Operator Theory
Advances and Applications
278

Raul E. Curto, William Helton,
Huaxin Lin, Xiang Tang,
Rongwei Yang, Guoliang Yu,
Editors

# Operator Theory, Operator Algebras and Their Interactions with Geometry and Topology

## Ronald G. Douglas Memorial Volume

Birkhäuser

Birkhäuser

# Operator Theory: Advances and Applications

**Volume 278**

**Founded in 1979 by Israel Gohberg**



More information about this series at http://www.springer.com/series/4850

Raul E. Curto • William Helton • Huaxin Lin •
Xiang Tang • Rongwei Yang • Guoliang Yu
Editors

# Operator Theory, Operator Algebras and Their Interactions with Geometry and Topology

## Ronald G. Douglas Memorial Volume

Birkhäuser

*Editors*
Raul E. Curto
Mathematics
University of Iowa
Iowa City, IA, USA

William Helton
Department of Mathematics
University of California San Diego
La Jolla, CA, USA

Huaxin Lin
Department of Mathematics
University of Oregon
Eugene, OR, USA

Xiang Tang
Department of Mathematics
Washington University in St. Louis
St. Louis, MO, USA

Rongwei Yang
Mathematics Department
University of Albany
Albany, NY, USA

Guoliang Yu
Department of Mathematics
Texas A&M University
College Station, TX, USA

ISSN 0255-0156 ISSN 2296-4878 (electronic)
Operator Theory: Advances and Applications
ISBN 978-3-030-43379-6 ISBN 978-3-030-43380-2 (eBook)
https://doi.org/10.1007/978-3-030-43380-2

Mathematics Subject Classification: 47-XX



This book is published under the imprint Birkhäuser, www.birkhauser-science.com, by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland


International Workshop on Operator Theory and its Applications
July 23-27, 2018 on the campus of East China Normal University in Shanghai

# Preface

The IWOTA 2018 was held at East China Normal University, Shanghai from July 23–27. To honor Ronald G. Douglas' expansive and profound contributions to mathematics, in particular his monumental contribution to operator theory research in China for the past 30 years, the organizers had planned to take this occasion to celebrate his 80th birthday. Sadly, Ron passed away at Brazos Valley Hospice in Bryan, Texas on February 27. The organizers thus decided to make this Proceeding of IWOTA 2018 a special memorial volume. In addition to papers pertinent to the themes of the conference, this volume collected papers from some of his collaborators and former students. Included also is an article by physicist Michael R. Douglas which gives a personal account of his father's influence.

Ron was born on December 10, 1938 in Osgood, Indiana. He earned his doctorate at Louisiana State University in 1962. His first paper was in measure theory and it was published in 1964 by Michigan Mathematical Journal. Ron's early career research centered mostly around classical operator theory topics such as invariant subspaces, Toeplitz operators, operator model theory, and $C^*$-algebras, and he soon emerged as one of the leaders in these fields. But in Ron's view, there is indeed no fence within mathematics. His work on operator theory extended naturally to complex geometry. The definition of Cowen–Douglas operator was announced at a symposium at Williamstown, Massachusetts in 1975. This notion made it possible to use geometric tools such as holomorphic bundle and curvature to study unitary equivalence of operators. The foundation of Brown–Douglas–Fillmore theory was laid in a 1977 joint paper. Although its original intention was to use topological methods to classify essentially normal operators, it in fact gave a simple analytic version of K-homology. This work turned out to be fundamental to noncommutative geometry and topology. His later work focused mainly on multivariable operator theory and in particular its analytic framework Hilbert modules in function spaces. The notion of Hilbert module was announced in a conference at Timişoara and Herculane, Romania in 1984. In response to Shunhua Sun's invitation, Ron gave a series of lectures on this topic at Sichuan University, China in 1985. A more extensive treatment was later carried out in a book coauthored with V. I. Paulsen in 1989. This framework greatly propelled the development of multivariable operator theory. A difficult problem in this field is the Arveson–Douglas conjecture which connects essentially normal Hilbert modules with algebraic geometry, differential equations, index theory, and K-homology.

Each of Ron's aforementioned visionary work has nurtured a large community of scholars. This volume contains a number of papers and surveys in these fields. In this regard, it serves as a testimony that Ron's mathematical ideas are still very much alive. In closing, we would like to thank the contributors to this volume and the many unnamed reviewers of the articles. Special thanks go to Huaxin Lin, Guoliang Yu, the local organizers Xiaoman Chen, Kunyu Guo, Qin Wang, Yi-Jun Yao, and the many volunteer helpers from East China Normal University and Fudan University, without whom this grand scale conference would not have been possible. This IWOTA 2018 is also indebted to East China Normal University, Fudan University, and the U.S. National Science Foundation (award no. 1800780) for their financial support.

St. Louis, MO, USA — Xiang Tang
Albany, NY, USA — Rongwei Yang
January 28, 2020

# Contents

# Following in the Footsteps of Ronald G. Douglas

**Michael R. Douglas**

**Abstract** Many of you knew my father Ronald Douglas as a mentor, a collaborator or a colleague. While I never wrote a paper with him, he was a powerful influence on my own career. Some of my most important works, such as those on Dirichlet branes and noncommutative geometry, turned out to have strong connections with his work. In this talk I will reminisce a bit and describe a few of these works and connections.

As I reflect on my father's life, I realize in how many ways I followed in his footsteps. I was one of his three children, growing up in Ann Arbor and then Stony Brook, and our father showed us how attractive the academic life could be—bringing back gifts and photos from conferences in exotic countries, hosting dinner parties for visiting friends and colleagues from around the world. Just as appealing were the simple things: his home office filled with books, some of which he had written, or his freedom to come home early from work when we needed him, say to help with a difficult project for school.

We made several long family trips which had a huge influence on me: especially a sabbatical semester in Newcastle-upon-Tyne in 1973, and a summer in France in 1970 which included a month in Les Houches. There our mother (his first wife Nancy) would take us walking in the mountains, while our father attended the well known summer school which that year was on statistical mechanics and quantum field theory. Arthur Jaffe has some nice reminiscences of that meeting in [1]. Later as a young string theorist and mathematical physicist, when I would meet older colleagues, I was often told that it was not for the first time, they remembered me from when I was little, many from that meeting.

This trip was also the seed of what would become a lifelong relationship with France. A year in 1990 visiting Volodya Kazakov and Edouard Brezin at the

M. R. Douglas (✉)
Simons Center for Geometry and Physics, Stony Brook University, Stony Brook, NY, USA
e-mail: mdouglas@scgp.stonybrook.edu



R. E. Curto et al. (eds.), *Operator Theory, Operator Algebras and Their Interactions with Geometry and Topology*, Operator Theory: Advances and Applications 278, https://doi.org/10.1007/978-3-030-43380-2_1

Laboratoire de Physique Theorique of the ENS in Paris, my collaboration with Alain Connes and Albert Schwarz at the IHES, my many visits there between 1999 and 2008 as the Louis Michel chair, and my current role as chairman of the Friends of IHES, all were in some way fulfilling that early attraction to French and European culture, to physics and to mathematics.

Although I was fascinated by mathematics, I chose to major in theoretical physics in college, in part just to avoid following too closely in my father's footsteps. Operators were something I used in quantum mechanics, but my mathematical education did not include operator algebras. In grad school I did take a course in mathematical methods with Barry Simon, but that focused on topology as used to study solitons and instantons. I had by then heard the initials BDF, but I have to admit that when I first tried to read the paper, and for many years after that, I did not understand any of it. Still, later on my own research would turn out to have many points of contact with his, both by choice and by chance.

