Technische Universität München
TUM School of Natural Sciences

Reduced Control Systems

&

Quantum Control Theory

Emanuel Malvetti

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitz:		Prof. Dr. Bernd Reif
Prüfer der Dissertation:	1.	Prof. Dr. Steffen J. Glaser
	2.	Prof. Dr. Robert König
	3.	Prof. Dr. Ugo Boscain (Sorbonne Université)

Die Dissertation wurde am 09.07.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 26.09.2024 angenommen.

Fir meng Elteren

Abstract

This thesis develops a method of reduced control systems and applies it to quantum control theory. Control systems subject to fast control on certain degrees of freedom admit a reduction to the remaining degrees of freedom without loss of information. A rigorous treatment of reduced control systems is made possible by generalizing methods from the perturbation theory of linear operators to symmetric Lie algebras. The method can be applied to great effect in quantum control theory and we study two instances in detail.

First, we address open Markovian quantum systems with fast unitary control. Such systems model noise and dissipation, present in all quantum systems, and which pose one of the main obstacles to the realization of fault-tolerant quantum computation. The reduced control system describes the evolution of the eigenvalues of the mixed quantum state, and hence also of the purity of the state. The results presented here have implications for the study of quantum thermodynamics and optimal cooling procedures.

Second, we study closed bipartite quantum systems with fast local unitary control. These systems allow for the study of entanglement between two quantum mechanical systems, one of the distinguishing features of quantum mechanics, and one of the core resources in emerging quantum technologies. In this case, the reduced control system describes the evolution of the singular values associated with the bipartite quantum state. As an application, we study optimal entanglement generation and quantum speed limits on the evolution of entanglement.

Resümmee

Dës Thees developpéiert eng Method vu reduzéierte Kontrollsystemer, a went se op d'Quantekontrolltheorie un. Kontrollsystemer bei deene gewësse Fräiheetsgrader séier gesteiert kënne ginn, erlaben et, se op déi reschtlech Fräiheetsgrader ze reduzéieren. Eng riguréis Behandlung vu reduzéierte Kontrollsystemer gëtt erméiglecht duerch eng Generaliséierung vu Methoden aus der Stéirungstheorie vu linearen Operatoren op symmetresch Lie Algebren. Dës Method ka mat grousser Wierkung op d'Quantekontrolltheorie applizéiert ginn, a mär studéieren zwee Beispiller am Detail.

Fir d'éischt behandele mer oppe Markovianesch Systemer déi séier, unitär kontrolléiert kënne ginn. Esou Systemer modelléieren d'Gedauschs an d'Dissipatioun, déi an alle Quantesystemer existéieren, an déi d'Haaptbarriär zur Realiséierung vu feelertolerante Quantecomputer poséieren. De reduzéierte Kontrollsystem beschreift d'Entwécklung vun den Eegewäerter vum gemëschte quantemechaneschen Zoustand, an domadden och vun der Rengheet vum Zoustand. D'Resultater, déi hei presentéiert ginn, si vu Bedeitung fir d'Etüd vun der Quantenthermodynamik a fir optimal Killprozesser.

Duerno widme mer eis zouene bipartite Quantesystemer déi séier lokalunitär kontrolléiert kënne ginn. Dës Systemer erlaben eis, quantemechanesch Verwuerelung tëscht zwee Systemer z'ënnersichen. Dës Verwuerelung ass eng vun deene charakteristesche Proprietéite vun der Quantemechanik, an eng vun deenen Haaptresource vun emergente Quantentechnologien. An dësem Fall beschreift de reduzéierte Kontrollsystem d'Evolutioun vun de Singulärwäerter déi zum bipartite Quantenzoustand gehéieren. Als Applikatioun ënnersiche mer, wéi een optimal Verwuerelung generéiert, a wéienge Vitesslimitatiounen d'Entwécklung vun der Verwuerelung ënnerläit.

Contents

Markovian Systems with Unitary Control

Bipartite Systems with Local Unitary Control

Epilogue

Appendix

Publications

I hereby declare that the content of my thesis is an original work, and it is based on the following publications, which have already been submitted or are planned to be submitted to scientific journals:

[1] E. Malvetti, G. Dirr, F. vom Ende, and T. Schulte-Herbrüggen. *Analytic, Differentiable and Measurable Diagonalizations in Symmetric Lie Algebras*. arXiv:2212.00713. 2022 (cit. on pp. xxvii, 1).

[2] E. Malvetti, G. Dirr, F. vom Ende, and T. Schulte-Herbrüggen. *Reduced Control Systems on Symmetric Lie Algebras*. arXiv:2307.13664, submitted to SIAM J. Control Optim., extended abstract of this work published in: Extended Abstracts presented at the 25th International Symposium on Mathematical Theory of Networks and Systems MTNS 2022. 2023 (cit. on pp. xxvii, 1, 72).

[3] F. vom Ende, E. Malvetti, G. Dirr, and T. Schulte-Herbrüggen. "Exploring the Limits of Controlled Markovian Quantum Dynamics with Thermal Resources". In: *Open Syst. Inf. Dyn.* 30.1 (2023), p. 2350005 (cit. on pp. xxvii, xxviii, 79, 81, 116, 123, 165, 217).

[4] E. Malvetti, F. vom Ende, G. Dirr, and T. Schulte-Herbrüggen. *Reachability, Coolability, and Stabilizability of Open Markovian Quantum Systems with Fast Unitary Control*. arXiv:2308.00561, accepted for publication in SIAM J. Control Optim. 2023 (cit. on pp. xxvii, xxviii, 79).

[5] E. Malvetti. *Optimal Control of a Markovian Qubit with Unitary Control*. arXiv:2404.09279. 2024 (cit. on pp. xxvii, xxviii, 79).

[6] E. Malvetti. *Provably Time-Optimal Cooling of Markovian Quantum Systems*. arXiv:2403.05285, accepted for publication in Proc. IEEE Conf. Decision Control. 2024 (cit. on pp. xxvii, xxviii, 72, 79, 139).

[7] E. Malvetti. *Bipartite Quantum Systems with Fast Local Unitary Control*. arXiv:2401.07024. 2024 (cit. on pp. xxvii, xxviii, 169).

[8] E. Malvetti and L. Van Damme. *Optimal Control of Bipartite Quantum Systems*. arXiv:2405.20034. 2024 (cit. on pp. xxvii, xxviii, 169).

[9] E. Malvetti. *Computing Common Eigenvectors and Simultaneous Triangulation*. arXiv:2309.14344. 2023 (cit. on pp. xxvii, xxviii, 79).

[10] F. vom Ende and E. Malvetti. "The Thermomajorization Polytope and Its Degeneracies". In: *Entropy* 26.2 (2024) (cit. on pp. xxvii, xxviii, 116, 215, 217, 219).

[11] E. Malvetti, C. Arenz, G. Dirr, and T. Schulte-Herbrüggen. *Randomized Gradient Descents on Riemannian Manifolds: Almost Sure Convergence to Global Minima in and Beyond Quantum Optimization*. arXiv:2405.12039. 2024 (cit. on pp. xxvii, xxviii, 215, 221).

Acknowledgments

Special thanks go to my close colleagues, Dr. Thomas Schulte-Herbrüggen, Dr. Frederik von Ende, and Dr. Gunther Dirr, for their guidance and support, our weekly meetings, the many interesting discussions at the blackboard, and the overall fruitful collaboration. Thanks also go to my other coauthors, Dr. Christian Arenz and Dr. Léo Van Damme.

I also wish to thank Prof. Dr. Steffen Glaser for giving me the opportunity to write this thesis in his group, and the remaining (current and former) members of the Glaser group, Dr. Amit Devra and Drs. in spe Amanda Nicotina, Armin Römer, Dennis Huber, Moritz Oberhauser, and Shuo Sun for the pleasant and productive work environment, for the many beers we've shared, and for listening to my presentations during the weekly seminars and giving honest feedback. Their company was especially important at the beginning of my PhD, during the most restrictive time of the pandemic.

I am also grateful to the remaining members of the BNMRZ for the collegial atmosphere and the many BBQs, and to the department of chemistry for giving me the opportunity to be a safety hazard in the lab. I would also like to thank Ms. Namalee Danwatte, Dr. Raimund Marx, and Mr. Lucas Schlett for their administrative support. My journey up to this thesis has no doubt been influenced immensely by countless teachers, professors and supervisors during my prior education, and I will not attempt to list all of them here.

Last, but certainly not least, I wish to thank my friends for their company on the occasions we get to meet, my family for always supporting me in my endeavors, even if they do not understand what I'm doing, and my partner for her endless patience.

This work was funded by the *Excellence Network of Bavaria* (ENB) under the International PhD Programme of Excellence *Exploring Quantum Matter* (ExQM), by the *Munich Center for Quantum Science and Technology* (MCQST), and by the *Munich Quantum Valley* (MQV) of the Bavarian State Government with funds from the *Hightech Agenda Bayern Plus*.

Introduction

"This proposal, like all proposals for quantum computation, relies on speculative technology, does not in its current form take into account all possible sources of noise, unreliability and manufacturing error, and probably will not work."
— Rolf Landauer, see [Llo99]

Quantum Information Technology

The idea of a quantum computer has captivated the interest of both experts and laypeople alike. However, the public's understanding of this area of research is fraught with misconceptions and hype. They will not magically solve all of humanity's problems — even if they existed. Therefore, the purpose of this section is to help the reader, as well as the author, to navigate the rapidly developing field of quantum information technology, and to motivate the study of quantum control theory, which is the subject of this thesis.

Brief History

The early 1980s saw the first attempts at formulating the idea of quantum computation. In 1980, Benioff [Ben80] introduced a quantum mechanical model of a Turing machine. In 1982, Feynman, motivated by the observation that classical computers could not efficiently simulate quantum mechanical systems, proposed the idea of a quantum computer [Fey82] for this purpose. In 1985, Deutsch further developed the idea and formally defined the universal quantum computer [Deu85].

Deutsch–Jozsa [Deu85, DJ92], Bernstein–Vazirani [BV97], and Simon [Sim94] showed that, in principle, quantum computers could achieve better computational complexities than their classical counterparts for certain problems, which were, however, not of great practical importance. Lloyd showed that quantum computers can efficiently simulate quantum systems whose dynamics are determined by local interactions [Llo96].

In 1984, Bennet and Brassard introduced quantum key distribution [BB14]. The idea of quantum computation gained traction with the discovery of quantum algorithms by Shor [Sho94, Sho97] and Grover [Gro96, Gro97] which could threaten current cryptographic schemes.

The first realizations of small-scale noisy quantum computers happened around the turn of the century using NMR, e.g., [JM98, CGK98], although this technology might not be capable of genuine quantum computation [SC99, Bra+99]. Recently, claims of quantum supremacy or quantum advantage have been made for the tasks of Boson sampling [Zho+20, Mad+22] and evaluating random quantum circuits [Aru+19]. However, the results for random quantum circuits were beaten by improved classical algorithms [PCZ22], and there is no guarantee that the Boson sampling results will not be beaten in the future. Although achieving supremacy in the near future for certain tasks is realistic, this should be regarded as a scientific achievement rather than an indication of imminent commercialization of quantum computers, similarly to recent advances in nuclear fusion [US 22]. In fact, so far there are no practical

applications of quantum computers [Bro23]. Indeed, a more useful measure is quantum practicality, which is likely still decades away, and only achievable for large enough problems with significant quantum speedups [HHT23].

At the moment of writing, massive public and private investment is driving research into quantum technology [Gib19]. Many quantum computing businesses have sprung up, although few of them are profitable, and revenues come mostly from consulting other companies about quantum technologies and licensing access to their machines to universities and businesses curious about near-term applications — which, however, seem elusive. In the, not unlikely, event that quantum computers do not materialize in the near future, this might lead to a so-called quantum winter, similar to the AI winters, in which interest and funding suddenly reduce dramatically [HG22]. Some people even warn that quantum computing is impossible altogether [Dya19, Dya20].

Quantum information technology can be broadly classified into *quantum sensing*, *quantum communication*, and *quantum computing*. We will discuss the state and the prospects of each below. A comprehensive analysis can be found in [HG22], see also the European roadmap [Ací+18].

Quantum Sensing and Metrology

Quantum sensing is the most promising near-term application of quantum information technology, and in fact it is a necessary ingredient for quantum communication and computation. Quantum sensors rely on the delicateness of quantum systems, which is a major problem for other applications, and uses it to its advantage to measure very faint signals. It is the most mature quantum technology, and some quantum sensors are already commercialized. Their applications lie, among others, in medical imaging [TB16], targeted mining, and military intelligence gathering [HG22]. They can measure physical quantities like magnetic and electric fields, time and frequency, rotations, temperature and pressure [DRC17].

Atomic clocks, nuclear magnetic resonance (NMR), magnetic resonance imaging (MRI), and positron emission tomography (PET) belong to the "first generation" of quantum technologies. Their functioning is based on quantum mechanics, but they do not directly use quantum mechanical effects like entanglement or coherence. The "second generation" quantum sensors will use coherence (and must therefore be protected from noise) and some use entanglement to achieve sensitivity or precision beyond what is possible classically [DRC17]. Such devices include superconducting quantum interference devices (SQUIDs) for measuring magnetic fields [HN19], atomic vapor for measuring electric and magnetic fields, nitrogen-vacancy (NV) centers for measuring magnetic fields at room temperature, and quantum dots for chemical detection [Li+19].

Improvements in atomic clocks will not only allow for a more precise measurement of time, but also for position measurement, as in global navigation satellite systems (GNSSs), as well as for measurements of the strength of the gravitational field [Cho+10][1]. Another application is to use cold atom gyrometers to measure acceleration and rotation [Gar19]. Quantum sonar uses SQUID-based magnetometers to look for underground deposits, submarines [Ham17] or land mines [Gar+01]. A speculative technology called quantum radar might improve radar using entangled photons [Sha20, Ass+23].

Quantum Communication and Cryptography

The seminal papers by Shor [Sho94, Sho97] and Grover [Gro96, Gro97] showed that many of the standard cryptographic schemes are vulnerable to attacks by a quantum computer. There are two ways to counter this threat. Either use different classical cryptographic schemes that are thought to be resilient against quantum computers, so-called post-quantum cryptography, or use quantum cryptographic

[1] This uses gravitational time dilation, whereby a clock in a strong gravitational field runs more slowly compared to one in a weak gravitational field.

schemes. For most applications the natural choice will be to use post-quantum cryptography [NW17]. Such schemes are known and it is just a matter of implementing appropriate standards, see for instance [ETS20, NIS23].

The most popular method of quantum cryptography is quantum key distribution (QKD), which was initiated with the invention of conjugate coding by Wiesner [Wie83] and the BB84 protocol [BB14]. Roughly speaking, it relies on the fact that when an eavesdropper measures the transmitted quantum mechanical information, this will change the information in a detectable manner. This has led to the often-repeated misunderstanding that the security of quantum cryptography is guaranteed by the laws of physics. In reality, the security still strongly depends on the quality of the physical implementation, both at the preparation and at the detection [SK14].

Now for some caveats. It should be noted that we are still very far from successfully implementing Shor's algorithm to decrypt a message encrypted today using modern cryptography. An estimate puts the number of physical qubits required for factoring a 2048-bit RSA integer at about twenty million [GE21]. The largest number factorized to date using Shor's algorithm is 21 [Mar+12], and all such implementations make use of some prior knowledge of the solution [SSV13]. On the other hand, even if implementing Shor's algorithm was impossible, quantum cryptography might still be useful since it does not rely on our inability to efficiently solve a mathematical problem. There always remains a possibility that someone will find an efficient way (classical or quantum) of cracking post-quantum cryptography — it may even hold that $\mathsf{P} = \mathsf{NP}$.

Quantum Simulation and Computing

Currently, the most discussed quantum technology is quantum computing. Only small noisy quantum computers have been built so far, but there is hope that quantum error correction will be able to deal with the noise and make quantum computers scalable. It is important to remember that quantum computers are special-purpose devices. Quite a few algorithms are known, but often the speedup is only polynomial, like for Grover's algorithm, and the exponential speedup for code-breaking will be rendered useless by switching to already existing post-quantum cryptography. Many people hope for speedups of optimization algorithms, but instances of provable exponential speedup are scarce. The most promising applications for quantum computing are small-data problems with exponential speedup, in particular in chemistry and materials science [HHT23].

Analog Quantum Simulation

One of the most promising applications is that of quantum simulation [GAN14, Dal+22]. Potential applications include quantum chemistry and materials science. The basic idea is that simulating quantum systems on a classical computer is exponentially difficult, but a quantum computer should have no problem with such simulations. In fact, this was Feynman's original motivation for building a quantum computer [Fey82]. One should mention that powerful methods for classically simulating quantum many body systems exist. Tensor network methods, such as matrix product states (MPS) and projected entangled pair states (PEPS), achieve polynomial scaling for interesting families of quantum states that appear in practice [VMC08].

Broadly speaking, quantum simulators come in two varieties, which one might call digital and analog quantum simulators. A large-scale general-purpose quantum computer would be considered digital, and through appropriate programming it could efficiently simulate a variety of different quantum systems. Since such quantum computers still seem far away, a more practical approach is to build a quantum system that has approximately the same Hamiltonian as the system we wish to study. This analog ap-

proach is less flexible, but significantly more realistic, and in fact it can already be applied, see for instance [Cho+16, Sch+21].

Models of Quantum Computation

The notion of computation can be mathematically formalized in different ways. The most well-known method is using Turing machines [Tur37]. Many different models have been devised, and many have the same computing power as Turing machines. Indeed, the Church–Turing thesis states, roughly speaking, that Turing machines fully capture the idea of computability. In order to define a precise notion of quantum computation, several models have been invented so far. We give a concise overview of the most important models to date. Many of them turn out to be equivalent in the sense that they can simulate each other with only polynomial overhead.

The first attempts at formalizing quantum computation used quantum Turing machines [Ben80, Deu85] as a direct generalization of classical Turing machines. This model turned out to be rather cumbersome, and was soon replaced by the equivalent quantum circuit model [DP89, Chi93],[2,3] in which a quantum computation is represented by a sequence of elementary quantum operations, called quantum gates, applied to an initial state and with measurements.

The one-way quantum computer [RB01, RB02] applies single qubit measurements to a certain class of highly entangled multi-qubit states, called cluster states [BR01], and it turns out to be equivalent to the circuit model.

The final two models are adiabatic quantum computation and quantum annealing. In both cases the solution to the computational problem is encoded in the ground state of a Hamiltonian, and the machine will physically prepare this ground state, which can finally be measured to obtain the result.

We start with adiabatic quantum computation, see [AL18] for a recent review. A precise definition of adiabatic quantum computation is given in [AL18, Def. 1].[4] The idea was proposed by Farhi et al. in [Far+00], and they showed that the quantum circuit model can simulate adiabatic quantum computation with polynomial overhead. The reverse direction was shown by Aharonov et al. [Aha+07]. Adiabatic quantum computation rests on "the" Adiabatic Theorem, which roughly speaking states that a system with time-dependent Hamiltonian initially in the ground state will remain in the ground state if the Hamiltonian changes sufficiently slowly. The cost of such an algorithm is given by the time required for the evolution. Although the roots of the adiabatic theorem go back to the early days of quantum mechanics [Ein14], the first rigorous proof was given by Kato in [Kat50]. More precise estimates were given later. If $\Delta = E_1 - E_0$ denotes the energy gap, then the runtime can be bounded by $\mathcal{O}(\Delta^{-3})$ or $\tilde{\mathcal{O}}(\Delta^{-2})$, see [JRS07] and [EH12] respectively. Unfortunately, computing Δ itself is a hard problem (in general it is even undecidable [CPW15]), which makes it difficult to prove speedups compared to classical algorithms.

The idea of quantum annealing is to implement a Hamiltonian whose ground state solves the given problem, and to obtain the ground state by cooling the system [RCC89, KN98]. The hope is that the phenomenon of quantum tunneling might lead to a quantum advantage for computational optimization methods. Quantum annealers have been built [Joh+11, Boo+20], but so far they don't seem to have achieved quantum advantage [Alb+15].

Some concrete algorithms for adiabatic quantum computation and quantum annealing are presented in the following section.

[2] Initially quantum circuits were not required to be acyclic, but nowadays this is usually assumed.

[3] The quantum circuit only has to simulate a given finite number of steps of the quantum Turing machine, which of course might not terminate.

[4] In particular one requires the ground state of the initial Hamiltonian to be a product state and both Hamiltonians must satisfy some locality conditions.

Quantum Algorithms

Although a large number of quantum algorithms has been discovered, the applications are more limited than what is often portrayed. A detailed list can be found on the "Quantum Algorithm Zoo"[5] website, see also the survey [Mon16].

Quantum Fourier transform and the hidden subgroup problem. One class of quantum algorithms deals with the hidden subgroup problem (HSP) for finitely generated Abelian groups. The key ingredient in these algorithms is the quantum Fourier transform, a direct analog of the discrete Fourier transform. Special cases of this algorithm were discovered by different people, see [NC10], but the first full formulation was given in [BL95]. It subsumes Shor's algorithms for factoring primes (breaking RSA), the discrete logarithm problem [Sho94, Sho97] (breaking DH and DSA), Simon's algorithm [Sim94], and the idea can be applied to the discrete logarithm in elliptic curves (breaking ECC). Together, these algorithms break many of the most important currently used cryptographic methods. Current hardware is, however, nowhere near implementing these algorithms on realistic problems [GE21].

Search algorithms. Another class of quantum algorithms deals with unstructured search. This includes Grover's algorithm and amplitude amplification, both yielding a quadratic speedup. Although the speedup is modest, it could still be of practical interest, especially because it can be applied to any NP-problem simply by searching the space of all potential solutions. Contrary to popular belief, Grover is not likely to be useful for searching databases, since the time to load the entire database into some form of quantum storage would still be linear. The algorithm can be generalized to perform Monte Carlo integration, again yielding a quadratic speedup [Mon15]. This has applications in mathematical finance for derivative pricing [Gla03].

Digital quantum simulation. The most promising application of general-purpose quantum computers is most likely digital quantum simulation. A prime application would be in simulating quantum chemistry [Cao+19]. Instead of going through the laborious process of synthesizing and characterizing thousands of molecules in the lab, one could off-load the work to a quantum computer and only use the most promising candidates in the experiment. This has the potential to significantly speed up drug discovery. There are several approaches: Using quantum phase estimation (QPE) [Kit96, AL99] one can compute the eigenvalues of a unitary, and hence determine the lowest energy levels of the chemical Hamiltonian. Since this method requires a fault-tolerant quantum computer, the variational quantum eigensolver (VQE) [Per+14] has been proposed to determine eigenvalues on noisy near-term devices. So far only very small molecules have been simulated using this method, and it seems very challenging to achieve the accuracy needed for quantum chemistry applications. The standard approach for simulating dynamics is Trotterization [Chi+21]. A more recent approach for simulating time evolution is called qubitization [LC19], which uses an oracle and requires ancillary qubits. Just like for molecules, the properties of materials are also determined by quantum mechanics. On classical devices, ab initio computations typically have exponential or even factorial complexity. Approximation techniques also quickly become intractable, and fail for highly correlated materials [Ale+24b]. As in the chemistry case, QPE, VQE, Trotterization and qubitization can be used for materials.

Variational Algorithms. The computational models of quantum annealing and adiabatic quantum computation require the implementation of a Hamiltonian whose ground state encodes the solution to some problem. A large class of problems can be formulated via quadratic unconstrained binary optimization (QUBO), as surveyed in [Koc+14]. In particular, mathematical finance applications such as portfolio optimization [Ros+16] and optimal currency arbitrage [Ros16] have recently garnered attention [Her+22]. Indeed QUBO is equivalent to the Ising model, which gives a concrete way of implementing the corresponding Hamiltonian.

[5] quantumalgorithmzoo.org

The quantum approximate optimization algorithm (QAOA) [FGG14] is an algorithm that tries to find the ground state of a Hamiltonian. QAOA is conjectured to yield exponential speedups, but so far no such example is known, and its power is an area of research [WHT16]. In this sense it is often used as a heuristic algorithm. More generally, QAOA can be seen as an ansatz for a variational quantum algorithm (VQA) [Cer+21].

Quantum Machine Learning. With the promise to combine two exciting areas of research, quantum machine learning has recently attracted a lot of attention [SP18]. Depending on whether the data or the processor is quantum or classical, one obtains different types of quantum machine learning. Typically one refers to a quantum processor handling classical data. Quantum speedups for computational complexity are usually quadratic, deriving from Grover's algorithm, but often no speedup can be proven mathematically. It was shown that quantum and classical sample complexity are polynomially equivalent [SG04]. As for variational methods, one hopes that quantum machine learning might lead to practical speedups, even without theoretical guarantees.

Topological Data Analysis. Using persistent homology one can extract topological information from data sets which is hoped to improve resilience to noise. It was shown that using quantum machine learning algorithms[6] one can achieve exponential speedups over the best known classical algorithms for the computation of Betti numbers, which count the numbers of connected components, holes and voids in the topological structure of the data [LGZ16].

It is worth noting that research into quantum algorithms can lead to interesting results in classical computer science. Claims of quantum speedup sometimes motivate improvements in classical algorithms. In some instances the quantum algorithm can be "dequantized" [Tan19].

Fault Tolerance

The quantum algorithms presented above, as well as their computational complexity, tacitly assume perfectly isolated quantum systems which do not suffer from decoherence. In reality this is never the case, not only because it is difficult to perfectly isolate a quantum system from the environment, but also because it is undesirable. Indeed, applying a quantum gate, and even more so measuring the state, inherently requires coupling to the environment, and hence will always induce some noise. Moreover, in order to implement fast quantum operations — which is necessary to outpace the natural relaxation time of the system — one requires strong coupling to the environment, which increases the induced noise. Furthermore, achieving quantum advantage will likely require computations that take weeks to complete [HHT23], although provable quantum advantage exists even for constant-depth quantum circuits [BGK18].

One way to combat decoherence is to improve the precision of our quantum devices and in many cases cool them to extremely low temperatures. However, it is probably impossible to completely eliminate noise through engineering alone. For this reason, quantum error correction, which accepts the occurrence of errors and corrects them on the fly, was introduced. Since error correction schemes require us to encode one logical qubit into several physical qubits, and since they introduce additional operations to detect and correct errors when they happen, error correction somewhat paradoxically increases the amount of noise in the system. Hence the entire idea can only work if we can correct errors more rapidly than they occur. That this is indeed possible — at least under certain assumptions — is proven by the Threshold Theorem, which states roughly that errors can be corrected successfully if the error rate of the hardware is below some threshold.

[6] The buzzwords "quantum" and "topological" mix very well with most other buzzwords like "machine learning". Indeed one can combine all of them and study topological quantum machine learning [HMR21, KC23] on a topological quantum computer [Kit03] to simulate and detect topological quantum phase transitions [TC21], ... *avis aux amateurs!*

A classical version of the Threshold Theorem was proven by von Neumann [Neu56]. The first proofs in the quantum case were given in [Sho96, KLZ98, AB08]. Due to the Threshold Theorem, research into quantum error correction is very active [Lid13], and recent experimental progress yields a first step towards validating the theorem [Ach+23]. Unfortunately, the nature of the noise is not well understood and our models remain simplistic. Moreover, the existing proofs of the Threshold Theorem make strong assumptions that are probably not satisfied in real experiments [Lid13, Ch. 26].

Quantum Hardware

Currently a large number of disparate physical systems are being considered as potential hardware for quantum technologies. Each has its own advantages and drawbacks, but none of them has yet emerged as a clear favorite for quantum computing, or at least people cannot seem to agree on which one it is. It is likely that several of these technologies will have their own specialized applications in quantum computing, communication and metrology.

Below we will briefly discuss some popular choices, namely trapped atoms [HR06], superconducting circuits [Bla+21], nuclear magnetic resonance [Lev08], and optical systems [KL10]. Many more systems exist and are being explored, and it is possible that the architecture of future quantum computers has not been discovered yet.

Atoms

A popular way of implementing a qubit is using two electronic states of an individual atom. Atoms can be trapped in relatively large arrays, and as all atoms of the same species are identical, these qubits do not suffer from manufacturing errors. The two main approaches use either neutral atoms or ions, and we will briefly discuss them below. We note that both trapped ions and neutral atoms can also be used to build optical atomic clocks [Lud+15].

Ions Trapped ions are one of the most promising avenues for quantum computing [Bru+19]. In 1995 Cirac and Zoller [CZ95] proposed the trapped-ion quantum computer, which uses the internal states of an ion as a qubit. Several ions are confined in radiofrequency traps which provide long lifetimes on the order of hours and even up to months. Similarly, the coherence times are also very long relative to gate times, although the gate times on the order of microseconds make large computations rather slow. Moreover, single and two-qubit gates, as well as state preparation and measurements, can be implemented with very high fidelity. Depending on the electronic transition used for the qubit, single qubit operations are implemented using lasers, microwaves, or Raman transitions. Multi-qubit gates, such as the Cirac–Zoller gate [CZ95] or the Mølmer–Sørensen gate [SM99], use the Coulomb interaction, turning the shared motional modes of the ions into a "quantum bus". Traps with about 20–50 qubits are possible, but scaling up has proven difficult. Since ions can be moved around in the trap, and by creating two-dimensional arrays, one obtains a more scalable architecture, called quantum charge-coupled device (QCCD), which was introduced in [KMW02].

Cold Neutral Atoms A more recent approach is to use neutral atoms [Hen+20, Win+23] that are trapped in a two dimensional periodic optical lattice via the Stark effect. To prepare the system, a dilute gas is cooled using Doppler cooling (the apparatus itself typically works at room temperature). Optical tweezers or an optical lattice then allow to trap a small number of atoms, typically in a 2D lattice, although complex geometries in 3D are possible. Since not all traps will contain an atom, the trapped atoms are rearranged to form a contiguous array. It is possible to trap hundreds of atoms, and it is expected that this can be scaled to thousands, the main limit being available laser power, but also

the time it takes to rearrange the atoms. Trap lifetimes are on the order of tens of seconds, but may be significantly extended using a cryostat. The qubit is usually encoded via electronic spin states of the atoms. Single qubit gates are performed using laser or microwave pulses, and the same gate can be applied to several qubits simultaneously. Two qubit gates are implemented using Rydberg states, which are highly excited electronic states. The spatial range of the Rydberg blockade determines the connectivity of the qubits.

Aside from digital quantum computation using gates, trapped neutral atoms also lend themselves to analog quantum simulation, e.g., by simulating the Bose–Hubbard model which is a popular model for studying superconductivity.

Circuits

Another popular approach to quantum computation is using superconducting electronic circuits [Wen17]. In contrast to atoms, they are engineered devices, and they can be designed to have specific properties, but one also has to contend with manufacturing errors. They have to be cooled in a helium dilution refrigerator (to about $0.02\mathrm{K}$). Due to their macroscopic size they couple strongly to electromagnetic fields which allows for fast control but also limits their coherence times.

Just like a classical LC circuit, composed of an inductor and a capacitor, is described by a harmonic oscillator, a superconducting LC circuit behaves like a quantum harmonic oscillator. Since, in this case, the energy levels are equally spaced, one has to add a non-linearity to the circuit to be able to address individual transitions. This is done using a Josephson junction [Jos62, Jos74], which is composed of two superconductors separated by a thin insulating barrier.

Phase, Flux and Charge qubit Generally, a superconducting circuit is made up of a capacitor, an inductor and a Josephson junction (although the real circuits are more complicated). They are characterized by the charging energy E_C, the inductive energy E_I, and the Josephson energy E_J. More precisely, for different ratios E_J/E_C and E_L/E_J one obtains a number of different qubits, cf. [Wen17].

The macroscopic state of the system is defined by the phase ϕ across the Josephson junction. The dynamics of ϕ are determined by its potential whose shape is typically described as resembling a tilted washboard. Quantized energy levels of ϕ in the local minima of this potential were first observed in [MDC85]. This forms the basis of the implementation of a so-called phase qubit [Mar09]. Here one works in the regime $E_J/E_C \gg 1$. Using the two lowest energy levels of such a local minimum, one can implement a qubit. By changing the slope of the washboard potential one can make sure that the two lowest levels are protected from tunnelling, while the third level is likely to tunnel, which makes it possible to measure the qubit.

Another approach is that of the flux qubit [Orl+99], which works in the regime $E_J/E_C \ll 1$. The qubit states are given by currents in a superconducting loop intersected by Josephson junctions circulating in opposite directions. The readout can be performed using a SQUID. A main limitation of flux qubits is their shorter coherence times.

Charge qubits, also known as Cooper-pair boxes, represent the qubit states using charge states, i.e. the number of excess Cooper pairs on an island. At low temperature and gate voltage, one can consider only the two lowest energy states, with 0 or 1 excess Cooper pair. In this case one works in the regime $E_J < E_C$. Quantum coherence in such systems was first observed in [NPT99].

Transmon Today, the most used qubit is the transmon qubit, which is a type of charge qubit insensitive to charge noise [Koc+07]. The transmon regime is $E_J \gg E_C$, since the charge noise is exponentially suppressed in E_J/E_C, at the cost of a modest reduction in the relative anharmonicity, making control somewhat more difficult.

The circuit quantum electrodynamics (QED) architecture [Bla+21, Bla+04] places a superconducting qubit (usually a transmon), which can be thought of as an artificial atom, inside of a transmission line resonator that forms a microwave cavity. A key advantage of this approach is the ability to suppress radiative decay, and it can be used to measure and couple superconducting qubits. Today, circuit QED is one of the leading architectures for quantum computing, used to implement the GKP code [Cam+20] and even making claims of quantum supremacy [Aru+19].

Spins

Spin is an intrinsic, quantized property of all particles, such as electrons and atomic nuclei, behaving like a "tiny magnet". Spins can be controlled using magnetic fields and they interact with each other [Lev08, Kup23].

Nuclear Magnetic Resonance NMR is a comparatively old and well-established technology [Hah50]. MRI machines are commonplace today, used for imaging soft body tissues and organs, and NMR is being used daily for determining the structure of molecules and proteins. It is thus no surprise that the earliest quantum computing experiments were performed using NMR machines [JM98, CGK98]. Currently, NMR is not a leading candidate for general-purpose quantum computation [SC99, Bra+99], but its usefulness for imaging and sensing is clear. Optimal control is used to design pulse sequences that optimize sensitivity, selectivity, as well as resilience to instrumental imperfections, and it is central to the improvement of magnetic resonance technologies [Gla+15]. Moreover, NMR has long been used as a test bed for quantum control and quantum information processing tasks, and it has motivated the development of sophisticated quantum control methods.

Nitrogen-Vacancy Centers A perfect diamond lattice only contains carbon atoms. A nitrogen-vacancy (NV) center in a diamond is a defect, consisting of a nitrogen atom substituting for a carbon atom, next to a vacancy. One of the outstanding properties of NV centers is the long spin coherence time at room temperature. NV centers are a promising platform for nanoscale MRI [Bor+19], a technique that would enable the measurement of an individual biomolecule. Moreover, the robustness of NV centers makes them attractive for quantum computation [Web+10, CH13], and similarly tin vacancies can be used to build scalable architectures [Li+24].

Photons

Photons can encode quantum information, for instance in their polarization. They are the natural system used for quantum communication, where photons are transmitted through fiber optic cables or free space. Although single qubit rotations can be implemented using simple linear optical elements (polarizing beam splitters and birefringent waveplates), two qubit gates seem to require strongly non-linear elements, which have not been discovered so far. This makes it all the more surprising that quantum computation is possible using linear optical quantum computing [KLM01, Kok+07], which uses photons and only linear optical elements to process quantum information. This comes at the cost of making the CNOT gates non-deterministic, with a certain success probability, which can, however, be made very large. In this scheme, the overhead of making CNOT gates highly successful is prohibitive [OBr07]. The resolution to this obstacle is the use of one-way computation and cluster states [RB01, Nie04]. Another advantage of the photonic approach is that the components can be manufactured at scale [Ale+24a].

Quantum Control Theory

After this leisurely stroll through the world of quantum technology, what follows is a gentle introduction to quantum control theory. All of the quantum technology applications listed above need to be carefully controlled in order to achieve their goals. This is true for any technology, but quantum mechanical properties like coherence and entanglement are especially fragile and ephemeral, and need to be handled with utmost precision. Typically, quantum systems are controlled using lasers or magnetic fields that have to be modulated in time. Quantum control theory provides tools to efficiently compute the control inputs which transform a given initial state into a desired final state, and do so with some degree of optimality. Indeed, optimal control theory aims to minimize the resources, such as time or energy, needed to perform a given task. Robustness to noise is another important consideration.

This section gives a very brief introduction to the quantum mechanical setting of the thesis, in particular dissipative quantum dynamics, as described by the Lindblad master equation, and quantum entanglement are discussed. It then introduces the basic notions of control theory, and most importantly for us, bilinear control systems. Finally, these two topics are combined, yielding the mathematical framework which will be explored in this thesis.

Quantum Mechanics

Quantum mechanics is the physical theory underlying all of the technologies presented above. Not only does it yield extremely precise predictions of physical phenomena, but it is also fundamentally distinct from classical physics, because it is inherently probabilistic. Quantum theory has its own logic, making it at times counterintuitive, but also enabling fundamentally new paradigms such as quantum computing and quantum cryptography.

As the topic is too vast for us to introduce here, we refer to the many great books on quantum mechanics for background. For instance there exist introductions for physicists [SN21] and for mathematicians [Hal13]. The information theoretic approach, most relevant for us, is taken in [NC10, Aar13].

This thesis addresses two important quantum mechanical phenomena: dissipation and entanglement. Briefly, dissipation stems from interactions of the system with its environment and is one of the main obstacles towards realizing scalable quantum computers. Entanglement is a fundamental quantum mechanical property that is exploited in quantum devices to obtain advantages over classical ones. Being able to control these two effects is of paramount importance for the realization of quantum technologies.

Basic Formalism

In quantum mechanics, states are described using unit vectors in a (complex) Hilbert space. Working only in finite dimensional spaces, this means that the state is a unit vector $|\psi\rangle$ in $\mathbb{C}^n$. The evolution of the system is determined by the system's Hamiltonian H, which is a Hermitian matrix ($H = H^*$) representing the energy of the state, via the Schrödinger equation[7]

$$|\dot{\psi}(t)\rangle = -\mathrm{i}H\,|\psi(t)\rangle\,.$$

Since $-\mathrm{i}H$ is skew-Hermitian, it is an element of the unitary Lie algebra $\mathfrak{u}(n)$ and thus the evolution is unitary and preserves the norm of the state. In principle the equation can be solved easily as $|\psi(t)\rangle = \exp(-\mathrm{i}Ht)\,|\psi(0)\rangle$, where $\exp(-\mathrm{i}Ht) \in \mathrm{U}(n)$ is the corresponding unitary propagator. However, once the Hamiltonian depends on time, which is always the case in quantum control theory, it becomes much more difficult to determine analytical solutions.

[7] We set the reduced Planck constant to $\hbar = 1$.

Dissipation

So far we have talked about closed systems, that is, those which do not interact with their environment. In practice, all quantum systems are open since it is all but impossible to perfectly isolate a system from its environment [BP02]. For instance, an atom in a cavity will always interact with the cavity in some way. Moreover, a perfectly isolated system would be useless for any information processing task. For these reasons it is necessary to consider systems that are coupled to some, typically large, environment. If the systems are left to equilibrate, one enters the realm of quantum thermodynamics, where the environment is often called a (heat) bath. Describing the coupled system is usually infeasible, and so approximations have to be made. We will focus on the simplest one, which assumes that the system follows a "Markovian" evolution.

Open quantum systems cannot be described using pure states and unitary evolution. Rather, we have to use mixed states, which represent a statistical mixture of pure states. Such a mixture may arise by considering a large ensemble of systems, as happens in NMR, or by "tracing out" an environment correlated with the state. Mathematically, a mixed state is described by a density matrix ρ, that is, a positive semidefinite matrix of unit trace. The set of all density matrices is denoted $\mathfrak{pos}_1(n)$. Equivalently, it is a Hermitian matrix whose eigenvalues are non-negative and add up to 1. As such, they form a discrete probability distribution. A pure state $|\psi\rangle$ is represented by the density matrix $|\psi\rangle\langle\psi|$ (that is, the rank one projector onto the state $|\psi\rangle$), while the density matrix $\mathbb{1}/n$ is called the maximally mixed state. The evolution of an open system is described using quantum channels (also called quantum maps), which are completely positive trace preserving (CPTP) maps. This condition ensures that quantum states are mapped to quantum states, even if the system is part of a larger system.

The set of quantum channels forms a semigroup (in contrast to the unitary group) since channels can be composed, but in general the inverse (which might not even exist) fails to be a quantum channel. Just like Lie groups have a corresponding set of generators given by their Lie algebra, the semigroup of quantum channels has a set of generators forming a so-called Lie wedge [Dir+09]. This is called the Kossakowski–Lindblad Lie wedge $\mathfrak{w}_{\mathsf{KL}}$, and its elements $-L \in \mathfrak{w}_{\mathsf{KL}}$ are called Lindblad generators. They give rise to the Lindblad equation [GKS76, Lin76]:

$$\dot{\rho} = -L(\rho) = -\mathrm{i}[H_0, \rho] - \sum_{k=1}^{r} \Gamma_{V_k}(\rho), \qquad \text{where} \quad -\Gamma_V(\rho) = V\rho V^* - \tfrac{1}{2}(V^*V\rho + \rho V^*V),$$

where H_0 is a Hermitian matrix, representing the coherent part of the evolution, and the V_k are arbitrary complex matrices, called Lindblad terms, representing the dissipative part of the evolution.

Physically, the Lindblad equation is obtained by making certain assumptions about the interaction between the system and the environment. Typically one uses the Born approximation (or weak-coupling limit), which assumes that the interaction is relatively weak, the Markov approximation, which assumes that the environment equilibrates quickly, and the rotating wave approximation, which allows to ignore certain fast oscillating terms. For complicated systems it is a major challenge to derive accurate noise models, and often the noise will not be Markovian in the first place. An important approach in this context is to measure the noise present in a system and to find the best Markovian approximation.

Entanglement

Quantum entanglement is a kind of correlation between quantum mechanical systems which goes beyond classical correlation. Entanglement is an important resource in pure-state quantum computation [JL03, Vid03, DV07], quantum cryptography [BB14, Gis+02], and quantum sensing [DRC17]. For the realization of such technologies, it is necessary to be able to control the entanglement within a quantum system. In particular, the generation of sufficient entanglement, as well as the stabilization

of entangled states, is essential. At the same time many fundamental questions about entanglement remain unanswered, and our theoretical understanding of multipartite and mixed state entanglement is limited [BŻ17]. Indeed, quantifying the amount of entanglement in a given system is challenging and a large number of distinct entanglement measures are used for this purpose.

In this thesis we focus on a simple case, namely that of a closed, bipartite quantum system, meaning a system composed of two subsystems. Given two quantum systems with corresponding Hilbert spaces $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$, the Hilbert space corresponding to the composite system is given by the tensor product $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \cong \mathbb{C}^{d_1 d_2}$. Concretely, if $|i\rangle$ for $i = 1, \ldots, d_1$ is an orthonormal basis of $\mathbb{C}^{d_1}$ and $|j\rangle$ for $j = 1, \ldots, d_2$ is an orthonormal basis of $\mathbb{C}^{d_2}$, then the set of all $|i\rangle \otimes |j\rangle$ forms an orthonormal basis of the composite system. States of the form $|\psi\rangle \otimes |\phi\rangle$ are called product states, and they are not entangled at all. A bipartite pure state which is not a product state is always entangled. The amount of entanglement in such a system can be quantified using the Schmidt decomposition (which is equivalent to the complex singular value decomposition (SVD)). This result states that for any given bipartite pure state $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ there exists a local unitary transformation $U = V \otimes W \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ such that

$$U |\psi\rangle = \sum_{i=1}^{d_{\min}} \sigma_i |i\rangle \otimes |i\rangle ,$$

where the $\sigma_i \in \mathbb{R}$ are uniquely defined up to order and sign and they are called the singular values of $|\psi\rangle$. (Their squares are often called Schmidt values.) Clearly then the singular values are invariant under local unitary transformations and hence they may be used to quantify entanglement.

If the bipartite system is symmetric (or anti-symmetric) under exchange of the two subsystems, we say that it is bosonic (or fermionic). Interestingly, in these cases, results analogous to the Schmidt decomposition exist. They rely on less well-known matrix decompositions, called the Autonne–Takagi factorization and the Hua factorization respectively.

Control Theory

Control theory is a vast field, spanning from applied control engineering to highly abstract mathematics. In practice, a control system might be, for instance, an electrical, mechanical, chemical or biological system with certain controls or inputs (e.g., voltages, forces, steam flow rates, sweating) and outputs (e.g., current, velocity, chemical product, body temperature). The goal is to choose the inputs such that the system is steered into a desired state, or stabilized in such a state, and usually to do so in some kind of optimal way. For us, the control systems will be quantum mechanical in nature, such as nuclear spins or electronic states of an atom, and they can be steered using, for instance, magnetic fields or lasers.

Mathematically, control systems are generally[8] described using a first order differential equation of the form [Son98, CK00, Zab20]

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0 \in \mathbb{R}^n,$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in U \subseteq \mathbb{R}^m$. Here x is the state of the system and u is the control function. This setting is extremely general as f can be an almost arbitrary function. In fact, this mathematical formulation is also appropriate to study perturbations, uncertainty and stochastic noise, cf. [CK00].

Control systems can also be formulated in a more geometric way using differential inclusions [AC84, Smi02] of the form

$$\dot{x}(t) \in \mathsf{derv}(x(t)), \quad x(0) = x_0 \in \mathbb{R}^n,$$

[8] This assumes that we use continuous representations of time and the state space. Sometimes it is useful to discretize one or both of these variables, leading for instance to quantum circuit synthesis. More generally, one could also add some stochasticity to this equation.

where derv is now the set-valued function of achievable derivatives. Indeed, these two approaches can be related by setting $\mathsf{derv}(x) = \{f(x,u) : u \in U\}$. Then, under certain assumptions, Filippov's Theorem [Smi02, Thm. 2.3] shows that these two systems are indeed equivalent.

