths from $\mathbf{0}$ to $\mathbf{1}$ through A. The vertex A is the only branch point in X; it has branch set $B^A = \{a_1, a_2\}$. The subcomplex X_1^A arises from X by crossing out the two cells c_{11} and c_{12} – apart from the left boundary 1–cells. Likewise, for X_2^A, the cells c_{21} and c_{32} – apart from the lower boundary 1–cells – have to be deleted. The first subcomplex has a secondary branch point B_1 with branches b_{11} and b_{12}. Likewise, the second one has a secondary branch point B_2 with branches b_{21} and b_{22}.

The homotopy type of the trace space $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ will be identified in Example 4.12 below.

4.2.2 The index category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ Hence, it is necessary to consider secondary, and in general higher order branch points, as well:

- The original space X comes with a set $\mathrm{BP} = \{b^1,\dots,b^l\}$ of branch points and associated maximal branch cubes $\mathrm{BC} = \{c_j^i\}$ and a surjective map $p\colon \mathrm{BC} \downarrow \mathrm{BP}$, $p(c_j^i) = b^i$.
- For every section $s(1)\colon \mathrm{BP} \uparrow \mathrm{BC}$ of the map p, consider the subcomplex $X_{s(1)} = \bigcap_{b_i\in\mathrm{BP}}(\downarrow c^{i-}_{s(1)(b_i)} \cup \mathcal{C}(\bigcup_{p(j)=b_i} \downarrow c_j^{i-}) \subset X$. $X_{s(1)}$ is in fact a subcomplex of X since the branch cubes are all maximal. It is a proper subcomplex containing $\mathbf{x}_0$ and $\mathbf{x}_1$.
- Such a complex $X_{s(1)}$ may have (second order) branch points $b^i(s(1)) \in \mathrm{BP}_{s(1)}$ and branch cubes $c_j^i(s(1)) \in \mathrm{BC}_{s(1)}$ which come with a projection $p(s(1))\colon \mathrm{BC}_{s(1)} \downarrow \mathrm{BP}_{s(1)}$ and sections $s(2)\colon \mathrm{BP}_{s(1)} \uparrow \mathrm{BC}_{s(1)}$.

- Iterate: Given subsequent sections $s(1),\dots,s(r)$ define a proper subcomplex $X_{s(r)} \subset X_{s(r-1)}$ as
$$X_{s(r)} = \bigcap_{0\le k\le r}\ \bigcap_{b_i\in \mathrm{BP}(s(k))} \Big(\downarrow c^{i-}_{s(k)(b_i)} \cup \mathcal{C}\Big(\bigcup_{p(k)(j)=b_i} \downarrow c_j^{i-}; X_{s(k-1)}\Big)\Big),$$
with the convention that $p(s(0))\colon \mathrm{BC}(s(0)) \downarrow \mathrm{BP}(s(0))$ is the original projection map $p\colon \mathrm{BC} \downarrow \mathrm{BP}$. This complex $X_{s(r)}$ may give rise to sets of new branch cells $\mathrm{BC}_{s(r)}$ and branch points $\mathrm{BP}_{s(r)}$ with a projection map $p(s(r))\colon \mathrm{BC}_{s(r)} \downarrow \mathrm{BP}_{s(r)}$. Since these subcomplexes become smaller and smaller under iteration in the finite complex X, every such iteration will ultimately end in a subcomplex $X_{s(r)}$ *without branch points*.
- A subsequent sequence of sections $s(k)\colon \mathrm{BP}_{s(k)} \uparrow \mathrm{BC}_{s(k)}$, $k \le r$, is called *coherent and complete* if $X_{s(r)}$ satisfies property (NB); cf Section 3. The set of all such coherent and complete sequences will be called $\mathrm{CCS}(X)$.
- Given a coherent and complete sequence $s \in \mathrm{CCS}(X)$, we may associate the set of branch points $\mathrm{BP}(s) = \bigcup_{k=0}^r \mathrm{BP}_{s(k)}$ and branch cubes $\mathrm{BC}(s) = \bigcup_{k=0}^r \mathrm{BC}_{s(k)}$, the projection $p(s)\colon \mathrm{BC}(s) \downarrow \mathrm{BP}(s)$, and the "tautological" section $\overline{s}\colon \mathrm{BP}(s) \uparrow \mathrm{BC}(s)$; there are no branch cells to choose at depth s!

Definition 4.7 (1) The poset category $\mathcal{M}(X)(\mathbf{x}_0,\mathbf{x}_1)$ has as

- *objects* all pairs of the form (S,C) with $\varnothing \ne S \subset \mathrm{CCS}(X)$ and C a set of the form $C = \prod_{b_i\in\bigcup_{s\in S}\mathrm{BP}(s)} C_i$, $\overline{s}(b_i) \in C_i \subset \mathrm{BC}(b_i)$; with $\overline{s}$ denoting the tautological section from Section 4.2.2,
- *morphisms* $(S,C) \le (S',C') \Leftrightarrow S \subseteq S', \forall s\in S, b_i \in \mathrm{BP}(s)\colon C_i \subseteq C_i'$.

Note that the minimal objects (S,D) of this category are composed of a set S with precisely one element s and such that $b_i \in \mathrm{BP}(s) \Rightarrow |C_i| = 1$.

(2) To a section $s \in \mathrm{CCS}(X)$, a branch point $b_i \in \mathrm{BP}(s(k))$ and a branch cube $c_i \in \mathrm{BC}(s(k))$, we associate the subspace $(X_{s(k)})^{c_i}_{b_i} \subset X$. To an object (S,C) in $\mathcal{M}(X)(\mathbf{x}_0,\mathbf{x}_1)$ we associate the subspace $X_{(S,C)} := I \subset X$.

(3) The category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is the full subcategory of $\mathcal{M}(X)(\mathbf{x}_0,\mathbf{x}_1)$ whose objects (S,C) are characterized by the fact that $\vec{T}(X_{(S,C)})(\mathbf{x}_0,\mathbf{x}_1)$ is *nonempty*.

Proposition 4.8 (1) $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1) = \bigcup_{(S,C)} \vec{T}(X_{(S,C)})(\mathbf{x}_0,\mathbf{x}_1)$, *where the union extends over all objects of the category* $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$.

(2) *The subspace* $\vec{T}(X_{(S,C)})(\mathbf{x}_0,\mathbf{x}_1)$ *is contractible for every object* (S,C) *of the subcategory* $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$.

Proof (1) This follows from Lemma 4.5 by induction.

(2) According to Proposition 3.3, the trace space $\vec{T}(X_{s(r)})(\mathbf{x}_0,\mathbf{x}_1)$ is empty or contractible for every complete coherent sequence of branches and branch points. For every object (S,C) of the category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$, the space $X_{(S,C)}$ is a finite intersection of spaces of type $\vec{T}(X_{s(r)})(\mathbf{x}_0,\mathbf{x}_1)$, $s\in S$. Apply Lemma 4.1. □

Remark 4.9 An algebraic representation of the category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is not as straightforward as for semaphore models that resulted in a poset category of binary matrices (cf Raussen [18]). In the general case, one has to consider combinations of all branch point sequences (of different lengths) and the occuring branches – which may have different cardinalities. I plan to develop algorithmically the representation of $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ in a future paper.

4.3 Homotopy equivalences

We will now present a description of the homotopy type of $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ as a colimit of simple spaces (generalizing the prodsimplicial complex for traces arising from semaphore models; cf Raussen [18]). For simplicity, we will describe only the case with branch points of order at most two.

In this case, to every first order branch point $b^i\in \mathrm{BP}$, we associate the branch cells $\mathrm{BC}^i=p^{-1}(b^i)$; a section $s\colon \mathrm{BP}\uparrow \mathrm{BC}$ gives rise to second order branch points in $\mathrm{BP}(s)$ and, to every of the second order branch points $b^i(s)\in \mathrm{BP}(s)$, a set of branch cells $\mathrm{BC}^i(s)$. Then, the total complex corresponding to *all* objects (S,C) corresponds to

$$\Delta^{|\mathrm{BC}^1|-1}\times\cdots\times\Delta^{|\mathrm{BC}^l|-1}\times\prod_{s\colon \mathrm{BP}\uparrow\mathrm{BC}} C(\Delta^{|\mathrm{BC}^1(s)|-1}\times\cdots\times\Delta^{|\mathrm{BC}^{l(s)}(s)|-1}).$$

Here CX denotes the *cone* over X.

A subset $S=S_1\times\cdots\times S_l\subseteq \mathrm{BC}^1\times\cdots\times\mathrm{BC}^l$ corresponds to a product of simplices $\Delta_S=\Delta^{|S_l|-1}\times\cdots\times\Delta^{|S_1|-1}$. A subset $C(s)=\prod_{b^i(s)\in\mathrm{BP}(s)} C^i(s)$, $s\colon \mathrm{BP}\uparrow\mathrm{BC}$, of products of first and second order branch cells corresponds to

$$\Delta^C(s):=\Delta^{|C^1(s)|-1}\times\cdots\times\Delta^{|C^{l(s)}(s)|-1}\subseteq\Delta^{|\mathrm{BC}^1(s)|-1}\times\cdots\times\Delta^{|\mathrm{BC}^{l(s)}(s)|-1}.$$

An object (S,C) in $\mathcal{M}(X)(\mathbf{x}_0,\mathbf{x}_1)$ corresponds to

$$\Delta(S,C):=\Delta_S\times\prod_{s\in S} C(\Delta^C(s))\times\prod_{s\notin S} *_s,$$

with $*_s$ the cone point in $C(\Delta^C(s))$. Morphisms $(S,C)\le(S',C')$ correspond to inclusions $\Delta(S,C)\hookrightarrow\Delta(S',C')$.

Definition 4.10 The complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is defined as the colimit

$$\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1) := \operatorname{colim}_{\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)} \Delta(S,C).$$

Theorem 4.11 *The trace space $\vec{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is homotopy equivalent to*

(1) *the nerve $\Delta(\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1))$ of the poset category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$, and*

(2) *the complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$.*

Proof The proof of Theorem 4.11 is analogous to that of [18, Theorem 3.5]: The homotopy colimit of the functor associating the contractible spaces $\vec{T}(X)_{(S,C)}(\mathbf{x}_0,\mathbf{x}_1)$, resp. $\Delta(S,C)$ to an object (S,C) in $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ is homotopy equivalent to the functor associating the same point to every (S,C), ie to the nerve of that category. Homotopy colimit and colimit of the first two functors are also homotopy equivalent; by Lemma 4.5, this colimit is the entire trace space, resp. by definition the complex $\mathbf{T}(X)(\mathbf{x}_0,\mathbf{x}_1)$. □

4.4 Example: Trace spaces for 2–dimensional pre-cubical complexes

Example 4.12 First, we look at the case of the space X described in Example 4.6 and Figure 1. There are four coherent and complete sequences of sections:

$$\begin{aligned} s_1(A) &= a_1, & s_1(1)(B_1) &= b_{11}, \\ s_2(A) &= a_1, & s_2(1)(B_1) &= b_{12}, \\ s_3(A) &= a_2, & s_3(1)(B_2) &= b_{21}, \\ s_4(A) &= a_2, & s_4(1)(B_2) &= b_{22}, \end{aligned} \tag{4-1}$$

corresponding to minimal objects $(S_i = \{s_i\}, C_i)$ with C_i the one-element set given by the branches chosen by the section.

The only nontrivial intersection occurs for

$$\vec{T}(X_A^{a_1} \cap X_{B_1}^{b_{11}})(\mathbf{0},\mathbf{1}) \quad \text{and} \quad \vec{T}(X_A^{a_2} \cap X_{B_2}^{b_{22}})(\mathbf{0},\mathbf{1})$$

giving rise to the object $(S_{14} = \{s_1, s_4\}, C_{14} = \{a_1, a_2\} \times \{b_{11}\} \times \{b_{22}\})$. In this case, the complex $\mathbf{T}(X)(\mathbf{0},\mathbf{1})$ is thus a disjoint union of the cones on two vertices (each of this edges corresponds to one of the special horizontal and vertical traces) and product of an edge with the cone on a vertex.

Using Theorem 4.11, we can conclude: $\vec{T}(X)(\mathbf{0},\mathbf{1}) \simeq \mathbf{T}(X)(\mathbf{0},\mathbf{1}) \cong I \sqcup I \sqcup I^2$ is homotopy equivalent to a set of three disjoint points.

Example 4.13 The 2–dimensional complex X in Figure 2 below arises from gluing the boundaries $\partial\square^3$ of two 3–cubes $\square^3$ along a common face $\square^2$. Its trace space has previously been studied by Bubenik [3]. The complex has two branch points $\mathbf{x}_0$ and A and *no higher order* branch points.