The most direct influence came in the early 1990s. In 1988, my first year as a postdoc at Chicago, I started working with Steve Shenker on random matrix theory. As many of you know, the usual starting point for this theory is the discussion of matrix ensembles such as the Gaussian unitary ensemble, defined by the following integral over $N \times N$ hermitian matrices,

$$\int \prod_{1 \le i,j \le N} dM_{i,j} e^{-N \mathrm{tr}\, M^2},$$

where the measure is independent and uniform for each of the matrix elements. We were particularly interested in generalizations such as

$$Z[N, \lambda] \equiv \int \prod_{1 \le i,j \le N} dM_{i,j} e^{-N\left(\frac{1}{2}\mathrm{tr}\, M^2 + \frac{\lambda}{3}\mathrm{tr}\, M^3\right)}$$

where $\lambda$ is a real parameter.

Our interest in this was not because $M$ was an operator, or any other property of the matrix $M$. Rather, it was because of a combinatorial interpretation of the integral, first pointed out in the physics literature by Gerard 't Hooft. It is a generating function for the number of planar triangulations of a genus $g$ Riemann surface with $F$ faces, call this $Z_{g,F}$,

$$Z[N, \lambda] = \sum_{g,F} N^{2-2g} \lambda^F Z_{g,F}.$$

This is explained in many references such as [3], and very recently in section 4 of [4], so I will not repeat it here.

One can go on, as proposed by Migdal and collaborators in the mid-1980s, to regard the terms in this expansion at fixed $g$ as defining a discrete approximation to two-dimensional quantum gravity. The idea is that a planar triangulation can be thought of as defining a Riemannian metric on the surface, with curvature concentrated at each of the vertices. This is a very special class of metrics, but if one only looks at the total curvature for regions containing many triangles, since in two dimensions the only invariant of a metric is the curvature scalar, one can argue that a general metric can be approximated this way. Thus, a sum over all triangulations is some sort of approximation to an integral over all Riemannian metrics, which is what physicists mean by the term quantum gravity. The specific results will of course depend on the choice of triangles versus squares or some other class of diagrams, but if we just look at the asymptotics for $F$ large, it is plausible that some aspects of the results are universal, and thus can be thought of as properties of a random Riemannian metric. This turns out to be true, and has even been shown rigorously in some cases [5].

Now, two-dimensional quantum gravity can be thought of as a simplified "toy" model of the quantum version of Einstein's theory of general relativity, but it can also be thought of as a simplified model of the two-dimensional world-sheet of a string as defined in superstring theory. Our goal was to understand the latter and in particular to find a model in which one could compute results for all genus $g$, and perhaps resum them to get a "nonperturbative string theory." To this end, we developed what we called the double scaling limit of taking $N \rightarrow \infty$ and $\lambda \rightarrow \lambda_c$ (the location of the singularity controlling the large $F$ asymptotics) holding an appropriate combination fixed. In this limit, the universal quantities could be computed exactly in terms of a solution of the Painlevé I equation, related to the integrable KdV hierarchy as first argued in [6].

This work was quite influential, to the point where my father started hearing about it from other physicists. This brings me to the story of the second time I went to visit Chen Ning Yang at Stony Brook. My father knew him well of course and brought me to visit him when I was first deciding where to go to graduate school and what to study. Yang explained that although particle physics might look attractive, one had to keep in mind that ultimately it was based on experiments done at colliders, and that the progress in such experiments was becoming more and more difficult. In fact, he counseled me against going into the field. So I took his advice, and decided to go to Caltech to work with John Hopfield on his new theory of neural networks. This was fall 1983, but in the summer of 1984 came the famous paper of Green and Schwarz on anomaly cancellation in ten-dimensional superstrings. Soon most of my fellow graduate students were working on string theory, and I was caught up in the excitement as well. But I had not forgotten Yang's advice, and was rather worried about what he would say when I returned to visit him in 1990. But he had heard of my work too, and told me that perhaps I had been right not to follow his advice.

After many discussions on random matrix theory with my father, he suggested I talk to Dan Voiculescu. As you all know, Voiculescu had developed a framework called free probability theory, which axiomatized the key properties of random matrix integrals in the large $N$ limit. This theory has been very influential in mathematics, but what was more attractive about it for a physicist was that it led to very simple and intuitive calculational tools, such as the R-transform and Stransform for additive and multiplicative free convolutions. Inspired by this, several physicists including Rajesh Gopakumar and David Gross, Matthias Staudacher, as well as Miao Li and myself, used free probability both to simplify the existing random matrix works and to solve new problems, most notably a problem raised by Is Singer of characterizing the master field for two-dimensional Yang–Mills theory.

Dan Voiculescu also invited me to a workshop he organized at the Fields Institute in March 1995, which I attended along with Tony Zee, a theoretical physicist who had moved from quantum field theory into statistical mechanics and condensed matter theory. This is a good point for me to comment on the difficulties of communication between physicists and mathematicians. Mathematicians often complain that they can't understand physicists because they never define what they are talking about, but it is just as difficult to follow a talk based on precise definitions which one is seeing for the first time, or even worse which are not spelled out in the talk. Now as a string theorist and as my father's son, I had some experience in interpreting mathematics, but Tony found the talks impenetrable and I remember having to give him many translations. Still this was time well spent, as Tony went on to write many papers using these ideas with Edouard Brezin and others, and free probability theory is now a well established tool in condensed matter physics.

While I first learned about the connection between random matrix theory and free probability theory from my father, the most direct connection between the topics of our research came a bit later and as a surprise to both of us. Now already by the time of the workshop I just recalled, string theorists including myself were moving on to a new topic, duality in supersymmetric field theory, epitomized by the famous Seiberg–Witten solution of $N = 2$ super Yang–Mills theory. That summer came the 1995 Strings conference held at USC, at which Chris Hull and Paul Townsend proposed their unification of superstring dualities, and Edward Witten gave the first talk on M theory. This was the beginning of the second superstring revolution, the most exciting part of my scientific career. For the next three years, almost every month there would be a new discovery which would force us to completely rethink our concept of string theory and our research directions.

Arguably the most important of these discoveries was the central role of the Dirichlet brane, explained at the end of 1995 by the late Joe Polchinski [7]. Now from the beginning of string theory, people had studied both open strings, maps from an oriented interval into space-time, and closed strings, maps from the circle into space-time. And during the 1970s it was realized that quantizing the open strings produced Yang–Mills theory, while quantizing the closed strings produced general relativity. One sign that this made sense was that whereas a closed string is in some sense unique, an open string can have "charged quarks" at its ends which couple to a background $U(N)$ Yang–Mills connection. In plainer mathematical terms, let $V$ be the defining representation of the Yang–Mills gauge group $U(N)$, and let $U(P)$ be the holonomy for a path $P$ of the Yang–Mills connection acting on $V$. Then the "coupling" means that we redefine the operator expressing the motion of the string from an initial interval $I_i$ to a final interval $I_f$, by tensoring it with $U(P_0) \otimes U^\dagger(P_1)$, where $P_0$ is a path from the "left" end of $I_i$ to the left end of $I_f$, and $P_1$ is a path from the "right" end of $I_i$ to the right end of $I_f$. In the limit that the length of the interval goes to zero, we have $P_0 = P_1$ and this amounts to taking the holonomy in the adjoint representation. But in general it is different.