Bilinear Control Systems

In practice one often studies linear control systems, obtained by locally linearizing a complicated system, since the theory is well understood and mature [Zab20]. Non-linear control systems on the other hand are often hopelessly complex. Quantum control systems are naturally modeled using bilinear control systems [Jur97, Ell09], which lie somewhere in the middle. They are not as simple as linear control systems, but still present a beautiful mathematical structure, which can be exploited to derive powerful results. Whereas linear control systems are of the form $\dot{x} = Ax + ub$ with $u(t) \in \mathbb{R}$ and $x, b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n,n}$, bilinear control systems, on the other hand, are of the form

$$\dot{x}(t) = \Big(A + \sum_{j=1}^{m} u_j(t) B_j\Big) x(t), \quad x(0) = x_0 \in \mathbb{R}^n,$$

where $A, B_j \in \mathbb{R}^{n,n}$. Now one can define the operator lift of this system as

$$\dot{X}(t) = \Big(A + \sum_{j=1}^{m} u_j(t) B_j\Big) X(t), \quad X(0) = \mathbb{1} \in \mathbb{R}^{n,n}.$$

This creates the bridge to Lie group [Kna02] and Lie semigroup [HHL89, HN93] theory, see [Law99]. Indeed, if all matrices A, B_j belong to some matrix Lie algebra, the solutions of the operator lift belong to the corresponding matrix Lie group. This connection is extremely powerful as it brings a wealth of tools to control theory. Conversely, tools from control theory were also instrumental in the development of Lie semigroup theory [HHL89].

Optimal Control

In applications one is usually not only interested in how one can achieve a given task, but also in implementing the task in an optimal way, typically in the least amount of time possible. There are two main ways to do this. The Hamilton–Jacobi–Bellman (HJB) equation, which is a partial differential equation (PDE), yields a sufficient condition for optimality [Zab20, Thm. 9.1]. Another approach is the Pontryagin Maximum Principle (PMP), which yields a necessary condition for optimality, cf. [Zab20, Thm. 12.1]. Indeed, the two approaches are closely related [BDZ21].

Quantum Control

Whether or not general-purpose quantum computing is a realistic goal for the next few decades, technologies like quantum sensing, quantum simulation and quantum communication are likely to remain active and fruitful areas of research. In each case, one of the main challenges is to control extremely delicate quantum mechanical systems such that they perform the desired information processing tasks with the required accuracy. Quantum control theory provides us with the apposite toolbox for this challenge. In this section, we give a brief introduction to the theory and the problems it presents [DH08]. See also the roadmaps [Gla+15, Koc+22] and the book by d'Alessandro [DAl21].

Using bilinear control theory we may formulate control systems modeling quantum mechanical experiments. Most commonly one studies the controlled Schrödinger equation

$$|\dot{\psi}(t)\rangle = -\mathrm{i}\Big(H_0 + \sum_{j=1}^{m} u_j(t) H_j\Big) |\psi(t)\rangle\,.$$

Here H_0 is the (uncontrolled) drift Hamiltonian, representing the natural evolution of the system, and the other H_j are the control Hamiltonians representing the possible controls that can be applied to the system. The functions $u_j : [0, T] \to \mathbb{R}$ are the control functions (or amplitudes) which modulate the control Hamiltonians. Depending on the experimental setup, there might be bounds on the maximal control amplitudes, or on their rate of change. Often one assumes that they are piecewise constant, but this is not a serious restriction.

Analogously one defines the controlled Lindblad equation

$$\dot{\rho} = -L(\rho) - \mathrm{i} \sum_{j=1}^{m} u_j(t)[H_j, \rho],$$

where $-L$ is the drift Lindblad generator representing the uncontrolled evolution of a Markovian quantum system and the H_j and u_j are as above. Sometimes the noise is switchable, i.e. it can be turned on and off at will, but we will not make this assumption.

There are two general kinds of questions one can ask about a control system. First, what can we do with it, and second, how do we do it in the best possible way? More concretely, the first question is about concepts like controllability, reachability, accessibility, and so on, meaning which states can be reached from a given initial state. The second question is about finding explicit solutions to the control system connecting initial and target state, and typically doing so in an optimal way.

As discussed above, optimal control synthesis can be performed analytically for small systems using the Pontryagin Maximum Principle (PMP) or the Hamilton–Jacobi–Bellmann (HJB) equation. Whenever such solutions can be obtained, one also gains a deep understanding of the problem at hand. In most practical cases however numerical methods are necessary. This is because the analytic approach only works for simple systems, which must be quite small and idealized. For larger systems with experimental imperfections and constraints, numerical methods are indispensable.

Two well-known methods are called GRAPE (gradient ascent pulse engineering) [Kha+05] and Krotov's method [RNK12], see [Mac+11] for a systematic comparison. An alternative approach is CRAB (chopped random-basis quantum optimization) [DCM11, CCM11], which expands the pulses in a functional basis and considers only a small subset of basis functions.

What Is This Thesis About?

This thesis develops a method of reduced control systems and applies it to quantum control theory. In particular we study two fundamental properties of quantum systems: decoherence and entanglement.

Reduced control systems The reduced control system is derived in full generality in Part I. We consider bilinear (or control-affine) control systems which admit fast control on a Lie group action. This assumption implies that two states in the same orbit can be considered equivalent. Hence the Lie group action can be factored out to obtain a new control system defined on the quotient space. The main result is that this reduced control system is, in a precise sense, equivalent to the original, full control system. Importantly, we present a constructive method for lifting solutions from the reduced system to the full system. To prove this result we need to generalize results from the perturbation theory of linear operators to the setting of symmetric Lie algebras. The remainder of the thesis then focuses on applying these abstract mathematical tools to more concrete physical systems.

Markovian systems with unitary control The current NISQ era of quantum computation is characterized by noisy systems, which severely limit the coherence time of the qubits and hence the number

of gates that can be applied in a quantum circuit while still achieving reasonable fidelity. The noise stems from unwanted interactions of the system with its environment. This interaction can be modeled in many different ways. One common approach is to assume that the interaction is Markovian, and to model it using a Lindblad equation. Additionally, we assume that we have fast control over the unitary group, that is, that we can implement arbitrary unitary transformations in an arbitrarily short amount of time. This assumption is justified if the unitary control is much faster than the dissipation of the system. Now the idea is the following. Since we can almost instantaneously apply any unitary transformation to the state ρ of the system, all states in the unitary orbit of ρ are effectively equivalent. Since the unitary orbit of ρ is exactly the set of all density matrices with the same eigenvalues as ρ, which we denote by λ, it should be possible to define an equivalent reduced control system on the set of eigenvalue vectors. Since the eigenvalues are non-negative and add up to unity, this set is the standard simplex, denoted Δ^{n-1}. This reduced control system does indeed exist, and it has the form

$$\dot{\lambda}(t) = -L_{U(t)}\lambda(t), \quad \lambda(0) = \lambda_0 \in \Delta^{n-1},$$

where $-L_{U(t)}$ is a stochastic generator matrix depending on the Lindblad generator $-L$ and the unitary $U(t)$, which is the new control function of the reduced control system. Part II is dedicated to the derivation and the study of this system.

Bipartite systems with local unitary control The main interest in quantum information technology stems from the fact that it yields (or at least promises to yield) quantum advantages in various fields, such as increased computational efficiency, cryptographic security or enhanced sensitivity. In many cases quantum entanglement is necessary to obtain such an advantage, and hence being able to control entanglement is essential. We will consider closed bipartite quantum systems, since in this case entanglement is well-understood. The two subsystems interact via some given coupling Hamiltonian, and we assume that we can implement arbitrary local unitary transformations in an arbitrarily short amount of time. Again this means that all states in the local unitary orbit of the initial state $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ may be considered equivalent. In this case the states are characterized by their singular values σ (due to the Schmidt decomposition), and hence there exists an equivalent reduced control system on the singular values. Due to the normalization of the state, the singular values lie on the hypersphere $S^{d_{\min}-1}$, where $d_{\min} = \min(d_1, d_2)$, and the reduced control system takes the form

$$\dot{\sigma}(t) = -H_{V(t)\otimes W(t)}\sigma(t), \quad \sigma(0) = \sigma_0 \in S^{d_{\min}-1},$$

where $-H_{V(t)\otimes W(t)}$ is a rotation generator matrix depending on the coupling Hamiltonian H_0, and the local unitary $V(t) \otimes W(t)$ is the new control function. Analogous results are obtained in the bosonic and fermionic cases. Part III explores these systems in detail.

Roadmap

The thesis is based on the works by the author listed on p. ix. More precisely, the thesis is based on the works [1–9], with each chapter corresponding (roughly) to one paper. The papers [10, 11] are only tangentially related to the main theme, and hence only touched upon briefly in the appendix.

To conclude the introduction, a concise outline is given here, with more detailed outlines given at the beginning of each part and each chapter.

Part I lays the mathematical foundation of the thesis. Chapter 1 is based on [1] and generalizes important results from the perturbation theory of linear operators to various matrix diagonalizations using the formalism of symmetric Lie algebras. Chapter 2, based on [2], introduces the method of reduced control systems which is the central tool in this work.

Part II applies these tools to dissipative quantum systems. Chapter 3, based on parts of [4], considers open Markovian quantum systems subject to fast unitary control and defines the reduced control system on the eigenvalues of the density matrix. Chapter 4, based on [4], parts of parts of [3], and [9], uses the reduced control system to draw first consequences for reachability, coolability and stabilizability in such systems. Chapter 5 is based on [5] and focuses on the case of a single qubit where explicit solutions and optimal controls can be obtained. Chapter 6 is based on [6] and considers the task of optimal cooling of Markovian quantum systems.

Part III applies the reduced control system to study entanglement in bipartite quantum systems. Chapter 7, based on parts of [7], considers closed bipartite quantum systems subject to fast local unitary control and defines the equivalent reduced control system on the singular values of the bipartite pure state. Chapter 8, based on parts of [7], uses the reduced control system to obtain some results on controllability, stabilizability and speed limits. Chapter 9, based on [8], applies the theory to low dimensional systems and obtains explicit solutions for optimal controls for tasks such as entanglement generation.

Part IV concludes the thesis while giving an outlook on outstanding problems and ongoing work. Appendices A and B briefly touch upon related work [10, 11].

Lists of figures, tables and theorems, as well as abbreviations and symbols, and an index are provided at the end for easy lookup.

Part I

Reduced Control Systems

„Die Mathematiker sind eine Art Franzosen: redet man zu ihnen, so übersetzen sie es in ihre Sprache, und dann ist es alsobald ganz etwas anders.“
— Johann Wolfgang von Goethe, *Maximen und Reflexionen* (1907)

« La géométrie est l'art du raisonnement correct à partir de figures mal dessinées. »
— Henri Poincaré, see [Let19]

This first part lays the mathematical and control theoretic foundation of the entire thesis. It develops the tools used in subsequent parts to study concrete applications in quantum control theory.

The main goal is to introduce a method of reduced control systems, and to do so in a mathematically rigorous way. The idea is that a bilinear (or, more generally, control affine) control system subject to fast control over a linear Lie group action can be reduced to the corresponding quotient space under certain conditions. To make this idea rigorous, we work in the setting of symmetric Lie algebras. This might sound restrictive, but we will show that this setting is rather general, corresponding to convenient geometric properties of the Lie group action. It turns out that these symmetric Lie algebras generalize and unify several matrix diagonalizations, such as the eigenvalue decomposition of Hermitian matrices and the complex singular value decomposition, but also some less well-known ones such as the Autonne–Takagi factorization and the Hua factorization. As mentioned in the introduction, these matrix decompositions occur in some relevant quantum mechanical systems, and for this reason it is worthwhile to establish a general theory using symmetric Lie algebras.

Outline Chapter 1 generalizes results from the perturbation theory of linear operators in the setting of symmetric Lie algebras. Chapter 2 defines the reduced control system, the central object of this thesis, and proves the important Equivalence Theorem.

Acknowledgments Part I is based on [1, 2], which are joint works with Gunther Dirr, Frederik vom Ende and Thomas Schulte-Herbrüggen. In [1], the original version of the proof of the analytic diagonalization is due to Gunther Dirr, and the appendix on orbifolds was carefully proofread and partially rewritten by Frederik vom Ende.

CHAPTER 1

Diagonalization in Symmetric Lie Algebras

1.1 Introduction

As outlined in the introduction of the thesis, we want to define reduced control systems, where the dynamics of the density matrix of an open system are reduced to those of its eigenvalues, and the dynamics of a bipartite pure state are reduced to those of its singular values. To do this, we have to be able to answer the following kind of questions. Given a path of Hermitian matrices $\rho(t)$, we might ask if it is possible to diagonalize all $\rho(t)$ in a consistent way. More precisely, if ρ is continuous, measurable, real analytic or k times differentiable, can we choose eigenvalue functions $\lambda_i(t)$ with the same properties? Similarly, can we choose a function $U(t)$ of diagonalizing unitaries with nice properties? Many of these questions have been answered for the eigenvalue decompositions of real symmetric and complex Hermitian matrices, as well as for singular value decompositions. However the treatment is not uniform, and several other diagonalizations have not been studied in the same detail. In particular we need to answer these questions also for the Autonne–Takagi factorization and the Hua factorization. For this reason, we consider symmetric Lie algebras, which provide a unifying framework for many notions of diagonalization, see Table 1.3 for some examples. We will answer the questions posed above and several more in this general setting.

We start by recalling some known results in this direction, which we will then generalize to symmetric Lie algebras. For the symmetric or Hermitian eigenvalue decomposition, many results can be found in [Rel69, Kat80, Bau85]. For instance, if a path of Hermitian matrices is continuous or continuously differentiable, then one can choose the eigenvalues to be continuous or continuously differentiable respectively. Furthermore [Kat80] shows that a real analytic path of Hermitian matrices can be diagonalized in a real analytic way. The real singular value decomposition (SVD) is considered in [Bun+91]; there it is shown that a real analytic path has a real analytic SVD, and that a smooth path of full-rank matrices with distinct singular values has a smooth SVD. In [QR14] it is shown that a measurable function of positive definite matrices can be measurably diagonalized. There are further results which may generalize to symmetric Lie algebras, but we will not do so here. For example, many results for Hermitian matrices have been extended to the infinite dimensional setting, again see [Kat80]. In [Rai11] improved results on differentiability are shown for normal and Hermitian matrices, see in particular Table 1 therein. In [MKS05] the behaviour of eigenvalues of matrices depending on several variables is studied.

Symmetric Lie Algebras

Now let us briefly introduce the notion of a symmetric Lie algebra,[1] which will provide a unifying framework for many diagonalizations. All necessary details are given in Appendix 1.A. This connection is explored in [Kle06] where algorithms for computing such diagonalizations are proposed. A symmetric Lie algebra is a (real, finite dimensional) Lie algebra $\mathfrak{g}$ together with an involutive Lie algebra automorphism s. This yields a vector space decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ into $+1$ and -1 eigenspaces of s which we call Cartan-like decomposition since it generalizes the usual Cartan decomposition. Importantly we have the following commutator relations: $[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}$, $[\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}$, $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$. This means that the adjoint action $\mathrm{ad}_k(\cdot) = [k, \cdot]$ of $\mathfrak{k}$ leaves $\mathfrak{p}$ invariant. Given a connected Lie group $\mathbf{G}$ with Lie algebra $\mathfrak{g}$, and a Lie subgroup $\mathbf{K} \subseteq \mathbf{G}$ with Lie algebra $\mathfrak{k}$, we say that the pair $(\mathbf{G}, \mathbf{K})$ is associated to the symmetric Lie algebra $(\mathfrak{g}, s)$. One can show that the adjoint action Ad of $\mathbf{K}$ on $\mathfrak{g}$ leaves $\mathfrak{p}$ invariant, an in fact the orbits of $\mathbf{K}$ in $\mathfrak{p}$ do not depend on the choice of $\mathbf{K}$. For this reason we can assume without loss of generality that $\mathbf{K}$ is connected. We will only consider symmetric Lie algebras which are semisimple and orthogonal, which implies that the group $\mathrm{Ad}_{\mathbf{K}}$ is compact. If $\mathfrak{a} \subseteq \mathfrak{p}$ is a maximal Abelian subspace, then every point $x \in \mathfrak{p}$ can be mapped to $\mathfrak{a}$ by some $K \in \mathbf{K}$, that is $\mathrm{Ad}_K(x) \in \mathfrak{a}$. This generalizes the idea of diagonalization. However the resulting element $\mathrm{Ad}_K(x) \in \mathfrak{a}$ is not unique, since the elements of $\mathbf{K}$ which leave $\mathfrak{a}$ invariant can act non-trivially on $\mathfrak{a}$. The resulting group of transformations of $\mathfrak{a}$ is called the Weyl group, denoted $\mathbf{W}$, and it is a finite group generated by reflections. A convenient fact about Weyl groups is that they admit a (closed) Weyl chamber $\mathfrak{w} \subseteq \mathfrak{a}$, such that each orbit $\mathrm{Ad}_{\mathbf{K}}(x)$ intersects $\mathfrak{w}$ in exactly one point. If this point lies in the interior of $\mathfrak{w}$, then x is called regular. Note that even if we fix $\mathrm{Ad}_K(x) \in \mathfrak{a}$, the element $K \in \mathbf{K}$ need still not be unique.

We can now formulate more precisely the questions that we will answer in this chapter. Given a path $p : I \to \mathfrak{p}$ with certain nice properties, can we choose a corresponding path $a : I \to \mathfrak{a}$ with similarly nice properties? How do we deal with the non-uniqueness of a caused by the Weyl group? What about a corresponding path $K : I \to \mathbf{K}$?

Outline and Main Results

The previous section set the stage by giving a quick introduction to the relevant concepts of symmetric Lie algebras. To make the exposition self-contained and to fix terminology, we give a rigorous treatment of symmetric Lie algebras in Appendix 1.A. We also list several examples in Section 1.2 which make the setting much more concrete.

In Section 1.3 we consider functions $p : X \to \mathfrak{p}$ that are continuous and we show that by diagonalizing them in a given Weyl chamber, the result is also continuous. Indeed the same argument extends to stronger forms of continuity, like uniform, Hölder, Lipschitz and absolute continuity, cf. Proposition 1.3.1.

In Section 1.4 we consider paths $p : I \to \mathfrak{p}$ that are differentiable. In Proposition 1.4.5 we show that if p is differentiable at a point $t \in I$, then $a : I \to \mathfrak{a}$ can be chosen to be differentiable at t, and we can give an explicit formula for this derivative. Furthermore, if p is (continuously) differentiable on I, then a can be chosen to be (continuously) differentiable on I, see Theorem 1.4.9. To prove this, we study (continuously) differentiable paths in orbifolds in Appendix 1.B. Moreover we show that if the path p only contains regular elements and is C^k, then one can find C^k paths K and a diagonalizing p, see Proposition 1.4.12.

In Section 1.5 we consider paths $p : I \to \mathfrak{p}$ that are real analytic. In this case one can find K and a real analytic, and moreover a is determined uniquely up to a global Weyl group action. This is in stark

[1] For background on Lie algebras see, e.g., [Kna02].

contrast to the differentiable case. In the analytic case we can also give a useful differential equation defining K. This is done in Theorem 1.5.7. Some tedious calculations are relegated to Appendix 1.C.

In Section 1.6 we consider paths $p : \Omega \to \mathfrak{p}$ that are measurable, where Ω can be any measurable space. Then we can find K and a measurable, see Theorem 1.6.6. In fact we can generalize this result to a finite family of commuting $p_i : \Omega \to \mathfrak{p}$ which we can then simultaneously measurably diagonalize, see Theorem 1.6.12. For absolutely continuous paths, this allows us to simultaneously measurably diagonalize the path and a certain projection of the derivative, see Proposition 1.6.13.

In Section 1.7 we show how the classification of simple Lie algebras over $\mathbb{C}$ and $\mathbb{R}$ translates to a classification of diagonalizations, and we explain in what sense all semisimple, orthogonal, symmetric Lie algebras are composed of these irreducible ones appearing in the classification, see Theorem 1.7.10.

1.2 Examples and Counterexamples

In order to warm up to our setting, let us revisit the eigenvalue decomposition of Hermitian matrices with the above Lie algebraic setting in mind. This example is crucial, as Part II is based on it. Moreover, it is useful in understanding the results presented in this chapter as it is, kind of by construction, a well-explored special case. Therefore we will use it as a running example throughout the chapter.

Example 1.2.1 (Hermitian EVD). *The semisimple Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ admits the Cartan decomposition $\mathfrak{sl}(n, \mathbb{C}) = \mathfrak{su}(n) \oplus \mathfrak{herm}_0(n, \mathbb{C})$ by means of the automorphism $s(X) = -X^*$. This gives it the structure of a semisimple, orthogonal, symmetric Lie algebra, and a possible pair associated to it is given by $(\mathrm{SL}(n, \mathbb{C}), \mathrm{SU}(n))$. Keeping the idea of diagonalization in mind, a convenient choice of a maximal Abelian subspace of $\mathfrak{herm}_0(n, \mathbb{C})$ (i.e. the traceless Hermitian $n \times n$ matrices) is the subset of all diagonal matrices. These will automatically be real and traceless; we denote this set by $\mathfrak{d}_0(n, \mathbb{R})$. The corresponding Weyl group — which captures the non-uniqueness of the diagonalized element from $\mathfrak{d}_0(n, \mathbb{R})$ — is isomorphic to the symmetric group S_n acting on n elements. The action on $\mathfrak{d}_0(n, \mathbb{R})$ is given by permutation of the diagonal elements of the matrix. Then, a natural choice of Weyl chamber is the subset of $\mathfrak{d}_0(n, \mathbb{R})$ with the diagonal elements in non-increasing order. The adjoint action of $X \in \mathrm{SL}(n, \mathbb{C})$ on $Y \in \mathfrak{sl}(n, \mathbb{C})$ is given by conjugation, that is, $\mathrm{Ad}_X(Y) = XYX^{-1}$ and similarly, for $X, Y \in \mathfrak{sl}(n, \mathbb{C})$, it holds that $\mathrm{ad}_X(Y) = [X, Y] := XY - YX$. This holds for matrix Lie algebras in general.*

The real SVD also corresponds to a symmetric Lie algebra, although the connection is less obvious than in Example 1.2.1.

Example 1.2.2 (Real SVD). *The pair $(\mathrm{SO}(p, q), \mathrm{SO}(p) \times \mathrm{SO}(q))$ is associated to the semisimple orthogonal symmetric Lie algebra $\mathfrak{so}(p, q)$ with $\mathfrak{k} = \mathfrak{so}(p) \oplus \mathfrak{so}(q)$ and $\mathfrak{p}$ equal to the set of matrices of the form $\left(\begin{smallmatrix} 0 & B \\ B^\top & 0 \end{smallmatrix}\right)$ where $B \in \mathbb{R}^{p,q}$. A maximal Abelian subspace is given by such matrices with $B \in \mathfrak{d}(p, q, \mathbb{R})$ diagonal, and the Weyl group acts by permutations and sign flips, so it is isomorphic to the signed symmetric group $\mathbb{Z}_2 \wr S_{p \wedge q}$ (here $\wr$ denotes the wreath product and $\wedge$ the minimum). The Weyl chamber consists of all diagonal matrices with non-negative diagonal elements in non-increasing order. The connection to the SVD stems from the adjoint action which is $\mathrm{Ad}_{(V,W)} \left(\begin{smallmatrix} 0 & B \\ B^\top & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & VBW^\top \\ (VBW^\top)^\top & 0 \end{smallmatrix}\right)$.*

As a special case we obtain the following:

Example 1.2.3 (Polar decomposition of $\mathbb{R}^n$). *Choosing $p = n$ and $q = 1$ in Example 1.2.2 yields the polar decomposition of $\mathbb{R}^n$, meaning that $\mathfrak{p} \cong \mathbb{R}^n$ and $\mathfrak{k} \cong \mathfrak{so}(n)$. The maximal Abelian subspaces are exactly the lines through the origin, with the Weyl group being isomorphic to $\mathbb{Z}_2$.*

Some further examples, namely the complex SVD, the Autonne–Takagi factorization and the Hua factorization, are discussed in great detail in Appendix 7.A. These are the diagonalizations on which Part III is based.

Now let us explore the most elementary example, the orthogonal diagonalization of (traceless) real symmetric 2×2 matrices, in a bit more detail. In particular this turns out to be equivalent to the polar decomposition of $\mathbb{C}$. As such, this is a special case of Example 1.2.3, but using a different symmetric Lie algebra.

Example 1.2.4 (Polar decomposition). *Consider the semisimple Lie algebra $\mathfrak{sl}(2,\mathbb{R})$. Similarly to Example 1.2.1, the automorphism $s(X) = -X^\top$ yields the Cartan decomposition $\mathfrak{sl}(2,\mathbb{R}) = \mathfrak{so}(2,\mathbb{R}) \oplus \mathfrak{sym}_0(2,\mathbb{R})$ into the orthogonal Lie algebra and the space of symmetric traceless matrices, and this yields the structure of a semisimple, orthogonal, symmetric Lie algebra. A choice of associated pair is given by $(\mathrm{SL}(2,\mathbb{R}), \mathrm{SO}(2,\mathbb{R}))$. Again we choose the diagonal matrices as our maximal Abelian subspace. Consider the identifications*

$$\imath : \mathfrak{sym}_0(2,\mathbb{R}) \to \mathbb{C}, \qquad \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \mapsto a + \mathrm{i}b,$$
$$\jmath : \mathrm{SO}(2,\mathbb{R}) \to \mathrm{U}(1), \quad \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \mapsto e^{\mathrm{i}2\phi},$$

where the first is an $\mathbb{R}$-linear isomorphism and the second is a double cover. Note that $\jmath$ induces an isomorphism on the quotient $\mathrm{SO}(2,\mathbb{R})/\{\pm\mathbb{1}\} \to U(1)$.

Either way, the chosen maximal Abelian subspace of $\mathfrak{sym}_0(2,\mathbb{R})$ corresponds (w.r.t. $\imath$) to the real numbers, with the non-negative numbers as an obvious choice of a Weyl chamber. One readily verifies $\imath(OAO^\top) = \jmath(O)\imath(A)$, which shows that the eigenvalue decomposition of real symmetric traceless 2 by 2 matrices is equivalent to the polar decomposition of complex numbers.

Interestingly, already in this simple setting many counterexamples can be found. The nature of these examples is that they violate regularity, that is, problems may occur as soon as the diagonalization does not live *only* in the interior of a Weyl chamber.

Example 1.2.5 (Differentiability of eigenvalues). *In [KM03, Example p. 2] it is shown that there exists a path $\rho : \mathbb{R} \to \mathfrak{sym}_0(2,\mathbb{R}) \cong \mathbb{C}$ which is C^∞, but the eigenvalues cannot be chosen as C^2 functions. This can only happen because the eigenvalues coincide at some point; such degeneracies correspond precisely to the boundary of the Weyl chamber of non-increasingly sorted eigenvalues. However, by [KM03, Thm. (C) p. 1] the eigenvalues can still be chosen twice differentiable.*

Example 1.2.6 (Continuity of diagonalization). *The following is Example 5.3 in [Kat80], originally due to Rellich. Consider the path $\rho : \mathbb{R} \to \mathfrak{sym}_0(2,\mathbb{R}) \cong \mathbb{C}$ given by*

$$\rho(x) = e^{-1/x^2} \begin{pmatrix} \cos(2/x) & \sin(2/x) \\ \sin(2/x) & -\cos(2/x) \end{pmatrix}, \quad \rho(0) = 0.$$

This path is C^∞ on $\mathbb{R}$, and so are the eigenvalues $\lambda_\pm = \pm e^{-1/x^2}$. However, there does not exist a continuous path of orthogonal matrices diagonalizing ρ.

Finally, semisimple, orthogonal, symmetric Lie algebras are closely related to the classification of simple Lie algebras over $\mathbb{C}$ and $\mathbb{R}$. Indeed, this connection allows for a classification of different diagonalizations in a certain irreducible case, and we will show that all diagonalizations considered here are in some sense composed of these irreducible diagonalizations, see Section 1.7.

Setting Throughout Sections 1.3–1.6 we consider a semisimple, orthogonal, symmetric Lie algebra $(\mathfrak{g}, s)$ with Cartan-like decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and with an associated pair $(\mathbf{G}, \mathbf{K})$ as defined in the introduction, where we assume that $\mathbf{K}$ is connected. We fix a choice of maximal Abelian subspace $\mathfrak{a} \subseteq \mathfrak{p}$ and a (closed) Weyl chamber $\mathfrak{w} \subseteq \mathfrak{a}$. The corresponding Weyl group is denoted by $\mathbf{W}$. A key geometric fact is that since $(\mathfrak{g}, s)$ is orthogonal, there exists an inner product on $\mathfrak{g}$, and hence also on $\mathfrak{p}$ and $\mathfrak{a}$, which is invariant under the action of $\mathbf{K}$ and $\mathbf{W}$ respectively. In particular $\mathrm{Ad}_{\mathbf{K}}$ is a compact Lie group and it acts isometrically on $\mathfrak{p}$. For precise definitions we refer to Appendix 1.A.

1.3 Continuous Diagonalization

We start with a natural way to make the diagonalization unique. Indeed, a basic fact about Weyl group actions is that they admit a Weyl chamber which intersects every $\mathbf{W}$-orbit, and every $\mathbf{K}$-orbit[2], in exactly one point, see Lemma 1.A.54. We denote by $\pi : \mathfrak{p} \to \mathfrak{p}/\mathbf{K}$ and $\pi_{\mathfrak{a}} : \mathfrak{a} \to \mathfrak{a}/\mathbf{W}$ the respective quotient maps. They are continuous and open. Consider the following diagram.[3]

$$\begin{array}{ccccc} \mathfrak{w} & \xhookrightarrow{\iota} & \mathfrak{a} & \xhookrightarrow{\iota} & \mathfrak{p} \\ & \searrow^{\psi} & \downarrow{\scriptstyle \pi_{\mathfrak{a}}} & & \downarrow{\scriptstyle \pi} \\ & & \mathfrak{a}/\mathbf{W} & \xhookrightarrow{\phi} & \mathfrak{p}/\mathbf{K} \end{array} \tag{1.1}$$

The maps $\psi(x) = \mathbf{W}x$ and $\phi(\mathbf{W}x) = \mathbf{K}x$ are the unique maps which make the diagram commute. Furthermore one can show that ψ and ϕ are in fact isometries, where the quotients $\mathfrak{a}/\mathbf{W}$ and $\mathfrak{p}/\mathbf{K}$ are endowed with their quotient metric. This is shown in Lemma 1.A.55. This crucially uses that all $\mathbf{K}$-orbits in $\mathfrak{p}$ intersect $\mathfrak{a}$ orthogonally. These facts already suffice to prove some interesting results:

Proposition 1.3.1. *For a given function p with values in $\mathfrak{p}$ we denote by $\mathrm{a}^{\downarrow} = \psi^{-1} \circ \phi^{-1} \circ \pi \circ p$ the corresponding function with values in $\mathfrak{w}$. Then it holds that $\pi \circ \mathrm{a}^{\downarrow} = \pi \circ p$ and*

(i) if $p : X \to \mathfrak{p}$ is continuous, then so is $\mathrm{a}^{\downarrow} : X \to \mathfrak{w}$;

(ii) if $p : Y \to \mathfrak{p}$ is uniformly continuous, then so is $\mathrm{a}^{\downarrow} : Y \to \mathfrak{w}$;

(iii) if $p : Y \to \mathfrak{p}$ is α-Hölder continuous, then so is $\mathrm{a}^{\downarrow} : Y \to \mathfrak{w}$, with the same constant $0 < \alpha \leq 1$;

(iv) if $p : Y \to \mathfrak{p}$ is L-Lipschitz continuous, then so is $\mathrm{a}^{\downarrow} : Y \to \mathfrak{w}$, with the same constant $L > 0$;

(v) if $p : I \to \mathfrak{p}$ is absolutely continuous, then so is $\mathrm{a}^{\downarrow} : I \to \mathfrak{w}$.

Here X denotes any topological space, Y any metric space, and I an interval.

Proof. From the commutativity of Diagram (1.1) it follows that $\pi \circ \mathrm{a}^{\downarrow} = \phi \circ \psi \circ \mathrm{a}^{\downarrow} = \pi \circ p$. The remaining statements follow immediately from the fact that π is non-expansive (by definition of the quotient metric, see Lemma 1.A.53) and the fact that $\phi \circ \psi$ is an isometry (Lemma 1.A.55). □

Remark 1.3.2. *In the setting of Example 1.2.1, this result generalizes the idea of [Kat80, p. 109] of choosing the eigenvalues continuously by ordering them in non-increasing order.*

Now one might wonder about the existence of a continuous function in $\mathbf{K}$ diagonalizing p, however, Example 1.2.6 shows that, even under stronger assumptions, continuity of the diagonalizing group elements cannot be guaranteed.

[2] We always consider the adjoint action of $\mathbf{K}$ on $\mathfrak{p}$, and so we will often shorten $\mathrm{Ad}_{\mathbf{K}}(x)$ to $\mathbf{K}x$.

[3] Here $\hookrightarrow$ denotes an injection and $\twoheadrightarrow$ denotes a surjection. By ι we denote the inclusion.

1.4 Differentiable Diagonalization

In this section we are interested in differentiable paths $p : I \to \mathfrak{p}$. We already know from Proposition 1.3.1 (v) that if p is absolutely continuous, then a can also be chosen absolutely continuous, and hence almost everywhere differentiable, simply by choosing $a = a^{\downarrow}$ to take values in the Weyl chamber $\mathfrak{w}$. However it is clear that this cannot work to give us a everywhere differentiable, as can be seen by choosing $p : I \to \mathfrak{a}$ differentiable and crossing several distinct Weyl chambers. Forcing a to take values in $\mathfrak{w}$ would introduce "kinks" in the path when p passes from one Weyl chamber to a different one. In this section we show that if $p : I \to \mathfrak{p}$ is (continuously) differentiable, then one can also choose $a : I \to \mathfrak{a}$ (continuously) differentiable, see Theorem 1.4.9. Then Example 1.2.5 shows that the analogous result for C^2 paths does not hold, and by Example 1.2.6 there might not even exist a continuous choice of $K : I \to \mathbf{K}$ diagonalizing p. Moreover we show that problems with the differentiability of a only occur at non-regular points. Indeed, Proposition 1.4.12 proves that if p is C^k and takes regular values, then we can find C^k paths K and a diagonalizing p.

Preliminaries

We start with an important geometric fact about the $\mathbf{K}$-orbits in $\mathfrak{p}$. For this we define the *commutant* of x in $\mathfrak{p}$ by $\mathfrak{p}_x = \{y \in \mathfrak{p} : [x, y] = 0\}$. Note that if $x \in \mathfrak{a}$ then $\mathfrak{a} \subseteq \mathfrak{p}_x$ with equality if and only if x is regular. It turns out that every $\mathbf{K}$-orbit in $\mathfrak{p}$ intersects the maximal Abelian subspace $\mathfrak{a}$ orthogonally, see Lemma 1.A.25. More precisely, for $x \in \mathfrak{a}$, the tangent space $T_x\mathfrak{p}$ splits into an orthogonal vector space sum of the tangent space to the orbit and $\mathfrak{p}_x$:

$$T_x\mathfrak{p} = T_x(\mathbf{K}x) \oplus \mathfrak{p}_x = \mathrm{ad}_{\mathfrak{k}}(x) \oplus \mathfrak{p}_x,$$

where we make liberal use of the canonical identification $T_x\mathfrak{p} \cong \mathfrak{p}$. We denote the orthogonal projection onto $\mathfrak{p}_x$ by $\Pi_x : \mathfrak{p} \to \mathfrak{p}_x$. Its kernel is then exactly $\mathrm{ad}_{\mathfrak{k}}(x)$. Similarly we denote by $\Pi_x^{\perp} = \mathbb{1} - \Pi_x$ the orthogonal projection onto $\mathrm{ad}_{\mathfrak{k}}(x)$ and with kernel $\mathfrak{p}_x$.

To gain some intuition let us consider a path $p : I \to \mathfrak{p}$ which admits a differentiable diagonalization, meaning that there exist differentiable $\mathrm{a} : I \to \mathfrak{a}$ and $K : I \to \mathbf{K}$ such that $p = \mathrm{Ad}_K(\mathrm{a})$.

Lemma 1.4.1. *Let $\mathrm{a} : I \to \mathfrak{a}$ and $K : I \to \mathbf{K}$ be differentiable and let $p = \mathrm{Ad}_K(\mathrm{a})$. Then*[4]

$$p' = \mathrm{Ad}_K(\mathrm{a}') - \mathrm{ad}_p(K'K^{-1}).$$

In particular it must hold that

$$\mathrm{ad}_p(K'K^{-1}) = -\Pi_p^{\perp} p' \tag{1.2}$$

$$\mathrm{a}' = \mathrm{Ad}_K^{-1}(\Pi_p p') = \Pi_{\mathrm{a}}(\mathrm{Ad}_K^{-1}(p')). \tag{1.3}$$

Proof. The first statement follows from a simple computation. Recall that $(K^{-1})' = -K^{-1}K'K^{-1}$. Then

$$p' = (K\mathrm{a}K^{-1})' = K\mathrm{a}'K^{-1} + K'\mathrm{a}K^{-1} - K\mathrm{a}K^{-1}K'K^{-1} = K\mathrm{a}'K^{-1} + [K'K^{-1}, p].$$

Conveniently, the two terms on the right hand side respect the orthogonal splitting of $\mathfrak{p}$ into kernel and image of ad_p, since $[p, \mathrm{Ad}_K(\mathrm{a}')] = \mathrm{Ad}_K([\mathrm{a}, \mathrm{a}']) = 0$, which proves the second statement. For the last equality we used Lemma 1.A.24 (iii). □

[4] We use a simplified notation in this lemma and its proof. For instance we use $K'K^{-1}$ as a shorthand for $r_K^{\star}(K')$, the pull back along the right multiplication r_K by K. If $\mathbf{K}$ is a matrix Lie group then both of these expressions are well defined and equal.

Lemma 1.4.1 already tells us much about the structure of differentiable paths $p : I \to \mathfrak{p}$, however it has some problems. It might seem that given p, we can find K by solving a differential equation obtained from (1.2), and then we can determine a by solving (1.3). Unfortunately, even if p is C^∞, there might not exist a diagonalizing function K which is continuous, see Example 1.2.6. Another problem is that in general the right hand side of (1.3) need not lie in $\mathfrak{a}$. Nevertheless, in Proposition 1.4.5 we will show that if p is differentiable at a point, then a can also be chosen differentiable at that point, and Formula (1.3) will return in a slightly modified form. Formula (1.2) will return in Section 1.5 where the much stronger condition of p being real analytic will guarantee that the solution K exists (and is itself real analytic). Similarly in Proposition 1.4.12 we will use Formula (1.2) to show that for regular C^k paths we can find a C^k diagonalization.

Before we can prove the main results of this section, we need to introduce some technical tools. For a point $x \in \mathfrak{p}$ we denote by $\mathbf{K}_x$ the *stabilizer* (also called *isotropy subgroup*) of x in $\mathbf{K}$. Similarly, for $y \in \mathfrak{a}$ we write $\mathbf{W}_y$ for the stabilizer of y in $\mathbf{W}$. With this we can define a number of quotient spaces. The details of the following facts can be found in Appendix 1.A.

For $x \in \mathfrak{a}$, there exists a homeomorphism $\phi_x : \mathfrak{a}/\mathbf{W}_x \to \mathfrak{p}_x/\mathbf{K}_x$, given by $\mathbf{W}_x z \mapsto \mathbf{K}_x z$, which is induced by the inclusion of $\mathfrak{a}$ in $\mathfrak{p}_x$.[5] Furthermore, it holds that if $y \in \mathrm{Ad}_{\mathbf{K}}(x)$ then there is a well-defined homeomorphism $\phi_{x,y} : \mathfrak{p}_x/\mathbf{K}_x \to \mathfrak{p}_y/\mathbf{K}_y$ induced by any $K \in \mathbf{K}$ with $\mathrm{Ad}_K(x) = y$. Summarizing, one can say that the diagram

$$\begin{array}{ccccc} \mathfrak{a} & \xhookrightarrow{\iota} & \mathfrak{p}_x & \xhookrightarrow{\mathrm{Ad}_K} & \mathfrak{p}_y \\ \downarrow{\scriptstyle \pi_{\mathfrak{a},x}} & & \downarrow{\scriptstyle \pi_x} & & \downarrow{\scriptstyle \pi_y} \\ \mathfrak{a}/\mathbf{W}_x & \xhookrightarrow{\phi_x} & \mathfrak{p}_x/\mathbf{K}_x & \xhookrightarrow{\phi_{x,y}} & \mathfrak{p}_y/\mathbf{K}_y \end{array} \tag{1.4}$$

commutes.

Although the quotients encountered here have singularities and hence are not manifolds, they can still be given the structure of an orbifold. In fact, the orbifolds that we deal with will have a single linear chart. The relevant facts about such orbifolds are proven in Appendix 1.B. In order to find a differentiable path $a : I \to \mathfrak{a}$, we need to be able to make sense of differentiable paths in such orbifolds. We start by defining the tangent bundle $T(\mathfrak{a}/\mathbf{W}) := (T\mathfrak{a})/\mathbf{W}$, where the action of $\mathbf{W}$ on $T\mathfrak{a}$ is given by $w \cdot (x, v) = (w \cdot x, w \cdot v)$. We denote the corresponding quotient map by

$$D\pi_{\mathfrak{a}} : T\mathfrak{a} \to T(\mathfrak{a}/\mathbf{W}) \tag{1.5}$$

If $x \in \mathfrak{a}$, then the tangent space in $\mathfrak{a}/\mathbf{W}$ at the point $\pi_{\mathfrak{a}}(x)$ is denoted by $T_{\pi_{\mathfrak{a}}(x)}(\mathfrak{a}/\mathbf{W})$ and it turns out to be homeomorphic to $(T_x\mathfrak{a})/\mathbf{W}_x$. Since one can canonically identify $T_x\mathfrak{a}$ and $\mathfrak{a}$, the commutative Diagram (1.4) shows that we have the homeomorphisms

$$T_{\pi_{\mathfrak{a}}(x)}(\mathfrak{a}/\mathbf{W}) \cong \mathfrak{a}/\mathbf{W}_x \cong \mathfrak{p}_x/\mathbf{K}_x.$$

Hence we can define the differential of the quotient map $\pi_{\mathfrak{a}}$ at a point x as the map

$$D\pi_{\mathfrak{a}}(x) : T_x\mathfrak{a} \to T_{\pi_{\mathfrak{a}}(x)}(\mathfrak{a}/\mathbf{W}), \quad v \mapsto \pi_{\mathfrak{a},x}(v).$$

Let us briefly recall what it means for a path $\xi : I \to \mathfrak{a}/\mathbf{W}$ to be differentiable in the orbifold sense, as defined in Definition 1.B.3. We say that ξ is differentiable at $t \in I$ if there exists a function $a : I \to \mathfrak{a}$ satisfying $\pi_{\mathfrak{a}} \circ a = \xi$, called a lift of ξ, which is differentiable at t. The derivative of ξ at t is given

[5] In particular, setting $x = 0$ one gets the homeomorphism $\phi : \mathfrak{a}/\mathbf{W} \to \mathfrak{p}/\mathbf{K}$ which we used in Section 1.3.

by $D\xi(t) := D\pi_{\mathfrak{a}}(a(t), a'(t))$ and it is well-defined. If ξ is differentiable at every $t \in I$, then we say that ξ is differentiable and if additionally $D\xi : I \to T(\mathfrak{a}/\mathbf{W})$ is continuous, then ξ is continuously differentiable or C^1. In the following proofs we will show that if p is (continuously) differentiable, then so is $\xi := \phi^{-1} \circ \pi \circ p$. Then we use Proposition 1.B.8 to show that there exists a (continuously) differentiable lift $a : I \to \mathfrak{a}$ of ξ.

Differentiable Diagonalization at a Point

First we consider differentiability at a single point. The following results will be quite useful. For the proof of the first lemma we use the concept of a "slice" for the action of $\mathbf{K}$ on $\mathfrak{p}$. A slice at a point x is an embedded submanifold containing x and intersecting the orbit through x in a complementary way, see Definition 1.A.50. Such slices exist in very general settings, but in our case we can even choose a slice in $\mathfrak{p}_x$ intersecting the orbit orthogonally. The main idea is then to "project" the path p onto the slice while keeping each point in its original orbit.

Lemma 1.4.2. *Let I be an open interval and let $p : I \to \mathfrak{p}$ be differentiable at $t_0 \in I$. Then there exists $K \in \mathbf{K}$ such that $x := \mathrm{Ad}_K^{-1}(p(t_0)) \in \mathfrak{a}$ and such that $v := \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))) \in \mathfrak{a}$. Moreover, there is a subinterval $I' \subseteq I$ containing t_0 and a function $\tilde{p} : I' \to \mathfrak{p}_x$ satisfying $\tilde{p}(t_0) = x$ and $\tilde{p}'(t_0) = v$ and $\pi \circ \tilde{p} = \pi \circ p$ on I'.*

Proof. By definition of the projection, $p(t_0)$ and $\Pi_{p(t_0)}(p'(t_0))$ commute, and hence by Lemma 1.A.26 there is some $K \in \mathbf{K}$ such that $x = \mathrm{Ad}_K^{-1}(p(t_0)) \in \mathfrak{a}$ and $v = \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))) \in \mathfrak{a}$. By the chain rule $\mathrm{Ad}_K^{-1}(p)$ is differentiable at t_0 and by linearity of Ad_K it holds that $(\mathrm{Ad}_K^{-1} \circ p)'(t_0) = \mathrm{Ad}_K^{-1}(p'(t_0))$.

By Lemma 1.A.51 there exists a slice S_x at x for the action of $\mathbf{K}$ on $\mathfrak{p}$, which is contained in $\mathfrak{p}_x$. Let $\mathfrak{k}_x := \mathfrak{k} \cap \ker \mathrm{ad}_x$ and let $\mathfrak{k}_x^\perp$ be the orthogonal complement of $\mathfrak{k}_x$ in $\mathfrak{k}$. Let O be an open neighborhood of the origin in $\mathfrak{k}_x^\perp$ and consider the map $\sigma : O \times S_x \to \mathfrak{p} : (k, y) \mapsto \mathrm{Ad}_{e^k}(y)$. Since

$$D\sigma(0, x) : \mathfrak{k}_x^\perp \oplus \mathfrak{p}_x \to \mathfrak{p}, \quad (l, z) \mapsto [l, x] + z, \tag{1.6}$$

Lemma 1.A.16 shows that $D\sigma(0, x)$ is bijective and by the inverse function theorem, and potentially by shrinking O and S_x, we may assume that σ is a diffeomorphism onto its image, denoted V. Hence $x \in V$ and σ can be seen as a chart for V. On V we define the smooth map $\kappa = \sigma \circ \mathrm{pr}_2 \circ \sigma^{-1}$, where pr_2 sets the first coordinate to 0. Then $\kappa(x) = x$ and (1.6) shows that $D\kappa(x) = \Pi_x$. By continuity of p at t_0, there is an open interval $I' \subseteq I$ containing t_0 such that the image of $\mathrm{Ad}_K^{-1}(p)$ on I' lies in V. Set $\tilde{p} = \kappa \circ \mathrm{Ad}_K^{-1} \circ p|_{I'}$, then $\pi \circ \tilde{p} = \pi \circ p$ on I', and $\tilde{p}(t_0) = x$, and $\tilde{p}'(t_0) = D\kappa(x)(\mathrm{Ad}_K^{-1}(p'(t_0)) = \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))) = v$ by Lemma 1.A.24 (iii). □

For the next lemma we use Diagram (1.4) as well as the fact that the stabilizer subgroup $\mathbf{W}_x$ still has the properties of a Weyl group and hence admits a (closed) Weyl chamber, which we denote $\tilde{\mathfrak{w}}$. In this step we go from the commutant $\mathfrak{p}_x$ to $\mathfrak{a}$ by diagonalizing in the appropriate Weyl chamber of $\mathbf{W}_x$. This makes sure that the path remains differentiable.