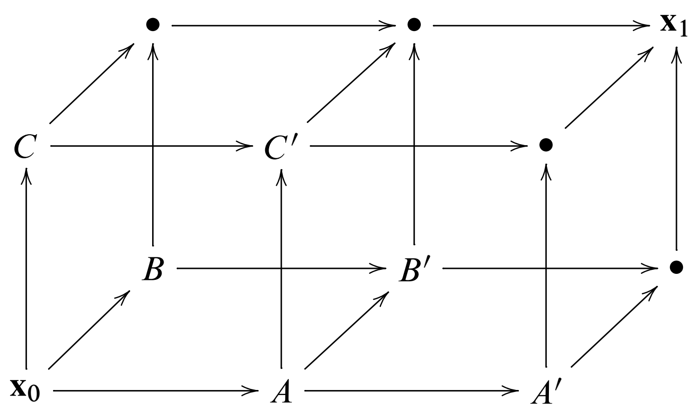


Figure 2: The complex X: Boundaries of two cubes glued together at common square $AB'C'\bullet$

Bubenik's "necklace" model yields a simplicial complex consisting of 16 triangles, 46 edges and 29 vertices which is seen to be homotopy equivalent to $S^1 \vee S^1$. The complex $\mathbf{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ – a prodsimplicial complex in the terminology of Kozlov [14], since there are no higher order branch points – homotopy equivalent to $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ (cf Theorem 4.11 above) is shown in Figure 3.

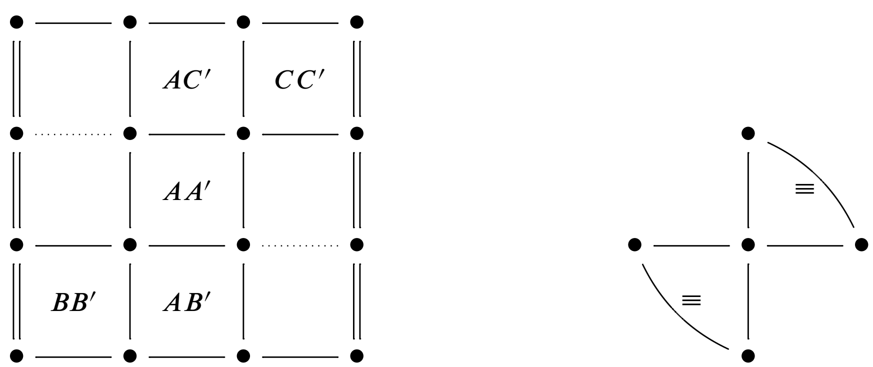


Figure 3: Prodsimplicial complex homotopy equivalent to the trace space for X – and a homotopy equivalent complex

It consists of the five named squares in the nine square decomposition of a 2–torus $\Delta^1 \times \Delta^1$ – identify boundary edges as usual – to which a full triangle has been attached along the circle on the vertical (left=right) triangle (marked by double lines =). The nine vertices correspond to 3×3 combinations of paths staying "under" the three branches (2–cubes) corresponding to the two branch points x_0 and A in Figure 2.

The labels in the five marked squares refer to paths staying under *two* branches for each of the branch points. They are marked by labels referring to traces through the points mentioned. For example, AC' denotes the space of all traces from $\mathbf{x_0}$ to A (front *and* bottom of the first cube) and then C' (front *and* left of the second cube) and from C' as an arbitrary d-path on the top square of the second cube to $\mathbf{x_1}$. It is easily seen, that there are no d-paths corresponding to the remaining four squares BC$'$, BA', CA' and CB'.

Furthermore, there is a full triangle (marked with =): Every trace entering the interior of the "left" square $x_0B \bullet C$ in Figure 2 leaves the past of the union of the branches corresponding to branch point A; such a path is therefore contained in *all three* sets X_j^A, giving rise to a full triangle $\Delta^0 \times \Delta^2$.

The complex $\mathbf{T}(X)(x_0, x_1)$ in Figure 3 consists thus of six 2–cells (five of type $\Delta^1 \times \Delta^1$, one of type $\Delta^0 \times \Delta^2$), 16 edges (all but the two stipled ones) and of all 9 vertices; it has Euler characteristic -1. A contraction of the full triangle can be extended to a contraction of the entire space to a union of two full triangles (shown on the right hand side of Figure 3) with three vertices (opposite vertices are identified). That simplicial 2–complex contracts to a 1–complex $S^1 \vee S^1$.

From Theorem 4.11, we may conclude: $\vec{T}(X)(\mathbf{x_0}, \mathbf{x_1}) \simeq S^1 \vee S^1$.

5 Trace spaces for general pre-cubical complexes

In this section we outline, how the methods previously explained can be adapted to trace spaces in a general pre-cubical complex X that may allow directed loops using suitable *coverings* of the complex X:

5.1 Nonlooping length coverings

We exploit the d-map (directed map) $s\colon X \to \vec{S}^1 \cong \mathbf{R}_{/\mathbf{Z}}$ introduced in Raussen [17]: just glue the maps $s(x_1, \dots, x_n) = \sum x_i \bmod 1$ on individual cubes. Consider the pullback $\widetilde{X}$ in the pullback diagram

$$\begin{array}{ccc} \widetilde{X} & \xrightarrow{S} & X \times \mathbf{R} \\ \pi \big\downarrow & & \big\downarrow \mathrm{id} \times \exp \\ X & \xrightarrow{\mathrm{id} \times s} & X \times S^1. \end{array}$$

The map π is a covering map with unique path lifting. Since exp can be interpreted as a semi-cubical map, $\widetilde{X}$ can be conceived as a semi-cubical complex: every cube e in X is replaced by infinitely many cubes $(e,n), n \in \mathbf{Z}$ with boundary maps given as $\partial_-(e,n) = (\partial_- e, n), \partial_+(e,n) = (\partial_+ e, n+1)$.

The directed paths on $\widetilde{X}$ are those that project to directed paths in X under the projection map π. Note the maps exp and s – and hence π and $\pi_2 \circ S$ – preserve the signed L_1–arc lengths from Section 2.2. Moreover, the L_1–length $l_1^{\pm}(p)$ of a path p in X (cf Section 2.2) with lift $\widetilde{p}$ can be expressed via the d-map $\widetilde{S}\colon \widetilde{X} \to X \times \mathbf{R}$ in the pullback diagram as follows:

Lemma 5.1 (1) $l_1^{\pm}(p) = \pi_2(S(\widetilde{p}(1))) - \pi_2(S(\widetilde{p}(0)))$.

(2) $\widetilde{X}$ *has* only trivial *directed loops.*

Proof (1) This is clearly true locally in any cell as long as start and end point have an L_1–distance less than one. Sum up and cancel!

(2) A directed path in $\widetilde{X}$ projects to a directed path in X with positive L_1–length, unless it is constant. Apply (1). □

Another method to construct this covering is to consider the homotopical length map

$$\pi_1(X) \stackrel{l_1^{\pm}}{\to} \mathbf{Z} \to 0$$

(cf Proposition 2.2) from the nondirected classical fundamental group of the cubical complex X. Consider the cover $\widetilde{X} \downarrow X$ with fundamental group $\pi_1(\widetilde{X}) = K \trianglelefteq \pi_1(X)$ the kernel of the homotopical length map. It can be given the structure of a pre-cubical complex, and every element $\mathbf{x}$ in X has lifts $\mathbf{x^n} \in \widetilde{X}$, $n \in \mathbf{Z}$. The projection map $\pi\colon \widetilde{X} \downarrow X$ preserves signed L_1–arc length. A path in $\widetilde{X}$ is directed if its projection to X is. There are no nontrivial directed loops in $\widetilde{X}$ – these need to have L_1–length 0!

Example 5.2 (1) Consider a torus $T = \partial\Delta^1 \times \partial\Delta^1$ as a pre-cubical set consisting of nine 2–cubes. The length cover $\widetilde{T}$ can then be modeled as an infinite strip of width 3 with identifications $(x,3) \sim (x+3,0)$; see Figure 4.

The subcomplex in Figure 4 between an initial vertex **0** and a second final vertex **1** at length distance $3n, n > 0$, has exactly one branch vertex • (3 to the left $\sim$ below the final vertex). The algorithm deriving the homotopy type of all d-paths of length $3n$ between these two vertices (consisting of $n+1$ contractible components corresponding to pairs (k,l) of nonnegative integers with $k+l=n$) runs through many higher order branch points and removes only few cells at a time.

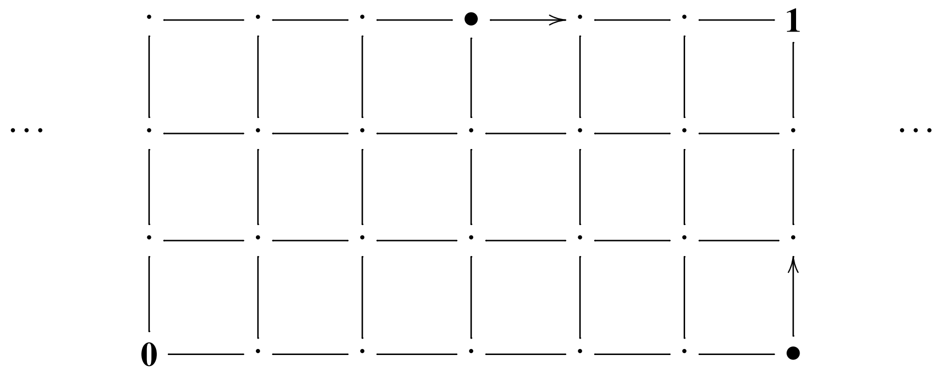


Figure 4: Length cover of a torus

(2) Now consider the space $\bar{T}$ arising from removing a (middle) cell in T. The length cover of $\bar{T}$ arises from $\tilde{T}$ by removing every third cell (marked X) in the middle strip; cf Figure 5. The lower corner vertices of the removed cells are all branch vertices.

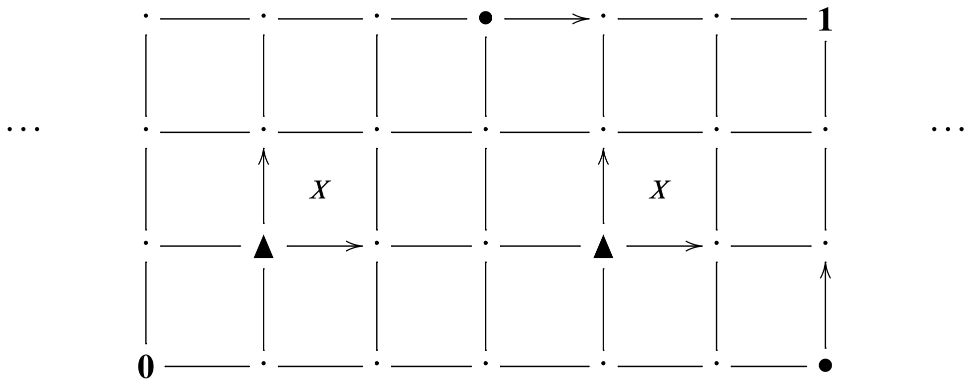


Figure 5: Length cover of torus with removed cell

For this space, no higher order branch points arise; the 2^n contractible components of the trace space correspond to the sequences of length n on the letters r, u – right and up.

5.2 A decomposition of the trace space

For a general pre-cubical complex X with length cover $\tilde{X}$ we obtain:

Proposition 5.3 *For every pair of points $\mathbf{x}_0, \mathbf{x}_1 \in X$, trace space $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ is homeomorphic to the disjoint union $\bigsqcup_{n\in\mathbf{Z}} \vec{T}(\tilde{X})(\mathbf{x}_0^0, \mathbf{x}_1^n)$.*

Proof An inverse to the projection map $\Pi\colon \coprod_{n\in\mathbf{Z}} \vec{T}(\widetilde{X})(\mathbf{x}_0^0, \mathbf{x}_1^n) \to \vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ induced by the covering projection $\pi\colon \widetilde{X} \downarrow X$ is given by unique lifts of the directed paths representing traces. Remark that many of the spaces $\vec{T}(\widetilde{X})(\mathbf{x}_0^0, \mathbf{x}_1^n)$ may be empty for specific $n \in \mathbf{Z}$. □

Since the covering $\widetilde{X}$ has *only trivial loops*, Proposition 5.3 allows us to apply the methods from Section 4 in order to describe the homotopy type of trace spaces $\vec{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ in an arbitrary pre-cubical complex X. It is of course desirable to exploit periodicity properties in the comparison of spaces $\vec{T}(\widetilde{X})(\mathbf{x}_0^0, \mathbf{x}_1^n)$ for different values of n.

Remark 5.4 Simple semaphore models with loops can be constructed from spaces of the form $X = T^n \setminus F$ with $T^n = (S^1)^n$ an n–torus and F a collection of forbidden hyperrectangles. For such a space, one may consider the (sub)covering

$$\begin{array}{ccc} \widetilde{X} & \hookrightarrow & \mathbf{R}^n \\ \downarrow & & \downarrow \exp \\ X & \hookrightarrow & T^n \end{array}$$

of the universal covering of X – a far bigger gadget. It has the property that (d-)paths, that are not homotopic in the torus T^n, lift to (d-)paths with different end points. The methods from [18] can be applied to $\widetilde{X}$ immediately. It is probably easier to get hold on periodicity properties in this setting. This line is currently being investigated by several colleagues; see Fajstrup [7] and Fajstrup et al [8].

6 Implementation issues

6.1 A directed graph associated with a cubical complex

To a cubical complex X, one may associate – forgetting dimensions and the pre-cubical relations – a *directed* graph $\vec{\Gamma}(X)$: the vertices are the cubes in X: $V(\vec{\Gamma}(X)) = \bigcup_n X_n$; to every vertex = cube $c \in X_n$, we associate arcs from c to $\partial_i^1 c$ and from ∂_i^0 to c. The past $\downarrow c \subset X$ is then the union of all *predecessors* of c regarded as a vertex in $\vec{\Gamma}(X)$; likewise, its future $\uparrow c \subset X$ is the union of all *successors* of c. Both can be determined recursively. Moreover, for a set C of cells, $\downarrow C = \bigcup_{c\in C} \downarrow c$.