In the physical applications of string theory, the strings are very small and this difference is rather subtle. But one can vary the construction to make it much more evident. For example, one could take the two ends of the open string to couple to two different Yang–Mills connections. Next, although the original definition of the open string allowed its ends to move anywhere in space-time, it is also consistent to constrain the end to a single point, or to an affine subspace of Minkowski space-time. In the terms of the path integral formalism this amounts to putting Dirichlet boundary conditions on the coordinates of the embedding map and thus the nomenclature. And once one considers a more general metric on space-time, one generalizes the constraint from an affine subspace to an arbitrary submanifold. Thus, the full definition of an open string requires maing a choice for each of the two ends $i = 0, 1$ of the string, of data $(\Sigma_i, A_i)$ where $\Sigma_i$ is a submanifold of space-time and $A_i$ is a connection on $\Sigma_i$.

All this had been pointed out by Dai, Leigh and Polchinski in 1989 [8], but what Polchinski showed in 1995 was that the Dirichlet brane could also be interpreted in closed string theory, as the natural object carrying Ramond–Ramond charge. Without going deeply into the physics, each string theory (and M theory) has a finite list of gauge fields. The philosophy of superstring duality then states that each of these gauge fields is associated to two fundamental objects, one carrying electric charge under the field and the other carrying magnetic charge. These fundamental objects were then the key to understanding the strong coupling behavior of the theory. Using arguments from supersymmetry, one could compute the mass of every fundamental object as a function of parameters, and then whichever was the lightest object would be "the" fundamental object in that regime. As an example, in the closed superstring theories, one of the gauge fields is the so-called "Neveu–Schwarz two-form field," for which the closed string is the fundamental object. And consistent with the philosophy, one finds that if the string coupling is weak, all of the other candidate fundamental objects have large mass. Now these other masses are proportional to an inverse power of the string coupling, so for strong coupling a different object will be the lightest. Which one depends on the theory. In the IIa superstring, one finds that the lightest object at strong coupling is a particle which is electrically charged under the "Ramond–Ramond one-form field," and treating it as fundamental leads to the identification of the strong coupling limit of IIa theory as M theory. But when this argument was made, there was no understanding of what this special particle might actually be in string theory terms. Polchinski showed that it is in fact the Dirichlet brane constrained to live at a point in space and move along a world-line in time, the so-called D-particle.

Hopefully the reader will not need to follow the details to see that the discovery of such a simple and intuitive idea led to another revolution in our understanding. My own contributions to these developments largely focused on the geometric interpretation of Dirichlet branes. It turned out that by just knowing how to calculate with Dirichlet branes, and following one's nose, one could rederive and extend many important mathematical results relating noncommutative algebra and geometry. This line of work began with Witten's "Small instantons in string theory" [9] and my [10], which rederived the ADHM construction of instanton moduli spaces this way, and with my joint work [11] with Greg Moore, which rederived and generalized the Kronheimer–Nakajima construction of four-dimensional self-dual metrics and their instanton moduli spaces.

Why do Dirichlet branes lead to noncommutative geometry? There should be a purely conceptual explanation of this point, but let me give the original argument in terms of the algebra of coordinates on space-time. In the physics of string theory, rather than work directly with the strings, one often proceeds through an intermediate step of "effective field theory," in which one identifies the subset of all of the degrees of freedom which are needed to describe the problem at hand. This is closely related to the idea of separating fast and slow variables in the analysis of ODE's and PDE's, and to the renormalization group. In particular, the relation I described earlier between open strings and Yang–Mills theory is an example; the effective theory of open strings is Yang–Mills theory. Following the physical arguments which lead to a Yang–Mills connection in the case of open strings moving in all of space-time, and modifying them to the case of a Dirichlet brane associated with the submanifold $\Sigma$ of space-time, we find that the counterpart of the Yang–Mills connection is a map from $\Sigma$ to its normal bundle $N\Sigma$, describing deformations of the embedding of $\Sigma$. Combining this with the "coupling" argument we gave earlier, this map to the normal bundle is tensored with an adjoint action of a gauge group for a Yang–Mills connection on $\Sigma$, becoming a map

$$X : \Sigma \to N\Sigma \otimes \mathrm{End}(V),$$

an intrinsically noncommutative object. While a general map of this type would contain far more data than a deformation of an embedding, the Yang–Mills equations also generalize to a flatness condition on the deformation,

$$[\delta X, \delta X] \sim 0.$$

In the simplest cases, say of Dirichlet branes in Minkowski space-time, this is zero and we conclude that the deformation lives in a diagonal subgroup of $\mathrm{End}(V)$, in other words it is like a direct sum of $N$ independent deformations. This is the case that reduces to ordinary commutative geometry. But in more general problems the right hand side is more interesting, and one finds that the Dirichlet branes realize a noncommutative geometry.

Now I had lectured at the 1995 Les Houches lectures organized by Alain Connes and Krzysztof Gawedzki, though not about Dirichlet branes as this was the summer and these developments were still to come. But our writeups were not due until 1996, so after the Dirichlet brane revolution I decided that this was a much more interesting subject and I devoted part of my Les Houches write-up to it [12]. I had been intrigued by Connes' lectures there about his work relating noncommutative gauge theory and the Standard Model, and in my writeup I had a passage comparing and contrasting the two pictures, that of Connes and that from Dirichlet branes, for how noncommutative geometry related to physics. When Alain saw this, I think he was happy to see a string theorist giving his work the attention it deserved, and perhaps this entered into the discussions which led the IHES to offer me a permanent position late in 1996. By then I was convinced that noncommutative geometry had a deep relationship to Dirichlet branes, and I happily accepted the position on a trial basis. This led to a visit in the fall of 1997 and my collaboration with Alain and Albert Schwarz in which we explained how M theory could be compactified on the noncommutative torus. This was hugely influential and is still my most cited paper.

Although for family reasons I did not take up the permanent position, I continued to visit the IHES frequently and during these visits I enjoyed discussions with Maxim Kontsevich and his many visitors. Maxim had a somewhat different concept of noncommutative geometry, based on algebraic geometry and concepts such as the derived category of coherent sheaves. This was extremely difficult for a physicist to get any purchase on, but as I continued to develop the geometry of Dirichlet branes I found myself learning more and more of this mathematics, including quiver algebras, tilting equivalences, and deformation theory. One point where Maxim's intuition was of immediate guidance was the role of the superpotential, which in his terms was a reduction of the holomorphic Chern–Simons action. But the real prize in the story was the role of the derived category of coherent sheaves, which Maxim had brought in to formulate his homological mirror symmetry conjecture. Since the Dirichlet brane theory was in some sense a generalization of Yang–Mills, I and other physicists had brought in all of the successful approaches to Yang–Mills, including the Donaldson–Uhlenbeck–Yau theorem and the necessary prerequisite of stability of holomorphic bundles. Gradually we realized that the right approach to understanding Dirichlet branes on Calabi–Yau manifolds was to generalize the concepts entering this theorem. Thus holomorphic bundles became coherent sheaves and then the derived category of coherent sheaves, and we were able to see how to get all of these generalizations out of the physical constructions. On the other hand there was no counterpart in the derived category of the stability condition of DUY. During 1999–2000 I asked many mathematicians about this, and the universal opinion was that it did not and could not exist, because there was no concept of subobject in a derived category. Still, the physics said it had to exist.

The resolution of this contradiction involved a great deal of additional input from string theory, which led to the formulation of Π-stability [14], a definition of stability which made sense for a derived category. With Paul Aspinwall we showed that this formulation passed several nontrivial consistency checks, but at this point the development was becoming too difficult for our physics techniques. Happily we were able to the explain the ideas to mathematicians, as in my ICM lecture [15] and most importantly at the M theory workshop we organized at the Newton Institute in the winter of 2002. There Tom Bridgeland took up the mantle and was able to turn these ideas into rigorous mathematics, now generally referred to as a Bridgeland stability condition [16]. Much of this story is explained in our book [17].