Lemma 1.4.3. *Let $x \in \mathfrak{a}$ be given and let $\tilde{\mathfrak{w}} \subseteq \mathfrak{a}$ be a Weyl chamber for the action of $\mathbf{W}_x$ on $\mathfrak{a}$. Then*

(i) there is a continuous map $\omega : \mathfrak{p}_x \to \tilde{\mathfrak{w}}$ satisfying $\pi_x \circ \omega = \pi_x$,

(ii) for any sequence y_n in $\mathfrak{p}_x$ converging to some $y \in \tilde{\mathfrak{w}}$ there is a subsequence y_n' and a sequence $K_n' \in \mathbf{K}_x$ such that $\mathrm{Ad}_{K_n'}^{-1}(y_n') \in \tilde{\mathfrak{w}}$ converge to y and there is some $K \in \mathbf{K}_x \cap \mathbf{K}_y$ such that $\mathrm{Ad}_{K_n'} \to \mathrm{Ad}_K$.

Proof. Recall from Corollary 1.A.57 that we have a homeomorphism $\phi_x : \mathfrak{a}/\mathbf{W}_x \to \mathfrak{p}_x/\mathbf{K}_x$ induced by the inclusion $\mathfrak{a} \hookrightarrow \mathfrak{p}_x$. By Lemma 1.A.59 the action of $\mathbf{W}_x$ on $\mathfrak{a}$ admits a (closed) Weyl chamber $\tilde{\mathfrak{w}}$. The proof of Lemma 1.A.55 also yields a homeomorphism $\psi_x : \tilde{\mathfrak{w}} \to \mathfrak{a}/\mathbf{W}_x$ induced by the inclusion $\tilde{\mathfrak{w}} \hookrightarrow \mathfrak{a}$. Combining this we define

$$\omega = \psi_x^{-1} \circ \phi_x^{-1} \circ \pi_x.$$

Since $\psi_x = \pi_{\mathfrak{a},x}$ on $\tilde{\mathfrak{w}}$ and $\phi_x \circ \pi_{\mathfrak{a},x} = \pi_x$ on $\mathfrak{a}$, it holds that $\pi_x \circ \omega = \pi_x$. This proves (i).

This shows that for every element $y_n \in \mathfrak{p}_x$ there is some $K_n \in \mathbf{K}_x$ with $\mathrm{Ad}_{K_n}^{-1}(y_n) \in \tilde{\mathfrak{w}}$. The same point in $\tilde{\mathfrak{w}}$ can be obtained using the continuous map ω applied to y_n. Since $y \in \tilde{\mathfrak{w}}$ it holds that

$$\mathrm{Ad}_{K_n}^{-1}(y_n) = \omega(y_n) \to \omega(y) = y.$$

The existence of a subsequence K_n' with the desired properties follows from the compactness of $\mathrm{Ad}_{\mathbf{K}}$. This proves (ii). □

Corollary 1.4.4. *Let $x \in \mathfrak{a}$ be given and let $\tilde{p} : I' \to \mathfrak{p}_x$ be differentiable at $t_0 \in I'$ satisfying $x = \tilde{p}(t_0)$ and $v := \tilde{p}'(t_0) \in \mathfrak{a}$. Then*

(i) there exists $\mathrm{a} : I' \to \mathfrak{a}$ differentiable at t_0 with $\mathrm{a}(t_0) = x$ and $\mathrm{a}'(t_0) = v$ and $\pi \circ \mathrm{a} = \pi \circ \tilde{p}$, and

(ii) for any sequence $t_n \to t_0$ in I' there is a subsequence t_n' and elements $K_n' \in \mathbf{K}_x$ and $K \in \mathbf{K}_x \cap \mathbf{K}_v$ such that $\mathrm{Ad}_{K_n'}^{-1}(\tilde{p}(t_n')) = \mathrm{a}(t_n')$ and such that $\mathrm{Ad}_{K_n'} \to \mathrm{Ad}_K$.

Proof. Let $\tilde{\mathfrak{w}}$ be a Weyl chamber for $\mathbf{W}_x$ containing v and let $\omega : \mathfrak{p}_x \to \tilde{\mathfrak{w}}$ denote the map from Lemma 1.4.3 (i). Define the path

$$\mathrm{a} : I' \to \mathfrak{a}, \quad t \mapsto \begin{cases} \omega\Big(\frac{\tilde{p}(t)-x}{t-t_0}\Big)(t-t_0) + x & \text{if } t \neq t_0 \\ x & \text{if } t = t_0. \end{cases}$$

Note that for $t > t_0$, a lies in $\tilde{\mathfrak{w}}$, and for $t < t_0$, a lies in $-\tilde{\mathfrak{w}}$. Then by continuity of ω it holds that

$$\frac{\mathrm{a}(t) - x}{t - t_0} = \omega\Big(\frac{\tilde{p}(t) - x}{t - t_0}\Big) \to \omega(v) = v,$$

as $t \to t_0$. By Lemma 1.4.3 (i) there exists for every $t \in I' \setminus \{t_0\}$ some element $K_t \in \mathbf{K}_x$ such that

$$\frac{\mathrm{a}(t) - x}{t - t_0} = \mathrm{Ad}_{K_t}^{-1}\Big(\frac{\tilde{p}(t) - x}{t - t_0}\Big), \tag{1.7}$$

and hence $\mathrm{a}(t) = \mathrm{Ad}_{K_t}^{-1}(\tilde{p}(t))$ which shows that $\pi \circ \mathrm{a} = \pi \circ \tilde{p}$. Hence a satisfies all the desired properties and this proves (i). Now let any sequence $t_n \to t_0$ in I' be given and set $K_n = K_{t_n}$. Then $K_n \in \mathbf{K}_x$ and $\mathrm{Ad}_{K_n}^{-1}(\tilde{p}(t_n)) = \mathrm{a}(t_n)$. By Eq. (1.7) and the compactness of $\mathrm{Ad}_{\mathbf{K}}$, there is a subsequence of Ad_{K_n} converging to some Ad_K with $K \in \mathbf{K}_x \cap \mathbf{K}_v$. This proves (ii). □

Now we are ready to prove the first main result of this section, which shows that if p is differentiable at some point, then one can also choose a to be differentiable at that point. Moreover the derivative of a is then unique up to some Weyl group action.

Proposition 1.4.5. *Let I be an open interval and let $p : I \to \mathfrak{p}$ be differentiable at some $t_0 \in I$. Then there is $\mathrm{a} : I \to \mathfrak{a}$ which is differentiable at t_0 and satisfies $\pi \circ p = \pi \circ \mathrm{a}$. Moreover there is some $K \in \mathbf{K}$ such that*

$$\mathrm{a}(t_0) = \mathrm{Ad}_K^{-1}(p(t_0)) \quad \textit{and} \quad \mathrm{a}'(t_0) = \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))),$$

and any other path $b : I \to \mathfrak{a}$ which is differentiable at t_0 and satisfies $\pi \circ p = \pi \circ b$ also satisfies that $b(t_0) = w \cdot \mathrm{a}(t_0)$ and $b'(t_0) = w \cdot \mathrm{a}'(t_0)$ for some $w \in \mathbf{W}$.

Remark 1.4.6. *If we write $\xi = \phi^{-1} \circ \pi \circ p$, then this proposition shows that ξ is differentiable at t_0 in the orbifold sense, as defined in Definition 1.B.3. The derivative of ξ is then*

$$D\xi(t_0) = D\pi_{\mathfrak{a}}(\mathrm{Ad}_K^{-1}(p(t_0)), \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0)))),$$

for any $K \in \mathbf{K}$ such that $\mathrm{Ad}_K^{-1}(p(t_0)) \in \mathfrak{a}$ and $\mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))) \in \mathfrak{a}$.

Proof. By Lemma 1.4.2 we find some $K \in \mathbf{K}$ such that $x := \mathrm{Ad}_K^{-1}(p(t_0)) \in \mathfrak{a}$ and at the same time $v := \mathrm{Ad}_K^{-1}(\Pi_{p(t_0)}(p'(t_0))) \in \mathfrak{a}$, as well as some open interval $I' \subseteq I$ containing t_0 and a path $\tilde{p} : I' \to \mathfrak{p}_x$ with $\tilde{p}(t_0) = x$ and $\tilde{p}'(t_0) = v$ satisfying $\pi \circ \tilde{p} = \pi \circ p$ on I'. Then by Corollary 1.4.4 (i) we obtain $\mathrm{a} : I' \to \mathfrak{a}$ satisfying the desired properties. The uniqueness of $(\mathrm{a}(t_0), \mathrm{a}'(t_0))$ up to Weyl group action follows immediately from Lemma 1.B.5 (i). □

Example 1.4.7. *Let us illustrate this result in the setting of Example 1.2.1. Let $\rho : I \to \mathfrak{herm}_0(n, \mathbb{C})$ be a path of traceless Hermitian matrices which is differentiable at some $t_0 \in I$. Let $\rho(t_0) = \sum_{j=1}^m \mu_j P_j$ be the eigendecomposition of $\rho(t_0)$. Then it holds that $\Pi_{\rho(t_0)}(\rho'(t_0)) = \sum_{j=1}^m P_j \rho'(t_0) P_j$. Using a unitary change of basis, we can assume that both $\rho(t_0)$ and $\Pi_{\rho(t_0)}(\rho'(t_0))$ are diagonal. Then, by Proposition 1.4.5 there exist eigenvalue functions $\lambda_i : I \to \mathbb{R}$ which are differentiable at t_0 and satisfy $\lambda_i(t_0) = \rho_{i,i}(t_0)$ and $\lambda_i'(t_0) = \rho_{i,i}'(t_0)$. The formula for $\lambda_i'(t_0)$ coincides with that of [Kat80, Ch. II, Thm. 5.4] and [Rel69, Ch. I.§5, Thm. 1].*

Continuously Differentiable Diagonalization

After considering differentiability at a single point, we want to extend the result to the entire path, both in the differentiable and in the continuously differentiable case. Most of the heavy lifting will be done in Appendix 1.B. We have shown that if p is differentiable at a point, then so is ξ in the sense of orbifolds, see Remark 1.4.6. By definition, this means that if p is everywhere differentiable, then so is ξ. The following technical lemma extends this to continuous differentiability.

Lemma 1.4.8. *Let $p : I \to \mathfrak{p}$ be continuously differentiable. Then $\xi : I \to \mathfrak{a}/\mathbf{W}$ given by $\xi = \phi^{-1} \circ \pi \circ p$ is continuously differentiable in the sense of orbifolds.*

Proof. By Proposition 1.4.5 we know that ξ is differentiable on I in the sense of orbifolds, and its derivative is denoted by $D\xi : I \to T(\mathfrak{a}/\mathbf{W})$. Let $t_0 \in I$ be arbitrary. We want to show that $D\xi$ is continuous at t_0. By Lemma 1.4.2 we obtain a path $\tilde{p}$ on an open interval $I' \subseteq I$ containing t_0 which satisfies $\pi \circ \tilde{p} = \pi \circ p$. From the definition of $\tilde{p}$ it is clear that it is C^1. Hence in the following we will work with $\tilde{p}$ instead of p. Note that by Remark 1.4.6 we know that $D\xi(t_0) = D\pi_{\mathfrak{a}}(x, v)$ where $x = \tilde{p}(t_0)$ and $v = \tilde{p}'(t_0)$.

First we show that $\Pi_{\tilde{p}(t)}(\tilde{p}'(t))$ is continuous at t_0. Let $\mathrm{a} : I' \to \mathfrak{a}$ be the function given by Corollary 1.4.4 (i). Let t_n be any sequence in I' converging to t_0. Then by Corollary 1.4.4 (ii) there is a subsequence t_n' and a sequence of elements $K_n' \in \mathbf{K}_x$ such that $\mathrm{Ad}_{K_n'}^{-1}(\tilde{p}(t_n')) = \mathrm{a}(t_n')$ as well as some

$K \in \mathbf{K}_x \cap \mathbf{K}_v$ such that $\mathrm{Ad}_{K'_n} \to \mathrm{Ad}_K$. To simplify notation we write $x_n = \mathrm{Ad}^{-1}_{K'_n}(\tilde{p}(t'_n))$ and $v_n = \mathrm{Ad}^{-1}_{K'_n}(\tilde{p}'(t'_n))$. Since a is continuous at t_0 it holds that $x_n \to x$, and since $\tilde{p}'(t'_n) \to v$ it holds that $v_n \to v$. If $\Pi_{\mathfrak{a}} : \mathfrak{p} \to \mathfrak{a}$ denotes the orthogonal projection onto $\mathfrak{a}$, then for any $z \in \mathfrak{a}$ it holds that $\Pi^{\perp}_z \circ \Pi_{\mathfrak{a}} = 0$ since $\mathfrak{a} \subseteq \mathfrak{p}_z$. Since $v \in \mathfrak{a}$ and since Π_{x_n} is an orthogonal projection, it holds that

$$\Pi_{x_n} v_n = v_n - \Pi^{\perp}_{x_n} v_n = v_n - \Pi^{\perp}_{x_n} \Pi^{\perp}_{\mathfrak{a}} v_n \to v.$$

Hence by Lemma 1.A.24 (iii) it holds that

$$\Pi_{\tilde{p}(t'_n)}(\tilde{p}'(t'_n)) = \mathrm{Ad}_{K'_n}(\Pi_{x_n}(v_n)) \to v,$$

as desired. Since for every sequence t_n we have found a subsequence t'_n, this shows that $\Pi_{\tilde{p}(t)}(\tilde{p}'(t))$ is continuous at t_0.

Now we show that $D\xi$ is also continuous at t_0. This is done using a similar method. Again let $t_n \to t_0$ be given. Consider the sequence $\Pi_{\tilde{p}(t_n)}(\tilde{p}'(t_n))$ which converges to v as shown above and note that by Lemma 1.A.52 the sequence lies in $\mathfrak{p}_x$. Then by Lemma 1.4.3 (ii) applied to this sequence, there exists a subsequence t'_n and elements $K'_n \in \mathbf{K}_x$ satisfying $v_n := \mathrm{Ad}^{-1}_{K'_n}(\Pi_{\tilde{p}(t'_n)}(\tilde{p}'(t'_n))) \in \tilde{\mathfrak{w}}$ and $v_n \to v$, as well as some $K \in \mathbf{K}_x \cap \mathbf{K}_v$ such that $\mathrm{Ad}_{K'_n} \to \mathrm{Ad}_K$. For each n, by Corollary 1.A.49, we find some $L_n \in \mathbf{K}_x \cap \mathbf{K}_{v_n}$ such that $x_n := \mathrm{Ad}^{-1}_{L_n} \mathrm{Ad}^{-1}_{K'_n}(\tilde{p}(t'_n)) \in \mathfrak{a}$. By Remark 1.4.6 it holds that $D\xi(t'_n) = D\pi_{\mathfrak{a}}(x_n, v_n)$. Moreover it holds that $x_n \to x$. Hence

$$D\xi(t'_n) = D\pi_{\mathfrak{a}}(x_n, v_n) \to D\pi_{\mathfrak{a}}(x, v) = D\xi(t_0)$$

by continuity of the quotient map $D\pi_{\mathfrak{a}}$. This concludes the proof. □

So far we have shown that if $p : I \to \mathfrak{p}$ is (continuously) differentiable, then $\xi : I \to \mathfrak{a}/\mathbf{W}$ is (continuously) differentiable in the sense of orbifolds. At this point it is not at all clear that a corresponding (continuously) differentiable path $a : I \to \mathfrak{a}$ must also exist. That this is the case is shown in detail in Appendix 1.B in the more general setting of orbifolds.

Theorem 1.4.9 (Differentiable Diagonalization). *Let $p : I \to \mathfrak{p}$ be (continuously) differentiable, then there exists a (continuously) differentiable path $\mathrm{a} : I \to \mathfrak{a}$ satisfying $\pi \circ p = \pi \circ \mathrm{a}$. Moreover, for every $t \in I$, there is some $K \in \mathbf{K}$ such that $\mathrm{Ad}^{-1}_K(p(t)) \in \mathfrak{a}$ and $\mathrm{Ad}^{-1}_K(\Pi_{p(t)}(p'(t))) \in \mathfrak{a}$, and for any such K it holds that*

$$(\mathrm{a}(t), \mathrm{a}'(t)) = w \cdot \mathrm{Ad}^{-1}_K(p(t), \Pi_{p(t)}(p'(t)))$$

for some Weyl group element $w \in \mathbf{W}$.

Proof. The differentiable case follows from Proposition 1.4.5 combined with Proposition 1.B.8 (ii). The continuously differentiable case follows from Lemma 1.4.8 combined with Proposition 1.B.8 (iii). □

Remark 1.4.10. *Considering the Cartan decomposition $\mathfrak{sl}(n, \mathbb{C}) = \mathfrak{su}(n) \oplus \mathfrak{herm}_0(n)$, Theorem 1.4.9 generalizes a well-known result by Rellich (see [Rel69, Ch. I.§5, Thm. 1], as well as [Rel69, Ch. I.§5, Theorem, pp. 44-45], or, for a simpler proof, [Kat80, Ch. II, Thm. 6.8]) showing that for a C^1 path of Hermitian matrices, the eigenvalues can be chosen as C^1 functions.*

Remark 1.4.11. *Let us mention some counterexamples to different generalizations of this result. [Kat80, Example 5.9] shows that for a C^1 path of diagonalizable (but not symmetric) matrices, the eigenvalues need not be C^1 (but they are differentiable). Due to Example 1.2.5, even if p is C^∞ we cannot guarantee that a can be chosen C^2. Note also that the diagonalizing unitary may have to be discontinuous, see Example 1.2.6. On the other hand, slight improvements of Theorem 1.4.9 might be possible by generalizing results from [Rai11].*

Higher Derivatives in the Regular Case

The following result shows that as long as we don't run into non-regular points, i.e., points with non-trivial stabilizer in the Weyl group, a C^k path can always be diagonalized in a C^k fashion. In Section 1.5 we will show that real analytic paths always have a real analytic diagonalization, even without the exclusion of non-regular points. First we need to define an inverse of ad_x by restricting the domain and codomain. Indeed we get a well-defined inverse $\operatorname{ad}_x^{-1} : \mathfrak{p}_x^\perp \to \mathfrak{k}_x^\perp$, since $\mathfrak{p}_x^\perp = \mathfrak{p} \cap \operatorname{im} \operatorname{ad}_x$. Note that this is essentially a restriction of the Moore–Penrose pseudo-inverse of ad_x. Recall that $\Pi_x^\perp$ is the orthogonal projection onto $\mathfrak{p}_x^\perp$.

Proposition 1.4.12. *Let I be an open interval and let $p : I \to \mathfrak{p}$ be regular and C^k for $0 \leq k \leq \infty$. Then there exists a C^k path $K : I \to \mathbf{K}$ such that $\mathrm{a}(t) := \operatorname{Ad}_K^{-1}(t)(p(t)) \in \mathfrak{a}$ for all $t \in I$. Moreover, for $k \geq 1$ we can choose K to satisfy*

$$K'(t) = h(t)K(t), \quad h(t) = -\operatorname{ad}_{p(t)}^{-1}(\Pi_{p(t)}^\perp(p'(t))), \quad \operatorname{Ad}_{K(t_0)}^{-1}(p(t_0)) \in \mathfrak{a}, \quad t_0 \in I. \tag{1.8}$$

Furthermore, any continuous path $b : I \to \mathfrak{a}$ satisfying $\pi \circ b = \pi \circ p$ satisfies $b = w \cdot \mathrm{a}$ for some fixed Weyl group element $w \in \mathbf{W}$.

Proof. First we consider the continuous case $k = 0$. Let $\mathfrak{p}_0$ denote the set of all regular points of $\mathfrak{p}$. By Lemma 1.6.4 (proven later) this is a trivial smooth fiber bundle over the open Weyl chamber $\mathfrak{w}_0$ with fiber $\mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$, where $Z_{\mathbf{K}}(\mathfrak{a}) = \mathbf{K}_{\mathfrak{a}}$ is the centralizer (stabilizer) of $\mathfrak{a}$ in $\mathbf{K}$. Hence we can project p to give continuous paths in $\mathfrak{w}_0$ and $\mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$. It remains to continuously lift the path in $\mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$ to $\mathbf{K}$. For any $t \in I$ one can find a local continuous lift in a neighborhood of t by working in any local trivialization of the bundle $\pi_{\mathbf{K}} : \mathbf{K} \to \mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$. Then such local lifts can be glued together to a global continuous lift, analogously to the proof of Lemma 1.B.7.

Now consider $k \geq 1$. By the above, there is some continuous $L : I \to \mathbf{K}$ such that $\operatorname{Ad}_{L(t)}^{-1}(p(t)) \in \mathfrak{a}$. Define h as in (1.8). Then by Lemma 1.A.24 (ii) and (iv) it holds that

$$h = -(\operatorname{Ad}_L \circ \operatorname{ad}^{-1}_{\operatorname{Ad}_L^{-1}(p)} \circ \Pi_{\mathfrak{a}}^\perp \circ \operatorname{Ad}_L^{-1})(p'). \tag{1.9}$$

This shows that h is continuous. Define K as in (1.8). In order to verify that $\operatorname{Ad}_K^{-1}(p)$ lies in $\mathfrak{a}$, we compute

$$(\operatorname{Ad}_K^{-1}(p))' = \operatorname{Ad}_K^{-1}(p') + [\operatorname{Ad}_K^{-1}(p), K^{-1}K'] = \operatorname{Ad}_K^{-1}(p') + \operatorname{Ad}_K^{-1}([p,h]) = \Pi_{\operatorname{Ad}_K^{-1}(p)}(\operatorname{Ad}_K^{-1}(p')),$$

and hence the part of the derivative tangent to the fibers is always zero. Together with the assumption $\operatorname{Ad}_{K(t_0)}^{-1}(p(t_0)) \in \mathfrak{a}$, this shows that $\operatorname{Ad}_K^{-1}(p)$ remains in $\mathfrak{a}$ at all times. It remains to show that we indeed have the desired level of differentiability. Note that if $h \in C^{j-1}$, this implies that $K \in C^j$, and by replacing L by K in (1.9), we see that $K \in C^j$ implies that $h \in C^{\min(j,k-1)}$, which by induction implies that $K \in C^k$. Finally, the uniqueness claim is clear, since a continuous lift of $\pi \circ p$ in $\mathfrak{a}$ must lie in a single open Weyl chamber and is uniquely defined within this Weyl chamber. □

Remark 1.4.13. *The homogeneous space $\mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$ is reductive with $\mathfrak{k} = \mathfrak{z}_{\mathfrak{k}}(\mathfrak{a}) \oplus \mathfrak{z}_{\mathfrak{k}}(\mathfrak{a})^\perp$, since $Z_{\mathbf{K}}(\mathfrak{a})$ is compact. This induces a connection of the principal bundle $\pi : \mathbf{K} \to \mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$, called the canonical connection, cf. [KN96, Ch. II, Thm 11.1]. In Proposition 1.4.12 we implicitly used this connection to lift differentiable paths from $\mathbf{K}/Z_{\mathbf{K}}(\mathfrak{a})$ to $\mathbf{K}$.*

Remark 1.4.14. *Proposition 1.4.12 generalizes [Bun+91, Thm. 2] which shows that a C^1 path of real m by n matrices of full-rank and with distinct singular values has a C^1 singular value decomposition.*

1.5 Analytic Diagonalization

In this section we show that for a real analytic path $p : I \to \mathfrak{p}$ there exists a real analytic path $K : I \to \mathbf{K}$ such that $\mathrm{Ad}_K^{-1}(p)$ lies in $\mathfrak{a}$. Clearly the path $\mathrm{Ad}_K^{-1}(p)$ is real analytic, and in fact it is the unique real analytic path in $\mathfrak{a}$ which is a lift of $\pi \circ p$, up to a global Weyl group action. This is the content of Theorem 1.5.7. This result stands in stark contrast to the previous section since even a C^∞ path p cannot guarantee the existence of a continuous diagonalizing K, see Example 1.2.6, or a C^2 choice of diagonal a, see Example 1.2.5.

Preliminaries

We will start with the well-known matrix case, and then lift the diagonalization to the symmetric Lie algebra via the adjoint representation. Consider a finite dimensional real inner product space V and its complexification $V^{\mathbb{C}}$. For an operator $A \in \mathfrak{gl}(V)$ we call $A_c \in \mathfrak{gl}(V^{\mathbb{C}}) \cong \mathfrak{gl}(V)^{\mathbb{C}}$ its complexification. To avoid confusion, we use $\widehat{\mathrm{ad}}$ to denote the adjoint maps on $\mathfrak{gl}(V)$ and $\mathfrak{gl}(V^{\mathbb{C}})$. We endow $\mathfrak{gl}(V)$ and $\mathfrak{gl}(V^{\mathbb{C}})$ with the Hilbert-Schmidt inner product $\langle A, B\rangle = \mathrm{tr}(A^*B)$.

Lemma 1.5.1. *Let V be a real inner product space and $A \in \mathfrak{gl}(V)$. If A is semisimple,[6] then so are A_c, and $\widehat{\mathrm{ad}}_A$, and $\widehat{\mathrm{ad}}_{A_c}$. Similarly, if A is normal, then so are A_c, and $\widehat{\mathrm{ad}}_A$, and $\widehat{\mathrm{ad}}_{A_c}$.*

Proof. Let A be semisimple. Recall that an operator is semisimple if and only if its minimal polynomial is square free. This shows that A_c is semisimple if and only if A is. Then it is easy to see that diagonalizability of A_c implies diagonalizability of $\widehat{\mathrm{ad}}_{A_c}$. But since $\widehat{\mathrm{ad}}_{A_c}$ is the complexification of $\widehat{\mathrm{ad}}_A$, the above shows that $\widehat{\mathrm{ad}}_A$ is semisimple.

If A is normal, then clearly A_c is too. An elementary computation shows that $\mathrm{tr}\left((\widehat{\mathrm{ad}}_A(B))^*C\right) = \mathrm{tr}\left(B^*\,\widehat{\mathrm{ad}}_{A^*}(C)\right)$, which implies that $(\widehat{\mathrm{ad}}_A)^* = \widehat{\mathrm{ad}}_{A^*}$ for $A \in \mathfrak{gl}(V)$ and thus

$$\left[(\widehat{\mathrm{ad}}_A)^*, \widehat{\mathrm{ad}}_A\right] = \widehat{\mathrm{ad}}_{[A^*,A]} = 0,$$

and so $\widehat{\mathrm{ad}}_A$ is also normal. The proof for A_c is identical. □

Lemma 1.5.2. *Let V be a finite dimensional real inner product space and let $V^{\mathbb{C}}$ be its complexification. Let $A \in \mathfrak{gl}(V)$ be semisimple and let A_c denote the complexification. Then it holds that*

$$\mathfrak{gl}(V) = \mathrm{im}\,\widehat{\mathrm{ad}}_A \oplus \ker \widehat{\mathrm{ad}}_A, \quad \mathfrak{gl}(V^{\mathbb{C}}) = \mathrm{im}\,\widehat{\mathrm{ad}}_{A_c} \oplus \ker \widehat{\mathrm{ad}}_{A_c},$$

and

$$\mathrm{im}\,\widehat{\mathrm{ad}}_A = \mathfrak{gl}(V) \cap \mathrm{im}\,\widehat{\mathrm{ad}}_{A_c}, \quad \ker \widehat{\mathrm{ad}}_A = \mathfrak{gl}(V) \cap \ker \widehat{\mathrm{ad}}_{A_c}.$$

Moreover, if A is normal, then the decompositions are orthogonal.

Proof. The decomposition into kernel and image holds for every semisimple operator. It is clear that if $X \in \mathfrak{gl}(V)$ then $[A_c, X_c] = 0$ if and only if $[A, X] = 0$. Let $Z \in \mathfrak{gl}(V^{\mathbb{C}})$ and assume that $\widehat{\mathrm{ad}}_{A_c}(Z) \in \mathfrak{gl}(V)$. Then $[A, Z+\overline{Z}]/2 = [A_c, Z]$. If A is normal, then clearly the decomposition $\mathfrak{gl}(V) = \mathrm{im}\,\widehat{\mathrm{ad}}_A \oplus \ker \widehat{\mathrm{ad}}_A$ is orthogonal. Using Lemma 1.5.1 the same is true for the complexification. □

[6] Recall that a linear operator is semisimple if each invariant subspace has an invariant complement. Over $\mathbb{C}$ this is equivalent to being diagonalizable.

Hence, for semisimple $A \in \mathfrak{gl}(V)$, we can define the projection $\widehat{\Pi}_A$ onto $\ker \widehat{\mathrm{ad}}_A$ and along $\operatorname{im} \widehat{\mathrm{ad}}_A$, and its complement $\widehat{\Pi}_A^- = \mathbb{1} - \widehat{\Pi}_A$. If A is normal, then the projection is orthogonal and we write $\widehat{\Pi}_A^\perp = \mathbb{1} - \widehat{\Pi}_A$. We use the same notation for the complexification A_c. Lemma 1.5.2 also shows that for A semisimple, the maps $\widehat{\mathrm{ad}}_A|_{\operatorname{im} \widehat{\mathrm{ad}}_A}$ and $\widehat{\mathrm{ad}}_{A_c}|_{\operatorname{im} \widehat{\mathrm{ad}}_{A_c}}$ are bijective. We will write the inverse maps as $\widehat{\mathrm{ad}}_A^{-1}$ and $\widehat{\mathrm{ad}}_{A_c}^{-1}$ and leave the restriction implicit.

More explicitly, for semisimple A_c with eigenvalues λ_i and eigenprojections P_i for $i = 1, \dots, n$, we can express the inverse of $\widehat{\mathrm{ad}}_{A_c}$ on the image of $\widehat{\Pi}_{A_c}^-$ by

$$\left(\widehat{\mathrm{ad}}_{A_c}^{-1} \circ \widehat{\Pi}_{A_c}^-\right)(B) = \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \frac{P_k B P_l}{\lambda_k - \lambda_l}. \tag{1.10}$$

Corollary 1.5.3. *For $A, B \in \mathfrak{gl}(V)$ and $C \in \operatorname{im} \widehat{\mathrm{ad}}_A$ it holds that $\Pi_{A_c}^\perp(B_c) = \Pi_A^\perp(B)$ and $\widehat{\mathrm{ad}}_{A_c}^{-1}(C_c) = \widehat{\mathrm{ad}}_A^{-1}(C)$.*

Proof. This follows from Lemma 1.5.2. □

Now we consider a real analytic path of operators $A : I \to \mathfrak{gl}(V)$. We can always find a simply connected open set $\mathbb{G} \subseteq \mathbb{C}$ containing I and such that there is an analytic continuation A_c of A on $\mathbb{G}$. Outside of a discrete set of exceptional points[7] in $\mathbb{G}$, the number of eigenvalues λ_i and the dimensions of the corresponding eigenprojectors P_i are constant. In fact, by [Kat80, Thm. 1.8], the eigenvalues λ_i, eigenprojectors P_i, and the eigennilpotents D_i are branches of analytic functions with only algebraic singularities at some of the exeptional points. If $A(t)$ is semisimple for all $t \in I$, then A_c is also semisimple on $\mathbb{G}$, since its eigen-nilpotents must vanish identically.

Lemma 1.5.4. *Let V be a finite dimensional complex Hilbert space and $A : I \to \mathfrak{gl}(V)$ be a real analytic curve of normal operators. Moreover, let $t_0 \in I$ be an exceptional point and let $A_c(z)$ denote the analytic extension of $A(t)$ to an open disk $\mathbb{D}_r$ of radius r about t_0 such that no other exceptional points are contained in $\mathbb{D}_r$. Then, on the punctured disc $\dot{\mathbb{D}}_r$ the following identity holds:*

$$-\left(\widehat{\mathrm{ad}}_{A_c(z)}^{-1} \circ \widehat{\Pi}_{A_c(z)}^-\right)(A_c'(z)) = \frac{1}{2} \sum_{k=1}^{n} [P_k'(z), P_k(z)]\,, \tag{1.11}$$

where $P_k(z)$ are the corresponding eigenprojections of $A_c(z)$. In particular, (1.11) shows that the expression can be continued analytically to $z = t_0$.

Proof. This follows from a long but straightforward computation involving resolvents given in Appendix 1.C. □

The motivation for (1.11) comes from Lemma 1.4.1 and Proposition 1.4.12. The right hand side is the formula derived in [Kat80, p. 105], whereas the left hand side is written in terms of Lie algebraic quantities. A priori it is not even clear that the left hand side should be continuous.

Corollary 1.5.5. *Let V be a real inner product space, and let I be an open interval containing 0. Let $A : I \to \mathfrak{gl}(V)$ be a real analytic curve of normal operators, and let $U : I \to \mathrm{GL}(V)$ be the solution of the ordinary differential equation*

$$U'(t) = -\left(\widehat{\mathrm{ad}}_{A(t)}^{-1} \circ \widehat{\Pi}_{A(t)}^\perp\right)(A'(t)) \cdot U(t), \quad U(0) = \mathbb{1} \in \mathrm{GL}(V). \tag{1.12}$$

[7] An exceptional point in I is a point at which two eigenvalues meet, without being permanently degenerate. Any compact set in $\mathbb{C}$ contains only finitely many exceptional points.

Then it holds for all $t, s \in I$ that

$$[\mathrm{Ad}^{-1}_{U(t)}(A(t)), \mathrm{Ad}^{-1}_{U(s)}(A(s))] = 0.$$

Proof. We start by complexifying. Let $\mathbb{G} \subseteq \mathbb{C}$ be a simply connected open set containing I such that there is an analytic continuation A_c of A on $\mathbb{G}$. Let $U_c : \mathbb{G} \to \mathrm{GL}(V^{\mathbb{C}})$ be the solution of the ordinary differential equation[8]

$$U_c'(z) = -\big(\widehat{\mathrm{ad}}^{-1}_{A_c(z)} \circ \widehat{\Pi}^-_{A_c(z)}\big)(A_c'(z)) \cdot U_c(z), \quad U_c(0) = \mathbb{1} \in \mathrm{GL}(V^{\mathbb{C}}).$$

By Lemma 1.5.4 the solution U_c satisfies

$$U_c'(z) = \frac{1}{2}\sum_{k=1}^{n} [P_k'(z), P_k(z)] \cdot U_c(z)$$

where the P_k are the eigenprojections of A_c. By [Kat80, Ch. II §4.5] this implies that $P_k(z) = \mathrm{Ad}_{U_c(z)}(P_k(0))$. Hence all $\mathrm{Ad}^{-1}_{U_c(z)}(A(z))$ for $z \in \mathbb{G}$ commute. Now we define by restriction $U := U_c|_I$. By Corollary 1.5.3 it is clear that U satisfies (1.12) and of course $[\mathrm{Ad}^{-1}_{U(t)}(A(t)), \mathrm{Ad}^{-1}_{U(s)}(A(s))] = 0$ still holds. □

Real Analytic Diagonalization

The main idea is to go from a semisimple, orthogonal, symmetric Lie algebra to concrete matrices via the adjoint representation and to then use the previous result. The following elementary properties of Lie algebra homomorphisms will be useful for this transition step. We use an arbitrary homomorphism ϕ here instead of ad to avoid confusion with the other uses of ad.

Lemma 1.5.6. *Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism and let ad and $\widehat{\mathrm{ad}}$ denote the respective adjoint maps. Then for $x, y \in \mathfrak{g}$ it holds that*

$$\phi(\mathrm{ad}_x(y)) = \widehat{\mathrm{ad}}_{\phi(x)}(\phi(y)), \tag{1.13}$$

and hence ϕ maps $\mathrm{im}\,\mathrm{ad}_x$ to $\mathrm{im}\,\widehat{\mathrm{ad}}_{\phi(x)}$ and $\ker \mathrm{ad}_x$ to $\ker \widehat{\mathrm{ad}}_{\phi(x)}$. Now assume that x and $\phi(x)$ are semisimple[9]. Let $\Pi^-_x : \mathfrak{g} \to \mathfrak{g}$ denote the projection onto $\mathrm{im}\,\mathrm{ad}_x$ along $\ker \mathrm{ad}_x$ and analogously for $\widehat{\Pi}^-_{\phi(x)} : \mathfrak{h} \to \mathfrak{h}$. Then

$$\phi(\Pi^-_x(y)) = \widehat{\Pi}^-_{\phi(x)}(\phi(y)). \tag{1.14}$$

Proof. Eq. (1.13) holds by definition, and it immediately implies that ϕ maps $\mathrm{im}\,\mathrm{ad}_x$ to $\mathrm{im}\,\widehat{\mathrm{ad}}_{\phi(x)}$ and $\ker \mathrm{ad}_x$ to $\ker \widehat{\mathrm{ad}}_{\phi(x)}$. Now assume that x and $\phi(x)$ are semisimple. If $y \in \mathrm{im}\,\mathrm{ad}_x$ and $z \in \ker \mathrm{ad}_x$ we see that $\phi(\Pi^-_x(y+z)) = \phi(y)$ and $\widehat{\Pi}^-_{\phi(x)}(\phi(y+z)) = \phi(y)$ by the previous observation. □

Theorem 1.5.7 (Real Analytic Diagonalization). *Let I be an open interval containing 0 and let $p : I \to \mathfrak{p}$ be a real analytic path. Then there exists a real analytic path $K : I \to \mathbf{K}$ such that $\mathrm{a}(t) := \mathrm{Ad}^{-1}_{K(t)}(p(t)) \in \mathfrak{a}$ for all $t \in I$. Moreover, such K can be obtained as the solution to*

$$K'(t) = k(t)K(t), \quad k(t) = -\mathrm{ad}^{-1}_{p(t)}(\Pi^{\perp}_{p(t)}(p'(t))), \quad \mathrm{Ad}^{-1}_{K(0)}(p(0)) \in \mathfrak{a}.$$

[8] The solution exists and is unique and holomorphic. See for instance [Kat80, p. 100].

[9] An element of a Lie algebra is semisimple if its ad-representation is a semisimple operator.

Furthermore, any real analytic path $b : I \to \mathfrak{a}$ satisfying $\pi \circ b = \pi \circ \rho$ satisfies $b = w \cdot \lambda$ for some fixed Weyl group element $w \in \mathbf{W}$.

Proof. Let $\widehat{p}(t) = \mathrm{ad}_{p(t)}$. This is a real analytic path in $\mathfrak{gl}(\mathfrak{g})$. By the proof of Lemma 1.A.29 $\widehat{p}$ is normal on I. By Corollary 1.5.5 the function $\widehat{K} : I \to \mathrm{GL}(\mathfrak{g})$ solving the differential equation

$$\widehat{K}'(t) = -\big(\widehat{\mathrm{ad}}^{-1}_{\widehat{p}(t)} \circ \widehat{\Pi}^{\perp}_{\widehat{p}(t)}\big)(\widehat{p}'(t)) \cdot \widehat{K}(t), \quad \widehat{K}(0) = \mathbb{1} \in \mathrm{GL}(\mathfrak{g})$$

diagonalizes $\widehat{p}$ in the sense that

$$[\mathrm{Ad}^{-1}_{\widehat{K}(t)}(\widehat{p}(t)), \mathrm{Ad}^{-1}_{\widehat{K}(s)}(\widehat{p}(s))] = 0 \tag{1.15}$$

for all $t, s \in I$. Now we need to translate this back to $\mathfrak{g}$. First we define the function $k : I \to \mathfrak{k}$ by

$$k(t) = -\mathrm{ad}^{-1}_{p(t)}(\Pi^{\perp}_{p(t)}(p'(t))), \tag{1.16}$$

and we claim that

$$\mathrm{ad}_{k(t)} = -\widehat{\mathrm{ad}}^{-1}_{\widehat{p}(t)}(\widehat{\Pi}^{\perp}_{\widehat{p}(t)}(\widehat{p}'(t))). \tag{1.17}$$

Indeed, this follows from Lemma 1.5.6 with $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ as ϕ. To apply the lemma we need to verify that both $p(t)$ and $\widehat{p}(t)$ are semisimple as Lie algebra elements, which follows in both cases from the above observation that $\widehat{p}(t)$ is a normal operator. So we may compute:

$$\widehat{\mathrm{ad}}_{\widehat{p}(t)}(\mathrm{ad}_{k(t)}) \overset{(1.13)}{=} \mathrm{ad}_{\mathrm{ad}_{p(t)}(k(t))} \overset{(1.16)}{=} -\mathrm{ad}_{\Pi^{\perp}_{p(t)}(p'(t))} \overset{(1.14)}{=} -\widehat{\Pi}^{\perp}_{\widehat{p}(t)}(\widehat{p}'(t)).$$

Now we define $K : I \to \mathbf{K}$ by

$$K'(t) = k(t)K(t), \quad K(0) = \mathbb{1},$$

and we claim that

$$\widehat{K}(t) = \mathrm{Ad}_{K(t)}.$$

Indeed, both $\widehat{K}(t)$ and $\mathrm{Ad}_{K(t)}$ lie in $\mathrm{GL}(\mathfrak{g})$, they satisfy $\widehat{K}(0) = \mathrm{Ad}_{K(0)}$, and since

$$(\mathrm{Ad}_{K(t)})' = \mathrm{ad}_{k(t)} \mathrm{Ad}_{K(t)},$$

they satisfy the same differential equation by (1.17).[10] This implies that $\mathrm{Ad}^{-1}_{\widehat{K}}(\mathrm{ad}_p) = \mathrm{ad}_{\mathrm{Ad}^{-1}_K(p)}$, and hence

$$0 \overset{(1.15)}{=} [\mathrm{Ad}^{-1}_{\widehat{K}(t)}(\widehat{p}(t)), \mathrm{Ad}^{-1}_{\widehat{K}(s)}(\widehat{p}(s))] = [\mathrm{ad}_{\mathrm{Ad}^{-1}_{K(t)}(p(t))}, \mathrm{ad}_{\mathrm{Ad}^{-1}_{K(s)}(p(s))}] = \mathrm{ad}_{[\mathrm{Ad}^{-1}_{K(t)}(p(t)), \mathrm{Ad}^{-1}_{K(s)}(p(s))]},$$

and by semisimplicity of $\mathfrak{g}$ this means that $[\mathrm{Ad}^{-1}_{K(t)}(p(t)), \mathrm{Ad}^{-1}_{K(s)}(p(s))] = 0$. Since $\mathrm{Ad} : \mathbf{G} \to \mathrm{GL}(\mathfrak{g})$ has discrete kernel, K is real analytic as continuous lift of $\widehat{K}$. Since all $\mathrm{Ad}^{-1}_{K(t)}(p(t))$ commute, by Lemma 1.A.26 there exists $L \in \mathbf{K}$ such that $\mathrm{Ad}^{-1}_{K(t)L}(p(t)) \in \mathfrak{a}$ for all $t \in I$. Then $\tilde{K}(t) = K(t)L$ also satisfies $\tilde{K}'(t) = k(t)\tilde{K}(t)$. Finally, uniqueness of the diagonalized path up to global Weyl group action follows from Lemma 1.B.11. □

Remark 1.5.8. *This theorem generalizes the well-known analytic diagonalization of Hermitian matrices, see [Kat80, Thm. 6.1], and also [Bun+91, Thm. 1] which is the special case for the real singular value decomposition.*

[10] Note that in the case where $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$, it even holds that $K = \widehat{K}$.

1.6 Measurable Diagonalization

The main result of this section is Theorem 1.6.6, which shows that any measurable[11] function $p : \Omega \to \mathfrak{p}$, where Ω is any measurable space, can be diagonalized as $p(\omega) = \mathrm{Ad}_{K(\omega)} \circ a(\omega)$ with measurable functions $K : \Omega \to \mathbf{K}$ and $a : \Omega \to \mathfrak{a}$. This generalizes the analogous result [QR14, Thm. 2.1] for the unitary diagonalization of positive definite matrices. We then further generalize this result in Theorem 1.6.12 to show that a finite family of commuting measurable functions can be simultaneously measurably diagonalized.

In order to show that a single measurable function $p : \Omega \to \mathfrak{p}$ can be measurably diagonalized, we will describe a stratification of $\mathfrak{p}$ into embedded submanifolds with a simple structure. This stratification originates from an intuitive partition of the Weyl chamber $\mathfrak{w}$.