6.2 Steps in the determination of a trace complex

In this section, we collect a few ideas on how to start the design of an algorithm determining the complex $\mathbf{T}(X)(\mathbf{x}_0, \mathbf{x}_1)$ associated to a nonlooping finite pre-cubical complex X and two vertices $\mathbf{x}_0, \mathbf{x}_1 \in X_0$:

- The lower corner $L(c) \in X_0$ of an n–cell $c \in X_n$ can be determined as $L(c) = (\partial_0^0)^n(c)$. Altogether, this recipe defines a map $L\colon X \to X_0$.
- A *maximal* cell $c \in X$ has *no coface* under ∂_*^0; maximal is to be understood with respect to a lower vertex. The set $M(X)$ of maximal cells is thus of the form $M(X) = \bigcup_n X_n \setminus \bigcup_{0\le i\le n} \partial_i^0(X_{n+1})$. The restriction of the map L to $M(X)$ denoted by $L_M\colon M(X) \to X_0$ associates to a maximal cell c its lower corner vertex in X_0.
- A branch point $v \in X_0$ is characterized by $|L_M^{-1}(v)| > 1$. Given the map L_M, the set of branch points $\mathrm{BP} \subset X_0$ corresponds to vertices with more than one inverse image, and the set of branch cells $\mathrm{BC} = L^{-1}(\mathrm{BP}) \subset M(X) \subset X$ corresponds to the union of these inverse images.

Using the directed graph $\vec{\Gamma}(X)$ from Section 6.1, one can determine consecutively

- the pasts $\downarrow c_i$ for all branch cells $c_i \in \mathrm{BC}$,
- the unions $\bigcup_{L(c_j)=L(c_i)} \downarrow c_j$ for every branch point $c_i \in \mathrm{BC}$ and their complements $\mathcal{C}(\bigcup_{L(c_j)=L(c_i)} \downarrow c_j) = \bigcap_{L(c_j)=L(c_i)} \mathcal{C}(\downarrow c_j)$; such a complement is a pre-cubical complex since the cells c_i are maximal,
- the pre-cubical complex $X_i = \downarrow c_i \cup \mathcal{C}(\bigcup_{L(c_j)=L(c_i)} \downarrow c_j)$ for every branch cell c_i.

The next step is the investigation of higher order branch points and branch cells:

- A section $s\colon \mathrm{BP} \uparrow \mathrm{BC}$ fixes one maximal branch cell c_j^i for every branch point $b_i \in \mathrm{BP}$. Form the intersection subcomplexes $X_s \subset X$ via intersections of subgraphs of $\vec{\Gamma}(X)$.
- For each of these subcomplexes as point of departure, iterate to determine second order branch points and branches and associated subcomplexes. Iterate to determine higher order ones.

By recursion, we arrive at the set of *all* – ie, including higher order – branch points and branch cubes and thus to the objects of the category $\mathcal{M}(X)(\mathbf{x}_0, \mathbf{x}_1)$; moreover, for every such object (S, C) the associated nonbranching (!) subcomplex $X_{(S,C)} \subset X$.

To find out whether (S, C) is an object of $\mathcal{C}(X)(\mathbf{x}_0, \mathbf{x}_1)$,we have to investigate whether there is a d-path from $\mathbf{x}_0$ to $\mathbf{x}_1$ in $X_{(S,C)}$. Since this space is nonbranching, the future $\uparrow\mathbf{x}_0$ of $\mathbf{x}_0$ within it has a unique maximal element by Lemma 3.2. It is therefore enough to find out whether $\mathbf{x}_1$ is the *only* maximal vertex in $X_{(S,C)}$ or whether there is at least one other maximal vertex $\mathbf{v}$. Such a "deadlock" vertex $\mathbf{v}$ has no arrow with tail $\mathbf{v}$ in $\vec{\Gamma}(X_{(S,C)})$ in that subcomplex.

6.3 Final comments

Although each single of the steps to be taken is quite easy to implement, the number of steps can be enormous. In particular, if higher order branch points arise, the categories $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1) \subseteq \mathcal{M}(X)(\mathbf{x}_0,\mathbf{x}_1)$ may be huge, even for a HDA X of moderate size. As in the semaphore case in Raussen [18; 19], it is enough to determine the *minimal* "dead" objects (S,C) with deadlocks in $X_{(S,C)}$. Still, the determination of the category $\mathcal{C}(X)(\mathbf{x}_0,\mathbf{x}_1)$ describing the homotopy type of the trace space of a pre-cubical complex may need a lot of time and memory.

References

[1] **R Brown**, **P J Higgins**, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Algebra 22 (1981) 11–41 MR621285

[2] **R Brown**, **P J Higgins**, *On the algebra of cubes*, J. Pure Appl. Algebra 21 (1981) 233–260 MR617135

[3] **P Bubenik**, *Simplicial models for concurrency*, from: "Proceedings of the workshop on Geometric and Topological Methods in Computer Science", Electronic Notes in Theoretical Comp. Sci. 283 (2012) 3–12

[4] **E W Dijkstra**, *Co-operating sequential processes*, from: "Programming languages", (F Genuys, editor), Academic Press, New York (1968) 43–110

[5] **U Fahrenberg**, **M Raussen**, *Reparametrizations of continuous paths*, J. Homotopy Relat. Struct. 2 (2007) 93–117 MR2369163

[6] **L Fajstrup**, *Dipaths and dihomotopies in a cubical complex*, Adv. in Appl. Math. 35 (2005) 188–206 MR2152887

[7] **L Fajstrup**, *Trace spaces of directed tori with rectangular holes*, Tech. Report R-2011-08, Department of Math. Sciences, Aalborg University (2011) Available at http://vbn.aau.dk/files/55669400/R_2011_08.pdf

[8] **L Fajstrup**, **E Goubault**, **E Haucourt**, **S Mimran**, **M Raussen**, *Trace spaces: An efficient new technique for state-space reduction*, from: "European Symposium on Programming 2012", (H Seidl, editor), Lecture Notes in Comp. Sci. 7211, Springer, Berlin (2012) 274–294

[9] **L Fajstrup**, **M Raußen**, **E Goubault**, *Algebraic topology and concurrency*, Theoret. Comput. Sci. 357 (2006) 241–278 MR2242768

[10] **M Farber**, *Zeros of closed* 1*–forms, homoclinic orbits and Lusternik–Schnirelman theory*, Topol. Methods Nonlinear Anal. 19 (2002) 123–152 MR1921889

[11] **M Farber**, *Topology of closed one-forms*, Math. Surveys and Monogr. 108, Amer. Math. Soc. (2004) MR2034601

[12] **M Grandis**, *Directed homotopy theory, I*, Cah. Topol. Géom. Différ. Catég. 44 (2003) 281–316 MR2030049

[13] **M Grandis**, *Directed algebraic topology: Models of non-reversible worlds*, New Math. Monogr. 13, Cambridge Univ. Press (2009) MR2562859

[14] **D Kozlov**, *Combinatorial algebraic topology*, Algorithms and Computation in Math. 21, Springer, Berlin (2008) MR2361455

[15] **V Pratt**, *Modelling concurrency with geometry*, from: "Proc. of the 18th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", (D Wise, editor), ACM, New York (1991) 311–322

[16] **M Raussen**, *Reparametrizations with given stop data*, J. Homotopy Relat. Struct. 4 (2009) 1–5 MR2481616

[17] **M Raussen**, *Trace spaces in a pre-cubical complex*, Topology Appl. 156 (2009) 1718–1728 MR2521708

[18] **M Raussen**, *Simplicial models of trace spaces*, Algebr. Geom. Topol. 10 (2010) 1683–1714 MR2683749

[19] **M Raussen**, *Execution spaces for simple higher dimensional automata*, Appl. Algebra Engrg. Comm. Comput. 23 (2012) 59–84 MR2917117

[20] **C P Rourke**, **B J Sanderson**, *$\triangle$–sets I: Homotopy theory*, Quart. J. Math. Oxford Ser. 22 (1971) 321–338 MR0300281

[21] **D Sullivan**, *Infinitesimal computations in topology*, Inst. Hautes Études Sci. Publ. Math. (1977) 269–331 MR0646078

[22] **R J van Glabbeek**, *Erratum to: "On the expressiveness of higher dimensional automata"* [Theoret. Comput. Sci. 356 (2006) 265–290 MR2223695], Theoret. Comput. Sci. 368 (2006) 168–194 MR2274947

[23] **G Winskel**, **M Nielsen**, *Models for concurrency*, from: "Handbook of logic in computer science, Vol. 4", (S Abramsky, D M Gabbay, T S E Maibaum, editors), Oxford Univ. Press, New York (1995) 1–148 MR1365754

Department of Mathematical Sciences, Aalborg University
Fredrik Bajersvej 7G, DK-9220 Aalborg Øst, Denmark

raussen@math.aau.dk

http://people.math.aau.dk/~raussen

Received: 13 September 2011

J. Homotopy Relat. Struct.
DOI 10.1007/s40062-013-0045-4

Homology of spaces of directed paths on Euclidean cubical complexes

Martin Raussen · **Krzysztof Ziemiański**

Received: 28 February 2013 / Accepted: 6 June 2013

Abstract We compute the homology of the spaces of directed paths on a certain class of cubical subcomplexes of the directed Euclidean space $\mathbb{R}^n$ by a recursive process. We apply this result to calculate the homology and cohomology of the space of directed loops on the $(n-1)$-skeleton of the directed torus $\vec{T}^n$.

Keywords Directed paths · Cubical complex · Path space · Homology · Cohomology

Mathematics Subject Classification (2000) 55P10 · 55P15 · 55U10 · 68Q85

1 Introduction

One of the most important problems of directed algebraic topology is the determination of the homotopy type of spaces of directed paths $\vec{P}(X)_x^y$ between two points x, y of a directed space X. This problem seems to be difficult in general; however several

---

Dedicated to Hvedri Inassaridze on his 80th birthday

Communicated by Ronald Brown.

The authors gratefully acknowledge support from the European Science Foundation network "Applied and Computational Algebraic Topology" that allowed them to collaborate on this paper during mutual visits made possible by grants 4671 and 5432.

M. Raussen ( )
Department of Mathematical Sciences, Aalborg University,
Fredrik Bajersvej 7G, 9220 Aalborg Øst, Denmark
e-mail: raussen@math.aau.dk

K. Ziemiański
Faculty of Mathematics Informatics and Mechanics, Warsaw University,
Banacha 2, 02-097 Warszawa, Poland
e-mail: ziemians@mimuw.edu.pl

results were obtained recently. The first author gave in a series of papers [12–15] a description of the homotopy type $\vec{P}(X)_x^y$ in the case where X is a directed cube from which collections of homothetic rectangular areas are removed. An alternative description is given in the paper [18] of the second author. The relevant path spaces are shown to be homotopy equivalent to either a simplicial complex or a cubical complex. Even in greater generality, such path spaces have the homotopy type of a CW-complex.

In this paper, we present explicit calculations of the homology and cohomology of directed path spaces in important particular cases in which path spaces can be described as homotopy colimits over simple combinatorial categories; this makes it possible to apply inductive methods.

1.1 d-spaces

For a topological space X, let $P(X) = X^I$ denote the space of all paths in X endowed with the compact-open topology. A d-space [8,9] is a pair $(X, \vec{P}(X))$, where X is a topological space, and $\vec{P}(X) \subseteq P(X)$ is a family of paths on X that contains all constant paths and that is closed under non-decreasing reparametrizations and concatenations. The family $\vec{P}(X)$ is called *a d-structure* on X, and paths which belong to $\vec{P}(X)$ will be called *directed paths* or d-paths. For $x, y \in X$ define *the directed path space* from x to y as

$$\vec{P}(X)_x^y = \{\alpha \in \vec{P}(X) :\ \alpha(0) = x \ \wedge\ \alpha(1) = y\}. \tag{1.1}$$

The *directed real line* $\vec{\mathbb{R}}$ is the d-space with underlying space $\mathbb{R}$ and $\vec{P}(\vec{\mathbb{R}})$ the set of all *non-decreasing* paths. *Directed Euclidean space* $\vec{\mathbb{R}}^n$ is the product $\vec{\mathbb{R}} \times \cdots \times \vec{\mathbb{R}}$ with the product d-structure $\vec{P}(\vec{\mathbb{R}}^n) = \vec{P}(\vec{\mathbb{R}}) \times \cdots \times \vec{P}(\vec{\mathbb{R}})$. Finally, the *directed torus* $\vec{T}^n$ is the quotient $\vec{\mathbb{R}}^n/\mathbb{Z}^n$; a path on $\vec{T}^n$ is directed iff it lifts to a directed path on $\vec{\mathbb{R}}^n$.

1.2 Euclidean cubical complexes

An elementary cube in $\mathbb{R}^n$ is a product

$$[k_1, k_1 + e_1] \times \cdots \times [k_n, k_n + e_n] \subseteq \mathbb{R}^n,$$

where $k_i \in \mathbb{Z}$ and $e_i \in \{0, 1\}$; the dimension of a cube is the sum $\sum_1^n e_i$. A *Euclidean cubical complex* is defined to be a subset $K \subseteq \mathbb{R}^n$ that is a union of elementary cubes. The *d-skeleton of* K, denoted by $K_{(d)}$, is the union of all elementary cubes contained in K which have dimensions less than or equal to d. Euclidean space can be identified with the geometric realization of a suitable pre-cubical set such that realizations of cubes of this pre-cubical set are elementary cubes in $\mathbb{R}^n$, and Euclidean cubical sets are the geometric realizations of pre-cubical subsets of that pre-cubical set. Every cubical complex is provided with the directed structure inherited from $\vec{\mathbb{R}}^n$.