While the rich structure of algebraic geometry allowed us to go very far, so far this has only been for a small subset of the space-times possible in string theory, the Calabi–Yau manifolds. For more general spaces one needs to start with more general foundations, and the most general mathematical context in which one study the Dirichlet brane is K theory, as pointed out by Edward Witten and especially by Greg Moore. This brings me to what is probably the most direct connection between my father's and my own respective bodies of research.

In 1982, working with Paul Baum, my father published "Index Theory, Bordism and K-homology." [2]. The introduction states that it was completed during a visit to the IHES, I believe the family visit we made that summer. One of the stories we still tell about that visit is about Bastille Day, when we went into Paris to watch the fireworks. We had taken public transit (the RER), and as some of you will know, this stops running around 1 in the morning, so we took care to leave a bit early to catch our train. Unfortunately, so many other spectators had similar constraints that the Metro was jam-packed, and we only made it in time for the last train. And the last train did not go all the way to Bures, it stopped in Massy-Palaiseau, several miles away. By the time we realized where the taxi stop was, they were all taken. So we had to make the long hike home, under the moonlight. Still we made the best of it, singing and playing word games, until my sister sprained her ankle, and we had to carry her the rest of the way. The sun was just coming up as we arrived at the Ormaille.

Despite having to watch us, evidently my father found time to do some work, and the resulting paper is (I am told) a classic in K theory. I will not get any farther than the first definition, however, which is that for a cycle in K-homology [2]:

**Definition 1** A cycle for $K_0(X)$ is a triple $(\sigma_0, \sigma_1, T)$ where $\sigma_0$ and $\sigma_1$ are $*$-representations of the algebra of complex continuous bounded functions on $X$, and $T$ is a bounded intertwining operator.

Amazingly enough, in [13] Jeff Harvey and Greg Moore showed that one can derive this definition from Dirichlet branes as well. This uses the physical idea of "tachyon condensation," and since the map $T$ in the definition corresponds physically to a tachyon, Harvey and Moore even used the same notation for it. This is (more or less) the reduction to K theory of the derived category constructions I was working out at the same time.

So I was destined to walk in my father's footsteps after all. Fortunately he had chosen some very fruitful directions to walk in, and I am eternally grateful for that and for all that we shared together.

## References

1. Arthur Jaffe, "Raymond Stora and Les Houches 1970," arXiv:1604.08771.
2. P. Baum and R. G. Douglas. "Index theory, bordism, and K-homology." Contemp. math 10 (1982): 1–31.
3. M. R. Douglas, "Large N quantum field theory and matrix models," in *Free Probability Theory*, Volume 12 of Fields Institute Monographs, ed. Dan V. Voiculescu, AMS 1997, ISSN 1069-5265.
4. R. Dijkgraaf and E. Witten, "Developments in Topological Gravity," arXiv:1804.03275.

5. J. Miller and S. Sheffield, “Liouville quantum gravity and the Brownian map I: The QLE (8/3, 0) metric,” arXiv:1507.00719.
6. M. R. Douglas, “Strings in Less Than One-dimension and the Generalized KdV Hierarchies,” Phys. Lett. B **238**, 176 (1990). https://doi.org/10.1016/0370-2693(90)91716-O
7. J. Polchinski, “Dirichlet Branes and Ramond-Ramond charges,” Phys. Rev. Lett. **75**, 4724 (1995) https://doi.org/10.1103/PhysRevLett.75.4724 [hep-th/9510017].
8. J. Dai, R. G. Leigh and J. Polchinski, “New Connections Between String Theories,” Mod. Phys. Lett. A **4**, 2073 (1989). https://doi.org/10.1142/S0217732389002331
9. E. Witten, “Small instantons in string theory,” Nucl. Phys. B **460**, 541 (1996) https://doi.org/10.1016/0550-3213(95)00625-7 [hep-th/9511030].
10. M. R. Douglas, “Branes within branes,” NATO Sci. Ser. C **520**, 267 (1999) [hep-th/9512077].
11. M. R. Douglas and G. W. Moore, “D-branes, quivers, and ALE instantons,” hep-th/9603167.
12. M. R. Douglas, “Superstring dualities, Dirichlet branes and the small scale structure of space,” hep-th/9610041.
13. J. A. Harvey and G. W. Moore, “Noncommutative tachyons and K theory,” J. Math. Phys. **42**, 2765 (2001) https://doi.org/10.1063/1.1377270 [hep-th/0009030].
14. M. R. Douglas, “D-branes, categories and N=1 supersymmetry,” J. Math. Phys. **42**, 2818 (2001) https://doi.org/10.1063/1.1374448 [hep-th/0011017].
15. M. R. Douglas, “Dirichlet branes, homological mirror symmetry, and stability,” Proceedings of the International Congress of Mathematicians, Beijing, 2002 III, 395–408, 2002, Higher Ed. Press. [math-ag/0207021].
16. T. Bridgeland, “Stability conditions on triangulated categories,” Annals of Mathematics, 2007 pp.317–345, [math/0212237].
17. P. S. Aspinwall *et al.*, *Dirichlet branes and mirror symmetry*, 2009, AMS.

# Functional Models for Commuting Hilbert-Space Contractions

**Joseph A. Ball and Haripada Sau**

*Dedicated to the memory of Ron Douglas, a leader and dedicated mentor for the field*

**Abstract** We develop a Sz.-Nagy–Foias-type functional model for a commutative contractive operator tuple $\underline{T} = (T_1, \ldots, T_d)$ having $T = T_1 \cdots T_d$ equal to a completely nonunitary contraction. We identify additional invariants $\mathbb{G}_\sharp, \mathbb{W}_\sharp$ in addition to the Sz.-Nagy–Foias characteristic function $\Theta_T$ for the product operator $T$ so that the combined triple $(\mathbb{G}_\sharp, \mathbb{W}_\sharp, \Theta_T)$ becomes a complete unitary invariant for the original operator tuple $\underline{T}$. For the case $d \geq 3$ in general there is no commutative isometric lift of $\underline{T}$; however there is a (not necessarily commutative) isometric lift having some additional structure so that, when compressed to the minimal isometric-lift space for the product operator $T$, generates a special kind of lift of $\underline{T}$, herein called a *pseudo-commutative contractive lift* of $\underline{T}$, which in turn leads to the functional model for $\underline{T}$. This work has many parallels with recently developed model theories for symmetrized-bidisk contractions (commutative operator pairs $(S, P)$ having the symmetrized bidisk $\Gamma$ as a spectral set) and for tetrablock contractions (commutative operator triples $(A, B, P)$ having the tetrablock domain $\mathbb{E}$ as a spectral set).

The research of the second named author was supported by SERB Indo-US Postdoctoral Research Fellowship, 2017.