Remark 1.6.1. *Let $\mathfrak{w} \subseteq \mathfrak{a}$ be a closed Weyl chamber. Then $\mathfrak{w}$ is a polyhedral cone, that is, it is defined by a finite set of linear homogeneous inequalities on $\mathfrak{a}$. Indeed one can choose these inequalities such that each corresponds to a reflection in $\mathbf{W}$ whose hyperplane defines a facet of $\mathfrak{w}$. Let $\mathfrak{w}_s$, with $s \in S$ some index set, denote the finitely many* open faces *of $\mathfrak{w}$, that is, the relative interiors[12] of the closed faces. Then the $\mathfrak{w}_s$ form a partition of $\mathfrak{w}$.*

Lemma 1.6.2. *Let $\mathfrak{w} \subseteq \mathfrak{a}$ be a closed Weyl chamber and $x, y \in \mathfrak{w}$. Then x and y belong the the same open face of $\mathfrak{w}$ if and only if $\mathbf{K}_x = \mathbf{K}_y$, or equivalently, $\mathbf{W}_x = \mathbf{W}_y$.*

Proof. First we argue that if $x, y \in \mathfrak{w}$ have the same stabilizer in $\mathbf{W}$, then they belong to the same open face of $\mathfrak{w}$. Indeed, if they have the same stabilizer, then by Remark 1.6.1, they satisfy the same equalities in the inequality description of $\mathfrak{w}$ and hence they belong to the same open face. Next we show that if $x, y \in \mathfrak{w}_s$ for some open face $\mathfrak{w}_s$, then they have the same stabilizer in $\mathbf{K}$. By Kleiner's Lemma, see [AB15, Lemma 3.70], if $p : [0, 1] \to \mathfrak{p}$ is a geodesic segment realizing the distance between the orbits $\mathrm{Ad}_{\mathbf{K}}(p(0))$ and $\mathrm{Ad}_{\mathbf{K}}(p(1))$, then all points $p(t)$ for $t \in (0, 1)$ have the same stabilizer in $\mathbf{K}$. By Corollary 1.A.56, every line segment in $\mathfrak{w}$ is of this type. Since the $\mathfrak{w}_s$ are convex and relatively open, this shows that all points belonging to the same $\mathfrak{w}_s$ have the same stabilizer in $\mathbf{K}$. Finally it is clear that if x, y have the same stabilizers in $\mathbf{K}$, then the same is true in $\mathbf{W}$. This concludes the proof. □

Corollary 1.6.3. *Let $\mathfrak{w} \subseteq \mathfrak{a}$ be a closed Weyl chamber and let $\mathfrak{w}_s$ be an open face of $\mathfrak{w}$. If $x, y \in \mathfrak{w}_s$, then $\mathfrak{p}_x = \mathfrak{p}_y$.*

Proof. For $z \in \mathfrak{a}$, Corollary 1.A.47 with $A = \{z\}$ shows that $\mathfrak{p}_z = \mathrm{Ad}_{\mathbf{K}_z}(\mathfrak{a})$. For $x, y \in \mathfrak{w}_s$, by Lemma 1.6.2 it holds that $\mathbf{K}_x = \mathbf{K}_y$ and hence $\mathfrak{p}_x = \mathfrak{p}_y$. □

Lemma 1.6.4. *Let $\mathfrak{w} \subseteq \mathfrak{a}$ be a closed Weyl chamber and denote $\mathfrak{p}_s = \mathrm{Ad}_{\mathbf{K}}(\mathfrak{w}_s)$ and let $\mathbf{K}_s$ denote the stabilizer in K of the points in $\mathfrak{w}_s$. Then the map $\mathbf{K}/\mathbf{K}_s \times \mathfrak{w}_s \to \mathfrak{p}_s$, $(K\mathbf{K}_s, x) \mapsto \mathrm{Ad}_K(x)$ is a $\mathbf{K}$-equivariant diffeomorphism and $\mathfrak{p}_s$ is an embedded submanifold.*

Proof. First note that the $\mathbf{K}_s$ are well-defined due to Lemma 1.6.2. Clearly the map $\phi : \mathbf{K}/\mathbf{K}_s \times \mathfrak{w}_s \to \mathfrak{p}_s$, $(K\mathbf{K}_s, x) \mapsto \mathrm{Ad}_K(x)$ is well-defined, smooth and $\mathbf{K}$-equivariant and surjective. To see that it is injective, consider two points $(K_1\mathbf{K}_s, x)$ and $(K_2\mathbf{K}_s, y)$ mapped to the same point. Then $x = y$ since each orbit intersects $\mathfrak{w}$ in exactly one point by Corollary 1.A.39, and $K_1^{-1}K_2 \in \mathbf{K}_x = \mathbf{K}_s$. Hence $K_1\mathbf{K}_s = K_2\mathbf{K}_s$. We show that the differential $D\phi(K\mathbf{K}_s, x)$ is injective. By equivariance, it suffices to

[11] In this section, all topological spaces will be endowed with their Borel σ-algebra, that is, the smallest σ-algebra containing all open sets, except for Prop. 1.6.13 where we use the Lebesgue measure.

[12] The relative interior of a subset S of a vector space V is the topological interior of S seen as a subset of the affine hull of S, that is, the smallest affine subspace containing S, see [Zăl02, pp. 2–3].

consider $K = \mathbb{1}$. Then for $v \in T_x\mathfrak{w}_s$ we obtain $D\phi(\mathbb{1}\mathbf{K}_s, x)(0, v) = v \in \mathfrak{a}$ and for $w \in \mathfrak{k}/\mathfrak{k}_s \cong \mathfrak{k}_x^\perp$ we obtain $D\phi(\mathbb{1}\mathbf{K}_s, x)(w, 0) = [w, x] \in \mathfrak{a}^\perp$ by Lemma 1.A.16. Hence $D\phi(\mathbb{1}\mathbf{K}_s, x)(w, v)$ only if $v = 0$ and $w = 0$. Hence ϕ is an immersion. To see that $\mathfrak{p}_s$ is embedded, let $x \in \mathfrak{w}_s$. Consider a sequence of points $x_i \in \mathfrak{w}_s$ and $K_i \in \mathbf{K}$ such that $\mathrm{Ad}_{K_i}(x_i) \to x$. Then, since the quotient map $\pi_\mathfrak{a}$ is open, $x_i \to x$ and since the action is proper, a subsequence of K_i converges to some $K \in \mathbf{K}_s$. This shows that ϕ is an embedding. □

Lemma 1.6.5. *Let $\mathbf{G}$ be a Lie group, $\mathbf{H}$ a closed subgroup of $\mathbf{G}$, and Ω a measurable space. If $\gamma : \Omega \to \mathbf{G}/\mathbf{H}$ is measurable, then there exists a measurable lift $\tilde{\gamma} : \Omega \to \mathbf{G}$.*

Proof. Since the quotient map $\pi : \mathbf{G} \to \mathbf{G}/\mathbf{H}$, $g \mapsto g\mathbf{H}$ is a smooth submersion (cf. [Lee13, Thm. 21.17]), there is an open neighborhood U of e in $\mathbf{G}$ and there are charts $\sigma : \mathbb{R}^n \to U$ and $\tau : \mathbb{R}^{n-k} \to \pi(U) \subseteq \mathbf{G}/\mathbf{H}$ such that $\tau^{-1} \circ \pi \circ \sigma : \mathbb{R}^n \to \mathbb{R}^{n-k}$ is simply the projection onto the first $n - k$ coordinates. The sets of the form $\pi(gU)$ form an open cover of $\mathbf{G}/\mathbf{H}$ and hence there exists a countable subcover whose open sets are $W_i := \pi(g_i U)$ with $g_i \in \mathbf{G}$ for $i \in \mathbb{N}$. Then define the sets $A_1 = W_1$ and $A_i = W_i \setminus \bigcup_{m=1}^{i-1} A_m$ for $i \geq 2$, which form a countable partition of $\mathbf{G}/\mathbf{H}$ consisting of measurable sets. Let $\Omega_i = \{\gamma \in A_i\}$ be the preimages, which form a countable measurable partition of Ω. Then it suffices to find measurable lifts $\tilde{\gamma}_i : \Omega_i \to \mathbf{G}/\mathbf{H}$ of each restriction $\gamma_i := \gamma|_{\Omega_i}$. By definition, $g_i^{-1}\gamma_i$ takes image in $g_i^{-1}A_i \subseteq \pi(U)$. Using the chart τ this path can be seen as a measurable path in $\mathbb{R}^{n-k}$, which can be lifted to $\mathbb{R}^n$ using the inclusion $\iota : \mathbb{R}^{n-k} \to \mathbb{R}^n : x \to (x, 0, \ldots, 0)$. That is, we define $\tilde{\gamma}_i = g_i \circ \sigma \circ \iota \circ \tau^{-1} \circ g_i^{-1} \circ \gamma_i$ and this concludes the proof. □

Putting everything together we can now prove the first main result of this section.

Theorem 1.6.6 (Measurable Diagonalization). *Let Ω be a measurable space and let $p : \Omega \to \mathfrak{p}$ be measurable. Then there exist measurable functions $K : \Omega \to \mathbf{K}$ and $a : \Omega \to \mathfrak{a}$ such that $p(\omega) = \mathrm{Ad}_{K(\omega)}(a(\omega))$ for all $\omega \in \Omega$.*

Proof. Let the $\mathfrak{w}_s$ and $\mathfrak{p}_s$ be as in Lemma 1.6.4. Then the $\mathfrak{p}_s$ yield a finite partition of $\mathfrak{p}$ into measurable subsets, and the sets $\Omega_s := \{p \in \mathfrak{p}_s\} \subseteq \Omega$ yield a finite partition of Ω into measurable subsets, and it suffices to find measurable functions $K_s : \Omega_s \to \mathbf{K}$ and $a_s : \Omega_s \to \mathfrak{a}$ satisfying $\mathrm{Ad}_{K_s(\omega)}(a_s(\omega)) = p(\omega)$ for all $\omega \in \Omega_s$. Let $p_s = p|_{\Omega_s}$, then by Lemma 1.6.4 one can consider p_s as a measurable map $\Omega_s \to \mathbf{K}/\mathbf{K}_s \times \mathfrak{w}_s$. Then we define the measurable maps $a_s : \Omega_s \to \mathfrak{a}$ by $a_s = \mathrm{pr}_2 \circ p_s$ and $\tilde{K}_s : \Omega_s \to \mathbf{K}/\mathbf{K}_s$ by $\tilde{K}_s = \mathrm{pr}_1 \circ p_s$, and using Lemma 1.6.5 we obtain a corresponding measurable map $K_s : \Omega_s \to \mathbf{K}$. □

Remark 1.6.7. *This generalizes [QR14, Thm. 2.1] which shows that a measurable function of positive definite matrices can be unitarily diagonalized in a measurable way.*

We can further strengthen this result by showing that finitely many commuting measurable functions $p_i : \Omega \to \mathfrak{p}$ for $i = 1, \ldots, n$ can be simultaneously measurably diagonalized. The proof will be based on induction on i. The idea will be to diagonalize p_{i+1} using group elements which stabilize all the previously diagonalized paths. To do this we need to work with symmetric Lie subalgebras of $\mathfrak{g}$, which will in general not be semisimple, but still reductive.

First we will need two simple lemmas about Lie subgroups and restrictions of Lie group homomorphisms:

Lemma 1.6.8. *Let $\mathbf{G}$ be a Lie group with Lie subgroups $\mathbf{H}$ and $\mathbf{K}$ satisfying the inclusion $\mathbf{K} \subseteq \mathbf{H}$. Then $\mathbf{K}$ is a Lie subgroup of $\mathbf{H}$.*

Proof. The identity on $\mathbf{G}$ descends to the inclusion $\mathbf{K} \hookrightarrow \mathbf{H}$, and by Lemma 1.A.21 the latter is smooth, so its image is a Lie subgroup in $\mathbf{H}$ by [Lee13, Thm. 7.17]. □

Lemma 1.6.9. *Let X be a real, finite dimensional vector space and let $Y \subseteq X$ be a subspace. Let $\mathbf{G} \subseteq \mathrm{GL}(X)$ be a Lie subgroup containing only elements which leave Y invariant. Let $\mathbf{H} \subseteq \mathrm{GL}(Y)$ be any Lie subgroup such that for $g \in \mathbf{G}$ the restriction $g|_Y$ lies in $\mathbf{H}$. Then the restriction $\mathbf{G} \to \mathbf{H}$ which maps $g \to g|_Y$ is a Lie group homomorphism.*

Proof. Let $\mathrm{GL}(X, Y) \subseteq \mathrm{GL}(X)$ be the subgroup of elements which leave Y invariant. If P_Y is any idempotent linear map on X with image Y, then $\mathrm{GL}(X, Y) = \{g \in \mathrm{GL}(X) : gP_Y = P_Y g P_Y\}$; hence it is a closed subgroup and thus an embedded Lie subgroup, see [Lee13, Thm. 7.21]. Clearly $\mathbf{G}$ is a subgroup of $\mathrm{GL}(X, Y)$, and by Lemma 1.6.8 and the above, it is a Lie subgroup. Let $r : \mathrm{GL}(X, Y) \to \mathrm{GL}(Y)$ be the restriction map $g \mapsto g|_Y$, which is a Lie group homomorphism. By Lemma 1.A.21 it descends to a Lie group homomorphism $\mathbf{G} \to \mathbf{H}$, which concludes the proof. □

Now we can give the promised induction argument which will be the key ingredient for the following theorem.

Lemma 1.6.10. *Let Ω be a measurable space and let $A \subseteq \mathfrak{p}$ be any subset. If $p : \Omega \to \mathfrak{p}_A$ is measurable, then there exists a measurable function $K : \Omega \to \mathbf{K}_A$ such that $\mathrm{Ad}^{-1}_{K(\omega)}(p(\omega)) \in \mathfrak{a}$ for all $\omega \in \Omega$.*

Proof. By Lemma 1.A.43 and Lemma 1.A.44 the commutant $\mathfrak{g}_A$ is reductive and can be written as $\mathfrak{g}_A = \mathfrak{h} \oplus \mathfrak{z}$ where $\mathfrak{h} = [\mathfrak{g}_A, \mathfrak{g}_A]$ is the semisimple part and $\mathfrak{z}$ is the center of $\mathfrak{g}_A$. Moreover $(\mathfrak{h}, s|_{\mathfrak{h}})$ is a semisimple, symmetric Lie subalgebra of $\mathfrak{g}$, and by Lemma 1.A.7 it is orthogonal. Its Cartan-like decomposition is $\mathfrak{h} = \mathfrak{l} \oplus \mathfrak{q}$, where $\mathfrak{l} = \mathfrak{h} \cap \mathfrak{k}$ and $\mathfrak{q} = \mathfrak{h} \cap \mathfrak{p}$. It holds that $\mathfrak{p}_A = \mathfrak{q} \oplus (\mathfrak{z} \cap \mathfrak{p})$ since $\mathfrak{h}$ and $\mathfrak{z}$ are invariant under s. Let $\tilde{p} : \Omega \to \mathfrak{q}$ be the component of p in $\mathfrak{q}$. Let $\mathfrak{b}$ denote a maximal Abelian subspace of $\mathfrak{q}$ and note that without loss of generality $\mathfrak{b} \subseteq \mathfrak{a}$. The first step is to diagonalize $\tilde{p}$ using the previous theorem. Let $(\mathbf{H} = \mathrm{Int}(\mathfrak{h}), \mathbf{L} = \mathrm{Int}_{\mathfrak{l}}(\mathfrak{h}))$ be the canonical pair associated with $(\mathfrak{h}, s|_{\mathfrak{h}})$ as in Lemma 1.A.20. By Theorem 1.6.6 there exists a measurable path $\tilde{K} : \Omega \to L$ such that $\mathrm{Ad}^{-1}_{\tilde{K}(\omega)}(\tilde{p}(\omega)) \in \mathfrak{b}$ for all $\omega \in \Omega$. The next step is to lift the path $\tilde{K}$ to K_A. Since $\mathfrak{l} \subseteq \mathfrak{k}_A$, Lemma 1.6.8 shows that $\mathbf{L} = \mathrm{Int}_{\mathfrak{l}}(\mathfrak{h})$ is a Lie subgroup of $\mathrm{Int}_{\mathfrak{k}_A}(\mathfrak{h})$, and so we can consider the path $\tilde{K}$ to take values in $\mathrm{Int}_{\mathfrak{k}_A}(\mathfrak{h})$. Consider the adjoint representation of $(\mathbf{K}_A)_0$ on $\mathfrak{h}$, denoted by $\mathrm{Ad}\,|_{\mathfrak{h}} : (\mathbf{K}_A)_0 \to \mathrm{Int}_{\mathfrak{k}_A}(\mathfrak{h})$, which is surjective. We need to show that this is a Lie group homomorphism. Indeed, we can write this as a composition $(\mathbf{K}_A)_0 \to \mathrm{Int}_{\mathfrak{k}_A}(\mathfrak{g}_A) \to \mathrm{Int}_{\mathfrak{k}_A}(\mathfrak{h})$. By Lemma 1.A.22 the first map is a Lie group homomorphism. Since the adjoint representation of $\mathfrak{k}_A$ preserves $\mathfrak{h}$ and by Lemma 1.6.9 the second one is also a Lie group homomorphism. Hence by the Lie group Isomorphism Theorem [Lee13, Theorem 21.27], we can consider the path $\tilde{K}$ to take values in $(\mathbf{K}_A)_0 / \ker(\mathrm{Ad}\,|_{\mathfrak{h}})$, and by Lemma 1.6.5 we obtain a measurable path $K : \Omega \to (\mathbf{K}_A)_0$ satisfying $\mathrm{Ad}^{-1}_{K(\omega)}(\tilde{p}(\omega)) \in \mathfrak{b}$ for all $\omega \in \Omega$. Finally we show that K is the desired path. But this follows from the fact that $\mathbf{K}_A$ leaves $\mathfrak{h}$ and $\mathfrak{z}$ invariant, and hence also $\mathfrak{q}$ and $\mathfrak{z} \cap \mathfrak{p}$, and from the fact that $\mathfrak{z} \cap \mathfrak{p} \subseteq \mathfrak{a}$. □

Now let us describe a partition of $\mathfrak{a}$ which extends the decomposition of $\mathfrak{w}$ of Remark 1.6.1. In fact we may simply take all relatively open faces of all Weyl chambers, removing duplicates of course. This yields a partition of $\mathfrak{a}$ into finitely many subsets $\mathfrak{a}_r$ with $r \in R$ for some index set R. We generalize this partition of $\mathfrak{a}$ to the n-fold Cartesian product $\mathfrak{a}^n$. Consider some tuple $\mathbf{r} = (r_i)_{i=1}^n \in R^n$ of indices in R. We write $\mathfrak{a}_{\mathbf{r}} = \mathfrak{a}_{r_1} \times \ldots \times \mathfrak{a}_{r_n}$ and note that there are finitely many such sets and they are disjoint and cover $\mathfrak{a}^n$.

Corollary 1.6.11. *Let $\mathbf{r} \in R^n$ and $x, y \in \mathfrak{a}_{\mathbf{r}}$. Then x and y have the same stabilizer in $\mathbf{K}$ and the same commutant in $\mathfrak{p}$, i.e., $\mathbf{K}_{\{x_1,\ldots,x_n\}} = \mathbf{K}_{\{y_1,\ldots,y_n\}}$ and $\mathfrak{p}_{\{x_1,\ldots,x_n\}} = \mathfrak{p}_{\{y_1,\ldots,y_n\}}$.*

Proof. This follows from Lemma 1.6.2 and Lemma 1.6.3. □

The corollary shows that we may define the simultaneous stabilizer $\mathbf{K}_{\mathbf{r}} := \mathbf{K}_{\{x_1,\dots,x_n\}}$ and the simultaneous commutant $\mathfrak{p}_{\mathbf{r}} := \mathfrak{p}_{\{x_1,\dots,x_n\}}$. With this we can prove the second main result of this section.

Theorem 1.6.12 (Simultaneous Measurable Diagonalization). *Let Ω be a measurable space and let $p_i : \Omega \to \mathfrak{p}$ be measurable for $i = 1,\dots,n$. Assume that $[p_i(\omega), p_j(\omega)] = 0$ for all $\omega \in \Omega$ and for all $i, j \in \{1,\dots,n\}$. Then there exists a measurable function $K : \Omega \to \mathbf{K}$ such that $\mathrm{Ad}^{-1}_{K(\omega)}(p_i(\omega)) \in \mathfrak{a}$ for all $i = 1,\dots,n$ and for all $\omega \in \Omega$.*

Proof. We proceed by induction on i by showing that if there exists a measurable $K : \Omega \to \mathbf{K}$ such that $\mathrm{Ad}^{-1}_{K(\omega)}(p_j(\omega)) \in \mathfrak{a}$ for all $j \leq i$ and for all $\omega \in \Omega$, then there exists a measurable $\tilde{K} : \Omega \to \mathbf{K}$ such that $\mathrm{Ad}^{-1}_{\tilde{K}(\omega)}(p_j(\omega)) \in \mathfrak{a}$ for all $j \leq i+1$ and for all $\omega \in \Omega$. The base case $i = 1$ is exactly Theorem 1.6.6. Assume now that $1 \leq i < n$ and let K be such that $\mathrm{Ad}^{-1}_{K} \circ p_j \in \mathfrak{a}$ for all $\omega \in \Omega$ and $j \leq i$. Now consider any subset $\mathfrak{a}_{\mathbf{r}}$ with $\mathbf{r} \in S^i$ of the partition of $\mathfrak{a}^i$ defined above. Then the set $\Omega_{\mathbf{r}} = \{\mathrm{Ad}^{-1}_{K}(p_1,\dots,p_i) \in \mathfrak{a}_{\mathbf{r}}\}$ is measurable and it suffices to show that we can diagonalize $p_{i+1}|_{\Omega_{\mathbf{r}}}$. By Corollary 1.6.11, for all $\omega \in \Omega_{\mathbf{r}}$, the set $\{\mathrm{Ad}^{-1}_{K(\omega)}(p_j(\omega)) : j = 1,\dots,i\}$ will have the same stabilizer $\mathbf{K}_{\mathbf{r}}$ in $\mathbf{K}$ and the same commutant $\mathfrak{p}_{\mathbf{r}}$ in $\mathfrak{p}$. Hence for all $\omega \in \Omega_{\mathbf{r}}$ it holds that $\mathrm{Ad}^{-1}_{K(\omega)}(p_{i+1}(\omega)) \in \mathfrak{p}_{\mathbf{r}}$ and by Lemma 1.6.10 there exists a measurable path $\tilde{K} : \Omega_{\mathbf{r}} \to \mathbf{K}_{\mathbf{r}}$ which diagonalizes $\mathrm{Ad}^{-1}_{K(\omega)}(p_{i+1}(\omega))$ on $\Omega_{\mathbf{r}}$. This proves the induction step and concludes the proof. □

For our final result we specialize to the case where our measurable space is an interval I with the Lebesgue measure and where $p : I \to \mathfrak{p}$ is absolutely continuous.

Proposition 1.6.13. *Let $I \subseteq \mathbb{R}$ be an interval. Let $p : I \to \mathfrak{p}$ be absolutely continuous. Then there exists $K : I \to \mathbf{K}$ measurable such that $\mathrm{a}(t) = \mathrm{Ad}^{-1}_{K(t)}(p(t))$ and $b(t) = \mathrm{Ad}^{-1}_{K(t)}(\Pi_{p(t)}(p'(t)))$ lie in $\mathfrak{a}$ and $b(t) = \mathrm{a}'(t)$ for almost every $t \in I$. In fact we can ensure that $\mathrm{a} = \mathrm{a}^{\downarrow}$.*

Proof. First we show that $\Pi_{p(t)}(p'(t))$ is measurable. By Theorem 1.6.6 there is a measurable $K : I \to \mathbf{K}$ such that $\mathrm{a}(t) := \mathrm{Ad}^{-1}_{K(t)}(p(t)) \in \mathfrak{a}$. By Lemma 1.A.24 (iii) it suffices to show that $\Pi_{\mathrm{a}(t)}$ is measurable, and in fact we may show this on each $I_s = \{\mathrm{a} \in \mathfrak{w}_s\}$ with the partition from Remark 1.6.1. Indeed, on these sets, $\Pi_{\mathrm{a}(t)}$ is a constant linear projection, so it is clearly measurable. Since p and $\Pi_{p(t)}(p'(t))$ are measurable and commute almost everywhere by construction, there is, by Theorem 1.6.12, some measurable $\tilde{K} : I \to \mathbf{K}$ such that $\tilde{\mathrm{a}}(t) = \mathrm{Ad}^{-1}_{\tilde{K}(t)}(p(t))$ and $\tilde{b} = \mathrm{Ad}^{-1}_{\tilde{K}(t)}(\Pi_{p(t)}(p'(t)))$ are measurable and lie in $\mathfrak{a}$ almost everywhere. Now consider the path $a^{\downarrow}$ as defined in Proposition 1.3.1, which is absolutely continuous by item (v) of the same proposition. Then the path $(a^{\downarrow}, (a^{\downarrow})')$ in $T\mathfrak{a}$ is measurable almost everywhere. By Proposition 1.4.5 there exists for almost every $t_0 \in I$ some $\tilde{w} \in \mathbf{W}$ such that $(a^{\downarrow}(t_0), (a^{\downarrow})'(t_0)) = \tilde{w} \cdot (\tilde{\mathrm{a}}(t_0), \tilde{b}(t_0))$. By Lemma 1.B.10 there is some measurable $w : I \to \mathbf{W}$ such that $(a^{\downarrow}, (a^{\downarrow})') = w \cdot (\tilde{\mathrm{a}}, \tilde{b})$ almost everywhere. Let $L : I \to \mathbf{K}$ be a measurable lift of w, and define $K : I \to \mathbf{K}$ as $\tilde{K}L^{-1}$. Then K satisfies the desired properties and this concludes the proof. □

1.7 Classification of Diagonalizations

As illustrated in Section 1.2, the semisimple, orthogonal, symmetric Lie algebras correspond to various notions of diagonalization. In this section we first recall the classification of irreducible orthogonal symmetric Lie algebras and prove that every semisimple, orthogonal, symmetric Lie algebra is orbit

equivalent to a direct sum of irreducible orthogonal symmetric Lie algebras (and a trivial part). This is the content of Theorem 1.7.10. Then, in Table 1.3, we give a list of diagonalizations corresponding to the irreducible orthogonal symmetric Lie algebras, as described in [Kle06].

An orthogonal symmetric Lie algebra $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is called *irreducible* if it is semisimple, strongly effective and irreducible, as defined in Appendix 1.A. Then [Hel78, Ch. VIII, Thms. 5.3, 5.4] shows that there are exactly four types of irreducible orthogonal symmetric Lie algebras $(\mathfrak{g}, s)$:

I $\mathfrak{g}$ is a compact, simple Lie algebra over $\mathbb{R}$ and s is any involutive automorphism of $\mathfrak{g}$;

II $\mathfrak{g}$ is a compact Lie algebra, and it is the Lie algebra direct sum $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ of simple ideals, where s interchanges $\mathfrak{g}_1$ and $\mathfrak{g}_2$;

III $\mathfrak{g}$ is a non-compact, simple Lie algebra over $\mathbb{R}$, its complexification $\mathfrak{g}_{\mathbb{C}}$ is a simple Lie algebra over $\mathbb{C}$, and $\mathfrak{k}$ is compactly embedded in $\mathfrak{g}$;

IV $\mathfrak{g}$ is a complex simple Lie algebra considered as a real Lie algebra and s is the conjugation with respect to a maximal compactly embedded subalgebra.

Moreover there is a duality between types I and III and between types II and IV, given in Lemma 1.A.9. This shows that the problem of classifying all irreducible orthogonal symmetric Lie algebras is equivalent to the classification of all simple Lie algebras over $\mathbb{R}$ and $\mathbb{C}$. We have summarized these well-known results in Tables 1.1 and 1.2, where we omitted the exceptional Lie algebras for simplicity. For explanations of the notation see Remarks 1.7.2 and 1.7.11.

Label	$\mathfrak{g}$	$\mathfrak{k}$	$\mathfrak{p}$	$\mathfrak{a}$
A	$\mathfrak{sl}(n+1,\mathbb{C})$	$\mathfrak{su}(n+1)$	$\mathfrak{herm}_0(n+1,\mathbb{C})$	$\mathfrak{d}_0(n+1,\mathbb{R})$
B	$\mathfrak{so}(2n+1,\mathbb{C})$	$\mathfrak{so}(2n+1)$	$\mathrm{i}\,\mathfrak{asym}(2n+1,\mathbb{R})$	$\mathfrak{qd}(2n+1,\mathrm{i}\mathbb{R})$
C	$\mathfrak{sp}(n,\mathbb{C})$	$\mathfrak{sp}(n)$	$\left(\begin{smallmatrix} X & \bar{Y} \\ Y & -\bar{X} \end{smallmatrix}\right)$, $\begin{smallmatrix} X\in\mathfrak{herm}(n,\mathbb{C}) \\ Y\in\mathfrak{sym}(n,\mathbb{C}) \end{smallmatrix}$	$X \in \mathfrak{d}(n,\mathbb{R}), Y = 0$
D	$\mathfrak{so}(2n,\mathbb{C})$	$\mathfrak{so}(2n)$	$\mathrm{i}\,\mathfrak{asym}(2n,\mathbb{R})$	$\mathfrak{qd}(2n,\mathrm{i}\mathbb{R})$

Table 1.1: **Irreducible orthogonal symmetric Lie algebras (types II and IV).** We list the simple Lie algebras $\mathfrak{g}$ over $\mathbb{C}$ and a maximal compactly embedded subalgebra $\mathfrak{k}$. Then $s(X) = -X^*$ is the corresponding Cartan involution and $\mathfrak{p} = \mathrm{i}\mathfrak{k}$ is the -1 eigenspace. Moreover, $\mathfrak{a}$ is a maximal Abelian subspace of $\mathfrak{p}$, and its complexification is a Cartan subalgebra of $\mathfrak{g}$. These are the irreducible orthogonal symmetric Lie algebras of type IV. The corresponding compact irreducible orthogonal symmetric Lie algebras of type II are then $\mathfrak{k} \oplus \mathfrak{k}$ and s simply interchanges the terms. See also [Hel78, Ch. III §8, Ch. X]. For explanations of the notation see Remarks 1.7.2 and 1.7.11.

Remark 1.7.1. *The Lie algebras in Tables 1.1 and 1.2 can be represented in different but equivalent ways. We use the definitions given in [Hel78, p. 446]. Hence they are all real or complex matrix Lie algebras.*

[13] The name comes from the fact that the dual symmetric Lie algebra is $\mathfrak{su}(2n)$. Note that $\mathfrak{su}^*(2n)$ is isomorphic to $\mathfrak{sl}(n,\mathbb{H})$ via the standard embedding $\jmath$. Similarly the corresponding $\mathfrak{p}$ part equals $\jmath(\mathfrak{herm}_0(n,\mathbb{H}))$.

[14] Again the name stems from the dual symmetric Lie algebra $\mathfrak{so}(2n)$.

Label	$\mathfrak{g}$	$\mathfrak{k}$	$\mathfrak{p}$	$\mathfrak{a}$
AI	$\mathfrak{sl}(n,\mathbb{R})$	$\mathfrak{so}(n)$	$\mathfrak{sym}_0(n,\mathbb{R})$	$\mathfrak{d}_0(n,\mathbb{R})$
AII	$\mathfrak{su}^*(2n)$[13]	$\mathfrak{sp}(n)$	$\left(\begin{smallmatrix} X & -\bar{Y} \\ Y & \bar{X} \end{smallmatrix}\right)$, $\begin{smallmatrix} X\in\mathfrak{herm}_0(n,\mathbb{C}) \\ Y\in\mathfrak{asym}(n,\mathbb{C}) \end{smallmatrix}$	$X \in \mathfrak{d}_0(n,\mathbb{R}), Y = 0$
AIII	$\mathfrak{su}(p,q)$	$\mathfrak{s}(\mathfrak{u}(p)\oplus\mathfrak{u}(q))$	$\left(\begin{smallmatrix} 0 & Y \\ Y^* & 0 \end{smallmatrix}\right)$, $Y \in \mathbb{C}^{p,q}$	$Y \in \mathfrak{d}(p,q,\mathbb{R})$
BDI	$\mathfrak{so}(p,q)$	$\mathfrak{so}(p)\oplus\mathfrak{so}(q)$	$\left(\begin{smallmatrix} 0 & Y \\ Y^\top & 0 \end{smallmatrix}\right)$, $Y \in \mathbb{R}^{p,q}$	$Y \in \mathfrak{d}(p,q,\mathbb{R})$
CI	$\mathfrak{sp}(n,\mathbb{R})$	$\imath(\mathfrak{u}(n))$	$\left(\begin{smallmatrix} X & Y \\ Y & -X \end{smallmatrix}\right)$, $X,Y \in \mathfrak{sym}(n,\mathbb{R})$	$X \in \mathfrak{d}(n,\mathbb{R}), Y = 0$
CII	$\mathfrak{sp}(p,q)$	$\mathfrak{sp}(p)\oplus\mathfrak{sp}(q)$	$\jmath(\left(\begin{smallmatrix} 0 & Y \\ Y^* & 0 \end{smallmatrix}\right))$, $Y \in \mathbb{H}^{p,q}$	$Y \in \mathfrak{d}(p,q,\mathbb{R})$
DIII	$\mathfrak{so}^*(2n)$[14]	$\imath(\mathfrak{u}(n))$	$\left(\begin{smallmatrix} X & Y \\ Y & -X \end{smallmatrix}\right)$, $X,Y \in \mathrm{i}\mathfrak{asym}(n,\mathbb{R})$	$X \in \mathfrak{qd}(n,\mathrm{i}\mathbb{R}), Y = 0$

Table 1.2: **Irreducible orthogonal symmetric Lie algebras (types I and III).** We list the simple Lie algebras $\mathfrak{g}$ over $\mathbb{R}$ with a Cartan involution $s(X) = -X^*$, the corresponding Cartan decomposition $\mathfrak{k}\oplus\mathfrak{p}$, and a Cartan subalgebra $\mathfrak{a}$. These are the irreducible orthogonal symmetric Lie algebras of type III. The corresponding compact irreducible orthogonal symmetric Lie algebras of type I are easily obtained via duality. See also [Hel78, Ch. X §2.3]. For explanations of the notation see Remarks 1.7.2 and 1.7.11.

Remark 1.7.2. *Let $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. For $x \in \mathbb{K}$, $\overline{x}$ denotes the (complex or quaternionic) conjugate. For a matrix $X \in \mathbb{K}^{m,n}$, $\overline{X}$ denotes the elementwise conjugate, $X^\top$ denotes the transposed matrix, and $X^* = \overline{X}^\top$ denotes the Hermitian conjugate. Then $\mathfrak{sym}(n,\mathbb{K}) = \{X \in \mathbb{K}^{n,n} : X = X^\top\}$ denotes the set of all symmetric matrices. Similarly $\mathfrak{asym}(n,\mathbb{K}) = \{X \in \mathbb{K}^{n,n} : X = -X^\top\}$ denotes the set of all skew-symmetric matrices. Moreover $\mathfrak{herm}(n,\mathbb{K}) = \{X \in \mathbb{K}^{n,n} : X = X^*\}$ denotes the set of all Hermitian matrices. If we additionally assume that the matrices are traceless, we write $\mathfrak{sym}_0(n,\mathbb{K})$ and $\mathfrak{herm}_0(n,\mathbb{K})$. Finally, diagonal matrices are denoted by $\mathfrak{d}(m,n,\mathbb{K})$, or $\mathfrak{d}(n,\mathbb{K})$, and with subscript 0 if the diagonal elements add up to 0. Furthermore we define some useful matrices:*

$$I_{n,m} = \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix}, \quad J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

where I_n denotes the identity matrix of size n. Similarly we will write 0 for the zero matrix (where the size is clear from context). With this we can define the quasi-diagonal *matrices $\mathfrak{qd}(n,\mathbb{K})$ as $\mathfrak{d}(n/2,\mathbb{K})\otimes J_1$ for n even and with an additional row and column of zeros if n is odd. We will also use the standard embeddings*

$$\imath : \mathbb{C} \to \mathbb{R}^{2,2}, \quad x + \mathrm{i}y \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \tag{1.18}$$

and

$$\jmath : \mathbb{H} \to \mathbb{C}^{2,2}, \quad a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d = \alpha + \mathrm{j}\beta \mapsto \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \tag{1.19}$$

where $\alpha = a + \mathrm{i}b$ and $\beta = c - \mathrm{i}d$, which can analogously be defined to act on matrices.

Since the irreducible orthogonal symmetric Lie algebras can be fully classified, it is natural to ask under which conditions an orthogonal, symmetric Lie algebra can be decomposed in some sense into such irreducible pieces. We have seen that an effective, orthogonal, symmetric Lie algebra can be

decomposed into a Euclidean, a compact, and a non-compact part, cf. Lemma 1.A.10. Similarly, a semisimple, strongly effective, orthogonal, symmetric Lie algebra can be decomposed into a Euclidean part and a direct sum of irreducible orthogonal symmetric Lie algebras, see [Hel78, Ch. VIII, Prop. 5.2]. The case we are mostly interested in, semisimple, orthogonal, symmetric Lie algebras, lies between these two cases. In the following we will show that a semisimple, orthogonal, symmetric Lie algebra is still orbit equivalent to a direct sum of irreducible orthogonal symmetric Lie algebras (and a trivial part). This is the content of Theorem 1.7.10.

First we make the concept of orbit equivalence precise.

Definition 1.7.3. *Let X be a set and let $\mathbf{G}, \mathbf{H}$ be groups acting on X. We say that the actions are* orbit equivalent *if they have the same set of orbits, that is $X/\mathbf{G} = X/\mathbf{H}$.*

By Corollary 1.A.48 it holds for semisimple, orthogonal, symmetric Lie algebras that the orbits in $\mathfrak{p}$ do not depend on the choice of associated pair. Thus we will often simply choose to work with the canonical associated pair. This also allows us to define orbit equivalence for symmetric Lie algebras with the same $\mathfrak{p}$:

Definition 1.7.4. *Let $\mathfrak{g}_i = \mathfrak{k}_i \oplus \mathfrak{p}$ for $i = 1, 2$ be symmetric Lie algebras. They are* orbit equivalent *if the representations of $\mathbf{K}_i = \mathrm{Int}_{\mathfrak{k}_i}(\mathfrak{g}_i)$ on $\mathfrak{p}$ are orbit equivalent.*

Now we show how $\mathfrak{p}$ splits into irreducible representations and that the maximal Abelian subspace $\mathfrak{a}$ as well as the tangent space to the orbit $\mathrm{ad}_{\mathfrak{k}}(x)$ for x regular respect this decomposition.

Lemma 1.7.5. *Let $(\mathfrak{g}, s)$ be a semisimple, symmetric Lie algebra and let $V, W \subseteq \mathfrak{p}$ be $\mathrm{ad}_{\mathfrak{k}}$-invariant and orthogonal with respect to the Killing form B of $\mathfrak{g}$. Then they commute.*

Proof. Let $v \in V$ and $w \in W$. Then $B([v, w], [v, w]) = B(v, [w, [v, w]]) = 0$ and so $[v, w] = 0$. □

Lemma 1.7.6. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra with Cartan-like decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Then there is an orthogonal decomposition $\mathfrak{p} = \bigoplus_{i=1}^n \mathfrak{p}_i$ into irreducible components for the action of $\mathrm{ad}_{\mathfrak{k}}$ such that $\mathfrak{p}_i \subseteq \mathfrak{p}_\pm$ for each $i = 1, \dots, n$. If $\mathfrak{a}_i = \mathfrak{p}_i \cap \mathfrak{a}$, then it holds that $\mathfrak{a} = \bigoplus_{i=1}^n \mathfrak{a}_i$.*

Proof. We use the inner product from Lemma 1.A.15, in which all ad_k for $k \in \mathfrak{k}$ are skew-symmetric. First, using Corollary 1.A.13, $\mathfrak{p}$ splits into $\mathfrak{p}_- \oplus \mathfrak{p}_+$, the compact and non-compact parts, which are invariant under the action of $\mathfrak{k}$. Hence there exists an orthogonal decomposition $\mathfrak{p} = \bigoplus_{i=1}^n \mathfrak{p}_i$ into irreducible components for the action of $\mathfrak{k}$ such that each $\mathfrak{p}_i$ is contained in $\mathfrak{p}_-$ or $\mathfrak{p}_+$. Then the $\mathfrak{p}_i$ are also orthogonal with respect to the Killing form B of $\mathfrak{g}$ since for $x_i \in \mathfrak{p}_i$ and $x_j \in \mathfrak{p}_j$ it holds that $B(x_i, x_j) = \pm \langle x_i, x_j \rangle$. Hence, by Lemma 1.7.5, the $\mathfrak{p}_i$ commute. Now let $x \in \mathfrak{a}$, and $x = \sum_{i=1}^n x_i$ for $x_i \in \mathfrak{p}_i$. We show that all $x_i \in \mathfrak{a}_i \subseteq \mathfrak{a}$. Let $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ as in Lemma 1.A.20. Since all x_i commute with each other, there is $K \in \mathbf{K}$ such that $Kx_i \in \mathfrak{a}$ and hence $Kx = \sum_{i=1}^n Kx_i \in \mathfrak{a}$. By left multiplying K with an appropriate element from the normalizer $N_{\mathbf{K}}(\mathfrak{a})$ we may assume without loss of generality that $Kx = x$ and still $Kx_i \in \mathfrak{a}$. Hence by the invariance of the $\mathfrak{p}_i$ under $\mathbf{K}$ it holds that $Kx_i \in \mathfrak{p}_i$ and so $x_i = Kx_i$ and hence $x_i \in \mathfrak{a}_i$, as desired. □

Corollary 1.7.7. *We use the same notation as in Lemma 1.7.6. Let $x = \sum_{i=1}^n x_i \in \mathfrak{a}$ be regular. Then it holds that $\mathrm{ad}_{\mathfrak{k}}(x) = \bigoplus_{i=1}^n \mathrm{ad}_{\mathfrak{k}}(x_i)$.*

Proof. Regular elements exist by Lemma 1.A.18. By Lemma 1.A.16 it holds that $\mathfrak{a}^\perp = \mathrm{ad}_{\mathfrak{k}}(x)$, and by Lemma 1.7.6 we have the orthogonal decomposition $\mathfrak{p} = \bigoplus_{i=1}^n \mathfrak{p}_i$. Then for $y \in \mathfrak{p}_i$ it holds that $\langle y, \mathrm{ad}_{\mathfrak{k}}(x) \rangle = \langle y, \mathrm{ad}_{\mathfrak{k}}(x_i) \rangle$. Hence y is orthogonal to $\mathrm{ad}_{\mathfrak{k}}(x_i)$ if and only if $y \in \mathfrak{a}_i$, that is, we have

the orthogonal decomposition $\mathfrak{p}_i = \mathfrak{a}_i \oplus \mathrm{ad}_{\mathfrak{k}}(x_i)$. This shows that $\mathrm{ad}_{\mathfrak{k}}(x) = \mathfrak{a}^{\perp} = \bigoplus_{i=1}^{n} \mathrm{ad}_{\mathfrak{k}}(x_i)$, as desired. □

This local result on the splitting of the tangent space of an orbit can be generalized to the entire orbit, showing that the semisimple, orthogonal, symmetric Lie algebra is orbit equivalent to a direct sum of reductive ones with irreducible isotropy representations.

Lemma 1.7.8. *We use the same notation as in Lemma 1.7.6. For each i, $\mathfrak{g}_i = \mathfrak{k} \oplus \mathfrak{p}_i \subseteq \mathfrak{g}$ is a reductive, orthogonal, symmetric Lie subalgebra, and so is $(\mathfrak{g}', s') = \bigoplus_{i=1}^{n} (\mathfrak{g}_i, s|_{\mathfrak{g}_i})$. Since $\mathfrak{p} = \bigoplus_{i=1}^{n} \mathfrak{p}_i$ the isotropy representations of $(\mathfrak{g}, s)$ and of $(\mathfrak{g}', s')$ act on the same space, and in fact they are orbit equivalent.*

Proof. It is clear that $(\mathfrak{g}_i, s|_{\mathfrak{g}_i})$ is a symmetric Lie subalgebra and by Lemma 1.7.6 it holds that $\mathfrak{p}_i \subseteq \mathfrak{p}_{\pm}$. Hence it is reductive and orthogonal by Lemma 1.A.43. By Lemma 1.A.40, also $(\mathfrak{g}', s')$ is a reductive, orthogonal, symmetric Lie algebra. Let $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ and $\mathbf{K}' = \mathrm{Int}_{\mathfrak{k}'}(\mathfrak{g}')$ be the respective compact Lie groups acting on $\mathfrak{p}$. Note that on each $\mathfrak{p}_i$, $\mathbf{K}$ and $\mathbf{K}'$ generate the same orbits since $\mathbf{K}$ and $\mathbf{K}'$ are connected and $\mathrm{ad}_{\mathfrak{k}}|_{\mathfrak{p}_i} = \mathrm{ad}_{\mathfrak{k}'}|_{\mathfrak{p}_i}$. Hence, in $\mathfrak{p}$, each $\mathbf{K}$-orbit lies in some $\mathbf{K}'$-orbit. Let $x \in \mathfrak{a}$ be regular in $(\mathfrak{g}, s)$, then Corollary 1.7.7 shows that the tangent space at x of the $\mathbf{K}$ and $\mathbf{K}'$ orbits through x is the same for both isotropy representations. Since the orbits $\mathbf{K}x$ and $\mathbf{K}'x$ though x satisfy $\mathbf{K}x \subseteq \mathbf{K}'x$, and since they are connected and by the previous argument have the same dimension, they must coincide. This shows that the orbits of regular points of $(\mathfrak{g}, s)$ coincide. It remains to show the same for singular orbits. Since the regular points are Zariski open in $\mathfrak{p}$, they are dense in the standard topology. Now let $y, z \in \mathfrak{p}$ be non-regular for $(\mathfrak{g}, s)$ with distinct $\mathbf{K}$-orbits. Let N_y and N_z be disjoint tubular neighborhoods in $\mathfrak{p}$ for the action of $\mathbf{K}$, see [AB15, Thm. 3.57]. If there is $K \in \mathbf{K}'$ such that $Ky = z$ then K also maps some regular points in N_y to N_z which gives a contradiction. This concludes the proof. □

Lemma 1.7.9. *Let $(\mathfrak{g}, s)$ be a reductive, orthogonal, symmetric Lie algebra with $\mathrm{ad} : \mathfrak{k} \to \mathfrak{gl}(\mathfrak{p})$ irreducible. Then $\mathfrak{g}$ is a direct sum of symmetric Lie subalgebras $\mathfrak{g}', \mathfrak{g}'' \subseteq \mathfrak{g}$ where $(\mathfrak{g}', s|_{\mathfrak{g}'})$ is an irreducible orthogonal symmetric Lie algebra and $(\mathfrak{g}'', s|_{\mathfrak{g}''})$ has trivial isotropy representation.*

Proof. Since $\mathfrak{g}$ is reductive, by Lemma 1.A.42 it can be written as a direct sum of Lie subalgebras $[\mathfrak{g}, \mathfrak{g}]$ and $\mathfrak{z}$ where the former is semisimple and the latter is the center of $\mathfrak{g}$. Clearly they are symmetric Lie subalgebras of $\mathfrak{g}$, and by Lemma 1.A.7, or by Lemma 1.A.40, $[\mathfrak{g}, \mathfrak{g}]$ is orthogonal. Since $\mathrm{ad} : \mathfrak{k} \to \mathfrak{gl}(\mathfrak{p})$ is irreducible there are two possibilities: Either $[\mathfrak{g}, \mathfrak{g}] \cap \mathfrak{p}$ is zero, in which case we can set $\mathfrak{g}' = 0$ and $\mathfrak{g}'' = \mathfrak{g}$. Otherwise, $\mathfrak{z} \cap \mathfrak{p}$ is zero. Then since $[\mathfrak{g}, \mathfrak{g}]$ is semisimple it is the direct sum of simple ideals $\mathfrak{g}_i$ for $i = 1, \ldots, m$. For each i there is some j such that $s(\mathfrak{g}_i) = \mathfrak{g}_j$. Consider the semisimple, orthogonal, symmetric Lie subalgebras $\mathfrak{h}_i = \mathfrak{g}_i \oplus s(\mathfrak{g}_i)$ (without repetitions). Denote their Cartan-like decomposition by $\mathfrak{h}_i = \mathfrak{l}_i \oplus \mathfrak{q}_i$. Then, by irreducibility of $\mathfrak{g}$, all but one $\mathfrak{q}_i$ is zero. Without loss of generality we say that $\mathfrak{q}_1$ is non-zero. Then the adjoint action of $\mathfrak{l}_1$ on $\mathfrak{q}_1$ is irreducible and effective. Hence $\mathfrak{h}_1$ is an irreducible orthogonal symmetric Lie algebra so we set $\mathfrak{g}' = \mathfrak{h}_1$, and the remaining $\mathfrak{h}_i = \mathfrak{l}_i$ and $\mathfrak{z}$ together yield $\mathfrak{g}''$. □

Putting it all together, we obtain the main theorem of this section. This result is similar to [Dad85, Thm. 4], which considers more general polar actions.