Euclidan cubical complexes are special cases of general cubical complexes which are geometric realizations of general pre-cubical sets; cf [1] and the recent book [2].

1.3 Notation

Points on $\mathbb{R}^n$ will be denoted by bold letters, their coordinates by regular ones with suitable indices; for example $\mathbf{a} = (a_1, \ldots, a_n)$. Furthermore, we will write $\mathbf{0}$ for $(0, \ldots, 0)$ and $\mathbf{1}$ for $(1, \ldots, 1)$. Three kinds of comparison operators between points of $\mathbb{R}^n$ will be used:

$$\begin{aligned} \mathbf{a} \leqslant \mathbf{b} &\Leftrightarrow \forall_{i=1}^n \; a_i \leqslant b_i \\ \mathbf{a} < \mathbf{b} &\Leftrightarrow \forall_{i=1}^n \; a_i \leqslant b_i \;\wedge\; \mathbf{a} \neq \mathbf{b} \\ \mathbf{a} \ll \mathbf{b} &\Leftrightarrow \forall_{i=1}^n \; a_i < b_i. \end{aligned}$$

In analogy to the one-dimensional case, write $[\mathbf{a}, \mathbf{b}] := \{\mathbf{t} : \mathbf{a} \leqslant \mathbf{t} \leqslant \mathbf{b}\}$ for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Finally, let $|\mathbf{x}| := \sum_{i=1}^n |x_i|$ for $\mathbf{x} \in \mathbb{R}^n$ denote the l_1-norm; the l_1-metric $\mu(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ on $\mathbb{R}^n$ is compatible with standard Euclidean topology. Notice that $|\mathbf{x} - \mathbf{y}| = ||\mathbf{x}| - |\mathbf{y}||$ whenever $\mathbf{x} \leqslant \mathbf{y}$.

1.4 The main theorem

Let $\mathbf{k} \in \mathbb{Z}^n$, $n \geqslant 3$, and let $K \subseteq [\mathbf{0}, \mathbf{k}] \subseteq \mathbb{R}^n$ denote a Euclidean cubical complex that *contains* the $(n-1)$-skeleton $[\mathbf{0}, \mathbf{k}]_{(n-1)} \subseteq K$. A *cube sequence in K of length r* is a sequence

$$[\mathbf{a}^*] := [\mathbf{0} \ll \mathbf{a}^1 \ll \mathbf{a}^2 \ll \cdots \ll \mathbf{a}^r \leqslant \mathbf{k}], \tag{1.2}$$

where $\mathbf{a}^i \in \mathbb{Z}^n$ and such that $[\mathbf{a}^i - \mathbf{1}, \mathbf{a}^i] \not\subseteq K$. Let $CS_r(K)$ be the set of cube sequences of length r and define the graded abelian group $A_*(K)$ by

$$A_m(K) = \begin{cases} \mathbb{Z}[CS_{m/(n-2)}(K)] & \text{if } n-2 \text{ divides } m \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

The main theorem of this paper is the following

Theorem 1.1 *The graded abelian groups $H_*(\vec{P}(K))_{\mathbf{0}}^{\mathbf{k}}$ and $A_*(K)$ are isomorphic.*

1.5 An application: the space of paths on the $(n-1)$-skeleton of the directed torus $\vec{T}^n$

Assume that $n \geqslant 3$ and $d \geqslant 2$. Let $\vec{T}^n_{(d)} := \vec{\mathbb{R}}^n_{(d)}/\mathbb{Z}^n$ be the d-skeleton of $\vec{T}^n$. Every directed path $\alpha \in \vec{P}(\vec{T}^n_{(d)})_{\mathbf{0}}^{\mathbf{0}}$ represents a class $\mathbf{k} \in \pi_1(\vec{T}^n_{(d)})_{\mathbf{0}} \cong \mathbb{Z}^n$. By passing to the universal covering of $T^n_{(d)}$ we see that α lifts uniquely to a path $\tilde{\alpha} \in \vec{P}(\mathbb{R}^n_{(d)})_{\mathbf{0}}^{\mathbf{k}}$. Since $\tilde{\alpha}$ is directed, the class $\mathbf{k}$ is non-negative: $\mathbf{0} \leq \mathbf{k}$. Since also directed homotopies lift uniquely, we obtain a homeomorphism

$$\vec{P}(T^n_{(d)})_{\mathbf{0}}^{\mathbf{0}} \cong \coprod_{\mathbf{0} \leqslant \mathbf{k} \in \mathbb{Z}^n} \vec{P}(\mathbb{R}^n_{(d)})_{\mathbf{0}}^{\mathbf{k}}. \tag{1.4}$$

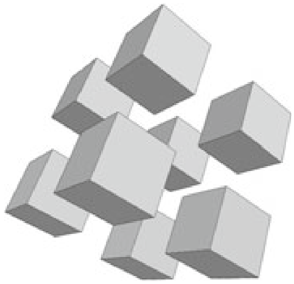


Fig. 1 The space $\mathbf{R}^3_{(2)} \cap [\mathbf{0}, \mathbf{2}]$ is homotopy equivalent to the complement of eight cubes within an outer cube $[\mathbf{0}, \mathbf{2}] \subset \mathbf{R}^3$. All figures courtesy to E. Haucourt and A. Lang

If $d = n - 1$, we can apply the main theorem to obtain an isomorphism (Fig. 1)

$$H_*(\vec{P}(T^n_{(n-1)})^{\mathbf{0}}_{\mathbf{0}}) \cong \bigoplus_{\mathbf{0} \leqslant \mathbf{k} \in \mathbb{Z}^n} A_*(\mathbb{R}^n_{(n-1)} \cap [\mathbf{0}, \mathbf{k}]). \tag{1.5}$$

The following proposition allows to calculate the Betti numbers of the components:

Proposition 1.2 *For* $\mathbf{k} = (k_1, \ldots, k_n) \geqslant \mathbf{0}$ *we get:*

$$\dim H_{r(n-2)}(\vec{P}(\mathbb{R}^n_{(n-1)})^{\mathbf{k}}_{\mathbf{0}}) = \dim A_{r(n-2)}(\mathbb{R}^n_{(n-1)} \cap [\mathbf{0}, \mathbf{k}]) = \binom{k_1}{r}\binom{k_2}{r} \cdots \binom{k_n}{r}.$$

Proof The map

$$CS_r(\mathbb{R}^n_{(n-1)} \cap [\mathbf{0}, \mathbf{k}]) \ni [\mathbf{a}^1 \ll \cdots \ll \mathbf{a}^r \leqslant \mathbf{k}]$$
$$\mapsto (\{a^1_1, \ldots, a^r_1\}, \ldots, \{a^1_n, \ldots, a^r_n\}) \in 2^{\{1,\ldots,k_1\}} \times \cdots \times 2^{\{1,\ldots,k_n\}}$$

is clearly a bijection. The conclusion follows. □

Remark An attempt to calculate the homology of $\vec{P}(\mathbb{R}^3_2)^{(k,l,m)}_{\mathbf{0}}$ using the poset description for the cell complex of the prod-simplicial complex homotopy equivalent to that path space according to Raussen [13] by "brute force"—even using sophisticated homology software—failed already for $k = l = m = 3$. The prod-simplicial complex in this case has dimension $klm(n-2)$; its homological dimension is only $\min\{k, l, m\}(n-2)$. This contrast was one of the motivations for looking for better descriptions of path spaces.

1.6 Euclidean cubical complexes and concurrency

One of the motivations for developing directed algebraic topology goes back to particular models in concurrency theory, the so-called higher dimensional automata, cf, e.g. [10, 17]. A particular class of higher dimensional automata arises from semaphore or mutex models: Each processor records on a time line when it accesses (P) and relinquishes (V) a number of shared objects; the forbidden region F associated to such a PV-program (cf [4]) consists of a union of isothetic hyperrectangles $R^i \subset \vec{I}^n$ within an n-cube $\vec{I}^n \subset \vec{\mathbb{R}}^n$; cf [7].

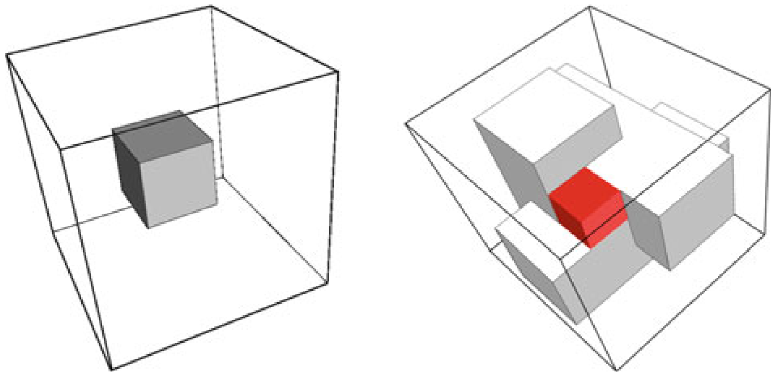


Fig. 2 *Left* Forbidden region corresponding to one semaphore of arity 2 in the presence of 3 processors. *Right* Forbidden (and unsafe) region corresponding to 3 dining philosophers, cf [5]

The particular Euclidean complexes whose path spaces we study correspond to PV-programs with the following two particular properties:

- All shared objects have arity $n-1$, ie, $n-1$ out of n but not all n processes can access the object at any given time (Fig. 2);
- The PV-program for every indiviudal processor is of type $(PV)(PV)\cdots(PV)$—a variety of shared objects is allowed. In particular, every access to a shared object is terminated before a new one is accessed. This has the consequence that the hyperretangles R^i and their projections to the axes do not overlap with each other.

No doubt that this represents a very particular case. On the other hand, our result seems to be the first non-trivial calculation of the homology of spaces of directed paths in closed form. Note that a description of a simplicial complex homotopy equivalent to directed paths in a torus with holes was obtained in Fajstrup [6].

The application in Sect. 1.5—which motivated this line of investigation—shows that it is also possible to consider programs with loops. The case considered here corresponds to n looped processors of type $(PaVa)^* \parallel \cdots \parallel (PaVa)^*$.

1.7 The case $n = 2$

Only for $n = 2$, the path spaces are, in general, not connected— and therefore the result of a distributed programme may depend on the schedule. The method described above still works, but there is a slight twist due to the fact that cube sequences regardless of their length all contribute (only) to dimension 0:

The space $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ is a space consisting of contractible components. The number of components is $\beta_0(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) = |CS(K)| + 1$, the number of *all* cube sequences in K augmented by one. The reason is that both sides in the equation above obey to the recursion formula $a_{k+1,l+1} = \begin{cases} a_{k+1,l} + a_{k,l+1} - a_{kl} & [(k,l),(k+1,l+1)] \subset K \\ a_{k+1,l} + a_{k,l+1} & [(k,l),(k+1,l+1)] \not\subset K \end{cases}$ with start values $a_{k,0} = a_{0,l} = 1$.

In the particular case dealt with in Sect. 1.5, we obtain: $\beta_0(\vec{P}(\mathbb{R}^2_1)_{\mathbf{0}}^{(k,l)}) = \binom{k+l}{k}$.

1.8 Outline of the paper

In Sect. 2 we construct, for an arbitrary Euclidean cubical complex K, a homotopy equivalence between $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ and a certain homotopy colimit of spaces which are homotopy equivalent to one of the "smaller" spaces $\vec{P}(K)_{\mathbf{0}}^{\mathbf{l}}$ for $\mathbf{l} < \mathbf{k}$. In Sect. 3, we construct a homomorphism $A_*(K) \to H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$. Then we prove, under the assumption that $[\mathbf{0}, \mathbf{k}]_{(n-1)} \subseteq K$, that this homomorphism is actually an isomorphism. The proof is inductive and uses the homotopy colimit description from Sect. 2. In Sect. 4, we determine the cohomology ring structure $H^*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$.

2 A recursive description of path spaces

In this section we construct a presentation of directed paths spaces on a Eulidean cubical complex as a homotopy colimit of path spaces of certain subcomplexes. Fix $\mathbf{0} \ll \mathbf{k} \in \mathbb{Z}^n$ and a cubical complex $K \subseteq [\mathbf{0}, \mathbf{k}]$.

2.1 A transversal section

Fix $\varepsilon \in (0, 1)$. Let $S(K) = \{\mathbf{x} \in K :\ |\mathbf{x}| = |\mathbf{k}| - \varepsilon\} \subset S = \{\mathbf{x} \in \mathbb{R}^n |\ |\mathbf{x}| = |\mathbf{k}| - \varepsilon\}$.