J. A. Ball (✉)
Department of Mathematics, Virginia Tech, Blacksburg, VA, USA
e-mail: joball@math.vt.edu

H. Sau
Tata Institute of Fundamental Research, Centre for Applicable Mathematics, Bangalore, India
e-mail: sau2019@tifrbng.res.in



R. E. Curto et al. (eds.), *Operator Theory, Operator Algebras and Their Interactions with Geometry and Topology*, Operator Theory: Advances and Applications 278, https://doi.org/10.1007/978-3-030-43380-2_2

**Keywords** Commutative contractive operator-tuples · Functional model · Unitary dilation · Isometric lift · Spectral set · Pseudo-commutative contractive lift

**Mathematics Subject Classification (2010)** Primary: 47A13; Secondary: 47A20, 47A25, 47A56, 47A68, 30H10

## 1 Introduction

A major development in the theory of nonnormal operator theory was the Sz.-Nagy dilation theorem (*any Hilbert-space contraction operator $T$ can be represented as the compression of a unitary operator to the orthogonal difference of two invariant subspaces*) and the concomitant Sz.-Nagy–Foias functional model for a completely nonunitary contraction operator (we refer to [42] for a complete treatment). Since then there have been many forays into extensions of the formalism to more general settings. Perhaps the earliest was that of Andô [9] who showed that any pair of commuting contractions can be dilated to a pair of commuting unitary operators, but the construction had no functional form like that of the Sz.-Nagy–Foias model for the single-operator case and did not lead to a functional model for a commutative contractive pair. Around the same time the Commutant Lifting Theorem due to Sz.-Nagy et al. [42] appeared, with a seminal special case due to Sarason [49]. It was soon realized that there is a close connection between the Andô Dilation Theorem and Commutant Lifting (see [47, Section 3]). However in the same paper of Parrott it was shown that Andô's result fails for $d$ commuting contractions as soon as $d \geq 3$. Arveson [11] gave a general operator-algebraic/function-algebraic formulation of the general problem which also revealed the key role of the property of complete contractivity as opposed to mere contractivity for representations of operator algebras.

Since the appearance of [21], much work has focused on the $d$-tuple of coordinate multipliers $M_{z_1}, \ldots, M_{z_d}$ on the Hardy space over the polydisk $H^2_{\mathbb{D}^d}$ as well as the coordinate multipliers $M_{\zeta_1}, \ldots, M_{\zeta_d}$ on the Lebesgue space over the torus $L^2_{\mathbb{T}^d}$ and variations thereof as models for commuting isometries, and the quest for Wold decompositions related to variations of these two simple examples. While the most definitive results are for the doubly-commuting case (see [40, 50, 52]), there has been additional progress developing models to handle more general classes of commuting isometries [29, 30, 56, 57]. One can then study examples of commutative contractive tuples by studying compressions of such commutative isometric tuples to jointly coinvariant subspaces (see e.g. the book of Douglas and Paulsen [35] for an abstract approach and work of Yang [61]). This work has led to a wealth of distinct new types of examples with special features, including strong rigidity results (see e.g. [36]). In case the commutative contractive tuple itself is doubly commuting, one can get a rather complete functional analogue of the Schäffer construction of the minimal unitary dilation (see [24, 58]).

More recent work of Agler and Young along with collaborators [1, 6, 7], inspired by earlier work of Bercovici et al. [18] having motivation from the notion of structured singular-value in Robust Control Theory (see [18, 38]), explored more general domains on which to explore the Arveson program: a broad overview of this direction is given in Sect. 2.2 below. Followup work by Bhattacharyya and collaborators (including the second author of the present manuscript) [22, 23, 25–28] as well as of Sarkar [51] found analogues of the Sz.-Nagy–Foias defect operator $D_T = (I - T^*T)^{\frac{1}{2}}$ and a more functional form for the dilation and model theory results established for these more general domains (specifically, the symmetrized bidisk $\Gamma$ and tetrablock domain $\mathbb{E}$ to be discussed below).

The goal of the present paper is to adapt these recent advances in the theory of $\Gamma$- and $\mathbb{E}$-function-theoretic operator theory to the original Andô-Parrott setting where the domain is the polydisk $\mathbb{D}^d$ and the associated operator-theoretic object is a operator-tuple $\underline{T} = (T_1, \dots, T_d)$ of commuting contraction operators on a Hilbert space $\mathcal{H}$. Specifically, we adapt the definition of *Fundamental Operators*, originally introduced in [25] for $\Gamma$-contractions and then adapted to $\mathbb{E}$-contractions in [22], to arrive at a definition of *Fundamental Operators* $\{F_{j1}, F_{j2} \colon j = 1, \dots, d\}$ for a commutative contractive operator tuple $\underline{T} = (T_1, \dots, T_d)$. We then show that the set of Fundamental Operators can be jointly Halmos-dilated to another geometric object which we call an *Andô tuple* as it appears implicitly as a key piece in Andô's construction of a joint unitary dilation in [9] for the pair case. While the set of Fundamental Operators is uniquely determined by $\underline{T}$, there is some freedom in the choice of Andô tuple associated with $\underline{T}$. With the aid of an Andô tuple, we are then able to construct a (not necessarily commutative) isometric lift for $\underline{T}$ which has the form of a Berger–Coburn–Lebow (BCL) model (as in [21]) for a commutative isometric operator-tuple. While any commutative isometric operator-tuple can be modeled as a BCL-model, there is no tractable characterization as to which BCL-models are commutative, except in the $d = 2$ case. For the $d = 2$ case it can be shown that there is an appropriate choice of the Andô tuple which leads to a commutative BCL-model—thereby giving a more succinct proof of Andô's original result. For the general case where $d \geq 3$, we next show how the noncommutative isometric lift constructed from the Andô tuple can be cut down to the minimal isometric-lift space for the single product operator $T = T_1 \cdots T_d$ to produce an analogue of the single-variable lift for the commutative tuple situation which we call a *pseudo-commutative contractive lift* of $\underline{T}$. For the case where $T = T_1 \cdots T_d$ is completely nonunitary, we model the minimal isometric-lift space for $T$ as the Sz.-Nagy–Foias functional-model space $\begin{bmatrix} \frac{H^2(\mathcal{D}_{T^*})}{\Delta_{\Theta_T} L^2(\mathcal{D}_T)} \end{bmatrix}$ for $T$ based on the Sz.-Nagy–Foias characteristic function $\Theta_T$ for $T$ and we arrive at a functional model for the whole commutative tuple $\underline{T} = (T_1, \dots, T_d)$ consistent with the standard Sz.-Nagy–Foias model for the product operator $T = T_1 \cdots T_d$. Let us mention that the basic ingredients of this model already appear in the work of Das et al. [31] for the pure-pair case ($d = 2$ and $T = T_1 T_2$ has the property that $T^{*n} \to 0$ strongly as $n \to \infty$). This leads to the identification of additional unitary invariants (in addition to the characteristic function $\Theta_T$) so that the whole collection $\{\mathbb{G}, \mathbb{W}, \Theta_T\}$

(which we call a *characteristic triple* for the commutative contractive tuple $\underline{T}$) is a complete unitary invariant for $\underline{T}$ for the case where $T = T_1 \cdots T_d$ is completely nonunitary. Here $\mathbb{G} = \{G_{j1}, G_{j2} \colon j = 1, \dots, d\}$ consists of the Fundamental Operators for the adjoint tuple $\underline{T}^* = (T_1^*, \dots, T_d^*)$ and $\mathbb{W} = \{W_{\sharp 1}, \dots, W_{\sharp d}\}$ consists of a canonically constructed commutative unitary tuple of multiplication operators on the Sz.-Nagy–Foias defect model space $\overline{\Delta_{\Theta_T} \cdot L^2(\mathcal{D}_{\Theta_T})}$ with product equal to multiplication by the coordinate $M_\zeta$ on $\overline{\Delta_{\Theta_T} \cdot L^2(\mathcal{D}_{\Theta_T})}$, all of which is vacuous for the case where $T = T_1 \cdots T_d$ is pure. Let us also mention that we obtain an analogue of the Sz.-Nagy–Foias canonical decomposition for a contraction operator, i.e.: any commutative contractive operator tuple $\underline{T}$ splits as an orthogonal direct sum $\underline{T} = \underline{T}_u \oplus \underline{T}_c$ where $\underline{T}_u$ is a commutative unitary operator-tuple and $\underline{T}_c$ is a commutative contractive operator tuple with $T = T_1 \cdots T_d$ completely nonunitary. As the unitary classification problem for commutative unitary tuples can be handled by the spectral theory for commuting normal operators (see [10, 34]), the results for the case where $T = T_1 \cdots T_d$ is completely nonunitary combined with the spectral theory for the commutative unitary case leads to a model theory and unitary classification theory for the general class of commutative contractive operator-tuples.