Theorem 1.7.10 (Classification of Diagonalizations). *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Then there are irreducible orthogonal symmetric Lie subalgebras $(\mathfrak{g}_i, s|_{\mathfrak{g}_i})$ for $i = 1, \ldots, n$ and a symmetric Lie subalgebra $(\mathfrak{g}', s_{\mathfrak{g}})$ with trivial isotropy representation such that $\mathfrak{p} = \bigoplus_{i=1}^{n} \mathfrak{p}_i \oplus \mathfrak{p}'$ and such that the isotropy representation of $(\mathfrak{g}, s)$ is orbit equivalent to the*

isotropy representation of $\bigoplus_{i=1}^n (\mathfrak{g}_i, s_i) \oplus (\mathfrak{g}', s')$.

Proof. First let $\mathfrak{p} = \bigoplus_{i=1}^n \mathfrak{p}_i$ be an orthogonal decomposition into irreducible subrepresentations for the action of $\mathrm{ad}_{\mathfrak{k}}$ on $\mathfrak{p}$. Then by Lemma 1.7.8, each $\mathfrak{g}_i = \mathfrak{k} \oplus \mathfrak{p}_i \subseteq \mathfrak{g}$ is a reductive, orthogonal, symmetric Lie subalgebra of $(\mathfrak{g}, s)$ and their direct sum $(\mathfrak{g}', s') = \bigoplus_{i=1}^n (\mathfrak{g}_i, s|_{\mathfrak{g}_i})$ has isotropy representation orbit equivalent to that of $(\mathfrak{g}, s)$. By Lemma 1.7.9 each $\mathfrak{g}_i$ is a direct sum of two symmetric Lie subalgebras $\mathfrak{g}_i'$, and $\mathfrak{g}_i''$, where the first is an irreducible orthogonal symmetric Lie algebra and the second has a trivial isotropy representation. Hence $(\mathfrak{g}', s') = \bigoplus_{i=1}^n (\mathfrak{g}_i', s|_{\mathfrak{g}_i'}) \oplus (\mathfrak{g}_i'', s|_{\mathfrak{g}_i'})$. This is the desired decomposition. □

Now that we understand in which sense an arbitrary semisimple, orthogonal, symmetric Lie algebra can be decomposed into irreducible parts, we have a look at these irreducible cases and their relation to diagonalizations. In Table 1.3 we give a list of diagonalizations corresponding to the irreducible orthogonal symmetric Lie algebras. Recall Section 1.2 for some detailed examples. Theorem 1.7.10 shows that the diagonalizations shown in Table 1.3 are essentially the only possibilities (omitting diagonalizations stemming from exceptional Lie algebras). This means that if we are given a semisimple, orthogonal, symmetric Lie algebra, we may think of it as a direct sum of irreducible parts. In particular we have a decomposition $\mathfrak{p} = \bigoplus_{i=1}^n \mathfrak{p}_i$, and similarly $\mathfrak{a} = \bigoplus_{i=1}^n \mathfrak{a}_i$. If $p : I \to \mathfrak{p}$ is a path, we can compute the diagonalization in each $\mathfrak{p}_i$ individually, that is we consider the paths $p_i(t) \in \mathfrak{p}_i$ and compute their diagonalizations $a_i(t) \in \mathfrak{a}_i$, which can be done in practice using algorithms for the various diagonalizations in Table 1.3. Note however that it is not straightforward to find a diagonalizing $K \in \mathbf{K}$ form the individual K_i.

Remark 1.7.11. *The Hamiltonian matrices are defined as* $\mathfrak{ham}(n, \mathbb{K}) = \{X \in \mathbb{K}^{2n,2n} : J_n A = -A^\top J_n\}$. *Note that* $\mathfrak{u}(n) = \mathrm{i}\,\mathfrak{herm}(n, \mathbb{C})$, *and* $\mathfrak{su}(n) = \mathrm{i}\,\mathfrak{herm}_0(n, \mathbb{C})$, *and* $\mathfrak{so}(n, \mathbb{K}) = \mathfrak{asym}(n, \mathbb{K})$, *and for* $\mathbb{K} = \mathbb{R}, \mathbb{C}$, *we have* $\mathfrak{sp}(n, \mathbb{K}) = \mathfrak{ham}(n, \mathbb{K})$. *Moreover let us define the* $*$*-Hamiltonian matrices* $\mathfrak{ham}^*(n, \mathbb{K}) = \{X \in \mathbb{K}^{2n,2n} : J_n A = -A^* J_n\}$.

1.A Symmetric Lie Algebras

In this appendix we give a rigorous introduction to symmetric Lie algebras and prove a number of auxiliary results which are used repeatedly in the main text. Our definitions follow the standard reference of Helgason [Hel78].

Basic Definitions

We start by considering symmetric Lie algebras, as defined in [Hel78, p. 229].

Definition 1.A.1 (Effective orthogonal symmetric Lie algebra). *Let* $\mathfrak{g}$ *be a real finite-dimensional Lie algebra and* s *an involutive*[15] *Lie algebra automorphism of* $\mathfrak{g}$. *Then the pair* $(\mathfrak{g}, s)$ *is called a* symmetric Lie algebra. *Now if* $\mathfrak{k} \subseteq \mathfrak{g}$ *denotes the fixed point set of* s, *we define the following:*

(i) If $\mathfrak{k}$ *is a compactly embedded*[16] *subalgebra of* $\mathfrak{g}$, *then* $(\mathfrak{g}, s)$ *is* orthogonal.

(ii) If $\mathfrak{k} \cap \mathfrak{z} = \{0\}$, *where* $\mathfrak{z}$ *denotes the center of* $\mathfrak{g}$, *then* $(\mathfrak{g}, s)$ *is* effective.

[15] A map f is involutive if $f \circ f$ is the identity.

[16] This means that the analytic subgroup of $\mathrm{GL}(\mathfrak{g})$ with Lie algebra $\mathrm{ad}_{\mathfrak{k}}$, denoted by $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$, is compact, see [Hel78, p. 130]. More details on this group will be given later in this section.

Name	Group	Matrices	Diagonal form	Type
real EVD	$\mathrm{SO}(n)$	$\mathfrak{sym}_0(n,\mathbb{R})$	$\mathfrak{d}_0(n,\mathbb{R})$	AI
complex EVD	$\mathrm{SU}(n)$	$\mathfrak{herm}_0(n,\mathbb{C})$	$\mathfrak{d}_0(n,\mathbb{R})$	A
quaternionic EVD	$\mathrm{Sp}(n)$	$\jmath(\mathfrak{herm}_0(n,\mathbb{H}))$	$\jmath(\mathfrak{d}_0(n,\mathbb{R}))$	AII
skew-symmetric EVD	$\mathrm{SO}(2n+1)$	$\mathfrak{asym}(2n+1,\mathbb{R})$	$\mathfrak{qd}(2n+1,\mathbb{R})$	B
skew-symmetric EVD	$\mathrm{SO}(2n)$	$\mathfrak{asym}(2n,\mathbb{R})$	$\mathfrak{qd}(2n,\mathbb{R})$	D
real SVD	$\mathrm{S}(\mathrm{O}(p)\times\mathrm{O}(q))$	$\mathbb{R}^{p,q}$	$\mathfrak{d}(p,q,\mathbb{R})$	BDI
complex SVD	$\mathrm{S}(\mathrm{U}(p)\times\mathrm{U}(q))$	$\mathbb{C}^{p,q}$	$\mathfrak{d}(p,q,\mathbb{R})$	AIII
quaternionic SVD	$\mathrm{Sp}(p)\times\mathrm{Sp}(q)$	$\jmath(\mathbb{H}^{p,q})$	$\jmath(\mathfrak{d}(p,q,\mathbb{R}))$	CII
real symmetric Hamiltonian EVD	$\imath(\mathrm{U}(n))$	$\mathfrak{sym}(2n,\mathbb{R})\cap\mathfrak{ham}(n,\mathbb{R})$	$I_{1,1}\otimes\mathfrak{d}(n,\mathbb{R})$	CI
Hermitian Hamiltonian EVD	$\mathrm{Sp}(n)$	$\mathfrak{herm}(2n,\mathbb{C})\cap\mathfrak{ham}(n,\mathbb{C})$	$I_{1,1}\otimes\mathfrak{d}(n,\mathbb{R})$	C
Hermitian $*$-Hamiltonian EVD	$\mathrm{S}(\mathrm{U}(n)\times\mathrm{U}(n))$	$\mathfrak{herm}(2n,\mathbb{C})\cap\mathfrak{ham}^*(n,\mathbb{C})$	$I_{1,1}\otimes\mathfrak{d}(n,\mathbb{R})$	AIII
Autonne–Takagi factorization	$\mathrm{U}(n)$	$\mathfrak{sym}(n,\mathbb{C})$	$\mathfrak{d}(n,\mathbb{R})$	CI
Hua factorization	$\mathrm{U}(2n+1)$	$\mathfrak{asym}(2n+1,\mathbb{C})$	$\mathfrak{qd}(2n+1,\mathbb{R})$	DIII
Hua factorization	$\mathrm{U}(2n)$	$\mathfrak{asym}(2n,\mathbb{C})$	$\mathfrak{qd}(2n,\mathbb{R})$	DIII

Table 1.3: **Various Matrix Diagonalizations.** We list several matrix diagonalizations and the corresponding symmetric Lie algebras. Note that some of these are isomorphic in the sense that they correspond to the same symmetric Lie algebra, but the isomorphism is not necessarily obvious. For more details on the explicit expressions we refer to [Kle06]. The eigenvalue decompositions of so-called perplectic matrices described in [MMD05] also fit into this setting, see [Kle06, Sec. 4.2]. For explanations of the notation see Remarks 1.7.2 and 1.7.11.

(iii) If $\mathfrak{k}$ does not contain a non-trivial ideal of $\mathfrak{g}$, then $(\mathfrak{g}, s)$ is strongly effective.

Remark 1.A.2. *Note that if $\mathfrak{g}$ is semisimple, then the center of $\mathfrak{g}$ is trivial, i.e., $\mathfrak{z} = \{0\}$, and hence $(\mathfrak{g}, s)$ is automatically effective.*

A key feature of symmetric Lie algebras is that they admit a *Cartan-like decomposition* $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ with special commutator relations given in (1.20) below. In fact such a decomposition automatically yields the structure of a symmetric Lie algebra:

Lemma 1.A.3. *Let $\mathfrak{g}$ be a real Lie algebra. The following statements hold:*

(i) If s is an involutive Lie algebra automorphism of $\mathfrak{g}$ (i.e. $(\mathfrak{g}, s)$ is a symmetric Lie algebra) then $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ (as a direct sum of subspaces) where $\mathfrak{k}$ and $\mathfrak{p}$ are the $+1$ and -1 eigenspaces of s, respectively. It holds that

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}. \tag{1.20}$$

(ii) If $\mathfrak{g}$ admits a vector space decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ satisfying (1.20), then the linear map s defined as $+\mathbb{1}$ on $\mathfrak{k}$ and $-\mathbb{1}$ on $\mathfrak{p}$ is a Lie algebra automorphism, and hence $(\mathfrak{g}, s)$ is a symmetric Lie algebra.

Proof. (i): Since s is a linear involution on $\mathfrak{g}$ it satisfies $s^2 - \mathbb{1} = 0$. The minimal polynomial of s divides $(x+1)(x-1)$ and hence it splits into distinct linear factors, showing that s is diagonalizable. Thus $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ where $\mathfrak{k}, \mathfrak{p}$ are the corresponding ± 1 eigenspaces. Since s is a Lie algebra automorphism it satisfies the relations (1.20).

(ii): By definition s is a linear involution. Now given $k, l \in \mathfrak{k}$ and $x, y \in \mathfrak{p}$

$$\begin{aligned} s([k+x, l+y]) &= s([k,l]) + s([k,y]) + s([x,l]) + s([x,y]) = [k,l] - [k,y] - [x,l] + [x,y] \\ &= [k,l] + [k,-y] + [-x,l] + [-x,-y] = [k-x, l-y] = [s(k+x), s(l+y)]\,, \end{aligned}$$

showing that s is a Lie algebra endomorphism. Since it is bijective, it is in fact an automorphism. □

Due to this result, we will often specify a symmetric Lie algebra by its Cartan-like decomposition, instead of by the corresponding automorphism.

Next let us clarify the different notions of effectivity:

Lemma 1.A.4. *Given a symmetric Lie algebra $(\mathfrak{g}, s)$ the following statements hold.*

(i) $(\mathfrak{g}, s)$ is effective if and only if the adjoint representation of $\mathfrak{k}$ on $\mathfrak{g}$ is faithful.

(ii) $(\mathfrak{g}, s)$ is strongly effective if and only if the adjoint representation of $\mathfrak{k}$ on $\mathfrak{p}$ is faithful.

Furthermore if $(\mathfrak{g}, s)$ is strongly effective then it is effective.

Proof. Let $\mathfrak{z}$ denote the center of $\mathfrak{g}$. (i): This is clear since $\mathfrak{k} \cap \mathfrak{z}$ is exactly the kernel of the adjoint representation of $\mathfrak{k}$ on $\mathfrak{g}$. (ii): If $\mathfrak{k}$ contains a non-trivial ideal of $\mathfrak{g}$, then this ideal lies in the kernel of the adjoint representation of $\mathfrak{k}$ on $\mathfrak{p}$, as can be seen from (1.20). Conversely, the kernel of the adjoint representation of $\mathfrak{k}$ on $\mathfrak{p}$ is an ideal of $\mathfrak{g}$. □

Remark 1.A.5. *There are slightly different definitions used in the literature. For instance, Kobayashi & Nomizu [KN96, Ch. XI Sec. 2] call effective what we call strongly effective. Our terminology is consistent with Helgason [Hel78] who, however, does not give a name to what we call strongly effective.*

Next let us have a look at orthogonality. This condition guarantees the existence of an adapted inner product on $\mathfrak{g}$.

Lemma 1.A.6. *Let $(\mathfrak{g}, s)$ be an orthogonal, symmetric Lie algebra. Then there exists an s-invariant inner product for which all of $\mathrm{ad}_{\mathfrak{k}}$ is skew-symmetric.*

Proof. Start with any inner product $(\cdot, \cdot)$ on $\mathfrak{g}$. Let $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$, which is compact by assumption and hence admits a Haar measure denoted $d\mu$. Now define the averaged inner product

$$\langle x, y \rangle = \sum_{m=0}^{1} \int_{\mathbf{K}} (s^m(Kx), s^m(Ky))\, d\mu(K)\,. \tag{1.21}$$

Since $\mathbf{K}$ preserves the decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ due to the commutator relations (1.20), s commutes with the action of $K \in \mathbf{K}$. Thus (1.21) is invariant under s and under $\mathbf{K}$. This shows that $\mathbf{K}$ is a subgroup of the orthogonal group of $(\mathfrak{g}, \langle \cdot, \cdot \rangle)$, and hence its Lie algebra $\mathrm{ad}_{\mathfrak{k}}$ is a subalgebra of the orthogonal Lie algebra, so each ad_k for $k \in \mathfrak{k}$ is skew-symmetric with respect to $\langle \cdot, \cdot \rangle$. □

We have the following converse result in the semisimple case. First recall that a Lie algebra $\mathfrak{g}$ over a field of characteristic 0 is semisimple if and only if the Killing form $B(x, y) = \mathrm{tr}(\mathrm{ad}_x \circ \mathrm{ad}_y)$ is non-degenerate, or, if and only if $\mathfrak{g}$ contains no proper non-trivial Abelian ideals.

Lemma 1.A.7. *Let $(\mathfrak{g}, s)$ be a semisimple, symmetric Lie algebra and let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathfrak{g}$ for which all of $\mathrm{ad}_{\mathfrak{k}}$ is skew-symmetric. Then $(\mathfrak{g}, s)$ is orthogonal.*

Proof. Let $\mathbf{G}$ be "the" simply connected Lie group with Lie algebra $\mathfrak{g}$, see [Lee13, Thm. 20.21], and let $\mathbf{K}$ be the connected Lie subgroup with Lie algebra $\mathfrak{k}$. By [Lee13, Thm. 20.19] there is a unique Lie group automorphism $\sigma : \mathbf{G} \to \mathbf{G}$ with $D\sigma(e) = s$. As $\mathfrak{k}$ is the fixed point set of s, $\mathbf{K}$ equals the identity component of the fixed point set of σ. In particular, $\mathbf{K}$ is closed in $\mathbf{G}$, and hence an embedded Lie subgroup. Since $\mathfrak{g}$ is semisimple, $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is faithful, and hence the adjoint representation $\mathrm{Ad} : \mathbf{G} \to \mathrm{Int}(\mathfrak{g})$ is a covering homomorphism by [Lee13, Thm. 21.31]. The image of $\mathbf{K}$ under Ad is exactly $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$. Now let $K \in \mathbf{K}$ be arbitrary. There is a neighborhood W of Ad_K in $\mathrm{Int}(\mathfrak{g})$ diffeomorphic to some neighborhood W' of K in $\mathbf{G}$. Since $\mathbf{K}$ is embedded, and by shrinking the neighborhoods, we can assume that W' is a slice chart for $\mathbf{K}$, and hence the same is true for W and $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$. By [Lee13, Thm. 5.8] this shows that $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is embedded in $\mathrm{Int}(\mathfrak{g})$ and hence closed. By [Hel78, Ch. II Coro. 6.5] $\mathrm{Int}(\mathfrak{g})$ is closed in $\mathrm{GL}(\mathfrak{g})$. By assumption, $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g}) \subseteq \mathrm{SO}(\mathfrak{g})$ with respect to the given inner product. Since $\mathrm{SO}(\mathfrak{g})$ is closed in $\mathrm{GL}(\mathfrak{g})$, by Lemma 1.6.8 $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is closed in $\mathrm{SO}(\mathfrak{g})$, and hence compact. This concludes the proof. □

One calls a Lie algebra $\mathfrak{g}$ compact if its Killing form is negative definite[17]. These notions lead to different classes of symmetric Lie algebras:

Definition 1.A.8. *Let $(\mathfrak{g}, s)$ be a symmetric Lie algebra.*

(i) If $[\mathfrak{p}, \mathfrak{p}] = 0$, then $(\mathfrak{g}, s)$ is called of Euclidean type.

(ii) If $\mathfrak{g}$ is semisimple, it is called of compact type *if $\mathfrak{g}$ is compact. Otherwise it is called of* of non-compact type.

(iii) If the adjoint representation of $\mathfrak{k}$ on $\mathfrak{p}$ is irreducible, then $(\mathfrak{g}, s)$ is called irreducible.

[17] This definition excludes for instance Abelian Lie algebras which have trivial Killing form.

There exists an important duality between symmetric Lie algebras of compact and of non-compact type. For a symmetric Lie algebra $(\mathfrak{g}, s)$ with Cartan-like decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ we can define its dual by $\mathfrak{g} = \mathfrak{k} \oplus \mathrm{i}\mathfrak{p}$ as a subspace of the complexification of $\mathfrak{g}$. We denote the dual by $(\mathfrak{g}^*, s^*)$. For the following result see [Hel78, Ch. V, Prop. 2.1].

Lemma 1.A.9. *If $(\mathfrak{g}, s)$ is a semisimple, orthogonal, symmetric Lie algebra of compact type, then $(\mathfrak{g}^*, s^*)$ is of non-compact type, and vice-versa.*

With these definitions in place one can understand the structure of effective, orthogonal, symmetric Lie algebras. Most importantly one gets a decomposition of the Lie algebra into a Euclidean, a compact, and a non-compact subalgebra:

Lemma 1.A.10. *Let $(\mathfrak{g}, s)$ be an effective, orthogonal, symmetric Lie algebra and let $\langle \cdot, \cdot \rangle$ be an s-invariant inner product on $\mathfrak{g}$ such that all ad_k for $k \in \mathfrak{k}$ are skew-symmetric. Then there exists a decomposition of $\mathfrak{g}$ into s-invariant ideals*

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+,$$

which are orthogonal with respect to the Killing form. Denoting the restrictions of s by s_0,s_-, and s_+, the pairs $(\mathfrak{g}_0, s_0)$, $(\mathfrak{g}_-, s_-)$, and $(\mathfrak{g}_+, s_+)$ are effective, orthogonal, symmetric Lie algebras of Euclidean, compact, and non-compact type, respectively. Moreover, there exists a vector space decomposition

$$\mathfrak{p}_- \oplus \mathfrak{p}_+ = \bigoplus_{i=1}^{m} \mathfrak{p}_i,$$

such that $\langle \cdot, \cdot \rangle = \lambda_i B$ on $\mathfrak{p}_i$ with $\lambda_i \neq 0$. It holds that $\mathfrak{p}_i \subseteq \mathfrak{p}_-$ if $\lambda_i < 0$, and $\mathfrak{p}_i \subseteq \mathfrak{p}_+$ if $\lambda_i > 0$. Moreover the $\mathfrak{p}_i$ are orthogonal to each other with respect to the inner product and the Killing form, and $[\mathfrak{p}_i, \mathfrak{p}_j] = 0$.

For a proof we refer to [Hel78, Ch. V, Thm. 1.1] and its proof.

Remark 1.A.11. *Using strong effectivity, one even obtains a decomposition into irreducible ideals, see [Hel78, Ch. VIII, Prop. 5.2].*

Corollary 1.A.12. *If in addition to the assumptions from Lemma 1.A.10 $\mathfrak{g}$ is semisimple, then $\mathfrak{p}_0 = \{0\}$ where $\mathfrak{g}_0 = \mathfrak{k}_0 \oplus \mathfrak{p}_0$.*

Proof. Since $[\mathfrak{p}_0, \mathfrak{p}_0] = 0$, it holds that $\mathfrak{p}_0$ is an Abelian subalgebra of $\mathfrak{g}$. Furthermore we have that $[\mathfrak{k}_0, \mathfrak{p}_0] \subseteq \mathfrak{p}_0$ and hence $\mathfrak{p}_0$ is an ideal in $\mathfrak{g}$. By semisimplicity of $\mathfrak{g}$ we get that $\mathfrak{p}_0 = \{0\}$. □

Combining some of the results above we find the following structure for semisimple, orthogonal, symmetric Lie algebras:

Corollary 1.A.13. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Then we have the following decomposition into ideals*

$$\mathfrak{g} = \mathfrak{k}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+,$$

where $\mathfrak{g}_-$ and $\mathfrak{g}_+$ are of compact and of non-compact type respectively.

Proof. This follows immediately from Remark 1.A.2, Lemma 1.A.10, and Corollary 1.A.12. □

Before using this decomposition of effective, orthogonal, symmetric Lie algebras to connect the s-invariant inner product to the Killing form in the semisimple case, we need the following simple lemma.

Lemma 1.A.14. *Let $(\mathfrak{g}, s)$ be an effective, orthogonal, symmetric Lie algebra and let B denote the Killing form on $\mathfrak{g}$. Then B is negative definite on $\mathfrak{k}$.*

Proof. By Lemma 1.A.6 there exists an inner product on $\mathfrak{g}$ such that all ad_k for $k \in \mathfrak{k}$ are skew-symmetric. Note that the trace and hence the Killing form are independent of the inner product on $\mathfrak{g}$. Hence $B(k,k) = -\operatorname{tr}(\mathrm{ad}_k^\top \circ \mathrm{ad}_k) \leq 0$ with equality only if $\mathrm{ad}_k = 0$, that is $k \in \mathfrak{z}$, and hence by effectivity $k = 0$. □

Now for the result in question:

Lemma 1.A.15. *Let a semisimple, orthogonal, symmetric Lie algebra $(\mathfrak{g}, s)$ be given. Define $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ as follows: Given any $x, y \in \mathfrak{g}$ there exist by the previous results unique $x_k, y_k \in \mathfrak{k}$, $x_+, y_+ \in \mathfrak{p}_+$, and $x_-, y_- \in \mathfrak{p}_-$ such that $x = x_k + x_+ + x_-$, $y = y_k + y_+ + y_-$. Then*

$$\langle x, y \rangle := -B(x_k, y_k) + B(x_+, y_+) - B(x_-, y_-)$$

is an s-invariant inner product on $\mathfrak{g}$ such that ad_k is skew-symmetric for all $k \in \mathfrak{k}$.

Proof. First it is easy to see that the definition yields a bilinear, symmetric form on $\mathfrak{g}$. Then consider the group $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$. Since s and all $K \in \mathbf{K}$ define automorphisms of $\mathfrak{g}$, they leave the Killing form B invariant, and they clearly respect the decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}_- \oplus \mathfrak{p}_+$. Hence, they leave $\langle \cdot, \cdot \rangle$ invariant, and all ad_k for $k \in \mathfrak{k}$ are skew-symmetric. It remains to show that $\langle \cdot, \cdot \rangle$ is positive definite. For this consider another inner product $(\cdot, \cdot)$ as guaranteed by Lemma 1.A.6. Now Lemma 1.A.10 and Corollary 1.A.12 guarantee the existence of a decomposition $\mathfrak{p} = \mathfrak{p}_+ \oplus \mathfrak{p}_- = \bigoplus_{i=1}^m \mathfrak{p}_i$, such that $(\cdot, \cdot) = \lambda_i B$ on $\mathfrak{p}_i$ with $\lambda_i \neq 0$. Moreover $\mathfrak{p}_i \subseteq \mathfrak{p}_-$ if $\lambda_i < 0$, and $\mathfrak{p}_i \subseteq \mathfrak{p}_+$ if $\lambda_i > 0$. Let us denote the corresponding index sets by I_- and I_+, respectively. Now given arbitrary $k \in \mathfrak{k}$ and $x_i \in \mathfrak{p}_i$, using that the $\mathfrak{p}_i$ are orthogonal to each other with respect to the Killing form we compute

$$\begin{aligned}
\langle k + \sum_{i=1}^m x_i, k + \sum_{j=1}^m x_j \rangle &= -B(k,k) + B\Big(\sum_{i \in I_+} x_i, \sum_{j \in I_+} x_j\Big) - B\Big(\sum_{i \in I_-} x_i, \sum_{j \in I_-} x_j\Big) \\
&= -B(k,k) + \sum_{i \in I_+} B(x_i, x_i) - \sum_{i \in I_-} B(x_i, x_i) \\
&= -B(k,k) + \sum_{i=1}^m \frac{1}{|\lambda_i|}(x_i, x_i),
\end{aligned}$$

so Lemma 1.A.14 shows that $\langle \cdot, \cdot \rangle$ is positive definite, as claimed. □

Unless otherwise noted, we will always assume that we are using this inner product. The following result is the reason why we will mostly focus on semisimple Lie algebras, and it will be used several times in the main text.

Lemma 1.A.16. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra and let $x, y \in \mathfrak{p}$. Then the following are equivalent:*

$$[x, y] = 0 \iff \langle [k, x], y \rangle = 0 \quad \text{for all } k \in \mathfrak{k},$$

where the inner product on $\mathfrak{g}$ is as in Lemma 1.A.6. Differently put, for every $x \in \mathfrak{p}$ we have an orthogonal vector space decomposition

$$\mathfrak{p} = (\mathfrak{p} \cap \ker \mathrm{ad}_x) \oplus (\mathfrak{p} \cap \mathrm{im}\, \mathrm{ad}_x).$$

Proof. Let $k \in \mathfrak{k}$ and $x, y \in \mathfrak{p}$. Using Lemma 1.A.10 we can write $\mathfrak{p} = \oplus_{i=1}^m \mathfrak{p}_i$. Note that due to semisimplicity and Corollary 1.A.12 it holds that $\mathfrak{p}_0 = \{0\}$ and hence it is omitted from the decomposition. Furthermore, we know that the $\mathfrak{p}_i$ are mutually orthogonal and invariant under ad_k for all $k \in \mathfrak{k}$. Writing $x = \sum_{i=1}^m x_i$ and $y = \sum_{i=1}^m y_i$ we compute

$$\langle [k, x], y\rangle = \sum_{i=1}^m \langle [k, x_i], y_i\rangle = \sum_{i=1}^m \lambda_i B([k, x_i], y_i) = \sum_{i=1}^m \lambda_i B(k, [x_i, y_i])$$

If $[x, y] = 0$ then $[x_i, y_i] = 0$ for $i = 1, \ldots, m$ and hence $\langle [k, x], y\rangle = 0$ for all $k \in \mathfrak{k}$. Conversely we can choose $k_j = [x_j, y_j]/\lambda_j \in \mathfrak{k}$ for $j = 1, \ldots, m$. Then

$$0 = \langle [k_j, x], y\rangle = \sum_{i=1}^m B([x_j, y_j], [x_i, y_i]) = B([x_j, y_j], [x_j, y_j])$$

implies that $[x_j, y_j] = 0$ by Lemma 1.A.14. Hence $[x, y] = 0$. □

The following example shows that the presence of a Euclidean part causes problems in the preceding lemma.

Example 1.A.17. *This is [Hel78, Example (c) p. 230]. Let $\mathfrak{p} \neq \{0\}$ be any real vector space and let $\mathfrak{k}$ be the Lie algebra of any compact subgroup of $\mathrm{GL}(\mathfrak{p})$. For $x, y \in \mathfrak{p}$ and $k \in \mathfrak{k}$ we set $[x, y] = 0$ and $[k, x] = -[x, k] = k \cdot x$. Then $\mathfrak{k} \oplus \mathfrak{p}$ defines an effective, orthogonal, symmetric Lie algebra of Euclidean type. This violates Lemma 1.A.16 since there is always some $x \in \mathfrak{p}$ and $k \in \mathfrak{k}$ such that $[k, x] \neq 0$.*

It turns out that maximal Abelian subspaces $\mathfrak{a} \subseteq \mathfrak{p}$ play an important role in understanding the structure of symmetric Lie algebras. A useful notation for $x \in \mathfrak{p}$ is

$$\mathfrak{p}_x = \{y \in \mathfrak{p} : [x, y] = 0\},$$

which denotes the *centralizer* (or *commutant*) of x in $\mathfrak{p}$. Clearly, for $x \in \mathfrak{a}$ it holds that $\mathfrak{a} \subseteq \mathfrak{p}_x$. If equality holds, then we call x *regular*. Hence, a regular element is contained in a unique maximal Abelian subspace.

Lemma 1.A.18. *Let $(\mathfrak{g}, s)$ be an effective, orthogonal, symmetric Lie algebra and let $\mathfrak{a} \subseteq \mathfrak{p}$ be a maximal Abelian subspace. Then there exists $x \in \mathfrak{a}$ such that $\mathfrak{p}_x = \mathfrak{a}$.*

Proof. First we show that we can reduce the problem to $(\mathfrak{g}, s)$ being of compact type. Let $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$ be the decomposition of Result 1.A.10. Then we can write $\mathfrak{a} = \mathfrak{a}_0 \oplus \mathfrak{a}_- \oplus \mathfrak{a}_+$ where $\mathfrak{a}_0 = \mathfrak{p}_0$ and $\mathfrak{a}_\pm$ is a maximal Abelian subspace of $\mathfrak{p}_\pm$. If $x, y \in \mathfrak{p}$ then $[x, y] = [x_0, y_0] + [x_-, y_-] + [x_+, y_+]$ and hence x and y commute if and only if $[x_i, y_i] = 0$ for all $i \in \{0, -, +\}$. Assume that we have found $x_i \in \mathfrak{p}_i$ satisfying $(\mathfrak{p}_i)_{x_i} = \mathfrak{a}_i$ for all i. Then setting $x = \sum_i x_i$ we get that $\mathfrak{p}_x = \mathfrak{a}$. Hence it suffices to consider the Euclidean, compact and non-compact types separately. However the Euclidean case is trivial, since $\mathfrak{a}_0 = \mathfrak{p}_0$ so any element of $\mathfrak{p}_0$ does the job. Using the duality described in Lemma 1.A.9 it suffices to consider the compact case, since $x, y \in \mathfrak{p}$ commute if and only if $\mathrm{i}x$ and $\mathrm{i}y$ commute in the dual. Hence for the remainder of the proof we assume that $(\mathfrak{g}, s)$ is of compact type.

Let G be a compact Lie group with Lie algebra $\mathfrak{g}$ and consider the torus $A = \exp(\mathfrak{a})$. Let $x \in \mathfrak{a}$ be the generator of a dense winding of A. If $y \in \mathfrak{p}$ satisfies $[x, y] = 0$, then $\exp(ty)$ commutes with all elements of A for all $t \in \mathbb{R}$. This implies that $[y, \mathfrak{a}] = 0$, and since $\mathfrak{a}$ is maximal Abelian, $y \in \mathfrak{a}$. Thus $\mathfrak{p}_x = \mathfrak{a}$, as desired. □

We have previously used the Lie group action of $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ on $\mathfrak{g}$. Let us make this a bit more precise.[18] By definition, $\mathrm{Int}(\mathfrak{g})$ is the connected Lie subgroup of $\mathrm{GL}(\mathfrak{g})$ with Lie algebra $\mathrm{ad}_{\mathfrak{g}}$. In fact this is the group of inner automorphisms of $\mathfrak{g}$. Furthermore $\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is the connected Lie subgroup of $\mathrm{Int}(\mathfrak{g})$ with Lie algebra $\mathrm{ad}_{\mathfrak{k}}$. We slightly generalize this idea by defining associated pairs following [Hel78, Def. p. 213].

Definition 1.A.19. *Consider a symmetric Lie algebra $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Let $\mathbf{G}$ be a connected Lie group with a Lie algebra isomorphism[19] $\phi : \mathfrak{g} \to \mathrm{Lie}(\mathbf{G})$ and let $\mathbf{K} \subseteq \mathbf{G}$ be a Lie subgroup with $\phi(\mathfrak{k}) = \mathrm{Lie}(\mathbf{K})$. Then we say that $(\mathbf{G}, \mathbf{K})$ is a* pair associated to $\mathfrak{g}$.

Lemma 1.A.20. *For any semisimple, symmetric Lie algebra $(\mathfrak{g}, s)$ there exists an associated pair $(\mathbf{G}, \mathbf{K})$ with $\mathbf{G}$ and $\mathbf{K}$ connected. In fact we can choose $\mathbf{G} = \mathrm{Int}(\mathfrak{g})$ and $\mathbf{K} = \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$, in which case the Lie algebra isomorphism is $\phi = \mathrm{ad}$.*

Proof. This follows from the definitions, and the fact that $\mathfrak{g}$ is semisimple, and thus $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a Lie algebra isomorphism. □

We call the pair described in 1.A.20 the *canonical pair associated to* $(\mathfrak{g}, s)$.

Lemma 1.A.21. *Let $\phi : \mathbf{G} \to \mathbf{H}$ be a Lie group homomorphism, and let $\tilde{\mathbf{G}} \subseteq \mathbf{G}$ and $\tilde{\mathbf{H}} \subseteq \mathbf{H}$ be Lie subgroups. Assume that $\phi(\tilde{\mathbf{G}}) \subseteq \tilde{\mathbf{H}}$. Then the restriction $\phi \circ \iota : \tilde{\mathbf{G}} \to \tilde{\mathbf{H}}$ is also a Lie group homomorphism, where $\iota : \tilde{\mathbf{G}} \to \mathbf{G}$ is the inclusion.*

Proof. Since ι is a smooth immersion, $\phi \circ \iota$ is smooth as a map from $\tilde{\mathbf{G}}$ to $\mathbf{H}$. The fact that it is still smooth as a map from $\tilde{\mathbf{G}}$ to $\tilde{\mathbf{H}}$ holds because all Lie subgroups are weakly embedded, see [Lee13, Thm. 19.25]. □

Lemma 1.A.22. *Consider any symmetric Lie algebra $(\mathfrak{g}, s)$ and any associated pair $(\mathbf{G}, \mathbf{K})$ with $\mathbf{G}$ and $\mathbf{K}$ connected and corresponding Lie algebra isomorphism $\phi : \mathfrak{g} \to \mathrm{Lie}(\mathbf{G})$. Then the maps $\psi : \mathbf{G} \to \mathrm{Int}(\mathfrak{g}), g \mapsto \phi^{-1} \circ \mathrm{Ad}_g \circ \phi$ and $\psi|_{\mathbf{K}} : \mathbf{K} \to \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ are surjective Lie group homomorphisms.*

Proof. The maps are well-defined and surjective due to the connectedness of $\mathbf{G}$ and $\mathbf{K}$. By [Hel78, p. 127] the maps $\mathbf{G} \to \mathrm{Int}(\mathrm{Lie}(\mathbf{G}))$ and $\mathrm{Int}(\mathrm{Lie}(\mathbf{G})) \to \mathrm{Int}(\mathfrak{g})$ are Lie group homomorphisms, and hence also ψ. By 1.A.21 then also $\psi|_{\mathbf{K}} : \mathbf{K} \to \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is a Lie group homomorphism. □

In the following we suppress the Lie algebra isomorphism ϕ from the notation. Then ψ becomes Ad, the adjoint representation. Note that for the canonical associated pair, the adjoint representation is essentially the identity map.

Lemma 1.A.23. *Let $(\mathfrak{g}, s)$ be a semisimple, symmetric Lie algebra and let $(\mathbf{G}, \mathbf{K})$ be an associated pair with $\mathbf{G}$ and $\mathbf{K}$ connected. Then the adjoint representation of $\mathbf{K}$ on $\mathfrak{g}$, written $\mathrm{Ad} : \mathbf{K} \to \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is a covering homomorphism.*

Proof. By Lemma 1.A.22, $\mathrm{Ad} : \mathbf{K} \to \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ is a surjective Lie group homomorphism, and by semisimplicity it induces a Lie algebra isomorphism. Hence it is a covering map. □

In the next lemma we collect some equivariance properties of the adjoint action of $\mathbf{K}$, but first we introduce some notation. By Lemma 1.A.16 there is, for each $x \in \mathfrak{p}$, an orthogonal decomposition of $\mathfrak{p}$ into kernel and image of ad_x. By $\Pi_x : \mathfrak{p} \to \mathfrak{p}$ we denote the orthogonal projection onto $\mathfrak{p}_x$, and by

[18] For more details we refer to [Hel78, Ch. II §5].

[19] Here $\mathrm{Lie}(\mathbf{G})$ denotes the Lie algebra of $\mathbf{G}$.

$\Pi_x^\perp = \mathbb{1} - \Pi_x$ the complementary projection. Moreover we can define an inverse of ad_x by restricting the domain and codomain. Then we get a well-defined inverse $\mathrm{ad}_x^{-1} : \mathfrak{p}_x^\perp \to \mathfrak{k}_x^\perp$. This is nothing but a restriction of the Moore-Penrose pseudo-inverse.

Lemma 1.A.24 (Equivariance properties). *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra with associated pair $(\mathbf{G}, \mathbf{K})$ with $\mathbf{K}$ connected. For $K \in \mathbf{K}$ and $x \in \mathfrak{p}$, it holds that:*

(i) $\mathrm{Ad}_K \circ \mathrm{ad}_x = \mathrm{ad}_{\mathrm{Ad}_K(x)} \circ \mathrm{Ad}_K$ *on* $\mathfrak{p}$*;*

(ii) $\mathrm{Ad}_K \circ \mathrm{ad}_x^{-1} = \mathrm{ad}_{\mathrm{Ad}_K(x)}^{-1} \circ \mathrm{Ad}_K$ *on* $\mathfrak{p} \cap \mathrm{im}\,\mathrm{ad}_x$*;*

(iii) $\mathrm{Ad}_K \circ \Pi_x = \Pi_{\mathrm{Ad}_K(x)} \circ \mathrm{Ad}_K$ *on* $\mathfrak{p}$*;*

(iv) $\mathrm{Ad}_K \circ \Pi_x^\perp = \Pi_{\mathrm{Ad}_K(x)}^\perp \circ \mathrm{Ad}_K$ *on* $\mathfrak{p}$.

Proof. (i): Clear since $\mathrm{Ad}_K([g,h]) = [\mathrm{Ad}_K(g), \mathrm{Ad}_K(h)]$ for all $g, h \in \mathfrak{g}$. Moreover this implies $\mathrm{Ad}_K(\mathfrak{p}_x) = \mathfrak{p}_{\mathrm{Ad}_K(x)}$ and using orthogonality $\mathrm{Ad}_K(\mathfrak{k}_x^\perp) = \mathfrak{k}_{\mathrm{Ad}_K(x)}^\perp$. (ii): Consider $y = \mathrm{ad}_x(k)$ for some $k \in \mathfrak{k}_x^\perp$, then clearly $\mathrm{Ad}_K(\mathrm{ad}_x^{-1}(y)) = \mathrm{Ad}_K(k)$. By the previous point $\mathrm{Ad}_K(y) = \mathrm{ad}_{\mathrm{Ad}_K(x)}(\mathrm{Ad}_K(k))$ and $\mathrm{Ad}_K(k) \in \mathfrak{k}_{\mathrm{Ad}_K(x)}^\perp$ and hence $\mathrm{ad}_{\mathrm{Ad}_K(x)}^{-1}(\mathrm{Ad}_K(y)) = \mathrm{Ad}_K(k)$. (iii): Let $z \in \mathfrak{p}$ and consider the orthogonal decomposition $z = [k, x] + y$ where $k \in \mathfrak{k}_x^\perp$ and $y \in \mathfrak{p}_x$. Then, for any $K \in \mathbf{K}$, it holds that $\mathrm{Ad}_K \circ \Pi_x(z) = \mathrm{Ad}_K(y)$. On the other hand, using the first point again, $\Pi_{\mathrm{Ad}_K(x)} \circ \mathrm{Ad}_K(z) = \Pi_{\mathrm{Ad}_K(x)}([\mathrm{Ad}_K(k), \mathrm{Ad}_K(x)] + \mathrm{Ad}_K(y)) = \mathrm{Ad}_K(y)$. (iv): Analogous to (iii). □

Lemma 1.A.25. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra with associated pair $(\mathbf{G}, \mathbf{K})$. Let $x \in \mathfrak{p}$, and let $O = \mathrm{Ad}_{\mathbf{K}}(x)$ denote the $\mathbf{K}$-orbit containing x. Then, using the canonical identification $T_x\mathfrak{p} \cong \mathfrak{p}$ it holds that $\mathrm{ad}_{\mathfrak{k}}(x)$ is the tangent space T_xO and $\mathfrak{p}_x$ is its orthogonal complement.*

Proof. This follows immediately from Lemma 1.A.16. □

Lemma 1.A.26. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra with associated pair $(\mathbf{G}, \mathbf{K})$. Let $\mathfrak{a} \subseteq \mathfrak{p}$ be some maximal Abelian subspace. Then every $\mathbf{K}$-orbit in $\mathfrak{p}$ intersects $\mathfrak{a}$ and all maximal Abelian subspaces of $\mathfrak{p}$ are conjugate by some element in $\mathbf{K}$.*

Proof. Without loss of generality we assume that we are using the canonical associated pair $(\mathbf{G}, \mathbf{K}) = (\mathrm{Int}(\mathfrak{g}), \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g}))$. Then Lemma 1.A.22 guarantees the existence of the desired group element in any pair.

Let $x \in \mathfrak{a}$ be regular, which exists by Lemma 1.A.18, and let $y \in \mathfrak{p}$. Consider the smooth function

$$f : \mathbf{K} \to \mathbb{R}, \quad K \mapsto B(Ky, x).$$

Let $K \in \mathbf{K}$ be a critical point of f, which exists since $\mathbf{K}$ is compact. Then for any $k \in \mathfrak{k}$ it holds that

$$0 = \frac{d}{dt}\Big|_{t=0} (f(e^{tk}K)) = B([k, Ky], x) = B(k, [Ky, x]).$$

Since this holds for all $k \in \mathfrak{k}$, Lemma 1.A.14 shows that $[Ky, x] = 0$ and since x is regular, $Ky \in \mathfrak{a}$. This shows that every $\mathbf{K}$-orbit in $\mathfrak{p}$ intersects $\mathfrak{a}$.

For any $K \in \mathbf{K}$ it holds that $K \cdot \mathfrak{a}$ is maximal Abelian with regular element Kx by Lemma 1.A.24 (i). If $\mathfrak{a}'$ is any other maximal Abelian subspace, we can choose K such that $Kx \in \mathfrak{a}'$, and thus $\mathfrak{a}' = \mathfrak{p}_{Kx} = K \cdot \mathfrak{p}_x = K \cdot \mathfrak{a}$. Hence all maximally Abelian subspaces are conjugate by some element in $\mathbf{K}$. □

Definition 1.A.27. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra with associated pair $(\mathbf{G}, \mathbf{K})$, and let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal Abelian subspace. Then we define the* normalizer of $\mathfrak{a}$ in $\mathbf{K}$

$$N_{\mathbf{K}}(\mathfrak{a}) = \{K \in \mathbf{K} : \mathrm{Ad}_K(\mathfrak{a}) = \mathfrak{a}\},$$

and the centralizer of $\mathfrak{a}$ in $\mathbf{K}$

$$Z_{\mathbf{K}}(\mathfrak{a}) = \{K \in \mathbf{K} : \mathrm{Ad}_K(x) = x \text{ for all } x \in \mathfrak{a}\}.$$

Since $Z_{\mathbf{K}}(\mathfrak{a})$ is clearly normal in $N_{\mathbf{K}}(\mathfrak{a})$ we can define the Weyl group $\mathbf{W} = \mathbf{W}_{\mathfrak{a}} = N_{\mathbf{K}}(\mathfrak{a})/Z_{\mathbf{K}}(\mathfrak{a})$ *which acts on $\mathfrak{a}$.*

Note that since all maximal Abelian subspaces are conjugate by Lemma 1.A.26, different choices of $\mathfrak{a}$ lead to isomorphic Weyl groups. We will see later that the Weyl group also does not depend on the choice of associated pair $(\mathbf{G}, \mathbf{K})$.