Proposition 2.1 *For every path $\alpha \in \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ there exists a unique $s(\alpha) \in S(K)$ belonging to the image of α. Moreover, the map $s : \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \to S(K)$ is continuous (with respect to the compact-open topology on $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$).*

Proof Since $|\alpha(0)| = 0$, $|\alpha(1)| = |\mathbf{k}|$, there exists t_α such that $\alpha(t_\alpha) = |\mathbf{k}| - \varepsilon$. If $|\alpha(t_\alpha)| = |\alpha(t'_\alpha)|$ for $t_\alpha < t'_\alpha$, then $\alpha(t_\alpha) \leq \alpha(t'_\alpha)$. This implies that $|\alpha(t_\alpha) - \alpha(t'_\alpha)| = ||\alpha(t_\alpha)| - |\alpha(t'_\alpha)|| = 0$. Therefore $s(\alpha)$ is uniquely determined. In the parlance of Raussen [11], the subset $S(X)$ is both achronal and unavoidable from $\mathbf{0}$ to $\mathbf{k}$.

To prove the continuity of s, it is sufficient to show that the inverse images of open balls $B(\mathbf{x}, r) = \{\mathbf{y} \in K :\ |\mathbf{x} - \mathbf{y}| < r\}$ are open in $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$: Fix $\mathbf{x} \in K$, $r > 0$, $\alpha \in s^{-1}(B(\mathbf{x}, r))$ and $t_\alpha \in I$ such that $\alpha(t_\alpha) = s(\alpha)$. Let $\beta \in \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ be a path such that $|\beta(t_\alpha) - s(\alpha)| < r'$, where $r' = (r - |\mathbf{x} - s(\alpha)|)/2$. Since $s(\beta)$ and $\beta(t_\alpha)$ are comparable, we have

$$|\beta(t_\alpha) - s(\beta)| = ||\beta(t_\alpha)| - |s(\beta)|| = ||\beta(t_\alpha)| - |s(\alpha)|| < r'.$$

Finally,

$$|\mathbf{x} - s(\beta)| \leqslant |\mathbf{x} - s(\alpha)| + |s(\alpha) - \beta(t_\alpha)| + |\beta(t_\alpha) - s(\beta)| < |\mathbf{x} - s(\alpha)| + 2r' < r,$$

i. e. $\beta \in s^{-1}(B(\mathbf{x}, r))$. Hence the set $\{\beta \in \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} :\ \beta(t_\alpha) \in B(s(\alpha), r')\}$ is an open neighbourhood of α contained in $s^{-1}(B(\mathbf{x}), r)$. □

2.2 A description of $S(K)$

The map

$$R : S \ni \mathbf{t} \mapsto \varepsilon^{-1}(\mathbf{k} - \mathbf{t}) \in \{\mathbf{x} : \ \mathbf{x} \geqslant 0 \ \wedge \ |\mathbf{x}| = 1\} = |\Delta^{n-1}|$$

is a homeomorphism from S to the standard simplex Δ^{n-1}. It maps $S(K)$ homeomorphically onto a simplicial subcomplex $\Delta_K \subset \Delta^{n-1}$.

The category Δ_{n-1}^{op} of subsimplices of Δ_{n-1} can be identified by an isomorphism of categories with the (inverse) poset category $\mathcal{J}_{n-1}$ of sequences $\mathbf{j} \in \{0, 1\}^n$ with $\mathbf{j} \neq \mathbf{0}$. Such a sequence $\mathbf{j}$ corresponds to the subsimplex

$$\Delta_{\mathbf{j}} = \{\mathbf{t} \in \Delta_{n-1} : \ \forall_{i=1}^{n} \quad j_i = 0 \Rightarrow t_i = 0\} \subseteq \Delta_{n-1}.$$

The morphism $\mathbf{j} \to \mathbf{j}'$ (for every $\mathbf{j} \geqslant \mathbf{j}'$) corresponds to the inclusion $\Delta_{\mathbf{j}} \subset \Delta_{\mathbf{j}'}$.

The restriction of this correspondence to the category of subsimplices of Δ_K provides an isomorphism between that category and the full subcategory $\mathcal{J}_K \subset \mathcal{J}_{n-1}$ with objects

$$\mathrm{Ob}(\mathcal{J}_K) := \{\mathbf{0} < \mathbf{j} \in \{0, 1\}^n : \ \mathbf{j} \subseteq \Delta_K\} = \{\mathbf{j} \in \{0, 1\}^n : \ [\mathbf{k} - \mathbf{j}, \mathbf{k}] \subseteq K\}. \quad (2.1)$$

Two cases will be of particular importance:

$$\mathcal{J}_K = \begin{cases} \mathcal{J}_{n-1} & [\mathbf{k} - \mathbf{1}, \mathbf{k}] \subset K; \\ \hat{\mathcal{J}}_{n-1} := \mathcal{J}_{n-1} \backslash \{\mathbf{1}\} & [\mathbf{k} - \mathbf{1}, \mathbf{k}] \cap K = \partial[\mathbf{k} - \mathbf{1}, \mathbf{k}]. \end{cases}$$

2.3 A cover of $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$

The geometric realization of $S(K)$ can be covered by stars of its vertices and this cover lifts to a cover of $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$. For every $\mathbf{j} \in \mathcal{J}_K$ define

$$F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} := (R \circ s)^{-1}(\mathrm{st}(\mathbf{j})) = \{\alpha \in \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} : \ \forall_{i=1}^{n} \ j_i = 1 \Rightarrow s(\alpha)_i < k_i\}.$$

The spaces $F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ clearly cover all of $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$. The cover $\{F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}\}$ is closed under intersections since $F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \cap F_{\mathbf{j}'} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} = F_{\mathbf{j} \cup \mathbf{j}'} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$, where $(\mathbf{j} \cup \mathbf{j}')_i = \max\{j_i, j_i'\}$. Moreover the category associated with this cover is precisely $\mathcal{J}_K$. As a consequence, cf [16, Proposition 4.1], the inclusions $F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \subseteq \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ induce a homotopy equivalence

$$\mathrm{hocolim}_{\mathbf{j} \in \mathcal{J}_K} F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \longrightarrow \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} = \mathrm{colim}_{\mathbf{j} \in \mathcal{J}_K} F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}. \quad (2.2)$$

The next step is to prove that $F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ is homotopy equivalent to $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}$. This will enable us to use the decomposition (2.2) for inductive calculations of path spaces. But first, we need some technical lemmas which will be presented in a greater generality.

2.4 Past deformation retractions

Definition 2.2 Let X be a d-space with a subspace $Y \subseteq X$. *A past deformation retraction* of X onto Y is a d-map $d : X \times \vec{I} \to X$ (preserving d-structures; cf [9]) such that

- $r(x) := d(x, 0) \in Y$ for every $x \in X$,
- $d(x, 1) = x$ for every $x \in X$,
- $d(y, t) = y$ for every $y \in Y$ and every $t \in I$.

Proposition 2.3 *If $d : X \times \vec{I} \to X$ is a past deformation retraction on Y, then for every $x \in X$ and $y \in Y$ the maps*

$$F : \vec{P}(Y)_y^{d(x,0)} \ni \alpha \mapsto \alpha * d(x, -) \in \vec{P}(X)_y^x$$
$$G : \vec{P}(X)_y^x \ni \alpha \mapsto d(\alpha, 0) \in \vec{P}(Y)_y^{d(x,0)}$$

are mutually inverse homotopy equivalences.

Proof A homotopy H between $Id_{\vec{P}(Y)_y^{d(x,0)}}$ and $G \circ F$ is given by the formula

$$H(\alpha, s)(t) = \begin{cases} \alpha(t(1 - s/2)^{-1}) & \text{for } 0 \leqslant t \leqslant 1 - s/2 \\ d(x, 0) & \text{for } 1 - s/2 \leqslant t \leqslant 1, \end{cases}$$

and a homotopy H' between $Id_{\vec{P}(X)_y^x}$ and $F \circ G$ by

$$H'(\beta, s)(t) = \begin{cases} d(\beta(t(1 - s/2)^{-1}), 1 - s) & \text{for } 0 \leqslant t \leqslant 1 - s/2 \\ d(x, 2t - 1) & \text{for } 1 - s/2 \leqslant t \leqslant 1. \end{cases}$$

□

2.5 $F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ up to homotopy

For $\mathbf{j} \in \mathcal{J}_K$ define

$$X_{\mathbf{j}} := \{\mathbf{t} \in K : \forall_{i:\mathbf{j}_i=1}\ t_i < k_i\} \tag{2.3}$$
$$K_{\mathbf{j}} := \{\mathbf{t} \in K : \mathbf{t} \leqslant \mathbf{k} - \mathbf{j}\} = K \cap [\mathbf{0}, \mathbf{k} - \mathbf{j}], \tag{2.4}$$

and let $\bar{X}_{\mathbf{j}}$ be the closure of $X_{\mathbf{j}}$ in K.

Proposition 2.4 *$K_{\mathbf{j}}$ is a past deformation retract of $\bar{X}_{\mathbf{j}}$.*

Proof Every $\mathbf{t} \in \bar{X}_{\mathbf{j}}$ belongs to some cube $[\mathbf{c}, \mathbf{d}] \subseteq K$, $\mathbf{c}, \mathbf{d} \in \mathbb{Z}^n$ whose interior is contained in $X_{\mathbf{j}}$. It implies that $c_i = k_i - 1$ for every i such that $d_i = k_i$ and $j_i = 1$. Define the retraction $r^{\mathbf{j}} : \bar{X}_{\mathbf{j}} \to K_{\mathbf{j}}$ by the formula

$$r_i^{\mathbf{j}}(\mathbf{t}) = \begin{cases} t_i & \text{for } t_i \leqslant k_i - j_i \\ k_i - j_i & \text{for } t_i \geqslant k_i - j_i, \end{cases}$$

and the deformation between $r^{\mathbf{j}}$ and identity by convex combination. Since both $\mathbf{t}$ and $r^{\mathbf{j}}(\mathbf{j})$ belong to the cube $[\mathbf{c}, \mathbf{d}]$ the map $r^{\mathbf{j}}$ is well-defined. □

Proposition 2.5 *The image $\alpha(I)$ associated to any path $\alpha \in F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ is contained in $\bar{K}_{\mathbf{j}}$.*

Proof Let $\alpha \in F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ and let $t_\alpha \in I$ satisfies $\alpha(t_\alpha) = s(\alpha)$. For $t \leqslant t_\alpha$, we have $\alpha(t)_i \leqslant \alpha(t_\alpha)_i < k_i$ whenever $\mathbf{j}_i = 1$. Then $\alpha([0, t_\alpha]) \subseteq K_{\mathbf{j}} \subseteq \bar{K}_{\mathbf{j}}$. If $t > t_\alpha$, then $\alpha(t) \in [\mathbf{k} - \mathbf{j}', \mathbf{k}]$, where $[\mathbf{k} - \mathbf{j}', \mathbf{k}]$ is a minimal cube containing $s(\alpha)$ in its interior. As a consequence, $\alpha([t_\alpha, 1]) \subseteq \bar{K}_{\mathbf{j}}$ and hence $F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \subseteq \vec{P}(\bar{K}_{\mathbf{j}})_{\mathbf{0}}^{\mathbf{k}}$. □

For every cube $[\mathbf{c}, \mathbf{d}] \subseteq K$ let $i_{\mathbf{c}}^{\mathbf{d}} : \vec{P}(K)_{\mathbf{0}}^{\mathbf{c}} \to \vec{P}(K)_{\mathbf{0}}^{\mathbf{d}}$ denote the concatenation with the linear path $t \mapsto (1 - t)\mathbf{c} + t\mathbf{d}$. Note that $i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}}(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}) \subseteq F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ for $\mathbf{j} \in \mathcal{J}_K$.

Proposition 2.6 *For every $\mathbf{j} \in \mathcal{J}_K$ the map $i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}} : \vec{P}(K_{\mathbf{j}})_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}} \to F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ is a homotopy equivalence. Moreover, for every morphism $\mathbf{j}' \to \mathbf{j}$ in $\mathcal{J}_K$ the diagram*

$$\begin{array}{ccc}
\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}'} & \xrightarrow{i_{\mathbf{k}-\mathbf{j}'}^{\mathbf{k}-\mathbf{j}}} & \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}} \\
\big\downarrow {\scriptstyle i_{\mathbf{k}-\mathbf{j}'}^{\mathbf{k}}} & & \big\downarrow {\scriptstyle i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}}} \\
F_{\mathbf{j}'}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} & \xrightarrow{\subseteq} & F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}
\end{array} \tag{2.5}$$

commutes up to homotopy.

Proof The map $i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}}$ is a homotopy equivalence by Proposition 2.3. The commutativity of the diagram is obvious from the definitions. □

2.6 Specific path spaces

2.6.1 Boundary of a cube

Let $\mathbf{k} = \mathbf{1}$ and let $K = [\mathbf{0}, \mathbf{1}]_{(n-1)}$. Then $S(K) \simeq \partial\Delta^{n-1}$ and $\mathcal{J}_K = \mathcal{J}_{n-1} \setminus \{\mathbf{1}\}$ with objects $\{\mathbf{j} \in \{0, 1\}^n : 0 < \sum_{i=1}^{n} j_i < n\}$. Furthermore, for every such $\mathbf{j}$, $F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{1}} \simeq \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}} = \vec{P}[\mathbf{0}, \mathbf{k} - \mathbf{j}]_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}$ is contractible since it contains $\{\mathbf{0}\}$ as a past deformation retract, cf Proposition 2.3. As a consequence,

$$\vec{P}(\mathbb{R}^n_{(n-1)})_{\mathbf{0}}^{\mathbf{1}} \cong \operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}(K)} F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}} \simeq N\mathcal{J}(K) \cong \partial\Delta^{n-1} \simeq S^{n-2}.$$

Remark This result is also an immediate consequence of Raussen [13, Corollary 4.12].