Let us mention that Bercovici et al. [14–16] have also recently obtained a wealth of structural information concerning commutative contractive tuples. This work also builds off the BCL-model for the commutative isometric case, but also derives additional insight concerning the BCL-model itself. There also appears the notion of *characteristic function* for a commutative contractive operato-tuple, but this is quite different from our notion of characteristic function (simply the Sz.-Nagy–Foias characteristic function of the single operator equal to the product $T = T_1 \cdots T_d$).

The paper is organized as follows. After the present Introduction, Sect. 2 on preliminaries provides (1) a reference for some standard notations to be used throughout, (2) a review of the rational dilation problem, especially in the context of the specific domains $\Gamma$ (symmetrized bidisk) and $\mathbb{E}$ (tetrablock domain), including some discussion on how these domains arise from specific examples of the structured singular value arising in Robust Control theory, (3) some background on Fundamental Operators in the setting of the symmetrized bidisk, along with some additional information (4) concerning Berger–Coburn–Lebow models for commutative isometric-tuples [21] and (5) concerning the Douglas approach [32] to the Sz.-Nagy–Foias model theory which will be needed in the sequel. Let us also mention that the present manuscript is closely related to our companion paper [13] where the results of the present paper are developed directly for the pair case ($\underline{T} = (T_1, T_2)$ is a commutative contractive operator-pair) from a more general point of view where additional details are developed. Finally this manuscript and [13] subsume the preliminary report [54] posted on arXiv.

**Acknowledgements** Finally let us mention that this paper is dedicated to the memory of Ron Douglas, a role model and inspiring mentor for us. Indeed it is his approach to the Sz.-Nagy–Foias model theory in [32] which was a key intermediate

step in our development of the multivariable version appearing here. In addition his recent work with Bercovici and Foias [14–16] has informed our work as well.

## 2 Preliminaries

### *2.1 Notation*

We here provide a reference for a core of common notation to be used throughout the paper.

Given an operator $A$ on a Hilbert space $\mathcal{X}$, we write

- $\nu(A)$ = *numerical radius* of $A$= $\sup\{|\langle Ax, x\rangle_{\mathcal{X}}| : x \in \mathcal{X}$ with $\|x\| = 1\}$.
- $\rho_{\text{spec}}(A)$ = *spectral radius* of $A$ = $\sup\{|\lambda| : \lambda \in \mathbb{C}$ and $\lambda I - A$ not invertible$\}$.
- If $T \in \mathcal{L}(\mathcal{X})$ with $\|T\| \leq 1$, then $D_T$ denotes the *defect operator* of $T$ defined as $D_T = (I - T^*T)^{\frac{1}{2}}$ and $\mathcal{D}_T = \overline{\text{Ran}}\, D_T$.
- Given the set of $d$ indices $\{j : 1 \leq j \leq d\}$, $(j)$ denotes the tuple of $d - 1$ indices $(1, \dots, j - 1, j + 1, \dots, d)$.
- For a $d$-tuple $(T_1, T_2, \dots, T_d)$ of operators and an index $j$ such that $1 \leq j \leq d$, $T_{(j)}$ denotes the operator $T_1 \cdots T_{j-1} T_{j+1} \cdots T_d$.

### *2.2 Domains with Motivation from Control: The Symmetrized Bidisk $\mathbb{G}$ and the Tetrablock $\mathbb{E}$*

The symmetrized bidisk $\mathbb{G}$ is the domain in $\mathbb{C}^2$ defined as

$$\mathbb{G} = \{(s, p) \in \mathbb{C}^2 : \exists\, (\lambda_1, \lambda_2) \in \mathbb{D}^2 \text{ such that } s = \lambda_1 + \lambda_2 \text{ and } p = \lambda_1\lambda_2\}. \tag{2.1}$$

The study of this domain from a function-theoretic and operator-theoretic point of view was initiated in a series of papers by Agler and Young starting in the late 1990s (see [3–8]) with original motivation from Robust Control Theory (see [38] and the papers of Bercovici et al. [17–20]). The control motivation can be explained as follows.

A key role is played by the notion of structured singular value introduced in the control literature by Packer and Doyle [45]. The *structured singular value* $\mu_{\mathbf{\Delta}}(A)$ of a $N \times N$ matrix over $\mathbb{C}$ with respect to an *uncertainty set* $\mathbf{\Delta}$ (to be thought of as the admissible range for an additional unknown variable $\Delta$ which is used to parametrize the set of possible true plants around the chosen nominal (oversimplified) model plant) is defined to be

$$\mu_{\mathbf{\Delta}}(A) = [\sup\{r \in \mathbb{R}_+ : I - \Delta A \text{ invertible for } \Delta \in \mathbf{\Delta} \text{ with } \|\Delta\| \leq r\}]^{-1}$$

After appropriate normalizations, it suffices to test whether $\mu_{\mathbf{\Delta}}(A) < 1$;

$$\mu_{\mathbf{\Delta}}(A) < 1 \Leftrightarrow I - A\Delta \text{ invertible for all } \Delta \in \mathbf{\Delta} \text{ with } \|\Delta\| \le 1.$$

In the control theory context, this appears as the test for internal stability not only for the nominal plant but for all other possible true plants as modeled by the uncertainty set $\mathbf{\Delta}$. In practice the uncertainty set is taken to be the set of all matrices having a prescribed block diagonal structure.

For the case of $2 \times 2$ matrices, there are three possible block-diagonal structures:

$$\mathbf{\Delta}_{\text{full}} = \left\{ \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} : z_{ij} \in \mathbb{C} \right\} = \text{ all } 2 \times 2 \text{ matrices.}$$

$$\mathbf{\Delta}_{\text{scalar}} = \left\{ \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix} : z \in \mathbb{C} \right\} = \text{ all scalar } 2 \times 2 \text{ matrices.}$$

$$\mathbf{\Delta}_{\text{diag}} = \left\{ \begin{bmatrix} z_1 & 0 \\ 0 & z_2 \end{bmatrix} : z_1, z_2 \in \mathbb{C} \right\} = \text{ all diagonal matrices.}$$

An easy exercise using the theory of singular-value decompositions is to show that

$$\mu_{\mathbf{\Delta}_{\text{full}}}(A) = \|A\|.$$

To compute $\mu_{\mathbf{\Delta}_{\text{scalar}}}(A)$, one can proceed as follows. Given $A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]$, from the definitions we see that

$$\begin{aligned} \mu_{\mathbf{\Delta}_{\text{scalar}}}(A) < 1 &\Leftrightarrow \det\left( \begin{bmatrix} 1 - za_{11} & -za_{12} \\ -za_{21} & 1 - za_{22} \end{bmatrix} \right) \ne 0 \text{ for all } z \text{ with } |z| \le 1 \\ &\Leftrightarrow 1 - (\operatorname{tr} A)z + (\det A)z^2 \ne 0 \text{ for all } z \text{ with } |z| \le 1. \end{aligned} \tag{2.2}$$