Root Space Decomposition

In order to understand the Weyl group action on $\mathfrak{a}$ defined above we need to understand the root space decomposition of a semisimple, orthogonal, symmetric Lie algebra. This is the goal of the present section. By $\mathfrak{g}_{\mathbb{C}}$ we denote the complexification of $\mathfrak{g}$.

Definition 1.A.28. *Let a symmetric Lie algebra $(\mathfrak{g}, s)$ and a maximal Abelian subspace $\mathfrak{a} \subseteq \mathfrak{p}$ be given. For any linear functional $\alpha \in \mathrm{Hom}_{\mathbb{R}}(\mathfrak{a}, \mathbb{C})$ we define*

$$\mathfrak{g}_{\mathbb{C}}^{\alpha} = \{x \in \mathfrak{g}_{\mathbb{C}} : \mathrm{ad}_y(x) = \alpha(y)x \text{ for all } y \in \mathfrak{a}\}.$$

If $\mathfrak{g}_{\mathbb{C}}^{\alpha}$ is non-trivial we call α a root *and $\mathfrak{g}_{\mathbb{C}}^{\alpha}$ the corresponding* root space. *The non-zero elements of a root space are called* root vectors. *We denote by Δ the set of all non-zero roots and by Δ_0 the set of all roots including zero.*[20]

Lemma 1.A.29. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Then*

$$\mathfrak{g}_{\mathbb{C}} = \bigoplus_{\alpha \in \Delta_0} \mathfrak{g}_{\mathbb{C}}^{\alpha},$$

and every non-zero root is either imaginary or real and supported on the compact or non-compact part of $\mathfrak{a}$ respectively. We write Δ_- for the compact roots and Δ_+ for the non-compact ones.

Proof. We use the inner product of Lemma 1.A.15. By Corollary 1.A.13 we have that $\mathfrak{g} = \mathfrak{k}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$, and each ad_x for $x = x_- + x_+ \in \mathfrak{p}$ preserves this decomposition. It acts trivially on $\mathfrak{k}_0$ and as $\mathrm{ad}_{x_\pm}$ on $\mathfrak{g}_\pm$. Furthermore ad_{x_-} is skew-symmetric and ad_{x_+} is symmetric[21] since

$$\langle \mathrm{ad}_x(y), k\rangle = -B([x,y],k) = B(y,[x,k]) = \pm\langle y, \mathrm{ad}_x(k)\rangle,$$

for $x, y \in \mathfrak{p}_\pm$ and $k \in \mathfrak{k}_\pm$. If $x, y \in \mathfrak{p}$ commute, then also ad_x and ad_y since, by the Jacobi identity,

$$\mathrm{ad}_x \circ \mathrm{ad}_y(z) = [x,[y,z]] = [[x,y],z] + [y,[x,z]] = \mathrm{ad}_y \circ \mathrm{ad}_x(z).$$

[20] Some authors don't consider 0 a root at all.

[21] These (skew-)symmetric operators on $\mathfrak{g}$ are of course still real and (skew-)symmetric on $\mathfrak{g}_{\mathbb{C}}$, and hence (skew-)Hermitian. In particular the eigenvalues are imaginary in the compact case and real in the non-compact case.

Hence all ad_x for $x \in \mathfrak{a}$ can be simultaneously (unitarily) diagonalized and we obtain a complete root space decomposition of $\mathfrak{g}_{\mathbb{C}}$. Now every root vector of $\mathfrak{g}_+$ relative to $\mathfrak{a}_+$ is a root vector of $\mathfrak{g}$ and similarly for $\mathfrak{g}_-$. Conversely, let α be a non-zero root and $x \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$. Then for $y \in \mathfrak{a}$ we have that

$$\mathrm{ad}_y(x) = [y_-, x_-] + [y_+, x_+] = \alpha(y)x.$$

If $x_- \neq 0$ and $x_+ \neq 0$ then $[y_-, x_-] = \alpha(y)x_-$ and $[y_+, x_+] = \alpha(y)x_+$ and so $\alpha(y)$ would have to be purely imaginary and real, leading to a contradiction. This shows that either α only takes imaginary values, and its root space lies in the compact part, or it only takes real values, and its root space lies in the non-compact part □

Corollary 1.A.30. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Then, in some basis, the Lie algebra $\mathrm{ad}_{\mathfrak{g}} = \{\mathrm{ad}_z : z \in \mathfrak{g}\}$ is a Lie algebra of real matrices closed under transposition on the vector space $\mathfrak{g}$ with inner product as in Lemma 1.A.15.*

Proof. By construction, all ad_k for $k \in \mathfrak{k}$ are skew-symmetric in the given inner product. The proof of Lemma 1.A.29 shows that ad_x is skew-symmetric or symmetric for $x \in \mathfrak{p}_-$ or $x \in \mathfrak{p}_+$ respectively. Hence for any $z \in \mathfrak{g}$, the transposed of ad_z with respect to the given inner product is also in $\mathrm{ad}_{\mathfrak{g}}$. Now any orthonormal basis on $\mathfrak{g}$ will do. □

We now have two involutions on $\mathfrak{g}_{\mathbb{C}}$, namely (the complexification of) s and complex conjugation. The next two lemmas show how they act on the root spaces.

Lemma 1.A.31. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Then $s(\mathfrak{g}_{\mathbb{C}}^{\alpha}) = \mathfrak{g}_{\mathbb{C}}^{-\alpha}$.*

Proof. Let $x \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$ then for all $y \in \mathfrak{a}$ it holds that $[y, s(x)] = s([s(y), x]) = s([-y, x]) = -s([y, x]) = -\alpha(y)s(x)$. Hence $s(x) \in \mathfrak{g}_{\mathbb{C}}^{-\alpha}$. This shows that $s(\mathfrak{g}_{\mathbb{C}}^{\alpha}) \subseteq \mathfrak{g}_{\mathbb{C}}^{-\alpha}$. But then $\mathfrak{g}_{\mathbb{C}}^{-\alpha} = s(s(\mathfrak{g}_{\mathbb{C}}^{-\alpha})) \subseteq s(\mathfrak{g}_{\mathbb{C}}^{\alpha})$, as desired. □

This shows that $\mathfrak{g}^{\alpha} = (\mathfrak{g}_{\mathbb{C}}^{\alpha} \oplus \mathfrak{g}_{\mathbb{C}}^{-\alpha}) \cap \mathfrak{g}$ is invariant under s and hence decomposes as $\mathfrak{g}^{\alpha} = \mathfrak{k}^{\alpha} \oplus \mathfrak{p}^{\alpha}$ where $\mathfrak{k}^{\alpha} = \mathfrak{g}^{\alpha} \cap \mathfrak{k}$ and $\mathfrak{p}^{\alpha} = \mathfrak{g}^{\alpha} \cap \mathfrak{p}$.

Lemma 1.A.32. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra and let $\alpha \in \Delta_{\pm}$ be a non-zero root. Then $\overline{\mathfrak{g}_{\mathbb{C}}^{\alpha}} = \mathfrak{g}_{\mathbb{C}}^{\pm\alpha}$.*

Proof. Let $x \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$ then for all $y \in \mathfrak{a}$ it holds that $[y, \overline{x}] = \overline{[\overline{y}, x]} = \overline{[y, x]} = \overline{\alpha(y)x} = \pm\alpha(y)\overline{x}$. □

Using Lemmas 1.A.31 and 1.A.32 we can find for each root a corresponding root vector which is composed of an element of $\mathfrak{k}^{\alpha}$ and an element of $\mathfrak{p}^{\alpha}$.

Lemma 1.A.33. *Let $\alpha \in \Delta_{\pm}$ be a non-zero root. Then there exist $x \in \mathfrak{k}^{\alpha}$ and $y \in \mathfrak{p}^{\alpha}$ such that $x + \sqrt{\pm 1}y \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$.[22] We call such x and y related. For any $a \in \mathfrak{a}$ they satisfy*

$$[a, x] = \sqrt{\pm 1}\alpha(a)y, \quad [a, y] = \pm\sqrt{\pm 1}\alpha(a)x.$$

Proof. First consider $\alpha \in \Delta_+$. Since $\mathfrak{g}_{\mathbb{C}}^{\alpha}$ is invariant under complex conjugation, it contains a real root vector. More explicitly, if $z \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$, then $z + \overline{z} \in \mathfrak{g}^{\alpha}$. Hence there are $x \in \mathfrak{k}^{\alpha}$ and $y \in \mathfrak{p}^{\alpha}$ such that $z + \overline{z} = x + y$. Now consider $\alpha \in \Delta_-$ and let $u \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$. We know that $s(\overline{u}) \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$. Let $v, w \in \mathfrak{g}$ such that $u = v + \mathrm{i}w$, then $u + s(\overline{u}) = v + \mathrm{i}w + s(v) - \mathrm{i}s(w)$. Let $z = (u + s(\overline{u}))/2$ then $z \in \mathfrak{g}_{\mathbb{C}}^{\alpha}$ and $z = x + \mathrm{i}y$ where $x = (v + s(v))/2 \in \mathfrak{k}$ and $y = w - s(w) \in \mathfrak{p}$. Moreover $x = (z + \overline{z})/2 \in \mathfrak{g}^{\alpha}$ and $y = \mathrm{i}(-z + \overline{z})/2 \in \mathfrak{g}^{\alpha}$. In both cases, the claimed equations follow from the fact that $[a, x + \sqrt{\pm 1}y] = \alpha(a)(x + \sqrt{\pm 1}y)$ by equating the $\mathfrak{k}_{\mathbb{C}}$ and $\mathfrak{p}_{\mathbb{C}}$ parts. This concludes the proof. □

[22] Here we use $\sqrt{\pm 1}$ as a shorthand for 1 or i depending on the sign.

Lemma 1.A.34. *Let $\alpha \in \Delta_\pm$ be a non-zero root and let $x \in \mathfrak{k}^\alpha$ and $y \in \mathfrak{p}^\alpha$ be related. Then $[x,y] \in \mathfrak{a}$ and if $\|y\| = 1$, then $\langle [x,y], a\rangle = \sqrt{\pm 1}\alpha(a)$, and hence $[x,y]$ only depends on α. In particular $\|[x,y]\|^2 = \sqrt{\pm 1}\alpha([x,y])$.*

Proof. First it is clear that $[x,y] \in \mathfrak{p}$. Let $a \in \mathfrak{a}$ be arbitrary, then $[a,[x,y]] = [[a,x],y] + [x,[a,y]] = 0$ by Lemma 1.A.33. In particular, if a is regular this shows that $[x,y] \in \mathfrak{a}$. Now we find

$$\langle [x,y], a\rangle = \pm B([x,y],a) = \pm B(y,[a,x]) = \pm\sqrt{\pm 1}\alpha(a)B(y,y) = \sqrt{\pm 1}\alpha(a),$$

as desired. □

Lemma 1.A.35. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra. Let $\alpha \in \Delta_\pm$ be a non-zero root. Let $x \in \mathfrak{k}^\alpha$ and $y \in \mathfrak{p}^\alpha$ be related unit vectors. Then, if $z = [x,y]$, it holds that*

$$\mathrm{ad}_x^{2n-1}(z) = (-\|z\|^2)^n y, \quad \mathrm{ad}_x^{2n}(z) = (-\|z\|^2)^n z.$$

Proof. The case $n = 1$ follows immediately from Lemma 1.A.33 and Lemma 1.A.34. Then induction yields the result. □

Note that the resulting expressions are the same for the compact and non-compact case.

So far we have made no reference to an associated pair. In the next lemma we use an associated pair to relate the Weyl group action with the roots, but we note that the choice is arbitrary.

Lemma 1.A.36. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra, and let $(\mathbf{G}, \mathbf{K})$ be an associated pair. Let $\alpha \in \Delta$ be a non-zero root. Then there exists $K \in N_{\mathbf{K}}(\mathfrak{a}) \cap \mathbf{K}_0$ such that K acts on $\mathfrak{a}$ as orthogonal reflection with respect to the kernel of α.*

Proof. Let $x \in \mathfrak{k}^\alpha$ and $y \in \mathfrak{p}^\alpha$ be related unit vectors and let $z = [x,y] \in \mathfrak{a}$. Then using Lemma 1.A.35 we compute

$$\mathrm{Ad}_{e^{tx}}(z) = e^{\mathrm{ad}_{tx}}(z) = \sum_{n\geq 0} \frac{\mathrm{ad}_{tx}^{2n}}{(2n)!} z + \sum_{n\geq 1} \frac{\mathrm{ad}_{tx}^{2n-1}}{(2n-1)!} y = \cos(t\|z\|)z - \|z\|\sin(t\|z\|)y$$

Setting $t = \pi/\|z\|$ we get $\mathrm{Ad}_{e^{tx}} z = -z$ and for $u \in \mathfrak{a}$ satisfying $\langle z, u\rangle = 0$ it holds that $[x,u] = -\sqrt{\pm 1}\alpha(u)y = -\langle z,u\rangle\, y = 0$. Hence $\mathrm{Ad}_{e^{tx}} u = u$. This concludes the proof. □

In fact these reflections generate the entire Weyl group. First we need an alternative characterization of regular elements.

Lemma 1.A.37. *The commutant of an element $x \in \mathfrak{a}$ in $\mathfrak{g}_\mathbb{C}$ is given by*

$$(\mathfrak{g}_\mathbb{C})_x = \sum_{\substack{\alpha\in\Delta_0 \\ \alpha(x)=0}} \mathfrak{g}_\mathbb{C}^\alpha.$$

In particular, x is regular if and only if $\alpha(x) \neq 0$ for all $\alpha \in \Delta$.

Proof. First note that 0 is a root and $\mathfrak{p}^0 = \mathfrak{a}$. Let $y \in \mathfrak{g}_\mathbb{C}$. Then due to the rootspace decomposition $y = \sum_{\alpha\in\Delta_0} y_\alpha$ it holds that

$$[x,y] = \sum_{\alpha\in\Delta} \alpha(x) y_\alpha.$$

This proves the form of the commutant $(\mathfrak{g}_{\mathbb{C}})_x$. In particular

$$\mathfrak{p}_x = \bigoplus_{\substack{\alpha\in\Delta_0' \\ \alpha(x)=0}} \mathfrak{p}^\alpha,$$

where Δ_0' contains 0 and exactly one root of each pair α and $-\alpha$. Hence $\mathfrak{p}_x = \mathfrak{a} = \mathfrak{p}^0$ if and only if $\alpha(x) \neq 0$ for all $\alpha \in \Delta$. □

The kernels of the non-zero roots define root hyperplanes in $\mathfrak{a}$. The complement of the union of all root hyperplanes is a disjoint union of open connected components, called *(open) Weyl chambers* (we denote the closure of a Weyl chamber by $\mathfrak{w}$). Lemma 1.A.37 then shows that the regular elements in $\mathfrak{a}$ are exactly the ones that lie in an open Weyl chamber.

Lemma 1.A.38. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra, and let $(\mathbf{G}, \mathbf{K})$ be an associated pair. The Weyl group is finite and generated by the orthogonal reflections about the root hyperplanes. Furthermore it acts simply transitively on the Weyl chambers.*

Proof. Let $\mathfrak{n}$ be the Lie algebra of $N_{\mathbf{K}}(\mathfrak{a})$. Then for any $k \in \mathfrak{n}$ and $x \in \mathfrak{a}$ it holds that $[k, x] \in \mathfrak{a}$. By Lemma 1.A.16 it holds that $\mathrm{ad}_{\mathfrak{k}}(x) \subset \mathfrak{a}^\perp$, and thus $[k, x] = 0$. Hence $N_{\mathbf{K}}(\mathfrak{a})$ and $Z_{\mathbf{K}}(\mathfrak{a})$ have the same Lie algebra and the Weyl group must be discrete, and by compactness it must be finite.

We will only sketch the remainder of the proof, see [Hel78, Ch. VII, Thm. 2.12] for details. The plan is to show that the subgroup of the Weyl group generated by the root reflections (recall Lemma 1.A.36) acts transitively on the Weyl chambers, and moreover the Weyl group acts simply transitively on the Weyl chambers. This then shows that the Weyl group is generated by the reflections. For transitivity, let $\mathbf{W}'$ denote the subgroup generated by the reflections, then for $x, y \in \mathfrak{a}$ we can find $w \in \mathbf{W}'$ such that $|x - w \cdot y|$ is minimal. Then x and $w \cdot y$ lie in the same Weyl chamber, since otherwise there is a reflection which reduces the distance. For simple transitivity, assume that $w \in W$ maps some Weyl chamber $\mathfrak{w}$ into itself. Then by averaging one finds a regular fixed point $x \in \mathfrak{w}$, i.e., $w \cdot x = x$. Using the duality of Lemma 1.A.9 we may assume that $(\mathfrak{g}, s)$ is of compact type and $\mathbf{G}$ is compact. Hence the closure of the one-parameter group generated by x is a torus $\mathbf{T}$. If Ad_K represents w, then K commutes with $\mathbf{T}$ and hence one can show that $K = e^k$ for some $k \in \mathfrak{k}$ and $[k, x] = 0$. Since x is regular, $[k, \mathfrak{a}] = 0$ and hence $w = \mathbb{1}$. □

Corollary 1.A.39. *Each $\mathbf{W}$-orbit intersects the closed Weyl chamber $\mathfrak{w}$ exactly once.*

Proof. By Lemma 1.A.38 each Weyl group orbit intersects $\mathfrak{w}$, and for $x \in \mathfrak{a}$ regular, this intersection is unique. Now consider $y \in \mathfrak{a}$ singular and let $x \in \mathfrak{w}$ be a regular point in some slice U about y (i.e., U has the property that, for all $w \in \mathbf{W}$, if $w \in \mathbf{W}_y$, then $wU = U$, otherwise $(wU) \cap U = \emptyset$, see Def. 1.A.50 for the general definition). Now assume that $y, wy \in \mathfrak{w}$ for some $w \in \mathbf{W}$ and consider wx which lies in wU. One can show that there is an element in $w' \in \mathbf{W}_{wy}$ such that $w'wx \in \mathfrak{w}$. By uniqueness, $w'wx = x$ and $w' = w^{-1}$ and hence $w \in \mathbf{W}_y$. Thus $wy = y$ as desired. □

Subalgebras and Quotients

In this section we will look at some further properties of symmetric Lie algebras. In particular, we are interested in subalgebras that appear as commutants and related quotient spaces. We start with a useful property of direct sums of symmetric Lie algebras.

Lemma 1.A.40. *Let $\mathfrak{g}_i = \mathfrak{k}_i \oplus \mathfrak{p}_i$ be symmetric Lie algebras for $i = 1, 2$. Then $\mathfrak{g}_1 \oplus \mathfrak{g}_2 = (\mathfrak{k}_1 \oplus \mathfrak{k}_2) \oplus (\mathfrak{p}_1 \oplus \mathfrak{p}_2)$ is a symmetric Lie algebra. Moreover, $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ is orthogonal if and only if both $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are orthogonal.*

Proof. See [Hel78, Ch. V, Lem. 1.6]. □

Next we consider symmetric subalgebras of a symmetric Lie algebra $(\mathfrak{g}, s)$, and we will see how certain properties are inherited. In general we say that a subset of $\mathfrak{g}$ is *symmetric*, if it is left invariant by s.

Definition 1.A.41. *Let $(\mathfrak{g}, s)$ be a symmetric Lie algebra and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a Lie subalgebra invariant under s. Then we say that $\mathfrak{h}$, or more precisely $(\mathfrak{h}, s|_{\mathfrak{h}})$, is a* symmetric Lie subalgebra of $(\mathfrak{g}, s)$.

The Cartan-like decomposition of the symmetric Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is given by $\mathfrak{h} = (\mathfrak{h} \cap \mathfrak{k}) \oplus (\mathfrak{h} \cap \mathfrak{p})$. The following results show that orthogonality is inherited, but semisimplicity has to be replaced by the slightly weaker condition of reductivity[23]. Let us first recall a basic fact about reductive Lie algebras.

Lemma 1.A.42. *Let $\mathfrak{g}$ be a reductive Lie algebra, then $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \oplus \mathfrak{z}$ where $\mathfrak{z}$ is the center of $\mathfrak{g}$ and $[\mathfrak{g}, \mathfrak{g}]$ is semisimple, and this decomposition of $\mathfrak{g}$ into a direct sum of a semisimple and an Abelian subalgebra is unique.*

Proof. Let $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{a}$ where $\mathfrak{s}$ and $\mathfrak{a}$ are semisimple and Abelian ideals in $\mathfrak{g}$. First we note that $\mathfrak{a}$ must equal the center of $\mathfrak{g}$. Indeed, it is clear that $\mathfrak{a} \subseteq \mathfrak{z}$. Conversely, let $s + a \in \mathfrak{z}$ where $s \in \mathfrak{s}$ and $a \in \mathfrak{a}$, then $0 = [s + a, \mathfrak{s}] = [s, \mathfrak{s}]$ and so $s = 0$. Finally $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{s}, \mathfrak{s}] = \mathfrak{s}$ since $\mathfrak{s}$ is semisimple. □

Lemma 1.A.43. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a symmetric Lie subalgebra. Moreover assume that $\mathfrak{h} \cap \mathfrak{p} = (\mathfrak{h} \cap \mathfrak{p}_-) \oplus (\mathfrak{h} \cap \mathfrak{p}_+)$, where $\mathfrak{p}_\pm$ denote the non-compact and compact parts. Then $(\mathfrak{h}, s|_{\mathfrak{h}})$ is a reductive, orthogonal, symmetric Lie algebra. Moreover setting $\mathfrak{h}' = [\mathfrak{h}, \mathfrak{h}]$, it holds that $(\mathfrak{h}', s|_{\mathfrak{h}'})$ is a semisimple, orthogonal, symmetric Lie algebra.*

Proof. This result generalizes [Kna02, Cor. 6.29]. We will work with the inner product on $\mathfrak{g}$ defined in Lemma 1.A.15. Consider the adjoint representation of $\mathfrak{h}$ on $\mathfrak{g}$. By semisimplicity this is a faithful representation, and the assumption on $\mathfrak{h} \cap \mathfrak{p}$ implies that $\mathrm{ad}_h|_{\mathfrak{g}}$ is closed under transposition, c.f. Corollary 1.A.30. By [Kna02, Prop. 1.56] a Lie algebra of real matrices closed under transposition is reductive, and so $\mathfrak{h}$ is reductive. By Lemma 1.A.42 $\mathfrak{h}'$ is semisimple. Clearly $\mathfrak{h}'$ is also invariant under s, and so by Lemma 1.A.7 it holds that $(\mathfrak{h}', s|_{\mathfrak{h}'})$ is orthogonal. Let $\mathfrak{z}$ denote the center of $\mathfrak{h}$. For $h \in \mathfrak{h}$ and $z \in \mathfrak{z}$ we see by $[s(z), h] = s([z, s(h)])$ that $\mathfrak{z}$ is a symmetric subalgebra, and it is trivially orthogonal. Hence also $\mathfrak{h} = \mathfrak{h}' \oplus \mathfrak{z}$ is orthogonal by Lemma 1.A.40. □

Centralizers (commutants) are a common source of symmetric Lie subalgebras, so we briefly give a more general definition and notation. Let $\mathfrak{g}$ be a Lie algebra and let $A, B \subseteq \mathfrak{g}$ be arbitrary subsets. Then the centralizer of A in B is defined as $B_A := \{b \in B : [b, a] = 0 \text{ for all } a \in A\}$. Of course these subsets have related stabilizer subgroups $\mathbf{G}_A$ and $\mathbf{K}_A$.

Lemma 1.A.44. *Let $(\mathfrak{g}, s)$ be a semisimple, orthogonal, symmetric Lie algebra with associated pair $(\mathbf{G}, \mathbf{K})$. Let $A \subset \mathfrak{p}$ be any subset and let $\mathfrak{g}_A$ be its commutant. Then $\mathfrak{g}_A$ is a reductive, orthogonal, symmetric Lie subalgebra with Cartan-like decomposition $\mathfrak{g}_A = \mathfrak{k}_A \oplus \mathfrak{p}_A$, and $\mathfrak{g}_A$ is invariant under the action of the stabilizer $\mathbf{G}_A$. Moreover $\mathfrak{p}_A = (\mathfrak{p}_A \cap \mathfrak{p}_-) \oplus (\mathfrak{p}_A \cap \mathfrak{p}_+)$ and for every $x \in \mathfrak{p}_A$, we get the orthogonal decomposition $\mathfrak{p}_A = \mathfrak{p}_{A,x} \oplus [\mathfrak{k}_A, x]$.*

[23] Recall that a Lie algebra (over a field of characteristic 0) is reductive if and only if it can be written as a direct sum of a semisimple and an Abelian Lie algebra.

Proof. That $\mathfrak{g}_A$ is a Lie subalgebra follows from the Jacobi identity. Let $x \in \mathfrak{g}_A$, since $[s(x), z] = -s([x, z]) = 0$ for $z \in A$, we see that $\mathfrak{g}_A$ is invariant under s. Similarly, if $G \in \mathbf{G}_A$ then $[\mathrm{Ad}_G(x), z] = \mathrm{Ad}_G([x, \mathrm{Ad}_G^{-1}(z)]) = 0$. Finally consider $x = x_- + x_+ \in \mathfrak{p}_A$. Then, for any $z \in A$ it holds that $[x_-, z] + [x_+, z] = 0$, and since $\mathfrak{k}_-$ and $\mathfrak{k}_+$ have zero intersection, this means that $[x_-, z] = [x_+, z] = 0$. By Lemma 1.A.43 this shows that $(\mathfrak{g}_A, s)$ is reductive and orthogonal. Let $x = x_- + x_+ \in \mathfrak{p}_A$ and $y = y_- + y_+ \in \mathfrak{p}_A$ and $k \in \mathfrak{k}_A$. Then, using the inner product and the Killing form on $\mathfrak{g}$, we compute $\langle [k, x], y\rangle = B(k, [x_+, y_+] - [x_-, y_-])$. Hence $\mathfrak{p}_{A,x} \subseteq \mathfrak{p}_A \cap [\mathfrak{k}_A, x]^\perp$. Since $[x_+, y_+] - [x_-, y_-] \in \mathfrak{k}_A$, the converse is true by Lemma 1.A.14. □

In order to better understand the action of $\mathbf{K}_A$ on $\mathfrak{g}$ we give the following powerful generalization of Lemma 1.A.26.

Lemma 1.A.45. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra, and let $(\mathbf{G}, \mathbf{K})$ be an associated pair. Let $\mathfrak{a} \subseteq \mathfrak{p}$ be some maximal Abelian subspace. Let $A \subset \mathfrak{a}$ be any subset. For any $y \in \mathfrak{p}_A$, there is some $K \in \mathbf{K}_A$ such that $\mathrm{Ad}_K(y) \in \mathfrak{a}$.*

Proof. As in the proof of Lemma 1.A.26 we may assume that we are working with the canonical pair $(\mathbf{G}, \mathbf{K}) = (\mathrm{Int}(\mathfrak{g}), \mathrm{Int}_{\mathfrak{k}}(\mathfrak{g}))$, and in particular $\mathbf{K}$ is compact. Now let $x \in \mathfrak{a}$ be regular, and let $y \in \mathfrak{p}_A$. Consider the smooth function

$$f : \mathbf{K}_A \to \mathbb{R}, \quad f(K) = B(Ky, x)$$

If K is a critical point for f, which exists since $\mathbf{K}_A$ is compact, then for every $k \in \mathfrak{k}_A$ it holds that

$$0 = B([k, Ky], x) = B(k, [Ky, x])$$

and since $[Ky, x] \in \mathfrak{k}_A$ this means that $[Ky, x] = 0$ by Lemma 1.A.14. Since x is regular, $Ky \in \mathfrak{a}$. □

Lemma 1.A.46. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra, and let $(\mathbf{G}, \mathbf{K})$ be an associated pair. Let $\mathfrak{a} \subseteq \mathfrak{p}$ be some maximal Abelian subspace. Let $A \subset \mathfrak{a}$ be any subset and let $K \in \mathbf{K}$ be such that $\mathrm{Ad}_K(A) \subseteq \mathfrak{a}$. Then there is some $w \in \mathbf{W}$ such that $w \cdot x = \mathrm{Ad}_K(x)$ for all $x \in A$.*

Proof. Let $K \in \mathbf{K}$ be as in the statement, and let $x \in \mathfrak{a}$ be regular. Then $\mathrm{Ad}_K^{-1}(x) \in \mathfrak{p}_A$ and by Lemma 1.A.45 there exists $L \in \mathbf{K}_A$ such that $\mathrm{Ad}_{LK^{-1}}(x) \in \mathfrak{a}$. Hence $KL^{-1} \in N_{\mathbf{K}}(\mathfrak{a})$ (note we took the inverse) and $\mathrm{Ad}_{KL^{-1}} = \mathrm{Ad}_K$ on A and hence the Weyl group element corresponding to KL^{-1} does the job. □

We immediately get some useful corollaries.

Corollary 1.A.47. *Let $A \subseteq \mathfrak{a}$ be any subset. Let $\mathfrak{p}_A$ be the centralizer of A in $\mathfrak{p}$ and $\mathbf{K}_A$ the stabilizer of A in $\mathbf{K}$. Then $\mathfrak{p}_A = \mathrm{Ad}_{\mathbf{K}_A}(\mathfrak{a})$.*

Proof. First let $K \in \mathbf{K}_A$, and $x \in A$, and $y \in \mathfrak{a}$. Then $[x, \mathrm{Ad}_K(y)] = \mathrm{Ad}_K([x, y]) = 0$. Hence $\mathrm{Ad}_{\mathbf{K}_A}(\mathfrak{a}) \subseteq \mathfrak{p}_A$. For the reverse inclusion let $x \in \mathfrak{p}_A$. Then by Lemma 1.A.45 there is $K \in \mathbf{K}_A$ such that $\mathrm{Ad}_K(x) \in \mathfrak{a}$. Hence $x \in \mathrm{Ad}_K^{-1}(\mathfrak{a}) \subseteq \mathrm{Ad}_{\mathbf{K}_A}(\mathfrak{a})$. □

Corollary 1.A.48. *For any $x \in \mathfrak{a}$ it holds that $\mathbf{W}x = \mathbf{K}x \cap \mathfrak{a}$. In particular the choice of $\mathbf{K}$, even if $\mathbf{K}$ is disconnected, does not affect the $\mathbf{K}$-orbits in $\mathfrak{p}$.*

Corollary 1.A.49. *Let $A \subset \mathfrak{a}$ and $B \subset \mathfrak{p}$ such that $A \cup B$ is Abelian. Then there is some $K \in \mathbf{K}_A$ such that $\mathrm{Ad}_K(B) \subseteq \mathfrak{a}$.*

We have encountered several Lie group actions, such as the action of $\mathbf{K}$ on $\mathfrak{p}$ or that of $\mathbf{W}$ on $\mathfrak{a}$. In the following we want to understand the structure of the corresponding quotient spaces $\mathfrak{p}/\mathbf{K}$ and $\mathfrak{a}/\mathbf{W}$. For this the concept of a slice is crucial. We recall the definition here, see [AB15, Def. 3.47].

Definition 1.A.50. *Let $\mathbf{G}$ be a Lie group, M a smooth manifold, and $\mu : \mathbf{G} \times M \to M$ a smooth action. A* slice at $x \in M$ for the action μ of $\mathbf{G}$ *is an embedded submanifold S_x of M containing x and satisfying the following properties:*

(i) $T_xM = D\mu_x(\mathfrak{g}) \oplus T_xS_x$ *and* $T_yM = D\mu_y(\mathfrak{g}) + T_yS_x$ *for* $y \in S_x$ *and where* $\mu_x(g) = g \cdot x$.

(ii) S_x *is invariant under the stabilizer* $\mathbf{G}_x = \{g \in \mathbf{G} : g \cdot x = x\}$.

(iii) if $y \in S_x$ *and* $g \in \mathbf{G}$ *such that* $g \cdot y \in S_x$, *then* $g \in \mathbf{G}_x$.

By [AB15, Thm. 3.49], proper Lie group actions on manifolds admit slices at every point of the manifold. The next result is an immediate consequence of Lemma 1.A.25 and links slices at x to the commutant $\mathfrak{p}_x$.

Lemma 1.A.51. *Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a semisimple, orthogonal, symmetric Lie algebra, and let $(\mathbf{G}, \mathbf{K})$ be an associated pair. Let $x \in \mathfrak{p}$ and let $S_x = \mathfrak{p}_x \cap B_\varepsilon(x)$ where $B_\varepsilon(x)$ is an ε-ball around x in $\mathfrak{g}$. Then for $\varepsilon > 0$ small enough, S_x is a slice at x for the adjoint action of $\mathbf{K}$ on $\mathfrak{p}$.*

Proof. This follows from the proof of the slice theorem, see [AB15, Thm. 3.49]. In that proof, for an isometric action of $\mathbf{K}$ on M, a slice about $x \in M$ is constructed by taking the exponential of an ε-ball around the origin in the subspace of T_xM orthogonal to the tangent space of the orbit $T_x(\mathbf{K}x)$. By Lemma 1.A.25 we have the orthogonal decomposition $T_x\mathfrak{p} = T_x(\mathbf{K}x) \oplus \mathfrak{p}_x$. Since $\mathfrak{p}$ is a vector space we can canonically identify the tangent space $T_x\mathfrak{p}$ with $\mathfrak{p}$ and the exponential function is simply $\exp_x(v) = x + v$. This concludes the proof. □

Corollary 1.A.52. *Let $x \in \mathfrak{p}$ and let S_x be a slice, then for all $y \in S_x$ it holds that $\mathfrak{p}_y \subseteq \mathfrak{p}_x$.*

Proof. Without loss of generality $x, y \in \mathfrak{a}$. By Corollary 1.A.47 it holds that $\mathfrak{p}_x = \mathrm{Ad}_{\mathbf{K}_x}(\mathfrak{a})$ and analogously for y. But if $y \in S_x$ then $\mathbf{K}_y \subseteq \mathbf{K}_x$ by the definition of a slice. □

The next results relate the quotients $\mathfrak{a}/\mathbf{W}$ and $\mathfrak{p}/\mathbf{K}$, showing that they are isometric. But first let us clarify what metric each space is endowed with. As usual, all subspaces of $\mathfrak{g}$ are given the $\mathbf{K}$ invariant inner product of Lemma 1.A.15, which induces a norm and a metric. Hence the actions of $\mathbf{K}$ on $\mathfrak{p}$ and of $\mathbf{W}$ on $\mathfrak{a}$ are isometric. The closed Weyl chamber $\mathfrak{w} \subset \mathfrak{a}$ inherits the metric on $\mathfrak{a}$. We will consider quotient spaces such as $\mathfrak{p}/\mathbf{K}$, and $\mathfrak{p}_x/\mathbf{K}_x$, and $\mathfrak{a}/\mathbf{W}$. Then the following lemma describes the relevant metric properties.

Lemma 1.A.53. *Let M be a complete Riemannian manifold and $\mathbf{G}$ a compact Lie group acting isometrically on M. Then the quotient $M/\mathbf{G}$ with the distance $d(\mathbf{G}x, \mathbf{G}y) := d(x, \mathbf{G}y)$ becomes a metric space and the quotient map $\pi : M \to M/\mathbf{G}$ is non-expansive.*

Proof. Since $\mathbf{G}$ acts by isometries, the distance is well defined. The axioms of a metric are easily verified. It is clear that $d(\mathbf{G}x, \mathbf{G}y) \leq d(x, y)$, showing that the quotient map π is non-expansive. □

Lemma 1.A.54. *The maps $\psi : \mathfrak{w} \to \mathfrak{a}/\mathbf{W}$, $x \mapsto \mathbf{W}x$ and $\phi : \mathfrak{a}/\mathbf{W} \to \mathfrak{p}/\mathbf{K}$, $\mathbf{W}x \mapsto \mathbf{K}x$ are bijections.*

Proof. The map ψ is bijective since by Corollary 1.A.39 every $\mathbf{W}$ orbit in $\mathfrak{a}$ intersects $\mathfrak{w}$ in exactly one point. The map $\phi : \mathbf{W}x \mapsto \mathbf{K}x$ is well defined since for any $y \in \mathbf{W}x$ there exists by definition of the Weyl group some element $K \in N_{\mathbf{K}}(\mathfrak{a})$ such that $y = \mathrm{Ad}_K(x)$. Injectivity of ϕ follows from Corollary 1.A.48 and surjectivity follows from Lemma 1.A.26. □

Lemma 1.A.55. *The maps $\psi : \mathfrak{w} \to \mathfrak{a}/\mathbf{W}$ and $\phi : \mathfrak{a}/\mathbf{W} \to \mathfrak{p}/\mathbf{K}$ of the previous lemma are isometries with respect to any $\mathbf{K}$-invariant inner product on $\mathfrak{p}$ and its restriction to $\mathfrak{a}$ and $\mathfrak{w}$.*

Proof. We start by showing that ψ is an isometry. Let $\|\cdot\|$ denote any norm induced by a $\mathbf{K}$-invariant inner product on $\mathfrak{p}$. This norm also induces the metric on $\mathfrak{w}$ and $\mathfrak{a}$. Since the action of $\mathbf{W}$ on $\mathfrak{a}$ is isometric, it holds that $d(\mathbf{W}x, \mathbf{W}y) = \min_{w\in\mathbf{W}} \|x - wy\|$ where x, y can always be chosen in $\mathfrak{w}$. However the minimum must be achieved by the identity in $\mathbf{W}$, since otherwise the segment connecting x and wy lies in more than one Weyl chamber. Reflecting this segment into $\mathfrak{w}$ yields a continuous, piecewise linear path in $\mathfrak{w}$ connecting x to y which must be longer than the line segment connecting x to y. Hence $d(\mathbf{W}x, \mathbf{W}y) = \|x - y\|$.

Next we want to show that ϕ is an isometry. Consider any $x, y \in \mathfrak{a}$ with x regular. First note that $d(\mathrm{Ad}_{\mathbf{K}}(x), \mathrm{Ad}_{\mathbf{K}}(y)) \leq d(\mathbf{W}x, \mathbf{W}y)$. Moreover, since $\mathbf{K}$ acts isometrically on $\mathfrak{p}$, it clearly holds that $d(\mathrm{Ad}_{\mathbf{K}}(x), \mathrm{Ad}_{\mathbf{K}}(y)) = d(\mathrm{Ad}_{\mathbf{K}}(x), y)$. But any geodesic in $\mathfrak{p}$ realizing the distance $d(\mathrm{Ad}_{\mathbf{K}}(x), y)$ must be a line segment starting at x and orthogonal to $\mathrm{Ad}_{\mathbf{K}}(x)$, see [Car92, Ch. 9, Example 1]. Hence it is contained in $\mathfrak{a}$ and so $d(\mathbf{K}x, \mathbf{K}y) \geq d(\mathbf{W}x, \mathbf{W}y)$ and hence they are equal. The proof for x non-regular is similar: again we may assume that $x \in \mathfrak{a}$, and we know that the segment realizing the distance $d(\mathrm{Ad}_{\mathbf{K}}(x), y)$ is orthogonal to $\mathrm{Ad}_{\mathbf{K}}(x)$, and hence contained in $\mathfrak{p}_x$, and so is y. By Lemma 1.A.45, there is some $K \in \mathbf{K}_x$ such that $\mathrm{Ad}_K(y) \in \mathfrak{a}$. This concludes the proof. □

Corollary 1.A.56. *Let $x, y \in \mathfrak{w}$ be distinct. Then the line segment connecting x to y is a geodesic segment in $\mathfrak{p}$ realizing the distance between the orbits $\mathrm{Ad}_{\mathbf{K}}(x)$ and $\mathrm{Ad}_{\mathbf{K}}(y)$.*

Proof. By Lemma 1.A.55 the distance in $\mathfrak{w}$ between x and y is the same as the distance in $\mathfrak{p}$ between the orbits $\mathrm{Ad}_{\mathbf{K}}(x)$ and $\mathrm{Ad}_{\mathbf{K}}(y)$. Since the straight line segment in $\mathfrak{w}$ realizes the distance between x and y in $\mathfrak{w}$, the same line segment considered in $\mathfrak{p}$ realizes the distance between the orbits. □

Similarly we can relate the quotients $\mathfrak{a}/\mathbf{W}_x$ and $\mathfrak{p}_x/\mathbf{K}_x$ for $x \in \mathfrak{p}$. In this case we just need a homeomorphism.

Corollary 1.A.57. *Let $x \in \mathfrak{a}$. The inclusion $\iota : \mathfrak{a} \hookrightarrow \mathfrak{p}_x$ descends to a homeomorphism $\phi_x : \mathfrak{a}/\mathbf{W}_x \to \mathfrak{p}_x/\mathbf{K}_x$, $\mathbf{W}_x y \mapsto \mathbf{K}_x y$.*

Proof. First we show that ϕ_x is a well defined bijection. Let $y, z \in \mathfrak{a}$ and let $w \in \mathbf{W}_x$. By definition of $\mathbf{W}_x$ there is some $K \in \mathbf{K}_x$ with $\mathrm{Ad}_K(y) = w \cdot y$. Hence ϕ_x is well-defined. If $\mathbf{K}_x y = \mathbf{K}_x z$, that is $y = \mathrm{Ad}_K(z)$ for some $K \in \mathbf{K}_x$, then by Lemma 1.A.46 there is some $w \in \mathbf{W}_x$ with $y = w \cdot z$, so ϕ_x is injective. Surjectivity follows from Lemma 1.A.45. By the definition of the quotient topology, ϕ_x is continuous, and it remains to show that the inverse is too.

We show that ϕ_x is open. Let $y \in \mathfrak{a}$. By Lemma 1.A.44, the centralizer of y in $\mathfrak{p}_x$ is the orthogonal complement of the tangent space at y of $\mathrm{Ad}_{\mathbf{K}_x}(y)$ in $\mathfrak{p}_x$. Setting $A = \{x, y\}$, this centralizer is denoted by $\mathfrak{p}_A$ and for ε small enough $S_y = \mathfrak{p}_A \cap B_\varepsilon(y)$ is a slice at y for the action of $\mathbf{K}_x$ on $\mathfrak{p}_x$, similarly to Lemma 1.A.51. Now let O be an open neighborhood of y in $\mathfrak{a}$. Then, if ε is chosen small enough, $S_y \subseteq \mathrm{Ad}_{\mathbf{K}_A}(O)$ and by the tubular neighborhood theorem, see [AB15, Thm. 3.57], $\mathrm{Ad}_{\mathbf{K}_x}(S_y)$ contains y in its interior. Hence the image of O in $\mathfrak{p}_x/\mathbf{K}_x$ contains the image of y in its interior. □

Moreover we can show that whenever $y = \mathrm{Ad}_K(x)$, the quotients $\mathfrak{p}_x/\mathbf{K}_x$ and $\mathfrak{p}_y/\mathbf{K}_y$ are isomorphic in a unique way.

Lemma 1.A.58. *Let $x \in \mathfrak{p}$ and let $y = \mathrm{Ad}_K(x)$ for some $K \in \mathbf{K}$. Then $\mathrm{Ad}_K : \mathfrak{p}_x \to \mathfrak{p}_y$ is a linear $\mathbf{K}_x$-$\mathbf{K}_y$-equivariant isomorphism. Hence it descends to a homeomorphsim $\phi_{x,y} : \mathfrak{p}_x/\mathbf{K}_x \to \mathfrak{p}_y/\mathbf{K}_y$ which does not depend on the choice of K. Furthermore if $z \in \mathfrak{p}$ belongs to the same $\mathbf{K}$-orbit as x and y, then $\phi_{yz} \circ \phi_{xy} = \phi_{xz}$, or equivalently ϕ_{xx} is the identity.*

Proof. First we show that $\mathrm{Ad}_K : \mathfrak{p}_x \to \mathfrak{p}_y$ is a linear isomorphism. Linearity and invertibility are clear, we just need to show that $\mathrm{Ad}_K(z) \in \mathfrak{p}_y$ for $z \in \mathfrak{p}_x$. But this follows from $[y, \mathrm{Ad}_K(z)] = \mathrm{Ad}_K([\mathrm{Ad}_K^{-1}(y), z]) = \mathrm{Ad}_K([x, z]) = 0$. Let $L \in \mathbf{K}_x$, then clearly $\mathrm{Ad}_K \circ \mathrm{Ad}_L = \mathrm{Ad}_{KLK^{-1}} \circ \mathrm{Ad}_K$, that is, Ad_K is equivariant. In particular, it maps orbits to orbits bijectively, and hence Ad_K induces a well-defined bijection $\phi_{x,y} : \mathfrak{p}_x/\mathbf{K}_x \to \mathfrak{p}_y/\mathbf{K}_y$. Note that $\phi_{x,y}$ does not depend on the choice of K, since any other choice differs from K by multiplication with a stabilizer element, which leaves the orbits unchanged. Continuity of $\phi_{x,y}$ follows from the definition of the quotient topology, and continuity of the inverse follows analogously. □

Lemma 1.A.59. *Let $x \in \mathfrak{a}$. Then the stabilizer subgroup $\mathbf{W}_x$ is generated by the reflections s_α corresponding to the roots $\Delta_x = \{\alpha \in \Delta : \alpha(x) = 0\}$. Hence Δ_x is a (possibly non-reduced) root system on its span and $\mathbf{W}_x$ is its Weyl group. In particular $\mathbf{W}_x$ acts simply transitively on its Weyl chambers.*

Proof. By [Hel78, Ch. VII, Thm. 2.16] Δ is a (generally non-reduced) root system and $\mathbf{W}$ is the corresponding Weyl group. By [Hel78, Ch. X Lem. 3.2] Δ contains a reduced root system $\Delta' \subseteq \Delta$ with the same Weyl group. By [Hum72, Sec. 10.3 Lemma B] it holds that $\mathbf{W}_x$ is generated by reflections with respect to elements $\alpha \in \Delta'_x \subseteq \Delta_x$. Since for $\alpha, \beta \in \Delta_x$ it holds that $s_{s_\alpha(\beta)}(x) = s_\alpha \circ s_\beta \circ s_\alpha^{-1}$ we know that Δ_x is invariant under its own reflections. Since $\Delta_x \subseteq \Delta$, the integrality condition is inherited and hence Δ_x is a possibly non-reduced root system on its span. □

Remark 1.A.60. *Corollary 1.A.57 and Lemma 1.A.58 show that for $x \in \mathfrak{a}$ we can identify $\mathfrak{a}/\mathbf{W}_x$ and $\mathfrak{p}_x/\mathbf{K}_x$ and $\mathfrak{p}_y/\mathbf{K}_y$ for all $y \in \mathrm{Ad}_{\mathbf{K}}(x)$. By Lemma 1.A.59, $\mathbf{W}_x$ is itself a Weyl group. Furthermore one can identify $\mathfrak{a}/\mathbf{W}_x$ with the orbifold tangent space $T_{\pi(x)}(\mathfrak{a}/\mathbf{W})$, see Definition 1.B.1.*

1.B Orbifolds

In this appendix we give the necessary background on orbifolds and prove some technical results which are used in the main text. For a general introduction to orbifolds see [ALJ07]. We only consider the local theory, that is, we work in a single linear orbifold structure chart. This section does not presuppose any knowledge of orbifolds. We note that the orbifolds encountered in the main text are of a special kind due to the Weyl group structure, but we will not make use of this assumption in this appendix.