For the remaining part of the paper, we fix a generator $x_{\mathbf{1}} \in H_{n-2}(\vec{P}([\mathbf{0}, \mathbf{1}]_{(n-1)})_{\mathbf{0}}^{\mathbf{1}}) \cong H_{n-2}(|N\mathcal{J}|)$. By shifting, we obtain the generators $x_{\mathbf{k}} \in H_{n-2}(\vec{P}([\mathbf{k}-\mathbf{1}, \mathbf{k}]_{(n-1)})_{\mathbf{k}-\mathbf{1}}^{\mathbf{k}})$ for $\mathbf{k} \in \mathbb{Z}^n$.

2.6.2 Connectivity of certain path spaces

Proposition 2.7 *If a subcomplex* $K \subseteq [\mathbf{0}, \mathbf{k}]$ *contains the* 2*-skeleton of* $[0, \mathbf{k}]$*, then* $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$ *is connected.*

Proof This is obvious if either $\mathbf{k} = \mathbf{0}$, or $n = 2$. Assume that the conclusion holds for all complexes $K' \subseteq [\mathbf{0}, \mathbf{k}']$, $\mathbf{k}' \in \mathbb{Z}^{n'}$, such that $n' \leqslant n$ or $n' = n$ and $\mathbf{k}' < \mathbf{k}$.
For $\mathbf{k} > \mathbf{0}$, $S(K) \subseteq \Delta^{n-1}$ contains the 1-skeleton of Δ^{n-1} and is therefore connected. Then

$$\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \simeq \operatorname{hocolim}_{\mathbf{j} \in \mathcal{J}_K} F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$$

is connected, because it is a homotopy colimit of connected spaces $F_{\mathbf{j}} \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} \simeq \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}} = \vec{P}(K \cap [\mathbf{0}, \mathbf{k} - \mathbf{j}])_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}$ (by the inductive assumption) over a connected category $\mathcal{J}_K$. □

3 Homology of the path space $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$

Fix $n \geqslant 3, \mathbf{k} \in \mathbb{Z}^n$. Let $K \subseteq [\mathbf{0}, \mathbf{k}] \subseteq \mathbb{R}^n$ denote a Euclidean cubical complex which contains the $(n-1)$-skeleton of $[\mathbf{0}, \mathbf{k}]$. We will define a homomorphism $\Phi_K : A_*(K) \to H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$ from the graded abelian group $A_*(K)$ defined in Sect. 1.4 into the homology of the path space and prove that it is an isomorphism.

3.1 The homomorphism Φ_K

3.1.1 Definitions

For every cube sequence $\mathbf{a}^* = [\mathbf{0} \ll \mathbf{a}^1 \ll \cdots \ll \mathbf{a}^r] \in CS_r(K)$ in K choose paths $\beta_i \in \vec{P}(K)_{\mathbf{a}^i}^{\mathbf{a}^{i+1}-\mathbf{1}}$, $i = 0, \ldots, r$ (we assume $\mathbf{a}^0 = \mathbf{0}$, $\mathbf{a}^{r+1} = \mathbf{k} + \mathbf{1}$). Let $c(\mathbf{a}^*)$ be the following concatenation map

$$c(\mathbf{a}^*) : \prod_{j=1}^{r} \vec{P}(K)_{\mathbf{a}^j - \mathbf{1}}^{\mathbf{a}^j} \ni (\alpha_j) \mapsto \beta_0 * \alpha_1 * \beta_1 * \cdots * \alpha_r * \beta_r \in \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}.$$

Then define Φ_K on generators by

$$\Phi_K([\mathbf{a}^1 \ll \cdots \ll \mathbf{a}^r]) := c(\mathbf{a}^*)_*(x_{\mathbf{a}^1} \times x_{\mathbf{a}^2} \times \cdots \times x_{\mathbf{a}^r}) \in H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}), \quad (3.1)$$

and extend as a homomorphism.

The element $\mathbf{x}_{\mathbf{a}^i} \in H_{n-2}(\vec{P}(K)_{\mathbf{a}^j-\mathbf{1}}^{\mathbf{a}^j}) \cong H_{n-2}(\vec{P}(\partial[\mathbf{a}^j - \mathbf{1}, \mathbf{a}^j])_{\mathbf{a}^j-\mathbf{1}}^{\mathbf{a}^j})$ is a generator chosen as in Sect. 2.6.1. According to Proposition 2.7, the map $c(\mathbf{a}^*)$ does not depend (up to homotopy) on the choice of the paths β_i, and this implies that Φ_K is well-defined.

3.1.2 Naturality

Let $\mathbf{l} \leqslant \mathbf{k}$ and let $L \subseteq [\mathbf{0}, \mathbf{l}]$ be a cubical complex such that $[\mathbf{0}, \mathbf{l}]_{(n-1)} \subseteq L \subseteq K \cap [\mathbf{0}, \mathbf{l}]$. With respect to the homomorphism given by

$$\varphi_L^K : A_*(L) \ni [\mathbf{a}^*] \mapsto \begin{cases} [\mathbf{a}^*] & \text{if } [\mathbf{a}^*] \text{ is a cube sequence in } K \\ 0 & \text{otherwise.} \end{cases} \in A_*(K),$$

the homomorphism Φ_K is natural in the following sense:

Proposition 3.1 *The diagram*

$$\begin{array}{ccc} A_*(L) & \xrightarrow{\varphi_L^K} & A_*(K) \\ \downarrow{\scriptstyle \Phi_L} & & \downarrow{\scriptstyle \Phi_K} \\ H_*(\vec{P}(L)_{\mathbf{0}}^{\mathbf{l}}) & \longrightarrow & H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \end{array} \qquad (3.2)$$

is commutative. The bottom map is induced by the concatenation with a fixed directed path $\alpha \in \vec{P}(K)_{\mathbf{l}}^{\mathbf{k}}$.

Proof Straightforward from the definitions. □

3.2 The main theorem

The main result of this section is the following

Theorem 3.2 *For every Euclidean cubical complex* $[\mathbf{0}, \mathbf{k}]_{(n-1)} \subseteq K \subseteq [\mathbf{0}, \mathbf{k}]$, $\mathbf{k} \in \mathbb{Z}^n$, $\mathbf{k} \geq \mathbf{0}$, $n > 2$, *the homomorphism* $\Phi_K : A_*(K) \to H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$ *is an isomorphism of graded abelian groups.*

The proof is by induction on $\mathbf{k}$. To start the induction, notice that if $\prod_i k_i = 0$, then both $H_*(K)_{\mathbf{0}}^{\mathbf{k}}$ and $A_*(K)$ are isomorphic to $(\mathbb{Z}, 0)$, the homology of a point, since $[\mathbf{0}, \mathbf{k}]_{(n-1)} \subseteq K$. Let us assume that Theorem 3.2 is valid for all Euclidean cubical complexes contained in $[\mathbf{0}, \mathbf{l}]$ for $\mathbf{l} < \mathbf{k}$.

Since K is assumed to contain the $(n-1)$-skeleton of $[\mathbf{0}, \mathbf{k}]$, there are only two cases to consider: either $[\mathbf{k}-\mathbf{1}, \mathbf{k}]$ is contained in K, or it is not; in that case $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \cap K = \partial[\mathbf{k}-\mathbf{1}, \mathbf{k}]$. For simplicity, we will write $\mathcal{J} = \mathcal{J}_{n-1}$, resp. $\hat{\mathcal{J}} = \hat{\mathcal{J}}_{n-1}$ for the relevant categories; cf Sect. 2.2. Let $\mathbf{Ab}_*$ be the category of graded abelian groups.

3.2.1 The case $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \subseteq K$

The objects of the category $\mathcal{J}$ are all n-tuples $\mathbf{0} < \mathbf{j} \in \{0, 1\}^n$, cf (2.1). For $\mathbf{j} \in \{0, 1\}^n$ denote $K_{\mathbf{j}} := K \cap [\mathbf{0}, \mathbf{k} - \mathbf{j}]$. Notice that for any morphism $\mathbf{j} \to \mathbf{j}'$ in $\mathcal{J}$, the homomorphisms

$$\begin{aligned} \varphi_{K_{\mathbf{j}}}^{K_{\mathbf{j}'}} &: A_*(K_{\mathbf{j}}) \to A_*(K_{\mathbf{j}'}) \\ (i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}-\mathbf{j}'})_* &: H_*(\vec{P}(K_{\mathbf{j}})_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}) \to H_*(\vec{P}(K_{\mathbf{j}'})_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}'}) \\ (incl)_* &: H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})) \to H_*(F_{\mathbf{j}'}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \end{aligned}$$

define functors $A_*(K_{(-)})$, $H_*(\vec{P}(K_{(-)})_{\mathbf{0}}^{\mathbf{k}-(-)}$, $F_{(-)}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} : \mathcal{J} \to \mathbf{Ab}_*$; compare Propositions 3.1 and 2.6.

Proposition 3.3 *If* $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \subseteq K$*, the homomorphism* Φ_K *is the composition*

$$A_*(K) \overset{\cong}{\leftarrow} \operatorname{colim}_{\mathbf{j}\in\mathcal{J}} A_*(K_{\mathbf{j}}) \xrightarrow{\operatorname{colim}_{\mathbf{j}} \Phi_{K_{\mathbf{j}}}} \operatorname{colim}_{\mathbf{j}\in\mathcal{J}} H_*(\vec{P}(K_{\mathbf{j}})_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}})$$
$$\xrightarrow{\operatorname{colim}_{\mathbf{j}}(i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}})_*} \operatorname{colim}_{\mathbf{j}\in\mathcal{J}} H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \xrightarrow{Q} H_*(\operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \overset{\cong}{\rightarrow} H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}).$$

with Q *the colimit of the maps* $Q_{\mathbf{j}} : H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \to H_*(\operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$, $\mathbf{j} \in \mathcal{J}$. *Moreover, all these homomorphisms are isomorphisms.*

Proof It is easy to check that the homomorphism $\operatorname{colim}_{\mathbf{j}\in\mathcal{J}} A_*(K_{\mathbf{j}}) \to A_*(K)$ induced by inclusions $\varphi_{K_{\mathbf{j}}}^{K}$ is an isomorphism – since $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \subseteq K$. Thus, $A_*(K)$ is generated by cube sequences $[\mathbf{a}^*]$ in $K_{\mathbf{j}}$. Now the conclusion follows from Proposition 3.1 applied for pairs $K_{\mathbf{j}} \subseteq K$. Furthermore, $\operatorname{colim}_{\mathbf{j}} \Phi_{K_{\mathbf{j}}}$ is an isomorphism by the induction hypothesis, $\operatorname{colim}_{\mathbf{j}}(i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}})_*$ by Proposition 2.6 and for the last isomorphism by (2.2). We are left to show that also Q is an isomorphism; this will be achieved in Proposition 3.6. □

Proposition 3.4 *The compositions*

$$A_*(K_{\mathbf{j}}) \xrightarrow{\Phi_{K_{\mathbf{j}}}} H_*(\vec{P}(K_{\mathbf{j}})_{\mathbf{0}}^{\mathbf{k}-\mathbf{j}}) \xrightarrow{(i_{\mathbf{k}-\mathbf{j}}^{\mathbf{k}})_*} H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$$

define a natural equivalence of functors $A_*(K_{(-)})$ *and* $H_*(F_{(-)}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$ *from* $\mathcal{J}$ *into the category of graded abelian groups.*

Proof Both homomorphisms are isomorphisms by the inductive hypothesis and Proposition 2.6. The naturality of the transformations is a consequence of Propositions 3.1 and 2.6. □

Proposition 3.5 *For every* $t \geqslant 0$*, the functor* $A_t(K_{(-)})$ *is a projective object in the category of functors* $\mathcal{J} \to \mathbf{Ab}$.

Proof If $n-2$ does not divide t, then by definition $A_t(K_{(-)}) = 0$ is projective. Assume that $t = (n-2)q$, $q \in \mathbb{Z}$. We find a presentation of $A_t(K_{(-)})$ as a direct sum of projective summands: Within the set $CS_q(K)$ of cube sequences in K (cf Sect. 1.4) let

$$X_{\mathbf{j}} = CS_q(K_{\mathbf{j}}) \setminus \bigcup_{\mathbf{j}<\mathbf{h}\in\mathcal{J}} CS_q(K_{\mathbf{h}}).$$

Next, define functors $M_{\mathbf{j}} : \mathcal{J} \to \mathbf{Ab}$ by

$$M_{\mathbf{j}}(\mathbf{h}) = \begin{cases} \mathbb{Z}[X_{\mathbf{j}}] & \text{if } \mathbf{h} \leqslant \mathbf{j} \\ 0 & \text{otherwise,} \end{cases};$$

the morphisms are identities whenever possible, and trivial otherwise. Immediately from the definitions we obtain that $CS_q(K_{\mathbf{j}}) = \bigcup_{\mathbf{h}\geqslant\mathbf{j}} X_{\mathbf{h}}$ and hence $A_t(K_{(-)}) \cong \bigoplus_{\mathbf{j}\in\mathcal{J}} M_{\mathbf{j}}$.