Thus the decision as to whether $\mu_{\mathbf{\Delta}_{\text{scalar}}}(A) < 1$ depends only on $\operatorname{tr} A$ and $\det A$, i.e., on $\operatorname{tr} A = \lambda_1 + \lambda_2$ and $\det A = \lambda_1 \lambda_2$ where $\lambda_1, \lambda_2$ are the eigenvalues of $A$. This suggests that we define a map $\pi_{\mathbb{G}} \colon \mathbb{C}^{2\times 2} \to \mathbb{C}^2$ by

$$\pi_{\mathbb{G}}(A) = (\operatorname{tr} A, \det A)$$

and introduce the domain

$$\begin{aligned} \mathbb{G}' = &\{x = (x_1, x_2) \in \mathbb{C}^2 \colon \exists\, A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right] \in \mathbb{C}^{2\times 2} \\ &\text{with } \pi_{\mathbb{G}}(A) = x \text{ and } \mu_{\mathbf{\Delta}_{\text{scalar}}}(A) < 1\}. \end{aligned} \tag{2.3}$$

Note next that the first form of the criterion (2.2) for $\mu_{\mathbf{\Delta}_{\text{scalar}}}(A) < 1$ can also be interpreted as saying that $A$ has no inverse-eigenvalues inside the closed unit disk, i.e., all eigenvalues of $A$ are in the open unit disk, meaning that $\rho_{\rm spec}(A) < 1$. In this way we see that the symmetrized bidisk $\mathbb{G}$ (2.1) is exactly the same as the domain $\mathbb{G}'$ given by (2.3). This equivalence gives the connection between the symmetrized bidisk and the structured singular value $A \mapsto \mu_{\mathbf{\Delta}_{\rm scalar}}(A)$.

Noting that similarity transformations

$$A \mapsto A' = SAS^{-1} \text{ for some invertible } S$$

preserve eigenvalues and using the fact that $\rho_{\rm spec}(A) < 1$ if and only if $A$ is similar to a strict contraction (known as Rota's Theorem [48] among mathematicians whereas engineers think in terms of $X = S^*S \succ 0$ being a solution of the Linear Matrix Inequality $A^*XA - X \prec 0$—see e.g. [38, Theorem 11.1 (i)]), we see that yet another characterization of the domain $\mathbb{G}$ is

$$\mathbb{G} = \{x = (s, p) \in \mathbb{C}^2 \colon \exists\, A \in \mathbb{C}^{2\times 2} \text{ with } \pi_{\mathbb{G}}(A) = x \text{ and } \|A\| < 1\}. \tag{2.4}$$

The fact that one can always write down a companion matrix $A$ whose characteristic polynomial $\det(zI - A)$ is equal to a given polynomial $1 - sz + ps^2$ leads us to one more equivalent definition of $\mathbb{G}$:

$$\mathbb{G} = \{(s, p) \in \mathbb{C}^2 \colon 1 - sz + pz^2 \neq 0 \text{ for } |z| \leq 1\}. \tag{2.5}$$

The closure of $\mathbb{G}$ is denoted by $\Gamma$.

A similar story holds for the tetrablock domain $\mathbb{E}$ defined as

$$\mathbb{E} = \{x = (x_1, x_2, x_3) \in \mathbb{C}^3 \colon 1 - x_1 z - x_2 w + x_3 zw \neq 0 \text{ whenever } |z| \leq 1, |w| \leq 1\} \tag{2.6}$$

(the analogue of definition (2.5) for the symmetrized bidisk $\mathbb{G}$) and its connection with the structured singular value $A \mapsto \mu_{\mathbf{\Delta}_{\rm diag}}(A)$. From the definitions we see that, for $A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]$,

$$\begin{aligned} \mu_{\mathbf{\Delta}_{\rm diag}}(A) < 1 &\Leftrightarrow \det\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} z & 0 \\ 0 & w \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) \neq 0 \text{ for } |z| \leq 1,\ |w| \leq 1 \\ &\Leftrightarrow 1 - za_{11} - wa_{22} + zw \cdot \det A \neq 0 \text{ whenever } |z| \leq 1,\ |w| \leq 1. \end{aligned}$$

This suggests that we define a mapping $\pi_{\mathbb{E}} \colon \mathbb{C}^{2\times 2} \to \mathbb{C}^3$ by

$$\pi_{\mathbb{E}}\left(\left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]\right) = (a_{11}, a_{22}, a_{11}a_{22} - a_{12}a_{21})$$

and we define a domain $\mathbb{E}$ by

$$\mathbb{E} = \{x = (x_1, x_2, x_3) \in \mathbb{C}^3 \colon \exists\, A \in \mathbb{C}^{2\times 2} \text{ with } \pi_{\mathbb{E}}(A) = x \text{ and } \mu_{\mathbf{\Delta}_{\rm diag}}(A) < 1\}. \tag{2.7}$$

If $x = (x_1, x_2, x_3)$ belongs to $\mathbb{E}$ as defined in (2.6) above, we can always take $A = \begin{bmatrix} x_1 & x_1x_2 - x_3 \\ 1 & x_2 \end{bmatrix}$ to produce a matrix $A$ with $\pi_{\mathbb{E}}(A) = (x_1, x_2, x_3)$ and then this $A$ has the property that $\mu_{\mathbf{\Delta}_{\rm diag}}(A) < 1$. Thus definitions (2.6) and (2.7) are equivalent. Among the many equivalent definitions of $\mathbb{E}$ (see [1, Theorem 2.2]), one of the more remarkable ones is the following variation of definition (2.7):

$$\mathbb{E} = \{x = (x_1, x_2, x_3) \in \mathbb{C}^3 \colon \exists\, A \in \mathbb{C}^{2\times 2} \text{ with } \pi_{\mathbb{E}}(A) = x \text{ and } \|A\| < 1\}. \tag{2.8}$$

That (2.7) and (2.8) are equivalent can be seen as a consequence of the $2s + f$ theorem in the control literature (with $s = 0$, $f = 2$ so that $2s + f = 2 \le 3$)—see [38, Theorem 8.27], but is also proved in [1] directly.

While the original motivation was the control theory connections, most of the ensuing research concerning the domains $\mathbb{G}$ and $\mathbb{E}$ focused on their role as new concrete domains to explore operator- and function-theoretic questions concerning general domains in $\mathbb{C}^d$. One such question is the *rational dilation problem* formulated by Arveson [11]. Let us assume that $K$ is a compact set in $\mathbb{C}^d$ (as is the case for $K$ equal to $\Gamma = \overline{\mathbb{G}}$ or $\overline{\mathbb{E}}$). Suppose that we are given a commutative tuple $\underline{T} = (T_1, \dots, T_d)$ of Hilbert space operators with Taylor joint spectrum contained in $K$ (if the Hilbert space $\mathcal{H}$ is finite-dimensional, one can take Taylor joint spectrum to mean the set of joint eigenvalues). If $r$ is any function holomorphic in a neighborhood of $K$ ( if $K$ is polynomially convex, one can take $r$ to be polynomial) any reasonable functional calculus can be used to define $r(\underline{T})$. We say that $\underline{T}$ *is a $K$-contraction* (sometimes also phrased as $K$ *is a spectral set for* $\underline{T}$), if for all $r \in \mathrm{Rat}(K)$ (rational functions holomorphic in a neighborhood of $K$) it is the case that the following *von Neumann inequality* holds:

$$\|r(\underline{T})\|_{\mathcal{B}(\mathcal{H})} \le \|r\|_{\infty, K} = \sup_{z \in K}\{|r(z)|\}$$