To kick things off we recall a basic topological concept in the context of group actions: given a finite dimensional real vector space V, and a finite group Γ acting linearly (and thus continuously) on V, we denote by V/Γ the usual quotient space endowed with its quotient topology. Moreover, $\pi : V \to V/\Gamma$, $x \mapsto [x]$ denotes the quotient map which is continuous by definition of the quotient topology and one can easily show that it is open.[24] With this we can define *tangent spaces* of points in V/Γ by "pulling over" the well-known concept of tangent spaces of manifolds:

Definition 1.B.1. *Let V be a finite dimensional real vector space, and Γ a finite group acting linearly on V. We define the tangent bundle of V/Γ to be $T(V/\Gamma) := (TV)/\Gamma$ where, when identifying T_xV*

[24] Let $\mathbf{G}$ be any group acting on a topological space X by homeomorphisms, then $\pi : X \to X/\mathbf{G}$ is open. Indeed let $U \subseteq X$ be any open subset. By definition of the quotient topology it holds that $\pi(U)$ is open if and only if $\pi^{-1}(\pi(U))$ is open. But $\pi^{-1}(\pi(U)) = \bigcup_{g \in \mathbf{G}} gU$ is clearly open.

with V as usual, the action of Γ on TV is given by $g \cdot (x, v) = (g \cdot x, g \cdot v)$. One can illustrate this with the following commutative diagram:

$$\begin{array}{ccc} TV & \xrightarrow{D\pi} & T(V/\Gamma) \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle \tilde{p}} \\ V & \xrightarrow{\pi} & V/\Gamma \end{array} \qquad \textit{or, respectively,} \qquad \begin{array}{ccc} (x,v) & \xrightarrow{D\pi} & [(x,v)] \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle \tilde{p}} \\ x & \xrightarrow{\pi} & [x] \end{array}$$

for all $x \in V$, $v \in T_xV$, where π and $D\pi$ are the respective quotient maps and $p, \tilde{p}$ are the respective footpoint maps.

One readily verifies that $\tilde{p}$ is well-defined. Moreover, one can show that for all $x \in V$ the tangent space (also called *tangent cone*) of $[x]$, that is, $\tilde{p}^{-1}([x])$, is homeomorphic[25] to $(T_xV)/\Gamma_x$ by means of the map $\tilde{p}^{-1}([x]) \to (T_xV)/\Gamma_x$, $[(x, v)] \mapsto \{g \cdot v : g \in \Gamma_x\}$. We denote the tangent space $\tilde{p}^{-1}([x]) \simeq (T_xV)/\Gamma_x$ by $T_{\pi(x)}(V/\Gamma)$ or $T_{[x]}(V/\Gamma)$.

Example 1.B.2. *Consider the manifold $\mathbb{R}$ with the group action of $\mathbb{Z}_2$ whose non identity element acts by $x \mapsto -x$. The quotient map is the absolute value: $\pi(x) = |x|$. Then the action on the tangent bundle $T\mathbb{R}$ is given by $(x, v) \mapsto (-x, -v)$ (which is not a reflection anymore). Note that the only fixed point of this action is $(0, 0)$. Then for $x \neq 0$ it holds that $T_{|x|}(\mathbb{R}/\mathbb{Z}_2) \cong \mathbb{R}$ and $T_{|0|}(\mathbb{R}/\mathbb{Z}_2) \cong \mathbb{R}/\mathbb{Z}_2$.*

After establishing the basic setting we can make sense of differentiating within the quotient:

Definition 1.B.3. *Let V be a finite dimensional real vector space, and Γ a finite group acting linearly on V. Let I be an open interval and let $\xi : I \to V/\Gamma$ be a continuous function. Given $t \in I$ we say that ξ is differentiable at t if there exists $\lambda : I \to V$ such that $\xi \equiv \pi \circ \lambda$ (such λ is called a "lift") and such that λ is differentiable at t. We then call*[26]

$$D\xi(t) := \{g \cdot (\lambda(t), \lambda'(t)) : g \in \Gamma\} \in T(V/\Gamma)$$

the derivative of ξ at t. We say that ξ is differentiable if it is differentiable at every $t \in I$, and we denote the derivative by $D\xi : I \to T(V/\Gamma)$. If additionally the derivative $D\xi$ is continuous, then we say that ξ is C^1.

Using the previously discussed homeomorphism one could equivalently define $D\xi(t)$ as the collection $\{g\lambda'(t) : g \in \Gamma_{\lambda(t)}\} \in (T_{\lambda(t)}V)/\Gamma_{\lambda(t)}$. Our definition however has the advantage that all derivatives live in the same space $T(V/\Gamma)$.

Remark 1.B.4. *Properly defining derivatives of maps between orbifolds is notoriously difficult. The situation is easier for us, since we only deal with paths. Our definition is quite general, as it is "pointwise". In Proposition 1.B.8 we show that it is equivalent to a seemingly stronger definition, which in turn is similar to the original definition given in [Sat56].*

Of course, we first have to make sure that the derivative of ξ is well-defined in the first place, that is, it does not dependent on the chosen lift λ:

Lemma 1.B.5. *Let a finite dimensional real vector space V as well as a finite group Γ acting linearly on V be given. Moreover, let I be an open interval and let $\xi : I \to V/\Gamma$ be continuous. Now for arbitrary $t \in I$ the following statements hold:*

[25] See for instance [ALJ07, p. 11]. As usual, $\Gamma_x := \{g \in \Gamma : gx = x\}$ denotes the stabilizer of x in Γ.

[26] Another way to write this would be $D\xi(t) = D\pi((\lambda(t), \lambda'(t)))$.

(i) If ξ is differentiable at t, then the derivative $D\xi(t)$ is a well-defined element of $T(V/\Gamma)$.

(ii) Given any two lifts λ, μ of ξ which are continuous at t, there exists a neighborhood I' of t in I such that the following hold: if $\mu(s) = g\lambda(s)$ for some $s \in I'$ and some $g \in \Gamma$, then $\mu(t) = g\lambda(t)$, and if $\mu(t) = h\lambda(t)$, then for each $s \in I'$ there is some $g \in \Gamma_{\mu(t)}$ such that $g\mu(s) = h\lambda(s)$.

(iii) Assume that ξ is differentiable at t. If λ is any lift of ξ which is continuous on $(t-\varepsilon, t+\varepsilon)$ for some $\varepsilon > 0$, then λ admits a left and a right derivative at t which are both elements of $D\xi(t)$.

Proof. (i): Let λ_1 and λ_2 be lifts of ξ that are differentiable at t. What we have to show now is that

$$\{g \cdot (\lambda_1(t), \lambda_1'(t)) : g \in \Gamma\} = \{\tilde{g} \cdot (\lambda_2(t), \lambda_2'(t)) : \tilde{g} \in \Gamma\}\,.$$

For this let $g \in \Gamma$ as well as any sequence $(t_n)_{n\in\mathbb{N}}$ in I which converges to t be given. By the lift property of λ_1, λ_2 there exists for all $n \in \mathbb{N}$ some $g_n \in \Gamma$ such that $g\lambda_1(t_n) = g_n\lambda_2(t_n)$.

Now because Γ is finite there exists a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ of $(g_n)_{n\in\mathbb{N}}$ which is constant, that is, equal to some $\tilde{g} \in \Gamma$. Using continuity of λ_1, λ_2, and the group action we compute

$$g\lambda_1(t) = \lim_{k\to\infty} g\lambda_1(t_{n_k}) = \lim_{k\to\infty} g_{n_k}\lambda_2(t_{n_k}) = \tilde{g}\lambda_2(t)\,.$$

For all $k \in \mathbb{N}$ this yields

$$\begin{aligned} g \cdot \frac{\lambda_1(t_{n_k}) - \lambda_1(t)}{t_{n_k} - t} &= \frac{g\lambda_1(t_{n_k}) - g\lambda_1(t)}{t_{n_k} - t} = \frac{g_{n_k}\lambda_2(t_{n_k}) - \tilde{g}\lambda_2(t)}{t_{n_k} - t} \\ &= \frac{\tilde{g}\lambda_2(t_{n_k}) - \tilde{g}\lambda_2(t)}{t_{n_k} - t} = \tilde{g} \cdot \frac{\lambda_2(t_{n_k}) - \lambda_2(t)}{t_{n_k} - t} \end{aligned}$$

so taking the limit $k \to \infty$ shows $g\lambda_1'(t) = \tilde{g}\lambda_2'(t)$ meaning $\tilde{g}$ is the group element we were looking for.

(ii): First of all, the lift property guarantees that one finds $h \in \Gamma$ which satisfies $\mu(t) = h\lambda(t)$. Now because Γ is finite there exists an open set U containing $\mu(t)$ (called "slice") with the property that, for all $g \in \Gamma$, if $g\mu(t) = \mu(t)$, then $gU = U$, and if $g\mu(t) \neq \mu(t)$ then $(gU) \cap U = \emptyset$[27]. Combining this with the lift property we get $\lambda(t) = h^{-1}\mu(t) \in h^{-1}U$. Therefore continuity yields $I' \subseteq I$ such that on all of I', μ lies in U and λ lies in $h^{-1}U$.

For the first part we have to show that, given any $s \in I'$ and any $g \in \Gamma$ such that $\mu(s) = g\lambda(s)$, one also has $\mu(t) = g\lambda(t)$. But by our previous continuity argument we know that $\mu(s) \in U$ as well as $\mu(s) = g\lambda(s) \in g(h^{-1}U) = gh^{-1}U$. Therefore $U \cap (gh^{-1}U) \neq \emptyset$ which by the slice property implies $\mu(t) = gh^{-1}\mu(t) = g(h^{-1}\mu(t)) = g\lambda(t)$, as desired.

For the second part note that since $h\lambda(s) \in U$ for all $s \in I'$, any $g \in \Gamma$ satisfying $g\mu(s) = h\lambda(s)$ must lie in $\Gamma_{\mu(t)}$.

(iii): Now let λ, μ be lifts of ξ where λ is arbitrary but continuous on $(t-\varepsilon, t+\varepsilon)$ for some $\varepsilon > 0$, and μ is differentiable at t (such μ exists because ξ is assumed to be differentiable). Again we start by using the lift property, that is, for all $s \in (t-\varepsilon, t+\varepsilon)$ there exists $g_s \in \Gamma$ such that $\mu(s) = g_s\lambda(s)$. On the other hand λ, μ are both continuous at t, so (ii) yields $\delta > 0$ such that for all $s \in (t-\min\{\delta,\varepsilon\}, t+\min\{\delta,\varepsilon\})$ one has $\mu(t) = g_s\lambda(t)$. With this in mind let us look at the map

$$\Lambda : (t, t+\min\{\delta,\varepsilon\}) \to V \times V, \quad s \mapsto \left(\lambda(s), \frac{\lambda(s) - \lambda(t)}{s - t}\right).$$

[27] This is indeed a special case of Definition 1.A.50.

Given s arbitrary from the domain of Λ we compute

$$\Lambda(s) = \left(\lambda(s), \frac{\lambda(s)-\lambda(t)}{s-t}\right) = g_s^{-1} \underbrace{\left(\mu(s), \frac{\mu(s)-\mu(t)}{s-t}\right)}_{\to(\mu(t),\mu'(t)) \text{ as } s\to t^+} . \tag{1.22}$$

As Γ is finite the set of possible accumulation points $\{g \cdot (\mu(t), \mu'(t)) : g \in \Gamma\} = D\xi(t)$ of $\Lambda(s)$ (as $s \to t^+$) is finite. But because Λ is bounded by (1.22) and continuous, its cluster set at any (locally connected) boundary point of its domain is either a point, or a continuum [CL66, p. 2], meaning it has exactly one accumulation point, denoted by $g_+ \cdot (\mu(t), \mu'(t))$. Similarly one obtains some g_-, and this concludes the proof. □

Remark 1.B.6. *Note that:*

(i) In general $g_+ \neq g_-$ so the left and right derivative of λ at t need not coincide.

(ii) In Lemma 1.B.5 (iii) it does not suffice for λ to be continuous only at the point of interest t. For this consider $\mathbb{R}/\mathbb{Z}_2$ (cf. Example 1.B.2) and the function $\xi : (-1,1) \to \mathbb{R}/\mathbb{Z}_2$, $t \mapsto [t] = |t|$, so ξ effectively describes a reflection of a 1D motion at the origin. Obviously, $\mu : (-1,1) \to \mathbb{R}$, $t \mapsto t$ is a lift of ξ which is differentiable (at the origin). Now an example of a lift of ξ which is continuous at 0 but does not admit a left- or right-derivative is given by

$$\lambda : (-1,1) \to \mathbb{R}/\mathbb{Z}_2, \quad t \mapsto \begin{cases} t & t \in \mathbb{Q} \\ -t & t \in \mathbb{R} \setminus \mathbb{Q} . \end{cases}$$

Continuity in $t = 0$ is as evident as the fact that neither $\lim_{t\to 0^+} \frac{\lambda(t)-\lambda(0)}{t}$ nor $\lim_{t\to 0^-} \frac{\lambda(t)-\lambda(0)}{t}$ exist because they both have the accumulation points 1 and -1.

The following result shows that if a path $\xi : I \to V/\Gamma$ admits local C^1-lifts, then these lifts can be stitched together to form a global C^1-lift. Note that this lift need not be unique, not even up to global group action. (e.g. $\xi : (-1,1) \to \mathbb{R}/\mathbb{Z}_2$, $x \mapsto [x^2]$ has four C^1-lifts).

Lemma 1.B.7. *Let V be a finite dimensional real vector space, and Γ a finite group acting linearly on V. Let I be an open interval and $(I_j)_{j\in J}$ be an arbitrary family of open intervals whose union is I. Given $\xi : I \to V/\Gamma$ the following statements hold:*

(i) If ξ admits a continuous lift on each interval I_j, then ξ has a continuous lift $\lambda : I \to V$.

(ii) If ξ admits a differentiable lift on each interval I_j, then ξ has a differentiable lift $\lambda : I \to V$.

(iii) If ξ admits a C^1-lift on each interval I_j, then ξ has a C^1-lift $\lambda : I \to V$.

Proof. We will only prove (iii), as the simpler cases (i) and (ii) can be shown analogously. First we will show that for every two open intervals with non-empty intersection we can "glue" the corresponding lifts together to a lift on their union. Let $j_1, j_2 \in J$ such that $I_{j_1} \cap I_{j_2} \neq \emptyset$ and neither interval fully contains the other. Then (w.l.o.g.) $I_{j_1} \cup I_{j_2} = \{t \in I_{j_1} : t \leq t_0\} \cup \{t \in I_{j_2} : t \geq t_0\}$ for an arbitrary but fixed $t_0 \in I_{j_1} \cap I_{j_2}$. Let λ_1 and λ_2 be C^1-lifts of ξ on I_{j_1}, I_{j_2}, respectively. Since their projections coincide on the open neighborhood $I_{j_1} \cap I_{j_2}$ of t_0, by well-definedness of the derivative we have

$$D\pi(\lambda_1(t_0), \lambda_1'(t_0)) = D\xi(t_0) = D\pi(\lambda_2(t_0), \lambda_2'(t_0)) \in T(V/\Gamma) .$$

Thus one finds $g \in \Gamma$ such that $\lambda_1(t_0) = (g \cdot \lambda_2)(t_0)$ and $\lambda_1'(t_0) = (g \cdot \lambda_2)'(t_0) = g \cdot \lambda_2'(t_0)$. Then the new path

$$\lambda_0 : I_{j_1} \cup I_{j_2} \to V, \quad t \mapsto \begin{cases} \lambda_1(t) & t \leq t_0, \\ g \cdot \lambda_2(t) & t > t_0, \end{cases}$$

is C^1 and hence λ_0 is a C^1-lift of $\xi|_{I_{j_1} \cup I_{j_2}}$.

Now consider a finite subset $J' \subseteq J$ such that $I' := \bigcup_{j \in J'} I_j$ is connected. Then we can construct a C^1 lift on I' as follows. Choose two elements j_1, j_2 of J' such that $I_{j_1} \cap I_{j_2} \neq \emptyset$, which exist since I' is connected. If one interval contains the other, discard the smaller one, otherwise replace both intervals with their union, on which we can construct a C^1 list by gluing as above. We continue doing this until only one interval is left, namely I'. This also implies that we can construct a C^1 lift on any compact set in I.

To construct a C^1 lift on the entire open interval I, consider first two non-empty closed intervals K_1, K_2 such that K_1 lies in the interior of K_2. Given a C^1 lift λ_1 on K_1, we want to find a C^1 lift λ_2 on K_2 which is an extension of the former. For this let D_1, D_2 be the connected components of $K_2 \setminus \operatorname{int}(K_1)$. As above we can find C^1 lifts on the closed intervals D_1 and D_2 and glue them as above to the given lift λ_1 on K_1, without modifying λ_1. Finally we extend this idea to a compact exhaustion of I by closed intervals K_i for $i \in \mathbb{N}$ using induction. This yields C^1 lifts λ_i on K_i satisfying $\lambda_{i+1}|_{K_i} = \lambda_i$. Hence we may define a lift $\lambda : I \to V$ by $\lambda|_{K_i} := \lambda_i$. Then λ is clearly C^1, as desired. □

The previous result can be further strengthened by only assuming that ξ has the corresponding property in each point. This will be the main result of this section. Given $\xi \in V/\Gamma$, we define the *degeneracy* of ξ as the size of the stabilizer of any lift of ξ, that is, as the number $\deg_\xi := |\Gamma|/|\xi|$ where $|\xi|$ is the cardinality of ξ when taken as a subset of V.

Proposition 1.B.8. *Let V be a finite dimensional real vector space, Γ a finite group acting linearly on V, and I an open interval. Given $\xi : I \to V/\Gamma$ the following statements hold:*

(i) If ξ is continuous, then ξ has a continuous lift $\lambda : I \to V$.

(ii) If ξ is differentiable, then ξ has a differentiable lift $\lambda : I \to V$.

(iii) If ξ is C^1, then ξ has a C^1-lift $\lambda : I \to V$.

Proof. By Lemma 1.B.7 it suffices to show that for every point $t_0 \in I$ there exists an open interval $I' \subseteq I$ containing t_0 on which a lift with the corresponding property exists.

(i) & (ii): We proceed by induction on the degeneracy $\deg_{\xi(t_0)}$ of $\xi(t_0)$. First assume that $\deg_{\xi(t_0)} = 1$. Let $\tilde{\lambda}$ be any lift of $\xi(t_0)$ and consider a slice U about $\tilde{\lambda}$. Then the restriction $\pi|_U$ of the quotient map is a homeomorphism and hence, for an open interval I' containing t_0 and such that $\xi(t) \in \pi(U)$ for $t \in I'$, the function $\lambda := (\pi|_U)^{-1} \circ \xi|_{I'}$ is a continuous lift of $\xi|_{I'}$. For (ii) we additionally need to show that at every $t \in I'$ the lift is differentiable. For this consider any other lift $\mu : I' \to V$ with $\mu(t) \in U$ which is differentiable at t. In a small enough neighborhood of t, μ takes values in U and hence coincides with λ, which is therefore differentiable at t.

Now assume that $\deg_{\xi(t_0)} > 1$ and that for any t with $\deg_{\xi(t)} < \deg_{\xi(t_0)}$ there exists a continuous resp. differentiable lift in a neighborhood of t. We will now show that there is a continuous resp. differentiable lift in a neighborhood of t_0. Again let $\tilde{\lambda}$ be any lift of $\xi(t_0)$ and let U be a slice about $\tilde{\lambda}$. Let $I' \subseteq I$ be an open interval containing t_0 such that $\xi(t) \in \pi(U)$ for all $t \in I'$. Let $J \subset I'$ be the open subset on which $\deg_{\xi(t)} < \deg_{\xi(t_0)}$ for all $t \in J$. Hence J is an at most countable union of

open intervals J_k, and by the induction hypothesis and using Lemma 1.B.7 (i) resp. (ii) there exists a continuous resp. differentiable lift on each J_k taking values in U, which we denote by $\lambda_k : J_k \to U$. Clearly $g_k\lambda_k$ is also a differentiable lift for all $g_k \in \Gamma_\lambda$ and still lies in U. Note that for all $t \in I' \setminus J$ there exists a unique lift μ_t in U. Hence we can define the function

$$\lambda : I' \to U, \quad \lambda(t) = \begin{cases} g_k\lambda_k(t) & t \in J_k \\ \mu_t & t \in I' \setminus J, \end{cases}$$

where the g_k will be determined later. First we show continuity. For $t \in J$ this is clear by construction, hence we consider $t \in I' \setminus J$. Let $t_n \to t$ be a sequence in I' and let $U' \subseteq U$ be any slice about t. Note that if $gU' \subseteq U$, the $gU' = U'$. Since the quotient map π is open, $\pi(U')$ is an open neighborhood of $\xi(t)$. By continuity of ξ this means that for n large enough, $\xi(t_n) \in \pi(U')$ and hence $\lambda(t_n) \in U'$. Note that this argument only uses that λ is a lift contained in U.

Next we show differentiability. By Lemma 1.B.5 (iii) λ admits left and right derivatives at each point and it remains to show that they agree. For $t \in J$ this is clear, hence consider $t \in I' \setminus J$. If t is an isolated point of $I' \setminus J$ connecting two intervals J_{k_1} and J_{k_2}, then one can choose g_{k_1} and g_{k_2} such that λ will be differentiable at t. Let P be the set of all isolated points of $I' \setminus J$ and let consider the open set $J \cup P$. This is again an at most countable union of open intervals J'_k and each J'_k contains at most countably many intervals J_k joint at their boundary points which lie in P. On each J'_k we can then define λ such that it is differentiable. If t is not an isolated point of $I' \setminus J$ then there exists a sequence t_n in $I' \setminus J$ converging to t such that $t_n \leq t$ for all n or $t_n \geq t$ for all n. Without loss of generality, assume that $t_n \geq t$ for all n. Hence the right derivative at t is invariant under the stabilizer $\Gamma_{\lambda(t)}$ and thus it is uniquely determined. This implies that it coincides with the left derivative and λ is differentiable at t.

(iii): By (ii) there exists a differentiable lift $\lambda : I \to V$ of ξ. We claim that λ must be C^1. Let ϕ_i be a basis of the dual space V^*. We show that each $\phi_i \circ \lambda$ is C^1 on I. Certainly $\phi_i \circ \lambda$ is differentiable on I. Assume that $\phi_i \circ \lambda'$ is discontinuous at $t_0 \in I$. Thus there is some $\varepsilon > 0$ such that for every neighborhood I' of t_0 there are $t_1, t_2 \in I'$ such that $|\phi_i \circ \lambda'(t_1) - \phi_i \circ \lambda'(t_2)| > \varepsilon$ and one can show that all values between $\phi_i \circ \lambda'(t_1)$ and $\phi_i \circ \lambda'(t_2)$ are taken, see [Kat80, p. 114]. Now let D_δ be a disk in V of radius $\delta > 0$ centred at $\lambda'(t_0)$. Then, since ξ' is continuous, there is a neighborhood J_δ of t_0 such that $\lambda'(t) \in \Gamma D_\delta$ for all $t \in J_\delta$. Choosing δ small enough we can ensure that $\phi_i(\Gamma D_\delta)$ does not contain an interval of length ε, which yields the desired contradiction. □

Remark 1.B.9. *Due to this result one could equivalently define a path $\xi : I \to V/\Gamma$ to be C^1 if it admits local C^1 lifts, which is more in line with the original definition of smooth functions on orbifolds.*

Finally we give two simple results that are used in the measurable and analytic diagonalizations respectively.

Lemma 1.B.10. *Let Ω be a measurable space, and let $\lambda, \mu : \Omega \to V$ be measurable (where V is endowed with its Borel σ-algebra). Assume that $\pi \circ \lambda = \pi \circ \mu$. Then there is some measurable function $\gamma : \Omega \to \Gamma$ such that $\lambda = \gamma \cdot \mu$.*

Proof. Let γ_i for $i = 1, \ldots, m$ be an enumeration of Γ. Define the sets $\Omega_1 = \{\omega \in \Omega : \lambda(\omega) = \gamma_1 \cdot \mu(\omega)\}$ and iteratively $\Omega_i = \{\omega \in \Omega : \lambda(\omega) = \gamma_i \cdot \mu(\omega)\} \setminus \bigcup_{j=1}^{i-1} \Omega_j$ for $i > 1$. These sets are measurable and form a partition of Ω. Now define $\gamma(\omega) = \gamma_i$ for $\omega \in \Omega_i$. This γ is measurable and satisfies the desired condition by construction. □

Lemma 1.B.11. *Let $\lambda : I \to V$ be real analytic. Then any real analytic path μ satisfying $\pi \circ \mu = \pi \circ \lambda$ satisfies $\mu = g \cdot \lambda$ for some $g \in \Gamma$.*

Proof. Consider the size of the stabilizer $|\Gamma_{\mu(t)}|$ as a function of t. Let $t_0 \in I$ be a point where this value is minimal. Then there is an open interval $J \subseteq I$ containing t_0 on which the stabilizer of $\mu(t)$ is constant. Hence if $\mu(t_0) = g \cdot \lambda(t_0)$, then in a neighborhood J' of t_0 it holds by Lemma 1.B.5 (ii) that $h(t) \cdot \mu(t) = g \cdot \lambda(t)$ for some $h(t) \in \Gamma_{\mu(t_0)}$. Hence on $J \cap J'$ it holds that $\mu(t) = h(t)\mu(t) = g \cdot \lambda(t)$. Since both paths are real analytic, they coincide by the identity theorem. □

1.C Resolvent Computations

Lemma 1.C.1. *Let V be a finite dimensional complex vector space and let $A \in \mathfrak{gl}(V)$ be diagonalizable. Moreover, let $\mathbb{D}_\varepsilon$ be a small open disc around the eigenvalue λ_k of A which does not contain any other eigenvalue λ_l, $l \neq k$. Then the identity*

$$\frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_\varepsilon} (\lambda - \mu)^{-1} R(\lambda, A) \mathrm{d}\lambda = \begin{cases} (\lambda_k - \mu)^{-1} P_k & \text{for } \mu \notin \overline{\mathbb{D}}_\varepsilon \,, \\ \sum_{\substack{l=1 \\ l \neq k}}^{n} (\mu - \lambda_l)^{-1} P_l & \text{for } \mu \in \mathbb{D}_\varepsilon \,, \end{cases} \tag{1.23}$$

is fulfilled, where $R(\lambda, A) = (\lambda - A)^{-1}$ denotes the resolvent of A and P_k, P_l stand for the corresponding eigenprojections.

Proof. The equality can easily be obtained via the representation $A = \sum_{l=1}^{n} \lambda_l P_l$. Indeed, it follows easily from

$$\frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_\varepsilon} (\lambda - \mu)^{-1} R(\lambda, A) \mathrm{d}\lambda = \sum_{l=1}^{n} \Big(\frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_\varepsilon} (\lambda - \mu)^{-1} (\lambda - \lambda_l)^{-1} \mathrm{d}\lambda \Big) P_l$$

as desired. □

With this we can give the proof promised in the main text:

Proof of Lemma 1.5.4. Let us consider the right-hand side of (1.11). By Kato [Kat80, Ch. II, Sec. §1.4 & Thm. 1.10] we know that the corresponding eigenprojections $P_k(z)$ can be chosen analytically on $\dot{\mathbb{D}}_r$ with analytic extension to t_0. Moreover, $P_k(z)$ has the well-known integral representation[28]

$$P_k(z) = \frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_{\varepsilon_k}} R(\lambda, A_c(z)) \mathrm{d}\lambda \,,$$

where $\varepsilon_k > 0$ has to be chosen such that $\mathbb{D}_{\varepsilon_k}$ contains only the k-th[29] eigenvalue of $A_c(z)$. Hence we obtain

$$\big[\dot{P}_k(z), P_k(z)\big] = \frac{1}{(2\pi \mathrm{i})^2} \int_{\partial \mathbb{D}_{\varepsilon'_k}} \int_{\partial \mathbb{D}_{\varepsilon_k}} \big[R\big(\lambda, A_c(z)\big) \dot{A}_c(z) R\big(\lambda, A_c(z)\big), R\big(\mu, A_c(z)\big) \big] \mathrm{d}\lambda \mathrm{d}\mu \,,$$

with $\varepsilon'_k > 0$ slightly larger than $\varepsilon_k > 0$. One could also choose $\varepsilon'_k > 0$ slightly smaller than $\varepsilon_k > 0$. Now the standard resolvent identity

$$R(\lambda, A) R(\mu, A) = (\mu - \lambda)^{-1} \big(R(\lambda, A) - R(\mu, A) \big)$$

[28] See for instance [Kat80, p. 67], but be aware that Kato defines the resolvent to be $R(\lambda, A) = (A - \lambda)^{-1}$ which explains the additional minus sign in his integral representation of the eigenprojection.

[29] Note that the numbering of the eigenvalues does not matter due to the summation on the right-hand side of (1.11).

yields

$$\begin{aligned}\big[\dot{P}_k(z), P_k(z)\big] =& \frac{1}{(2\pi\mathrm{i})^2}\int_{\partial\mathbb{D}_{\varepsilon'_k}}\int_{\partial\mathbb{D}_{\varepsilon_k}}(\lambda-\mu)^{-1}R(\lambda,A_c(z))\dot{A}_c(z)R(\mu,A_c(z))\mathrm{d}\lambda\mathrm{d}\mu \\ &- \frac{1}{(2\pi\mathrm{i})^2}\int_{\partial\mathbb{D}_{\varepsilon'_k}}\int_{\partial\mathbb{D}_{\varepsilon_k}}(\lambda-\mu)^{-1}R(\mu,A_c(z))\dot{A}_c(z)R(\lambda,A_c(z))\mathrm{d}\lambda\mathrm{d}\mu\,.\end{aligned}$$

First we investigate the second last term of the above identity, and using Lemma 1.C.1 we conclude

$$\frac{1}{(2\pi\mathrm{i})^2}\int_{\partial\mathbb{D}_{\varepsilon'_k}}\int_{\partial\mathbb{D}_{\varepsilon_k}}(\lambda-\mu)^{-1}R(\lambda,A_c(z))\dot{A}_c(z)R(\mu,A_c(z))\mathrm{d}\lambda\mathrm{d}\mu = \sum_{\substack{l=1\\l\neq k}}^{n}(\lambda_l-\lambda_k)^{-1}P_k(z)\dot{A}_c(z)P_l(z).$$

Similarly, we conclude for the last term

$$\frac{1}{(2\pi\mathrm{i})^2}\int_{\partial\mathbb{D}_{\varepsilon'_k}}\int_{\partial\mathbb{D}_{\varepsilon_k}}(\lambda-\mu)^{-1}R(\mu,A_c(z))\dot{A}_c(z)R(\lambda,A_c(z))\mathrm{d}\lambda\mathrm{d}\mu = \sum_{\substack{l=1\\l\neq k}}^{n}(\lambda_l-\lambda_k)^{-1}P_l(z)\dot{A}_c(z)P_k(z).$$

Summing up, we obtain

$$\frac{1}{2}\sum_{k=1}^{n}[\dot{P}_k(z),P_k(z)] = \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n}(\lambda_l-\lambda_k)^{-1}P_k(z)\dot{A}_c(z)P_l(z).$$

Using (1.10) this concludes the proof. □

CHAPTER 2

Reduced Control Systems

2.1 Introduction

This chapter introduces the main concept of Part I, namely the reduced control system, and proves the central Equivalence Theorem. We study control systems that admit fast controllability on certain degrees of freedom, represented by a Lie group action. Intuitively, one should be able to factor out these directions, and so our goal is to define an associated reduced control system on the remaining degrees of freedom, and to show that the two systems are essentially equivalent, in a sense which will be specified later. As a consequence, no loss of information is incurred by switching to the reduced control system.

In this chapter we continue working in the setting of symmetric Lie algebras, and so we will make ample use of the tools provided by Chapter 1. Due to the relation of symmetric Lie algebras to matrix diagonalizations (recall Section 1.2), the results from this chapter can then be applied directly to important examples of control systems from quantum control theory. This will be done in Parts II and III.

A similar idea has been considered in [AS04, Ch. 22] for commuting controls under the assumption that the reduced state space is again a manifold. In our setting the controls do not commute and the reduced state space has singularities, which are the source of most complications. The idea of considering a reduced state space, even if the reduced control system is not defined explicitly, has come up several times in quantum control theory. This chapter generalizes the ideas presented in [STK04, RBR18, Yua10] in a mathematically rigorous manner. If the reduced state space is a Riemannian symmetric space, strong results can be derived [KBG01, JLB23]. Unfortunately such systems are rare in practice [KBG01, KGB02, ZYK08]. Often the quotient spaces are rather complicated, and one contents oneself with finding diameters of such spaces to derive speed limits [JLB23, GR21].

Motivational Example

We give a simple example to motivate our work. Consider the closed unit disk $D \subset \mathbb{R}^2$ in the plane and let X be some complete and sufficiently smooth vector field on D, such that D is invariant under the flow of X. The compact Lie group $\mathrm{SO}(2)$ acts on the disk by rotations. The connection to symmetric Lie algebras and matrix diagonalizations is elaborated in Example 1.2.4. Now consider a control system on D with constant drift X and fast control on the action of $\mathrm{SO}(2)$. Without the drift term, this means that we can move arbitrarily quickly within the orbits of the group action, which in this case are simply the concentric circles about the origin. Including the drift term, this is still approximately true. Hence, points on the same orbit may be considered equivalent, and the question becomes how one can move

between orbits. This suggests that there should be a natural way to define a corresponding control system on the quotient space $D/\mathrm{SO}(2) \cong [0,1]$, which in our example is the set of all radii[1]. Moreover, we want this reduced control system to be equivalent to the original system in some precise sense, so that no information is lost.

Let us see what this reduced control system should look like in our simple example. Instead of working on the quotient space, which in general is not a manifold, we will look at a subspace of our state space which intersects all orbits a finite number of times, and does so orthogonally. Here we choose the intersection of the horizontal axis with the disk, i.e. the line segment $A = [-1,1] \cdot e_1 \subset \mathbb{R}^2$ where $e_1 = (1,0)$. This will be our new reduced state space. If we restrict the drift vector field X to the axis A and project the vector field orthogonally onto the axis, this yields some possible dynamics on the reduced space. Using the fast control we can rotate our horizontal axis A to any other axis, and obtain a different vector field on the reduced space. Collecting all of these vector fields defines the reduced control system.

We can plot these vector fields all together in a single graph, where the abscissa is the reduced state space, see Figure 2.1. In the example X is affine linear, and so are the restricted vector fields and hence the graph is a collection of lines. This can be seen as a set-valued function of *achievable derivatives*, denoted derv, and the reduced control system can be seen as the corresponding differential inclusion, as we will show below.

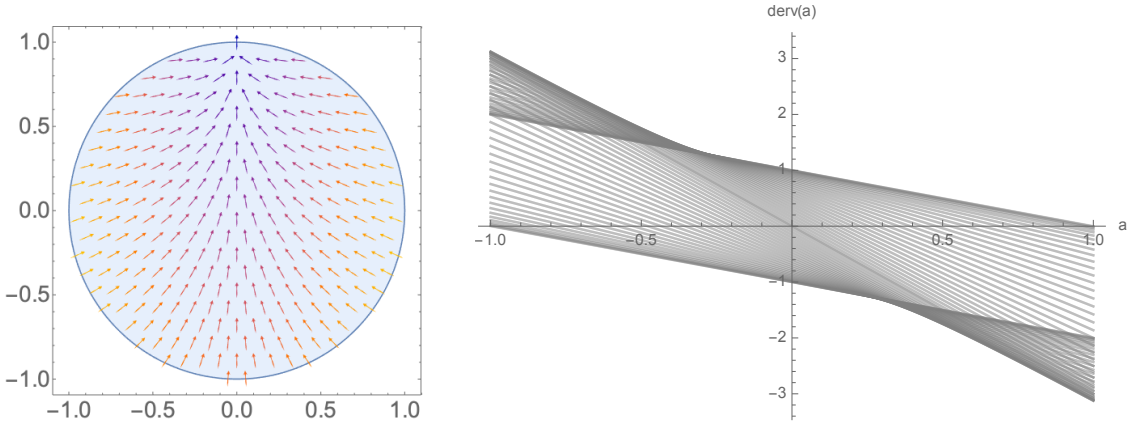

Figure 2.1: Left: A drift vector field X on the disk D. Right: A plot of the corresponding differential inclusion as a collection of affine linear vector fields defined on the interval $[-1,1] \cong A$.

We will come back to this example in Section 2.6 where we use it to illustrate how the method of reduced control systems can be used in practice.

Outline and Main Results

In Section 2.2 we introduce the control systems studied in this chapter, in particular we define the reduced control system, and we briefly address the operator lift of the original and reduced control systems. Some basic properties of the reduced control system are collected in Appendix 2.A.

We then go on to prove our main results in Section 2.3, establishing the equivalence of the reduced control system and the full one. We start out with a local equivalence result in Proposition 2.3.6, followed

[1]Note that the two boundary points of the quotient space have different meanings. Here 1 comes from the boundary of the disk, whereas 0 originates from the singular $\mathrm{SO}(2)$-orbit. This is important for defining the appropriate notion of differentiability in the quotient space, see Appendix 1.B.

by the global equivalence result, which will be separated into a projection, see Theorem 2.3.8, and a lift, see Theorem 2.3.14. The results are combined and slightly generalized in Theorem 2.3.16.

In Section 2.4 we explore how statements about important control theoretic concepts such as reachability, viability, controllability and accessibility can be determined using the reduced control system and lifted to the original one.

As an application, in Section 2.5 we give a majorization result which establishes the preorder induced by the Weyl group action as a kind of resource, see Theorem 2.5.3. For certain cost functions, this allows one to reduce the set of achievable derivatives to a significantly smaller set of optimal derivatives. This will be very useful when studying the problem of optimal cooling in Chapter 6.

Finally, the motivational example given above is worked out in detail in Section 2.6.

2.2 Full and Reduced Control Systems

For the remainder of the chapter we will be working with a semisimple orthogonal symmetric Lie algebra $(\mathfrak{g}, s)$ with Cartan-like decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and an associated pair $(\mathbf{G}, \mathbf{K})$ with $\mathbf{K}$ compact and connected[2]. Moreover $\mathfrak{a} \subseteq \mathfrak{p}$ denotes some choice of a maximal Abelian subspace, with Weyl group $\mathbf{W}$ and a closed Weyl chamber $\mathfrak{w}$. For a concise summary of symmetric Lie algebras, see Section 1.1. A detailed introduction is given in Appendix 1.A.

We start by defining the class of control-affine systems on $\mathfrak{p}$ that we want to study in the sequel. We are given a vector field X on $\mathfrak{p}$, called the *drift vector field*, and a set of *control directions* $k_1, \dots, k_m \in \mathfrak{k}$. The *full control system* we wish to study in this work is the following:

$$p'(t) = X(p(t)) + \sum_{j=1}^{m} u_j(t) \operatorname{ad}_{k_j}(p(t)), \quad p(0) = p_0 \in \mathfrak{p} \tag{F}$$

where ad_x denotes the adjoint operator of x, that is, $\operatorname{ad}_x(y) := [x, y]$. We will always consider solutions on an interval I of the form $[0, T]$ with $T \geq 0$, or of the form $[0, \infty)$. The *control functions* $u_j : I \to \mathbb{R}$ are required to be locally integrable, cf. [Son98, App. C]. A *solution* $p : I \to \mathfrak{p}$ is an absolutely continuous function satisfying (F) almost everywhere for some choice of control functions. Of course, when X is linear then (F) is in fact a bilinear control system [Jur97, Ell09].

The two key assumptions made throughout this chapter are:

(I) The control directions generate the full Lie algebra: $\langle k_1, \dots, k_m \rangle_{\mathsf{Lie}} = \mathfrak{k}$.

(II) No bounds are assumed on the control functions $u_j : I \to \mathbb{R}$, as they are only required to be locally integrable.

Under these assumptions, and if we neglect the drift X, we can move between any two points of a given $\mathbf{K}$-orbit in $\mathfrak{p}$ arbitrarily quickly, cf. [Ell09, Prop. 2.7]. We say that we have *fast and full control* on the Lie group $\mathbf{K}$ — and thus on its orbits in $\mathfrak{p}$. Some results will use the following strengthened version of Assumption (I):

(I') The control directions span the full Lie algebra: $\operatorname{span}(k_1, \dots, k_m) = \mathfrak{k}$.

Using this assumption makes it easier to compute the control functions in the full control system when one tries to lift a solution from the reduced to the full system, but it is not necessary.

[2] This is always possible, for instance by choosing $\mathbf{G} = \operatorname{Int}(\mathfrak{g})$ and $\mathbf{K} = \operatorname{Int}_{\mathfrak{k}}(\mathfrak{g})$, see Lemma 1.A.20.

Reduced Control System

Assumptions (I) & (II) imply that we can move into the maximal Abelian subspace $\mathfrak{a}$ at any time. This motivates us to define a reduced control system on $\mathfrak{a}$. First we introduce some concepts. For every $K \in \mathbf{K}$, we define its *induced vector field* on $\mathfrak{a}$

$$X_K := \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_K^\star(X) \circ \iota\,, \quad \mathfrak{X} := \{X_K : K \in \mathbf{K}\}$$

where $\Pi_{\mathfrak{a}} : \mathfrak{p} \to \mathfrak{a}$ is the orthogonal projection on $\mathfrak{a}$, and $\iota : \mathfrak{a} \hookrightarrow \mathfrak{p}$ is the inclusion[3]. By Ad_K we denote the adjoint action of K on $\mathfrak{p}$ and $\mathrm{Ad}_K^\star$ denotes the pullback action, that is, $\mathrm{Ad}_K^\star(X) = \mathrm{Ad}_K^{-1} \circ X \circ \mathrm{Ad}_K$. If X is linear, then so are all X_K.

Now we can define the *reduced control system* by

$$a'(t) = X_{K(t)}(a(t)), \quad a(0) = a_0 \in \mathfrak{a}\,, \tag{R}$$

where the control function $K : I \to \mathbf{K}$ is required to be measurable[4]. Again, a solution is an absolutely continuous function $a : I \to \mathfrak{a}$ which satisfies (R) almost everywhere.

Moreover, we define the set of *achievable derivatives* at $a \in \mathfrak{a}$ by

$$\mathsf{derv}(a) := \{X_K(a) : K \in \mathbf{K}\} := \mathfrak{X}(a) \subset T_a\mathfrak{a} \cong \mathfrak{a}\,.$$

Then we can also define a *differential inclusion* corresponding to (R) by

$$a'(t) \in \mathsf{derv}(a(t)), \quad a(0) = a_0 \in \mathfrak{a}\,, \tag{R$'$}$$

where $a : I \to \mathfrak{a}$ needs to be absolutely continuous and satisfy (R$'$) almost everywhere. In fact (R) and (R$'$) are equivalent, i.e. they have the same solutions. This follows from Filippov's theorem, see for instance [Smi02, Thm. 2.3], using only that X is continuous and that $\mathbf{K}$ is compact. The difference between (R) and (R$'$) is that the latter "forgets" about the controls, and leads to a more internal, geometric picture. We will switch between both viewpoints whenever it simplifies things.

Often it will be convenient to consider a *relaxed* version of the differential inclusion above given by

$$a'(t) \in \mathrm{conv}(\mathsf{derv}(a(t))), \quad a(0) = a_0 \in \mathfrak{a}\,, \tag{$\bar{\text{R}}$}$$

where conv denotes the convex hull. This will slightly enlarge the set of solutions, however, if X is Lipschitz, every solution to ($\bar{\text{R}}$) can still be approximated uniformly on compact time intervals by solutions to (R$'$), see [AC84, Ch. 2.4, Thm. 2]. The relationship between the full and reduced control systems is depicted in Figure 2.2.

Operator Lifts

Both the original and the reduced control system can be lifted to the operator level. For this section we use Assumption (II), but we drop Assumption (I). Let $X \in \mathfrak{gl}(\mathfrak{p})$ be a linear[5] vector field on $\mathfrak{p}$ and consider the *operator lift* of (F) on $\mathrm{GL}(\mathfrak{p})$ given by the following bilinear system:

$$L' = \Big(X + \sum_{j=1}^{m} u_j(t)\,\mathrm{ad}_{k_j}\Big)L\,, \quad L(0) = \mathbb{1} \in \mathrm{GL}(\mathfrak{p})\,. \tag{FL}$$

[3] In the following we will usually suppress the inclusion ι from the notation.

[4] With respect to the Lebesgue σ-algebra on I and the Borel σ-algebra on $\mathbf{K}$.

[5] For more general vector fields X, the operator lift of the control-affine system is typically defined on an (in general infinite dimensional) subgroup of the diffeomorphism group of $\mathfrak{p}$.