For an arbitrary functor $N : \mathcal{J} \to \mathbf{Ab}$ the set of transformations $\mathrm{Hom}_{\mathcal{J}}(M_{\mathbf{j}}, N)$ is naturally isomorphic to $\mathrm{Hom}(\mathbb{Z}[X_{\mathbf{j}}], N(\mathbf{j}))$. Therefore the projectivity of $\mathbb{Z}[X_{\mathbf{j}}]$ implies that the functors $M_{\mathbf{j}}$ are also projective. As a consequence, $A_t(K_{(-)}) \cong \bigoplus_{\mathbf{j}\in\mathcal{J}} M_{\mathbf{j}}$ is projective. □

Proposition 3.6 *The homomorphism*

$$Q : \mathrm{colim}_{\mathbf{j}\in\mathcal{J}}\, H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \longrightarrow H_*(\mathrm{hocolim}_{\mathbf{j}\in\mathcal{J}}\, F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$$

is an isomorphism.

Proof Following Bousfield and Kan [3, XII.5.7], there is a spectral sequence

$$E^2_{s,t} = \mathrm{colim}^s_{\mathcal{J}_k}\, H_t(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \Rightarrow H_*(\mathrm{hocolim}_{\mathbf{j}\in\mathcal{J}}\, F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}),$$

where colim^s stands for s-th left derived functor of colim. In fact, $E^2_{s,t} = 0$ for $s > 0$ since

$$H_t(F_{(-)}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \simeq A_t(K_{(-)})$$

is projective (Propositions 3.4 and 3.5). Hence the spectral sequence degenerates to the isomorphism $\mathrm{colim}_{\mathbf{j}\in\mathcal{J}}\, H_*(F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) \cong H_*(\mathrm{hocolim}_{\mathbf{j}\in\mathcal{J}}\, F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})$. □

Corollary 3.7 *Assume that* $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \subseteq K$ *and* $\Phi_{K_{\mathbf{j}}}$ *is an isomorphism for* $\mathbf{j} \in \mathcal{J}$. *Then* Φ_K *is an isomorphism, as well.*

Proof This follows immediately from Propositions 3.3 and 3.6. □

3.2.2 The case $[\mathbf{k}-\mathbf{1}, \mathbf{k}] \not\subseteq K$

Denote $L := K \cup [\mathbf{k}-\mathbf{1}, \mathbf{k}]$ and denote $\hat{\mathcal{J}} := \hat{\mathcal{J}}_{n-1}$, and $\mathcal{J} := \mathcal{J}_{n-1}$ with objects

$$\mathrm{Ob}(\hat{\mathcal{J}}) = \{\mathbf{j} \in \{0,1\}^n : \mathbf{0} < \mathbf{j} < \mathbf{1}\} \subseteq \{\mathbf{j} \in \{0,1\}^n : \mathbf{0} < \mathbf{j}\} = \mathrm{Ob}(\mathcal{J}).$$

Their nerves $N\hat{\mathcal{J}}, N\mathcal{J}$ have geometric realizations $S^{n-2} \cong \partial\Delta^{n-1} \cong |N\hat{\mathcal{J}}| \subset |N\mathcal{J}| \cong \Delta^{n-1}$. Consider the sequence of cofibrations

$$\begin{array}{ccc}
\vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\hat{\mathcal{J}}| & \longrightarrow & \vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\mathcal{J}| \\
\| & & \| \\
\vec{P}(L)_{\mathbf{0}}^{\mathbf{k-1}} \times |N\hat{\mathcal{J}}| & \longrightarrow & \vec{P}(L)_{\mathbf{0}}^{\mathbf{k-1}} \times |N\mathcal{J}| \\
\simeq\downarrow & & \simeq\downarrow \\
\operatorname{hocolim}_{\mathbf{j}\in\hat{\mathcal{J}}} \vec{P}(L)_{\mathbf{0}}^{\mathbf{k-1}} & \longrightarrow & \operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} \vec{P}(L)_{\mathbf{0}}^{\mathbf{k-1}} \\
\simeq\downarrow i_{\mathbf{k-1}}^{\mathbf{k}} & & \simeq\downarrow i_{\mathbf{k-1}}^{\mathbf{k}} \\
\operatorname{hocolim}_{\mathbf{j}\in\hat{\mathcal{J}}} F_{\mathbf{1}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} & \longrightarrow & \operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} F_{\mathbf{1}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} \\
\downarrow & & \downarrow \\
\operatorname{hocolim}_{\mathbf{j}\in\hat{\mathcal{J}}} F_{\mathbf{j}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} & \longrightarrow & \operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} F_{\mathbf{j}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} \\
\simeq\uparrow 2.10 & & \| \\
\operatorname{hocolim}_{\mathbf{j}\in\hat{\mathcal{J}}} F_{\mathbf{j}}\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} & \longrightarrow & \operatorname{hocolim}_{\mathbf{j}\in\mathcal{J}} F_{\mathbf{j}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} \\
\simeq\downarrow 2.3 & & \simeq\downarrow 2.3 \\
\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}} & \longrightarrow & \vec{P}(L)_{\mathbf{0}}^{\mathbf{k}}
\end{array} \tag{3.3}$$

In all the squares of the diagram apart from the middle one the vertical maps are homotopy equivalences; hence the cofibres are also homotopy equivalent. One can easily check, using the construction of the homotopy colimit, that the maps in the middle square induce a homeomorphism between cofibres of type $F_{\mathbf{1}}\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}} \times S^{n-1}$. The diagram above induces a transfomation between associated homology long exact sequences. In particular, the following diagram

$$\begin{array}{ccc}
H_{*+1}(\vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\mathcal{J}|, \vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\hat{\mathcal{J}}|) & \xrightarrow{1\times\partial_{\mathcal{J}}} & H_*(\vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\hat{\mathcal{J}}|) \\
\mu\downarrow\cong & & \nu\downarrow \\
H_{*+1}(\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}}, \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) & \xrightarrow{\partial_P} & H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}})
\end{array} \tag{3.4}$$

is commutative, where

$$\partial_{\mathcal{J}} : H_{*+1}(|N\mathcal{J}|, |N\hat{\mathcal{J}}|) \to H_*(|N\hat{\mathcal{J}}|)$$

is the differential of the long homology exact sequence of the pair $(|N\mathcal{J}|, |N\hat{\mathcal{J}}|)$ and where μ is an isomorphism. Let Ψ be the composition

$$A_{*-(n-2)}(K_{\mathbf{1}}) \xrightarrow{\Phi_{K_{\mathbf{1}}}} H_{*-(n-2)}\vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}}$$
$$\xrightarrow{\times\partial_{\mathcal{J}}^{-1}(x_{\mathbf{k}})} H_{*+1}(\vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\mathcal{J}'|, \vec{P}(K_{\mathbf{1}})_{\mathbf{0}}^{\mathbf{k-1}} \times |N\mathcal{J}|) \xrightarrow{\mu} H_{*+1}(\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}}, \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}),$$

where $x_{\mathbf{k}} \in H_{n-2}(|N\hat{\mathcal{J}}|)$ is a fixed generator. By the inductive assumption, $\Phi_{K_{\mathbf{1}}}$ is an isomorphism and so is Ψ. Define

$$\varrho : A_{*-(n-2)}(K_{\mathbf{1}}) \ni [\mathbf{a}^1 \ll \cdots \ll \mathbf{a}^r] \mapsto [\mathbf{a}^1 \ll \cdots \ll \mathbf{a}^r \ll \mathbf{k}] \in A_*(K), \qquad (3.5)$$

Proposition 3.8 *The diagram*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A_{*-(n-2)}(K_{\mathbf{1}}) & \xrightarrow{\varrho} & A_*(K) & \xrightarrow{\varphi_K^L} & A_*(L) & \longrightarrow & 0 \\ & & \downarrow{\Psi} & & \downarrow{\Phi_K} & & \downarrow{\Phi_L} & & \\ 0 & \longrightarrow & H_{*+1}(\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}}, \vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) & \xrightarrow{\partial_P} & H_*(\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}) & \xrightarrow{(K\subseteq L)_*} & H_*(\vec{P}(L)_{\mathbf{0}}^{\mathbf{k}}) & \longrightarrow & 0 \end{array}$$

is commutative, and it has exact rows. Moreover, all vertical homomorphisms are isomorphisms.

Proof Exactness of the upper row follows immediately from definitions. The right square commutes by Proposition 3.1. Since Φ_L is an isomorphism (by Proposition 3.3), the composition $\Phi_L \circ \varphi_K^L$ is surjective. Hence $(K \subseteq L)_*$ is also surjective and this implies exactness of the lower row. For every cube sequence $[\mathbf{a}^*]$ in $K_{\mathbf{1}}$, we have

$$\partial_P(\Psi([\mathbf{a}^*])) = \partial_P(\mu(\Phi_{K_{\mathbf{1}}}([\mathbf{a}^*]) \times \partial_{\mathcal{J}}^{-1}(x_{\mathbf{k}}))) = \nu(\Phi_{K_{\mathbf{1}}}([\mathbf{a}^*]) \times x_{\mathbf{k}})$$
$$= \nu(\Phi_{K_{\mathbf{1}}}([\mathbf{a}^*]) \times \Phi_{[\mathbf{k-1},\mathbf{k}]_{(n-1)}}([\mathbf{k}])) = \Phi_K([\mathbf{a}^* \ll \mathbf{k}]) = \Phi_K(\varrho([\mathbf{a}^*])).$$

Hence the left square commutes. Finally, since both Ψ and Φ_L are isomorphisms, Φ_K is an isomorphism by the Five Lemma. □

Proof of Theorem 3.2 Induction starts for $\mathbf{k} \not\gg \mathbf{0}$ as stated immediately after the statement of Theorem 3.2. Assume that the theorem holds for all $\mathbf{l} < \mathbf{k}$. By Corollary 3.7, Φ_K is an isomorphism whenever $[\mathbf{k-1}, \mathbf{k}] \subseteq K$ and by Proposition 3.8, it holds for any $K \subseteq [\mathbf{0}, \mathbf{k}]$. □

3.3 A generalization

The main Theorem 3.2 applies of course also to spaces that are homotopy equivalent to the path spaces $\vec{P}(K)_{\mathbf{0}}^{\mathbf{k}}$. To obtain such spaces consider a functor $Z : \mathbb{Z}_+^n \to \mathbf{Top}$ from the poset of non-negative n-dimensional vectors (regarded as a category) obeying to

$$Z(\mathbf{k}) \begin{cases} \text{contractible}, & \prod_i k_i = 0 \\ \simeq \operatorname{hocolim}_{j\in\mathcal{C}_{\mathbf{k}}} Z(\mathbf{k}-\mathbf{j}), \ \mathcal{C}_{\mathbf{k}} = \begin{cases} \mathcal{J}, & [\mathbf{k-1},\mathbf{k}] \subset K \\ \hat{\mathcal{J}}, & [\mathbf{k-1},\mathbf{k}] \not\subset K \end{cases}, & \prod_i k_i \neq 0. \end{cases} \qquad (3.6)$$

A particular simple such functor functor Z_0 can be constructed recursively by

$$Z_0(\mathbf{k}) = \begin{cases} * \text{ (a one point space)} & \prod_i k_i = 0 \\ \operatorname{hocolim}_{j \in \mathcal{C}_\mathbf{k}} Z(\mathbf{k} - \mathbf{j}) & \prod_i k_i \neq 0 \end{cases}$$

with $\mathcal{C}_\mathbf{k}$ as above.

Proposition 3.9 *Functors* $Z_i, Z_j : \mathbb{Z}^n_+ \to$ **Top** *obeying to* (3.6) *yield homotopy equivalent spaces* $Z_i(\mathbf{k}) \simeq Z_j(\mathbf{k})$ *for all* $\mathbf{k} \in \mathbb{Z}^n_+$.

Proof This can be seen inductively (for $Z_j = Z_0$) starting from constant maps $Z_j(\mathbf{k}) \to Z_0(\mathbf{k})$ for $\prod_i k_i = 0$ extending to

$$Z_i(\mathbf{k}) \simeq \operatorname{hocolim}_{j \in \mathcal{C}_\mathbf{k}} Z_i(\mathbf{k} - \mathbf{j}) \simeq \operatorname{hocolim}_{j \in \mathcal{C}_\mathbf{k}} Z_0(\mathbf{k} - \mathbf{j}) \simeq Z_o(\mathbf{k}).$$ □

In particular, the functor $Z(\mathbf{k}) = \vec{P}(K)_\mathbf{0}^\mathbf{k}$ obeys to (3.6), and hence $\vec{P}(K)_\mathbf{0}^\mathbf{k} \simeq Z_0(\mathbf{k})$ for all $\mathbf{k} \in \mathbb{Z}^n_+$. We shall now present a "sub"functor Z_1 also obeying to (3.6) that can serve to motivate the main Theorem 3.2:

Cube sequences between $\mathbf{0}$ and $\mathbf{k}$ are partially ordered by inclusion. A cube sequence is *maximal* if $\mathbf{a}^i \not\ll \mathbf{a}^{i+1} - \mathbf{1}$ for all i. Let $\bar{C}S(K)_\mathbf{0}^\mathbf{k}$ denote the set of maximal cube sequences between $\mathbf{0}$ and $\mathbf{k}$. For a maximal cube sequence $\mathbf{a}^*$, let $P(\mathbf{a}^*) = \prod_{j=1}^r \vec{P}(K)_{\mathbf{a}^j-\mathbf{1}}^{\mathbf{a}^j} \times \vec{P}(K)_{\mathbf{a}^j}^{\mathbf{a}^{j+1}-\mathbf{1}}$; by maximaity, the latter factor is always contractible, and hence $P(\mathbf{a}^*) \simeq \prod_{j=1}^r \vec{P}(K)_{\mathbf{a}^j-\mathbf{1}}^{\mathbf{a}^j} \simeq \prod_{j=1}^r (S^{n-2})^r$.