where $\mathcal{B}(\mathcal{H})$ is the Banach algebra of bounded linear operators on $\mathcal{H}$ with the operator norm. Let us say that operator tuple $\underline{U} = (U_1, \dots, U_d)$ is *$K$-unitary* if $\underline{U}$ is a commutative tuple of normal operators with joint spectrum contained in the distinguished boundary $\partial_e K$ of $K$. We say that $\underline{T}$ has a *$K$-unitary dilation* if there is a $K$-unitary operator-tuple $\underline{U}$ on a larger Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ such that $r(\underline{T}) = P_{\mathcal{H}} r(\underline{U})|_{\mathcal{H}}$ for all $r \in \mathrm{Rat}\, K$. If $\underline{T}$ has a $K$-unitary dilation $\underline{U}$, it follows that

$$\begin{aligned} \|r(\underline{T})\| &= \|P_{\mathcal{H}} r(\underline{U})|_{\mathcal{H}}\| \le \|r(\underline{U})\| = \sup_{z \in \partial_e K} |r(z)| \\ &\quad \text{(by the functional calculus for commutative normal operators)} \\ &= \sup_{z \in K} |r(z)| \text{ (by the definition of the distinguished boundary)} \end{aligned}$$

and it follows that $\underline{T}$ has $K$ as a spectral set. The *rational dilation question* asks: for a given compact set $K$, when is it the case that the converse direction holds, i.e., that $\underline{T}$ being a $K$-contraction implies that $\underline{T}$ has a $K$-unitary dilation $\underline{U}$? For the case of $K$ equal to the closed polydisk $\overline{\mathbb{D}}^d$, the rational dilation question is known to have an affirmative answer in case $d = 1$ (by the Sz.-Nagy dilation theorem [43]) as well as $d = 2$ (by the Andô dilation theorem [9]) but has a negative answer for $d \geq 3$ by the result of Parrott [47]. For the case of $K = \Gamma$ it is known that the rational dilation question has an affirmative answer ([7, 25]) while the case of $K = \overline{\mathbb{E}}$ was initially thought to be settled in the negative [46] but now appears to be still undecided [12].

It is known that existence of a $K$-unitary dilation for $\underline{T}$ is equivalent to the existence of a *$K$-isometric lift* for $\underline{T}$. Here a commutative operator-tuple $\underline{V} = (V_1, \ldots, V_d)$ defined on a Hilbert space $\mathcal{K}_+$ is said to be a *$K$-isometry* if there is a $K$-unitary $d$-tuple $\underline{U} = (U_1, \ldots, U_d)$ on a Hilbert space $\mathcal{K}$ containing $\mathcal{K}_+$ such that $\mathcal{K}_+$ is invariant for $\underline{U}$ and $\underline{U}$ restricted to $\mathcal{K}_+$ is equal to $\underline{V}$, i.e.,

$$U_j \mathcal{K}_+ \subset \mathcal{K}_+ \text{ and } U_j|_{\mathcal{K}_+} = V_j \text{ for } j = 1, \ldots, d.$$

Then we say that the $K$-contraction $\underline{T}$ has a $K$-isometric lift if there is a $K$-isometric operator-tuple $\underline{V}$ on a Hilbert space $\mathcal{K}_+$ containing $\mathcal{H}$ such that $\underline{V}$ is a lift of $\underline{T}$, i.e., for each $j = 1, \ldots, d$,

$$V_j^* \mathcal{H} \subset \mathcal{H} \text{ and } V_j^*|_{\mathcal{H}} = T_j^*.$$

It is known that a $K$-contraction $\underline{T}$ has a $K$-unitary dilation if and only if $\underline{T}$ has a $K$-isometric lift. In practice $K$-isometric lifts are easier to work with, so in the sequel we shall only deal with $K$-isometric lifts. This point has been made in a number of places (see e.g. the introduction in [12]).

We define a couple of terminologies here. To add flexibility to the construction of such lifts, we often drop the requirement that $\mathcal{H}$ be a subspace of $\mathcal{K}_+$ but instead require only an isometric identification map $\Pi\colon \mathcal{H} \to \mathcal{K}_+$. We summarize the precise language which we shall be using in the following definition.

**Definition 2.1** We say that $(\Pi, \mathcal{K}_+, \underline{S} = (S_1, \ldots, S_d))$ is a *lift* of $\underline{T} = (T_1, \ldots, T_d)$ on $\mathcal{H}$ if

- $\Pi\colon \mathcal{H} \to \mathcal{K}_+$ is isometric, and
- $S_j^* \Pi h = \Pi T_j^* h$ for all $h \in \mathcal{H}$ and $j = 0, 1, 2, \ldots, d$.

A lift $(\Pi, \mathcal{K}_+, \underline{S})$ of $\underline{T}$ is said to be *minimal* if

$$\mathcal{K}_+ = \overline{\operatorname{span}}\{S_1^{m_1} S_2^{m_2} \cdots S_d^{m_d} h \colon h \in \mathcal{H}, m_j \geq 0\}.$$

Two lifts $(\Pi, \mathcal{K}_+, \underline{S})$ and $(\Pi,' \mathcal{K}'_+, \underline{S}')$ of the same $(T_1, \ldots, T_d)$ are said to be *unitarily equivalent* if there is a unitary operator $\tau : \mathcal{K}_+ \to \mathcal{K}'_+$ so that

$$\tau S_j = S'_j \tau \text{ for each } j = 1, \ldots, d, \quad \text{and} \quad \tau \Pi = \Pi'.$$

It is known (see Chapter I of [42]) that when $K = \overline{\mathbb{D}}$, any two minimal isometric lifts of a given contraction are unitarily equivalent. However, minimality in several variables does not imply uniqueness, in general. For example, two minimal $\overline{\mathbb{D}}^2$-isometric lifts need not be unique (see [41]).

Instead of $\overline{\mathbb{E}}$-contraction, the terminology *tetrablock contraction* was used in [22]. We follow this terminology.

## 2.3 Fundamental Operators

For our study of commutative contractive tuples $\underline{T} = (T_1, \ldots, T_d)$, we shall have use for the following theorem concerning $\Gamma$-contractions. We refer back to Sect. 2.1 for other notational conventions.

**Theorem 2.2** *Let $(S, T)$ be a $\Gamma$-contraction on a Hilbert space $\mathcal{H}$. Then*

1. (See [25, Theorem 4.2].) *There exists a unique operator $F \in \mathcal{B}(\mathcal{D}_T)$ with $\nu(F) \leq 1$ such that*

$$S - S^*T = D_T F D_T.$$

2. (See [22, Lemma 4.1].) *The operator $F$ in part (1) above is the unique solution $X = F$ of the operator equation*

$$D_T S = X D_T + X^* D_T T.$$

This theorem has been a major influence on further developments in the theory of both $\Gamma$-contractions [26, 27] and tetrablock contractions [22, 28, 53]. The unique operator $F$ in Theorem 2.2 is called the *fundamental operator* of the $\Gamma$-contraction $(S, T)$.

## 2.4 Models for Commutative Isometric Tuples

The following result of Berger, Coburn and Lebow for commutative-tuples of isometries is a fundamental stepping stone for our study of commutative-tuples of contractions.

**Theorem 2.3** *Let $d \geq 2$ and $(V_1, V_2, \ldots, V_d)$ be a $d$-tuple of commutative isometries acting on $\mathcal{K}$. Then there exit Hilbert spaces $\mathcal{F}$ and $\mathcal{K}_u$, unitary operators $U_1, \ldots, U_d$ and projection operators $P_1, \ldots, P_d$ on $\mathcal{F}$, commutative unitary operators $W_1, \ldots, W_d$ on $\mathcal{K}_u$, such that $\mathcal{K}$ can be decomposed as*

$$\mathcal{K} = H^2(\mathcal{F}) \oplus \mathcal{K}_u \tag{2.9}$$