Concatenation defines maps $\bar{c}(\mathbf{a}^*) : P(\mathbf{a}^*) \to \vec{P}(K)_\mathbf{0}^\mathbf{k}$ assembling to $\bar{c} : \coprod_{\mathbf{a}^* \in \bar{C}S(K)_\mathbf{0}^\mathbf{k}} \vec{P}(\mathbf{a}^*) \to \vec{P}(K)_\mathbf{0}^\mathbf{k}$. The image defines a subspace $\vec{P}'(K)_\mathbf{0}^\mathbf{k} \subset \vec{P}(K)_\mathbf{0}^\mathbf{k}$ of d-paths through integral points that – alternatingly – have at least one coordinate in common or for which *every* coordinate is the successor of the previous one.

Corollary 3.10 *Inclusion* $\vec{P}'(K)_\mathbf{0}^\mathbf{k} \subset \vec{P}(K)_\mathbf{0}^\mathbf{k}$ *is a homotopy equivalence for all* $\mathbf{0} \leq \mathbf{k}$.

Proof It is easy to check that the functor $Z_1(\mathbf{k}) = \vec{P}'(K)_\mathbf{0}^\mathbf{k}$ obeys to (3.6). Apply Proposition 3.9. □

4 The cohomology ring of the path space $\vec{P}(K)_\mathbf{0}^\mathbf{k}$

Fix $\mathbf{k} \in \mathbb{Z}^n$ and a Euclidean cubical complex $K \subseteq [\mathbf{0}, \mathbf{k}]$ containing its $(n-1)$-skeleton. As proven in the previous section 3, the homology of the path space $\vec{P}(K)_\mathbf{k}^\mathbf{0}$ is isomorphic, as a graded group, to $A_*(K)$. Since this group is free, the cohomology of $\vec{P}(K)_\mathbf{k}^\mathbf{0}$ is isomorphic to its dual, i.e. there is a sequence of isomorphisms

$$\Phi^K : H^*(\vec{P}(K)_\mathbf{k}^\mathbf{0}) \cong \operatorname{Hom}(H_*(\vec{P}(K)_\mathbf{k}^\mathbf{0}), \mathbb{Z}) \xrightarrow{\operatorname{Hom}(\Phi_K, \mathbb{Z})} A^*(K) := \operatorname{Hom}(A_*(K), \mathbb{Z}). \tag{4.1}$$

Let $Z^*(K)$ denote the free graded exterior $\mathbb{Z}$-algebra with generators the cube sequences $[\mathbf{0} \ll \mathbf{1} \leq \mathbf{k}], [\mathbf{l} - \mathbf{1}, \mathbf{l}] \not\subset K$ (of length 1). Let $I(K)$ denote the ideal generated by products $\mathbf{l}_1\mathbf{l}_2$ with $\mathbf{l}_1 \not\ll \mathbf{l}_2$ and $\mathbf{l}_2 \not\ll \mathbf{l}_1$. Let $F^*(K)$ denote the quotient algebra $F^*(K) = Z^*(K)/I^*(K)$, a free abelian group with the cube sequences $[\mathbf{a}^*]$

in K as basis. Moreover, we can provide $A^*(K)$ with a ring structure via the $\mathbb{Z}$-module isomorphism $\Psi^K : A^*(K) \to F^*(K)$, $\phi \mapsto \sum_{\mathbf{a}^* \in CS(K)} \phi(\mathbf{a}^*)\mathbf{a}^*$.

Proposition 4.1 *The map* $\Psi^K \circ \Phi^K : H^*(\vec{P}(K)^{\mathbf{0}}_{\mathbf{k}}) \to F^*(K)$ *is a graded ring isomorphism.*

Proof Fix a cube sequence $\mathbf{a}^*$ in K, giving rise to inclusion [by concatenation, as in (3.1)] $c(\mathbf{a}^*) : \prod_{i=1}^r \vec{P}(K)^{\mathbf{a}^i}_{\mathbf{a}^i - \mathbf{1}} \to \vec{P}(K)^{\mathbf{k}}_{\mathbf{0}}$. To $\mathbf{a}^*$ corresponds

- a graded abelian group $i(\mathbf{a}^*) : A_*(\mathbf{a}^*) \subset A_*(K)$ generated by the set of *sub*-cube sequences of $\mathbf{a}^*$ with $i(\mathbf{a}^*)$ the inclusion homomorphism
- a free graded exterior algebra $F^*(\mathbf{a}^*) = Z^*(\mathbf{a}^*)$ with generators $\mathbf{a}^i$ and a projection homomorphism $p(\mathbf{a}^*) : F^*(K) \to F^*(\mathbf{a}^*)$ and
- an additive isomorphism $\Psi(\mathbf{a}^*) : A^*(\mathbf{a}^*) \to F^*(\mathbf{a}^*)$.

Moreover, the group isomorphism $\Phi(\mathbf{a}^*) : A_*(\mathbf{a}^*) \to H_*(\prod_{i=1}^r \vec{P}(K)^{\mathbf{a}^i}_{\mathbf{a}^i - \mathbf{1}})$ has a dual $\Phi(\mathbf{a}^*)^* : H^*(\prod_{i=1}^r \vec{P}(K)^{\mathbf{a}^i}_{\mathbf{a}^i - \mathbf{1}}) \to A^*(\mathbf{a}^*)$ fitting into the commutative diagram

$$\begin{array}{ccc} H^*(\vec{P}(K)^{\mathbf{k}}_{\mathbf{0}}) & \xrightarrow{c(\mathbf{a}^*)} & H^*(\prod_{i=1}^r \vec{P}(K)^{\mathbf{a}^i}_{\mathbf{a}^i-\mathbf{1}}) \\ \downarrow \Phi^K & & \downarrow \Phi(\mathbf{a}^*)^* \\ A^*(K) & \xrightarrow{i(\mathbf{a}^*)} & A^*(\mathbf{a}^*) \\ \downarrow \Psi^K & & \downarrow \Psi(\mathbf{a}^*) \\ F^*(K) & \xrightarrow{p(\mathbf{a}^*)} & F^*(\mathbf{a}^*) \end{array} \tag{4.2}$$

All vertical maps are isomorphisms of abelian groups, and all maps, apart from possibly Φ^K, are ring homomorphisms.

The assembly map $\bigoplus_{\mathbf{a}^* \in CS(K)} A_*(\mathbf{a}) \to A_*(K)$ is clearly surjective whence its dual $\bigoplus_{\mathbf{a}^* \in CS(K)} i(\mathbf{a}^*) : \bigoplus_{\mathbf{a}^* \in CS(K)} A^*(K) \to \bigoplus_{\mathbf{a}^* \in CS(K)} A^*(\mathbf{a}^*)$ injects. Hence, Φ^K is a ring isomorphism, as well. □

References

1. Brown, R., Higgins, P.J.: On the algebra of cubes. J. Pure Appl. Algebra **21**, 233–260 (1981)
2. Brown, R., Higgins, P.J., Sivera, R.: Nonabelian algebraic topology. Filtered spaces, crossed complexes, cubical homotopy groupoids. In: EMS Tracts in Mathematics, vol. 15. European Mathematical Society (EMS), Zürich (2011)
3. Bousfield, A.K., Kan, D.M.: Homotopy limits, completions and localizations. Lecture Notes in Mathematics, vol. 304. Springer, New York (1972)

4. Dijkstra, E.W.: Co-operating sequential processes. In: Genuys, F. (ed.) Programming Languages, pp. 43–110. Academic Press, New York (1968)
5. Dijkstra, E.W.: Hierarchical ordering of sequential processes. Acta Inform. **1**, 115–138 (1971)
6. Fajstrup, L.: Trace Spaces of Directed Tori with Rectangular Holes. Math. Struct. Comput. Sci. FirstView, 1–23. doi:10.1017/S096012951300169
7. Fajstrup, L., Goubault, É., Raussen, M.: Algebraic topology and concurrency. Theor. Comput. Sci. 357, 241–278 (2006) (revised version of Aalborg University, preprint, 1999)
8. Grandis, M.: Directed homotopy theory, I. The fundamental category. Cahiers Top. Geom. Diff. Categ. 44, 281–316 (2003)
9. Grandis, M.: Directed Algebraic Topology. Cambridge University Press, Cambridge (2009)
10. Pratt, V.: Modelling concurrency with geometry. In: Proceedings of the 18th ACM Symposium on Principles of Programming Languages, pp. 311–322 (1991)
11. Raussen, M.: On the classification of dipaths in geometric models for concurrency. Math. Struct. Comput. Sci. **10**, 427–457 (2000)
12. Raussen, M.: Trace spaces in a pre-cubical complex. Topol. Appl. **156**, 1718–1728 (2009)
13. Raussen, M.: Simplicial models for trace spaces. Algebra Geom. Topol. **10**, 1683–1714 (2010)
14. Raussen, M.: Simplicial models for trace spaces II: general higher dimensional automata. Algebra Geom. Topol. **12**, 1545–1565 (2012)
15. Raussen, M.: Execution spaces for simple higher dimensional automata. Appl. Algebra Eng. Commun. Comput. **23**, 59–84 (2012)
16. Segal, G.: Classifying spaces and spectral sequences. Inst. Hautes Etudes Sci. Publ. Math. **34**, 105–112 (1968)
17. van Glabbeek, R.J.: On the expressiveness of higher dimensional automata. Theor. Comput. Sci. **368**, 168–194 (2006)
18. Ziemiański, K.: A cubical model for path spaces in d-simplicial complexes. Topol. Appl. **159**, 2127–2145 (2012)

Appendix B: Co-author statements

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

We hereby confirm that Martin Raussen has made significant contributions to our joint paper

Detecting deadlocks in concurrent systems, L. Fajstrup, É. Goubault and M. Raussen. Lecture Notes in Computer Science, Vol. 1466, pp. 332-347, Springer, Berlin (1998).

Date: 2/9 -2013

Date: 01/09/2013

Lisbeth Fajstrup
Aalborg University

Éric Goubault
CEA Saclay and École Polytechnique

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

I hereby confirm that Martin Raussen has made significant contributions to our joint paper

Dihomotopy as a tool in state space analysis, É. Goubault and M. Raussen. Lecture Notes in Computer Science, Vol. 2286, pp. 16-37, Springer, Berlin (2002).

Date: 01/03/2013

Éric Goubault
CEA Saclay and École Polytechnique

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

We hereby confirm that Martin Raussen has made significant contributions to our joint paper

Components of the fundamental category, L. Fajstrup, M. Raussen, É. Goubault and E. Haucourt. Applied Categorical Structures, Vol. 12, No. 1, pp. 81-108 (2004).

Date: 2/9-2013

Lisbeth Fajstrup
Aalborg University

Date: 01/09/2013

Éric Goubault
CEA Saclay and École Polytechnique

Date: 30 august 2013

Emmanuel Haucourt
CEA Saclay

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

We hereby confirm that Martin Raussen has made significant contributions to our joint paper

Algebraic topology and concurrency, L. Fajstrup, M. Raussen and É. Goubault. Theoretical Computer Science, Vol. 357, No. 1-3, pp. 241-278 (2006).

Date: 4/9-2013

Date: 03/09/2013

Lisbeth Fajstrup
Aalborg University

Éric Goubault
CEA Saclay and École Polytechnique

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

I hereby confirm that Martin Raussen has made significant contributions to our joint paper

Geometric analysis of nondeterminacy in dynamical systems, R. Wisniewski and M. Raussen. Acta Informatica, Vol. 43, No. 7, pp. 501-519 (2007).

Date: ____________________

____

Rafal Wisniewski
Dept. of Electronic Systems
Aalborg University

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

I hereby confirm that Martin Raussen has made significant contributions to our joint paper

Reparametrizations of continuous paths, U. Fahrenberg and M. Raussen. Journal of Homotopy and Related Structures, No. 2, pp. 93-117 (2007).

Date: 3/9/2013

Ulrich Fahrenberg
IRISA, Rennes

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

We hereby confirm that Martin Raussen has made significant contributions to our joint paper

Trace Spaces: An Efficient New Technique for State-Space Reduction, L. Fajstrup, É. Goubault, E. Haucourt, S. Mimram and M. Raussen. Lecture Notes in Computer Science, Vol. 7211, pp. 274-294 (2012).

Date: 02-09-2013

Lisbeth Fajstrup
Aalborg University

Date: 01/09/2013

Éric Goubault
CEA Saclay and École Polytechnique

Date: 30 august 2013

Emmanuel Haucourt
CEA Saclay

Date: 02-09-2013

Samuel Mimram
CEA Saclay

AALBORG UNIVERSITY
DENMARK

Dept. of Mathematical Sciences
Fredrik Bajers Vej 7 G
9220 Aalborg
Denmark

www.math.aau.dk

Co-author Statement

I hereby confirm that Martin Raussen has made significant contributions to our joint paper

Homology of spaces of directed paths on Euclidean cubical complexes, M. Raussen and K. Ziemiański. Journal of Homotopy and Related Structures, DOI 10.1007/s40062-013-0045-4 (2013).

Date: 2013-08-30

Krzysztof Ziemiański

Krzysztof Ziemiański
University of Warsaw