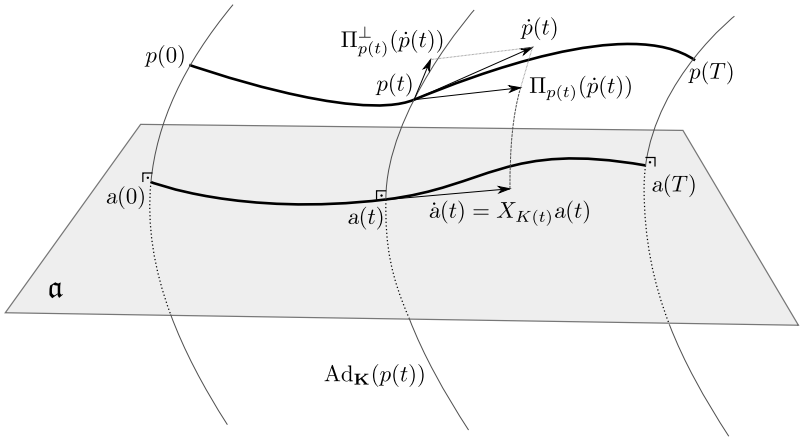

Figure 2.2: Relationship between the time evolution $p(t)$ of the full control system (F) on $\mathfrak{p}$, and the time evolution $a(t)$ of the reduced control system (R) on $\mathfrak{a}$. The solutions are related by $p(t) = \mathrm{Ad}_{K(t)}(a(t))$ for some $K(t)$ in $\mathbf{K}$. The derivative $p'(t)$ can always be split into a part orthogonal to the $\mathbf{K}$-orbit, using the orthogonal projection $\Pi_{p(t)}$ onto the commutant of $p(t)$, and a part tangent to the orbit, using the complementary projection $\Pi^{\perp}_{p(t)}$. A central result in this work is the Equivalence Theorem 2.3.16, which details the relation between the two control systems. Note that here we depict only the regular case.

Such right-invariant control systems defined on Lie groups are highly structured and allow for the application of Lie semigroup theory, see [Law99] for a concise introduction. This system is characterized by the set $\Omega := \{X + \mathrm{ad}_k : k \in \mathrm{span}(k_1, \ldots, k_m)\}$. The fact that Ω is not bounded causes some problems but will be remedied by passing to the reduced control system below.

For a rigorous formulation of the subsequent statements, we cannot avoid some further terminology and notation, see [Law99][6] for details: A *wedge* $\mathfrak{w}$ is a closed, convex cone, and its *edge* $E(\mathfrak{w})$ is the largest subspace contained in the wedge. A *Lie wedge* is a wedge $\mathfrak{w}$ in a Lie algebra closed under the adjoint action of its edge. A *Lie semigroup* is a closed, infinitesimally generated subsemigroup of a Lie group, cf. [Law99, Def. 4.7] and [HN93, p. 20]. The tangent cone at the identity of a closed subsemigroup is always a Lie wedge, and in this case the Lie wedge is called *global* or a *Lie saturate*, cf. [Law99, Prop. 6.2] and [HN93, Sec. 1.6].

Proposition 2.2.1. *Let* $\mathfrak{h} = \langle k_1, \ldots, k_m \rangle_{\mathsf{Lie}}$ *and let* $\mathbf{H} \subseteq \mathbf{K}$ *be the corresponding analytic subgroup. The following statements hold.*

(i) $\langle \Omega \rangle_{\mathsf{wedge}} = \mathbb{R}_+ X + \mathrm{span}(\mathrm{ad}_{k_1}, \ldots, \mathrm{ad}_{k_m})$ *and* $\langle \Omega \rangle_{\mathsf{LW}} \supseteq \langle \mathrm{ad}_{\mathfrak{h}}, \mathrm{Ad}^{\star}_{\mathbf{H}}(X) \rangle_{\mathsf{wedge}}$.

(ii) Assume that there is a Lie wedge $\mathfrak{v}$ *such that* $\mathrm{ad}_{\mathfrak{h}} \subseteq E(\mathfrak{v})$ *and such that* $X \in \mathfrak{v} \setminus E(\mathfrak{v})$. *Then* $\langle \Omega \rangle_{\mathsf{LW}} = \langle \mathrm{ad}_{\mathfrak{h}}, \mathrm{Ad}^{\star}_{\mathbf{H}}(X) \rangle_{\mathsf{wedge}}$.

(iii) If, in addition, $\mathrm{ad}_{\mathfrak{h}}$ *and* $\mathfrak{v}$ *are global*[7]*, then* $\langle \Omega \rangle_{\mathsf{LS}} = \langle \mathrm{ad}_{\mathfrak{h}}, \mathrm{Ad}^{\star}_{\mathbf{H}}(X) \rangle_{\mathsf{wedge}}$.

(iv) If $X, \mathrm{ad}_{h_i} \subseteq \mathfrak{l}$ *for some compact Lie algebra*[8] $\mathfrak{l}$*, then* $\langle \Omega \rangle_{\mathsf{LS}} = \langle X, \mathrm{ad}_{\mathfrak{h}} \rangle_{\mathsf{Lie}}$ *and, in particular,* $\overline{\mathsf{reach}_{\mathsf{FL}}(\mathbb{1})}$ *is a Lie subgroup.*

[6] For a more detailed treatment see also [HHL89, HN93], but beware of slight differences in terminology.
[7] Recall that this means that the Lie algebra $\mathrm{ad}_{\mathfrak{h}}$ generates a closed Lie subgroup in $\mathrm{GL}(\mathfrak{p})$.
[8] By this we mean that the Lie group generated by $\mathfrak{l}$ in $\mathrm{GL}(\mathfrak{p})$ is compact.

Proof. (i): The first part is clear since wedges are closed by definition. Since $\langle\Omega\rangle_{\mathsf{LW}}$ is a wedge, it contains the linear span of all ad_{k_i}, which must be contained in the edge $E(\langle\Omega\rangle_{\mathsf{LW}})$. The latter is a Lie algebra, hence $E(\langle\Omega\rangle_{\mathsf{LW}}) \supseteq \mathrm{ad}_{\mathfrak{h}}$. Since $\langle\Omega\rangle_{\mathsf{LW}}$ is a Lie wedge, for any $h \in \mathfrak{h}$ it contains $e^{\mathrm{ad}_{\mathrm{ad}_h}} X = \mathrm{Ad}_{e^h} X \,\mathrm{Ad}_{e^{-h}}$. (ii): The inclusion $\supseteq$ was shown in (i). As for the converse: let $\mathfrak{w} = \langle \mathrm{ad}_{\mathfrak{h}}, \mathrm{Ad}^{\star}_{\mathbf{H}}(X)\rangle_{\mathsf{wedge}}$. It is enough to show that $\mathfrak{w}$ is a Lie wedge. First we show that $\mathfrak{w} \cap E(\mathfrak{v}) = \mathrm{ad}_{\mathfrak{h}}$: If $w \in \mathfrak{w} \setminus \mathrm{ad}_{\mathfrak{h}}$, then $w = \mathrm{ad}_h + \lambda Y$ where $h \in \mathfrak{h}$, $\lambda > 0$ and $Y \in \mathrm{conv}(\mathrm{Ad}^{\star}_{\mathbf{H}}(X))$. But since $\mathfrak{v}$ is a Lie wedge, $Y \in \mathfrak{v} \setminus E(\mathfrak{v})$, and $w \notin E(\mathfrak{v})$. This proves the claim. Finally we can show that $\mathfrak{w}$ is a Lie wedge. Clearly $\mathfrak{w} \subseteq \mathfrak{v}$, and so $E(\mathfrak{w}) \subseteq E(\mathfrak{v})$. Hence, by the above claim, $E(\mathfrak{w}) = \mathrm{ad}_{\mathfrak{h}}$, and since $\mathfrak{w}$ is invariant under the action of $\mathrm{ad}_{\mathfrak{h}}$, it is a Lie wedge. (iii): It suffices to show that $\mathfrak{w}$ is global. For this we will use [HN93, Prop. 1.37]. Since $\mathfrak{v}$ and $E(\mathfrak{w}) = \mathrm{ad}_{\mathfrak{h}}$ (by (ii)) are global by assumption, we only need to show that $E(\mathfrak{v}) \cap \mathfrak{w} \subseteq E(\mathfrak{w})$. But this follows immediately from the claim above. (iv): This is a consequence of [Law99, Prop. 6.3]. □

The operator lift of (R), still assuming that X is linear and again using Assumption (I) (full control), is defined by

$$L'(t) = X_{K(t)} L(t), \quad L(0) = \mathbb{1} \in \mathrm{GL}(\mathfrak{a})\,. \tag{RL}$$

Remark 2.2.2. *Although the control systems on* $\mathfrak{p}$ *and* $\mathfrak{a}$ *are equivalent, the same is* not *true for the operator lifts on* $\mathrm{GL}(\mathfrak{p})$ *and* $\mathrm{GL}(\mathfrak{a})$. *More precisely, a reachable transformation in* $\mathrm{GL}(\mathfrak{a})$ *will in general not correspond to the restriction of some reachable transformation in* $\mathrm{GL}(\mathfrak{p})$. *In this sense the operator lift* (RL) *is a somewhat artificial construction which, however, turns out to be useful in practice, see for instance Chapter 8.*

Lemma 2.2.3. *Let* $X \in \mathfrak{gl}(\mathfrak{p})$ *be a linear vector field on* $\mathfrak{p}$. *Then it holds that*

$$\langle\mathfrak{X}\rangle_{\mathsf{wedge}} := \langle \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_{\mathbf{K}}(X) \circ \iota_{\mathfrak{a}}\rangle_{\mathsf{wedge}} = \Pi_{\mathfrak{a}} \circ \langle \mathrm{ad}_{\mathfrak{k}}, \mathrm{Ad}^{\star}_{\mathbf{K}}(X)\rangle_{\mathsf{wedge}} \circ \iota.$$

Proof. This follows immediately from the definition of the induced vector fields. □

This result relates the operator lifts (FL) and (RL) via the wedges they generate since the wedge $\langle \mathrm{ad}_{\mathfrak{k}}, \mathrm{Ad}^{\star}_{\mathbf{K}}(X)\rangle_{\mathsf{wedge}}$ on the right-hand side is related to $\langle\Omega\rangle_{\mathsf{LW}}$ and $\langle\Omega\rangle_{\mathsf{LS}}$ in the proposition above. Whether similar relations exist for $\langle\mathfrak{X}\rangle_{\mathsf{LW}}$ and $\langle\mathfrak{X}\rangle_{\mathsf{LS}}$ is an open question.

2.3 The Equivalence Theorem

The main goal of this chapter is to prove that the reduced control system (R) on $\mathfrak{a}$ is in some sense equivalent to the full control-affine system (F) on $\mathfrak{p}$. Instead of giving a general definition of equivalence in advance, each of our main results will contain the precise sense in which the equivalence in question is to be understood. First we will give a local equivalence result in Proposition 2.3.6, before proving the global case. This will be separated into a projection, see Theorem 2.3.8, and an approximate lift, see Theorem 2.3.14. The results are combined and slightly generalized to obtain the final Equivalence Theorem 2.3.16.

Before we begin, let us state the following important geometric consequences of the orthogonality of the symmetric Lie algebra. Let $x \in \mathfrak{p}$ and consider $\mathfrak{p}_x := \{y \in \mathfrak{p} : [x, y] = 0\}$, i.e. the commutant of x in $\mathfrak{p}$. A key fact is that the orbit of $\mathbf{K}$ through x, denoted by $\mathbf{K}x$, is orthogonal to $\mathfrak{p}_x$ at x. Since the tangent space of the orbit at x can be identified with $\mathrm{ad}_{\mathfrak{k}}(x)$, we can define the orthogonal projection $\Pi_x : \mathfrak{p} \to \mathfrak{p}$ with image $\mathfrak{p}_x$ and kernel $\mathrm{ad}_{\mathfrak{k}}(x)$, yielding the useful equation

$$\Pi_x(\mathrm{ad}_k(x)) = 0 \quad \text{for all } x \in \mathfrak{p},\ k \in \mathfrak{k}. \tag{2.1}$$

Local Equivalence

Here we show a local equivalence result which illustrates why the definition of the reduced control system is natural. Note that the global equivalence result proven later does not make use of this local result.

Unless stated otherwise, we make no assumption on the smoothness or boundedness of the drift X. We start with a simple but quite useful consequence of the orthogonality relation (2.1).

Lemma 2.3.1. *Let $p : [0,T] \to \mathfrak{p}$ be any path satisfying (F) at some $t_0 \in [0,T]$ and let $K \in \mathbf{K}$ be arbitrary. Then it holds that*

$$\mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) = \Pi_{\mathrm{Ad}_K^{-1}(p(t_0))} \circ \mathrm{Ad}_K^{\star}(X) \circ \mathrm{Ad}_K^{-1}(p(t_0))\,.$$

Proof. This is a simple computation:

$$\begin{aligned}\mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) &= \mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)} \circ X(p(t_0)) \\ &= \Pi_{\mathrm{Ad}_K^{-1}(p(t_0))} \circ \mathrm{Ad}_K^{-1} \circ X(p(t_0)) \\ &= \Pi_{\mathrm{Ad}_K^{-1}(p(t_0))} \circ \mathrm{Ad}_K^{\star}(X) \circ \mathrm{Ad}_K^{-1}(p(t_0))\,,\end{aligned}$$

where the first equality uses (F) & (2.1), the second one uses Lemma 1.A.24 (iii), and the third one uses the definition of the pullback. □

The following lemma is mostly a convenient restatement of Proposition 1.4.5 applied to solutions of (F).

Lemma 2.3.2. *Let $p : [0,T] \to \mathfrak{p}$ be any path satisfying (F) at some $t_0 \in [0,T]$. Then the following statements hold.*

(i) There is $\mathrm{a} : [0,T] \to \mathfrak{a}$ differentiable at t_0 which satisfies $\pi \circ p = \pi \circ \mathrm{a}$.

Now fix any a as in (i).

(ii) For any $b : [0,T] \to \mathfrak{a}$ differentiable at t_0 satisfying $\pi \circ p = \pi \circ b$ there is some $w \in \mathbf{W}$ such that[9] *$(b(t_0), b'(t_0)) = w \cdot (\mathrm{a}(t_0), \mathrm{a}'(t_0))$.*

(iii) For any $b : [0,T] \to \mathfrak{a}$ differentiable at t_0 satisfying $\pi \circ p = \pi \circ b$ there is some $K \in \mathbf{K}$ such that

$$b(t_0) = \mathrm{Ad}_K^{-1}(p(t_0)) \;\text{ and }\; b'(t_0) = \mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) = X_K(b(t_0))\,.$$

(iv) For any $K \in \mathbf{K}$ such that $\mathrm{Ad}_K^{-1}(p(t_0)) \in \mathfrak{a}$ and $\mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) \in \mathfrak{a}$ it holds that

$$(\mathrm{Ad}_K^{-1}(p(t_0)), \mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0))) = w \cdot (\mathrm{a}(t_0), \mathrm{a}'(t_0))$$

for some $w \in \mathbf{W}$. Moreover, it holds that

$$\mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) = X_K \circ \mathrm{Ad}_K^{-1}(p(t_0))\,.$$

Proof. (i) and (ii) follow immediately from Proposition 1.4.5. For (iii) we use the same proposition, together with Lemma 2.3.1, to obtain $b'(t_0) = \Pi_{b(t_0)} \mathrm{Ad}_K^{\star}(X)(b(t_0))$. Since $b'(t_0)$ is diagonal by assumption, this implies that $b'(t_0) = X_K(b(t_0))$. Finally, the first part of (iv) is a direct consequence of (ii), (iii), and Corollary 1.A.49. The second part follows from Lemma 2.3.1 as before. □

[9] We define $w \cdot (\mathrm{a}(t), \mathrm{a}'(t)) = (w \cdot \mathrm{a}(t), w \cdot \mathrm{a}'(t))$, which naturally extends the action of $\mathbf{W}$ to the tangent bundle $T\mathfrak{a}$.

As a converse we have the following lifting result:

Lemma 2.3.3. *Let $a_0 \in \mathfrak{a}$ and $K \in \mathbf{K}$ as well as $t_0 \in [0,T]$ be given. Assume that X is continuous on a neighborhood of $\mathrm{Ad}_K(a_0)$. Then there exists $p : [0,T] \to \mathfrak{p}$ which satisfies (F) on a neighborhood of t_0 and satisfies $p(t_0) = \mathrm{Ad}_K(a_0)$.*

Proof. By continuity of X on a neighborhood of $\mathrm{Ad}_K(a_0)$, Peano's Theorem [Tes12, Thm. 2.19], guarantees the existence of a solution $p : (t_0-\varepsilon, t_0+\varepsilon) \to \mathfrak{p}$ to $p'(t) = X(p(t))$ satisfying $p(0) = \mathrm{Ad}_K(a_0)$. This is clearly a solution to (F) at t_0 with all controls set to zero. □

These results motivate the following definition:

$$\widetilde{\mathsf{derv}}(a_0) = \{a'(t_0) : \begin{smallmatrix} a:[0,T]\to\mathfrak{a} \text{ differentiable at some } t_0\in[0,T],\ a(t_0)=a_0, \text{ and } \pi\circ a=\pi\circ p, \\ \text{where } p:[0,T]\to\mathfrak{p} \text{ satisfies (F) on a neighborhood of } t_0 \end{smallmatrix}\}, \tag{2.2}$$

which is the set of all possible derivatives at $a_0 \in \mathfrak{a}$ of solutions to (F). Now our previous results allow us to describe $\widetilde{\mathsf{derv}}(a_0)$ explicitly.

Lemma 2.3.4. *Assume that X is continuous. Then it holds for every $a_0 \in \mathfrak{a}$ that*

$$\widetilde{\mathsf{derv}}(a_0) = \{X_K(a_0) : K \in \mathbf{K} \text{ such that } \Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0) \in \mathfrak{a}\}.$$

Proof. "$\subseteq$": Let a, p and t_0 be as in (2.2). Lemma 2.3.2 (iii) and its proof show that there is some $K \in \mathbf{K}$ such that $a'(t_0) = \Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0)$. So $\Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0) \in \mathfrak{a}$ as desired.

"$\supseteq$": By Lemma 2.3.3 there exists $p : [0,T] \to \mathfrak{p}$ solving (F) in a neighborhood of t_0 such that $p(t_0) = \mathrm{Ad}_K(a_0)$. Lemma 2.3.1 shows that $\mathrm{Ad}_K^{-1} \circ \Pi_{p(t_0)}(p'(t_0)) = \Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0)$. Hence by Lemma 2.3.2 (i) and (iv) there is some $a : [0,T] \to \mathfrak{a}$ satisfying $\pi \circ a = \pi \circ p$ with $a(t_0) = a_0$ and $a'(t_0) = \Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0)$, as desired. □

For the next result we will make use of Kostant's famous convexity theorem [Kos73]. Recall that $\mathbf{K}_{a_0}$ and $\mathbf{W}_{a_0}$ denote the stabilizers of a_0 in $\mathbf{K}$ and $\mathbf{W}$, respectively.

Lemma 2.3.5. *Let $a_0 \in \mathfrak{a}$ and $K \in \mathbf{K}$ be given. Then there exists $\tilde{K} \in \mathbf{K}$ which satisfies $K\mathbf{K}_{a_0} = \tilde{K}\mathbf{K}_{a_0}$ such that $X_{\tilde{K}}(a_0) = \Pi_{a_0} \circ \mathrm{Ad}_{\tilde{K}}^\star(X)(a_0) \in \mathfrak{a}$. Moreover, for any such $\tilde{K}$ it holds that $\{X_{KL}(a_0) : L \in \mathbf{K}_{a_0}\} = \mathrm{conv}(\mathbf{W}_{a_0} \cdot X_{\tilde{K}}(a_0))$.*

Proof. For the existence of $\tilde{K}$ let us write $\tilde{K} = KL$ with corresponding element $L \in \mathbf{K}_{a_0}$. Then, using Lemma 1.A.24 (iii) we compute

$$\Pi_{a_0} \circ \mathrm{Ad}_{KL}^\star(X)(a_0) = \Pi_{a_0} \circ \mathrm{Ad}_L^{-1} \circ \mathrm{Ad}_K^\star(X)(a_0) = \mathrm{Ad}_L^{-1} \circ \Pi_{a_0} \circ \mathrm{Ad}_K^\star(X)(a_0)\,,$$

and hence by Lemma 1.A.45 there is some $L \in \mathbf{K}_{a_0}$ such that this expression lies in $\mathfrak{a}$. Next we compute, with $\tilde{K}$ as above, for arbitrary $M \in \mathbf{K}_{a_0}$

$$\begin{aligned} X_{\tilde{K}M}(a_0) &= \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_{\tilde{K}M}^\star(X)(a_0) = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_M^{-1} \circ \mathrm{Ad}_{\tilde{K}}^\star(X)(a_0) \\ &= \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_M^{-1} \circ \Pi_{a_0} \circ \mathrm{Ad}_{\tilde{K}}^\star(X)(a_0) = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_M^{-1} \circ X_{\tilde{K}}(a_0)\,. \end{aligned}$$

The result now follows from Lemma 1.A.44 and Kostant's convexity theorem. □

Now we are ready to prove a local equivalence result relating $\widetilde{\mathsf{derv}}$ and derv, showing that they are "almost the same".

Proposition 2.3.6. *Let $a_0 \in \mathfrak{a}$ and $K \in \mathbf{K}$ be arbitrary and assume that X is continuous on a neighborhood of $\mathrm{Ad}_K(a_0)$. It holds that*

$$\widetilde{\mathsf{derv}}(a_0) \subseteq \mathsf{derv}(a_0) = \bigcup_{v \in \widetilde{\mathsf{derv}}(a_0)} \mathrm{conv}(\mathbf{W}_{a_0} v) \subseteq \mathrm{conv}(\widetilde{\mathsf{derv}}(a_0))$$

so, in particular, $\mathrm{conv}(\mathsf{derv}(a_0)) = \mathrm{conv}(\widetilde{\mathsf{derv}}(a_0))$, *and if a_0 is regular, then it holds that* $\mathsf{derv}(a_0) = \widetilde{\mathsf{derv}}(a_0)$. *Moreover, assuming that $X(0) \in \mathfrak{a}$, it holds at the origin of $\mathfrak{a}$ that* $\mathbf{W}X(0) = \widetilde{\mathsf{derv}}(0) \subseteq \mathsf{derv}(0) = \mathrm{conv}(\mathbf{W}X(0))$.

Proof. The first inclusion is due to Lemma 2.3.4, the first equality follows from Lemma 2.3.5 together with Lemma 2.3.4, and the rest is straightforward. □

Proposition 2.3.6 tells us that the definition of the set of achievable derivatives $\mathsf{derv}(a)$ is "too large" whenever a is non-regular, but only in a negligible way since the convex hulls coincide (recall the relaxation result [AC84, Ch. 2.4 Thm. 2]). See Example 2.3.12 for a consequence of this fact.

Projection

Our main results describe the equivalence of the control-affine system (F) on $\mathfrak{p}$ and the reduced control system (R) on $\mathfrak{a}$. The first direction is projecting from $\mathfrak{p}$ to $\mathfrak{a}$. This means that given a solution $p : [0,T] \to \mathfrak{p}$ we are looking for a solution $a : [0,T] \to \mathfrak{a}$ satisfying $\pi \circ a = \pi \circ p$. Since semisimple orthogonal symmetric Lie algebras generally correspond to some kind of matrix diagonalization, this step could also be called diagonalization.

We start with a special case in which we are given a diagonalization of p.

Lemma 2.3.7. *Let $p : [0,T] \to \mathfrak{p}$ satisfy (F) such that there exist differentiable functions $\mathrm{a} : [0,T] \to \mathfrak{a}$ and $K : [0,T] \to \mathbf{K}$ with $p(t) = \mathrm{Ad}^{-1}_{K(t)}(\mathrm{a}(t))$. Then $\mathrm{a}'(t) = X_{K(t)}(\mathrm{a}(t))$ for all $t \in [0,T]$; in particular, a satisfies (R) everywhere.*

Proof. By differentiating the solution p and considering the part orthogonal to the orbit we obtain $a'(t) = \mathrm{Ad}^{-1}_{K(t)} \circ \Pi_{p(t)}(p'(t))$, see Lemma 1.4.1. Then the result follows from Lemma 2.3.2 (iv). □

Now let us consider the general case. The first difficulty is that a is not uniquely determined. This will be remedied by choosing a (closed) Weyl chamber $\mathfrak{w} \subset \mathfrak{a}$ and requiring that a take values in $\mathfrak{w}$. A consequence of this is that we may introduce kinks where the solution hits the boundary of $\mathfrak{w}$. Fortunately, this still leaves a absolutely continuous which allows us to show that a satisfies the differential inclusion almost everywhere.

Theorem 2.3.8 (Equivalence Theorem: Projection). *Let $p : [0,T] \to \mathfrak{p}$ be a solution to the control system (F) and let $a^{\downarrow} : [0,T] \to \mathfrak{w}$ be the unique path which satisfies $\pi \circ a^{\downarrow} = \pi \circ p$. Then $a^{\downarrow}$ is a solution to the reduced control system (R) (and hence also to (R̄)).*

Proof. By Proposition 1.3.1 (v) the path $a^{\downarrow}$ is absolutely continuous. Let $J \subseteq [0,T]$ be the subset on which both p and $a^{\downarrow}$ are differentiable. This set still has full (Lebesgue) measure. For $t_0 \in J$, by Lemma 2.3.2 (iii) it holds that $(a^{\downarrow})'(t_0) = X_K(a^{\downarrow}(t_0))$ for some $K \in \mathbf{K}$, which proves that $a^{\downarrow}$ satisfies the differential inclusion (R′) almost everywhere. By Filippov's theorem, see [Smi02, Thm. 2.3], $a^{\downarrow}$ is a solution to (R). □

Alternatively, one can prove Theorem 2.3.8 without making use of Filippov's theorem as follows:

By Proposition 1.6.13 there exists a (Lebesgue) measurable function $K : [0,T] \to \mathbf{K}$ such that $a^{\downarrow}(t) = \mathrm{Ad}_{K(t)}(p(t)) \in \mathfrak{a}$ and $(a^{\downarrow}(t))' = \mathrm{Ad}_{K(t)}(\Pi_{p(t)}(p'(t))) \in \mathfrak{a}$ almost everywhere. Hence by the proof of Lemma 2.3.7 it holds almost everywhere that $(a^{\downarrow})'(t) = X_{K(t)}(a^{\downarrow}(t))$ and so $a^{\downarrow}$ solves (R).

In Chapter 1 we proved several results which show that if $p : [0,T] \to \mathfrak{p}$ has a certain smoothness, then, in certain cases one can choose $\mathrm{a} : [0,T] \to \mathfrak{a}$ satisfying $\pi \circ \mathrm{a} = \pi \circ p$ with the same smoothness. This allows us to strengthen the result above in some instances.

Proposition 2.3.9. *Let $p : [0,T] \to \mathfrak{p}$ be a solution to the control system (F). Then there exists $\mathrm{a} : [0,T] \to \mathfrak{a}$ satisfying $\pi \circ \mathrm{a} = \pi \circ p$ and solving (R) such that:*

(i) if p is (continuously) differentiable, then a can be chosen (continuously) differentiable;

(ii) if p is C^ℓ, for $\ell = 2, \dots, \infty$, and regular, then a can be chosen C^ℓ;

(iii) if p is real analytic (denoted C^ω), then a can be chosen real analytic.

Moreover, in (ii) and (iii) we can choose a as before and $K : [0,T] \to \mathbf{K}$ such that $\mathrm{a} = \mathrm{Ad}^{-1}_{K(t)}(p(t))$ and such that K is C^ℓ (resp. real analytic). Then it holds that $\mathrm{a}'(t) = X_{K(t)}(\mathrm{a}(t))$, i.e. a solves (R) with control function K.

Proof. Item (ii) follows from Proposition 1.4.12, (iii) follows from Theorem 1.5.7, and (i) follows from Theorem 1.4.9, in each case using Lemma 2.3.7. In the cases (ii) and (iii), the same results provide $K : [0,T] \to \mathbf{K}$, and again Lemma 2.3.7 shows that $a'(t) = X_{K(t)}(a(t))$. □

Approximate Lift

The task of this section is the following: given a solution to the reduced control system (R), construct a solution to the full control system (F) which is, at least approximately, a lift of the former.

For regular solutions to the reduced control system we can construct an exact lift as well as a corresponding control function $k : [0,T] \to \mathfrak{k}$. In particular, if the control directions $k_1, \dots, k_m \in \mathfrak{k}$ of (F) span $\mathfrak{k}$ (which we called Assumption (I')), then one easily finds the corresponding control functions u_j.

To properly state the result, we have to define an appropriate inverse of $\mathrm{ad}_x : \mathfrak{k} \to \mathfrak{p}$ for $x \in \mathfrak{p}$. Note that the kernel of this map is exactly the commutant $\mathfrak{k}_x$, and due to orthogonality the image is $\mathfrak{p}_x^{\perp}$. Hence there is a unique inverse $\mathrm{ad}_x^{-1} : \mathfrak{p}_x^{\perp} \to \mathfrak{k}_x^{\perp}$. Indeed, this is nothing but a restriction of the Moore–Penrose pseudoinverse of ad_x.

Proposition 2.3.10. *Let X be C^ℓ and let $\mathrm{a} : [0,T] \to \mathfrak{a}$ be a solution to the reduced control system (R) with C^r ($r \geq 1$) control function $K : [0,T] \to \mathbf{K}$ such that a takes only regular values. If we set $p(t) = \mathrm{Ad}_{K(t)}(\mathrm{a}(t))$ and define $k : [0,T] \to \mathfrak{k}$ by*

$$k(t) = K'(t)K^{-1}(t) + \mathrm{ad}^{-1}_{p(t)} \circ \Pi^{\perp}_{p(t)} \circ X(p(t)) \,,$$

then k is of class $C^{\min(\ell, r-1)}$, and p satisfies $p'(t) = (\mathrm{ad}_{k(t)} + X)(p(t))$.

Proof. By differentiating[10] we get that

$$p'(t) = \mathrm{ad}_{K'(t)K^{-1}(t)}(p(t)) + \mathrm{Ad}_{K(t)}(a'(t)) = \mathrm{ad}_{K'(t)K^{-1}(t)}(p(t)) + \Pi_{p(t)} \circ X(p(t)) \,,$$

[10] If $K : I \to \mathbf{K}$ is differentiable at some $t \in I$, then $\frac{d}{dt}\,\mathrm{Ad}_{K(t)} = \mathrm{ad}_{K'(t)K(t)^{-1}} \circ \mathrm{Ad}_{K(t)} = \mathrm{Ad}_{K(t)} \circ \mathrm{ad}_{K(t)^{-1}K'(t)}$.

since $\mathrm{Ad}_{K(t)}(a'(t)) = \mathrm{Ad}_{K(t)} \circ \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{-1}_{K(t)} \circ X \circ \mathrm{Ad}_{K(t)}(a(t)) = \Pi_{p(t)} \circ X(p(t))$, where we used that $p(t)$ is regular to introduce $\Pi_{p(t)}$. Hence

$$\begin{aligned}(\mathrm{ad}_{k(t)} + X)(p(t)) &= \mathrm{ad}_{K'(t)K^{-1}(t)}(p(t)) - \Pi^{\perp}_{p(t)}(X(p(t))) + X(p(t)) \\ &= \mathrm{ad}_{K'(t)K^{-1}(t)}(p(t)) + \Pi_{p(t)}(X(p(t))) = p'(t)\,,\end{aligned}$$

as desired. □

The control Hamiltonian in Proposition 2.3.10 has two components. To the *direct control* $K'(t)K^{-1}(t)$, which one might naively expect to do the job, one has to add the *compensating control* $\mathrm{ad}^{-1}_{p(t)} \circ \Pi^{\perp}_{p(t)} \circ X(p(t))$ which deals with the orbital component of $X(p(t))$, cp. Lemma 1.4.1.

In practice one might find the lift p of a without knowing a corresponding control function K for (R). In this case any K diagonalizing p and $\Pi_p \circ X(p)$ will do:

Lemma 2.3.11. *Let $\mathrm{a} : [0,T] \to \mathfrak{a}$ be a regular solution to (R). Assume that $p : [0,T] \to \mathfrak{p}$ and $K : [0,T] \to \mathbf{K}$ satisfy $p(t) = \mathrm{Ad}_{K(t)}(\mathrm{a}(t))$ and $\mathrm{a}'(t) = \mathrm{Ad}^{-1}_{K(t)} \circ \Pi_{p(t)} \circ X(p(t))$. Then $\mathrm{a}'(t) = X_K(t)(\mathrm{a}(t))$.*

Proof. This is straightforward using the definition of X_K and regularity of a: $X_{K(t)}(\mathrm{a}(t)) = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{-1}_{K(t)} \circ X \circ \mathrm{Ad}_{K(t)}(\mathrm{a}(t)) = \mathrm{Ad}^{-1}_{K(t)} \circ \Pi_{p(t)} \circ X(p(t)) = \mathrm{a}'(t)$. □

If we allow for non-regular solutions, an exact lift might not even exist, as shown in the following example.

Example 2.3.12. *To see that approximating solutions cannot be avoided in general, consider a system where $X(0) \neq 0$. Then $p \equiv 0$ is not a solution of (F), but $\mathrm{a} \equiv 0$ is a solution of (R). Indeed, by Kostant's convexity theorem, and assuming that $X(0) \in \mathfrak{a}$, it follows from Proposition 2.3.6 that*

$$\mathsf{derv}(0) = \{\Pi_{\mathfrak{a}} \circ \mathrm{Ad}_K(X(0)) : K \in \mathbf{K}\} = \mathrm{conv}(\mathbf{W}X(0))\,,$$

and hence $\mathsf{derv}(0)$ contains the convex combination $\frac{1}{|\mathbf{W}|}\sum_{w\in\mathbf{W}} w \cdot X(0)$, which equals 0, the unique fixed point of a Weyl group action. Thus $\mathrm{a} \equiv 0$ is a solution to (R). Note also that $\widetilde{\mathsf{derv}}(0)$ contains exactly the vertices of $\mathsf{derv}(0)$.

For this reason, we have to look for an approximate lift in general. Before we prove the existence of such a lift, we need the following technical result.

Lemma 2.3.13. *Let $\mathbf{G}$ be a Lie group and $\mathbf{K}$ be a compact subgroup such that the norm on $\mathfrak{g}$ is invariant under the adjoint action of $\mathbf{K}$. If $\delta : [0,T] \to \mathfrak{g}$ is differentiable, then for every integrable $h : [0,T] \to \mathfrak{k}$ it holds that*

$$\|\delta(t)\| \leq \|\delta(0)\| + \int_0^t \|\mathrm{ad}_{h(s)}(\delta(s)) + \delta'(s)\|\, ds,$$

for all $t \in [0,T]$.

Proof. Let $\phi : [0,T] \to \mathbf{K}$ satisfy $\phi'(t) = \phi(t)h(t)$. We compute (cf. Footnote 10)

$$\begin{aligned}\|\delta(t)\| = \|\mathrm{Ad}_{\phi(t)}(\delta(t))\| &= \Big\|\mathrm{Ad}_{\phi(0)}(\delta(0)) + \int_0^t \frac{d}{ds}(\mathrm{Ad}_{\phi(s)}(\delta(s)))\, ds\Big\| \\ &= \Big\|\mathrm{Ad}_{\phi(0)}(\delta(0)) + \int_0^t \mathrm{Ad}_{\phi(s)} \circ \mathrm{ad}_{h(s)}(\delta(s)) + \mathrm{Ad}_{\phi(s)}(\delta'(s))\, ds\Big\| \\ &\leq \|\delta(0)\| + \int_0^t \|\mathrm{ad}_{h(s)}(\delta(s)) + \delta'(s)\|\, ds\,.\end{aligned}$$

This concludes the proof. □

Finally we can prove:

Theorem 2.3.14 (Equivalence Theorem: Approximate Lift). *Assume that X is locally Lipschitz and linearly bounded[a] and let $\mathrm{a} : [0,T] \to \mathfrak{a}$ be any solution to the reduced control system (R) with control function $K : [0,T] \to \mathbf{K}$. Then $p := \mathrm{Ad}_K(\mathrm{a})$, which is a lift of a to $\mathfrak{p}$, can be approximated by solutions to the original control system (F) to arbitrary precision. More precisely, for every $\varepsilon > 0$ there exists a solution $p_\varepsilon : [0,T] \to \mathfrak{p}$ to (F) such that $\| \mathrm{Ad}_K(\mathrm{a}) - p_\varepsilon\|_\infty \leq \varepsilon$.*

[a] By this we mean that $\|X(v)\| \leq C_1\|v\| + C_2$ for some $C_1, C_2 \geq 0$.

Proof. We start by proving the result under stronger assumptions, and then show that we can weaken the assumptions while maintaining uniform convergence on $[0,T]$.

First we assume that X and K are real analytic and that $\mathrm{a}(0)$ is regular. Moreover, we invoke Assumption (I'), meaning that the control directions $k_1, \dots, k_m$ span $\mathfrak{k}$. Then the solution a is also real analytic since it satisfies $\mathrm{a}'(t) = X_{K(t)}(\mathrm{a}(t))$ and the map $(\mathrm{a}, K) \mapsto X_K(\mathrm{a})$ is real analytic. Since the non-regular points in $\mathfrak{a}$ are formed by a finite union of hyperplanes, a will be regular with finitely many exceptions $t_1, \dots, t_n$ in $[0,T]$. We define the set $J_\varepsilon := [0,T] \setminus \bigcup_{i=1}^n (t_i - \varepsilon, t_i + \varepsilon)$, as well as $p(t) = \mathrm{Ad}_{K(t)}(\mathrm{a}(t))$ and the control function $k_\varepsilon(t) = K'(t)K^{-1}(t) + \mathbf{1}_{J_\varepsilon}(t)\,\mathrm{ad}^{-1}_{p(t)}\,\Pi^\perp_{p(t)}(X(p(t)))$. Note that k_ε is (piecewise, in time) real-analytic and bounded. Hence we can define p_ε as the solution of $p'_\varepsilon(t) = (\mathrm{ad}_{k_\varepsilon(t)} + X)(p_\varepsilon(t))$, with $p_\varepsilon(0) = p(0)$.

By Lemma 2.3.13 we find that $\|p_\varepsilon(t)\| \leq \int_0^t \|X(p_\varepsilon(s))\|ds + \|p_\varepsilon(0)\|$. Since X is linearly bounded and using Grönwall's inequality[11] we obtain $\|p_\varepsilon(t)\| \leq (\|\mathrm{a}(0)\| + tC_2)e^{tC_1}$. In particular there is some $R > 0$ independent of ε such that $\|p(t)\| \leq R$ and $\|p_\varepsilon(t)\| \leq R$ for all $t \in [0,T]$. Restricting to this compact domain, we may assume that X is in fact globally Lipschitz with constant L.

Setting $\delta = p - p_\varepsilon$ we obtain $\delta'(t) = \mathrm{ad}_{k_\varepsilon(t)}(\delta(t)) - X(p_\varepsilon(t)) - \mathrm{ad}_{k_\varepsilon}(t)(p(t)) + p'(t)$ and using Lemma 2.3.13 we get $\|\delta(t)\| \leq \int_0^t \| - X(p_\varepsilon(s)) - \mathrm{ad}_{k_\varepsilon(s)}(p(s)) + p'(s)\|ds$. Using that

$$\begin{aligned} p'(t) - \mathrm{ad}_{k_\varepsilon(t)}(p(t)) &= \mathrm{ad}_{K'(t)K^{-1}(t)}\,p(t) + \Pi_{\mathrm{Ad}_{K(t)}(\mathfrak{a})} \circ X(p(t)) \\ &\quad - \mathrm{ad}_{K'(t)K^{-1}(t)}\,p(t) + \mathbf{1}_{J_\varepsilon}(t)\Pi^\perp_{p(t)} \circ X(p(t)) \\ &= \mathbf{1}_{J_\varepsilon}(t)X(p(t)) + \mathbf{1}_{J^c_\varepsilon}(t)\Pi_{\mathrm{Ad}_{K(t)}(\mathfrak{a})} \circ X(p(t))\,, \end{aligned}$$

where $\mathbf{1}$ denotes the indicator function. We obtain

$$\begin{aligned} \|\delta(t)\| &\leq \int_0^t \|\mathbf{1}_{J_\varepsilon}(s)X(p(s)) - X(p_\varepsilon(s)) + \mathbf{1}_{J^c_\varepsilon}(s)(\Pi_{\mathrm{Ad}_{K(s)}(\mathfrak{a})}X(p(s)) - X(p_\varepsilon(s)))\|ds \\ &\leq \int_0^t L\|\delta(s)\|ds + 2\mu(J^c_\varepsilon)(C_1R + C_2) \end{aligned}$$

where $\|\cdot\|_\infty$ denotes the supremum norm and μ denotes the Lebesgue measure. Finally, we again apply Grönwall's inequality to obtain $\|\delta(t)\| \leq 2\mu(J^c_\varepsilon)(C_1R + C_2)e^{Lt}$ for all $t \in [0,T]$. Since $\mu(J^c_\varepsilon) \to 0$ as $\varepsilon \to 0$, this shows that p_ε converges uniformly to p on $[0,T]$.

Now we show that the result also holds under the more general assumptions. This will follow from a sequence of standard approximations. Let X, a, and K be as in the statement and use Assumptions (I) & (II). Let some $\varepsilon > 0$ be given. Again we define $p(t) = \mathrm{Ad}_{K(t)}(\mathrm{a}(t))$. Now let $K^{(m)}$ be

[11] Recall that Grönwall's inequality states that if $\alpha \geq 0$ is non-decreasing, β, u are continuous on $[0,T]$, and $u(t) \leq \alpha(t) + \int_0^t \beta(s)u(s)ds$ for all $t \in [0,T]$, then $u(t) \leq \alpha(t)\exp(\int_0^t \beta(s)ds)$ for all $t \in [0,T]$.

a sequence of real analytic controls converging uniformly to K and let $a_0^{(m)}$ be a sequence of regular points converging to $a(0)$. Let $a^{(m)}$ be the solution to (R) with initial point $a_0^{(m)}$ and control function $K^{(m)}$. Then by [Son98, Thm. 1] the $a^{(m)}$ converge uniformly to a, and setting $p^{(m)} = \mathrm{Ad}_{K^{(m)}}(a^{(m)})$ we find that the $p^{(m)}$ converge uniformly to p. In particular there is m such that $\|p - p^{(m)}\|_\infty \leq \frac{\varepsilon}{4}$. Now let $X^{(n)}$ be a sequence of real analytic and linearly bounded vector fields converging uniformly on compact subsets to X. Let $a^{(m,n)}$ be the corresponding solutions and $p^{(m,n)} = \mathrm{Ad}_{K^{(m)}}(a^{(m,n)})$. Then by [Kha02, Thm. 3.5] $p^{(m,n)} \to p^{(m)}$ and there is some n such that $\|p^{(m)} - p^{(m,n)}\|_\infty \leq \frac{\varepsilon}{4}$. Now we can use the result proven above to find a solution $p_\varepsilon^{(m,n)}$ to (F) using Assumption (I') such that $\|p^{(m,n)} - p_\varepsilon^{(m,n)}\|_\infty \leq \frac{\varepsilon}{4}$. Finally due to [Liu97] we can drop Assumption (I') and obtain solutions $p_\varepsilon^{(m,n,k)}$ to (F) such that for some k we have $\|p_\varepsilon^{(m,n)} - p_\varepsilon^{(m,n,k)}\|_\infty \leq \frac{\varepsilon}{4}$. Combining all these approximations then yields the result. □

Remark 2.3.15. *In general the control k_ε obtained by setting $\varepsilon = 0$ need not be integrable since the expression $\mathrm{ad}_{p(t)}^{-1}$ typically leads to singularities of order t^{-1} as p passes through a non-regular point. Nevertheless, in can happen that the controls do not explode even as we pass through a non-regular point, cf., e.g, the worked example in Section 2.6.*

Global Equivalence

Theorems 2.3.8 and 2.3.14 together prove the global equivalence of the full and reduced control systems. Since in applications the control system is typically not directly formulated on symmetric Lie algebras, but in some isomorphic manner, the following slight generalization of the complete equivalence result is very useful, see also Section 7.A.

Theorem 2.3.16 (Equivalence Theorem). *Let V be an n-dimensional real inner product space, let $\mathbf{L}$ be a compact Lie group with Lie algebra $\mathfrak{l}$ acting on V, and let Y be a linear vector field on V. Consider the following control system with controls $l_j \in \mathfrak{l}$:*

$$v' = \Big(Y + \sum_{i=j}^{m} u_j(t) l_j\Big) v \tag{2.3}$$

with fast and full control on $\mathbf{L}$. Moreover assume that we have a semisimple orthogonal symmetric Lie algebra $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ with associated pair $(\mathbf{G}, \mathbf{K})$ and maximal Abelian subspace $\mathfrak{a}$. Let $\imath : V \to \mathfrak{p}$ be a linear isometric isomorphism and $\jmath : \mathbf{L} \to \mathrm{Ad}_{\mathbf{K}}$ a surjective Lie group homomorphism such that

$$\jmath(L)\imath(v) = \imath(Lv). \tag{2.4}$$

Then the control system is equivalent in the sense of Theorems 2.3.8 and 2.3.14 to the following reduced control system on $W := \imath^{-1}(\mathfrak{a})$:

$$w'(t) = Y_{L(t)}(w(t)) \tag{2.5}$$

where $Y_L = \Pi_W \circ L^\star(Y) \circ \iota$, with $\Pi_W : V \to W$ the orthogonal projection onto W, $\iota : W \to V$ the inclusion, and $(\cdot)^\star$ the pullback.

Proof. Let $\jmath_\star := D\jmath(e)$ is the surjective Lie algebra homomorphism $\mathfrak{l} \to \mathrm{ad}_{\mathfrak{k}}$ corresponding to $\jmath$. By surjectivity there are $k_i \in \mathfrak{k}$ such that $\jmath_\star(l_i) = \mathrm{ad}_{k_i}$ for all $i = 1, \ldots, m$. By differentiating (2.4) we get

$\jmath_\star(l_i)\imath(v) = \imath(l_i(v))$. Let $X := \imath_\star(Y)$, where $\imath_\star$ denotes the pushforward, be the drift vector field on $\mathfrak{p}$. If $\jmath(L) = \mathrm{Ad}_K$, then

$$\imath^\star(X_K) = \imath^{-1} X_K \imath = \imath^{-1}\Pi_{\mathfrak{a}} \mathrm{Ad}_K^{-1} X \mathrm{Ad}_K \imath = \Pi_W \imath^{-1} \mathrm{Ad}_K^{-1} X \mathrm{Ad}_K \imath = \Pi_W (\imath L)^{-1} X \imath L = Y_L.$$

The remainder of the proof is split into two parts, the projection and the lift.

We begin with the projection. Let $v : [0,T] \to V$ be a solution to (2.3). We want to show that $v^\downarrow : [0,T] \to W$, defined by $\imath(v^\downarrow) = \imath(v)^\downarrow$, is a solution to (2.5). We get that

$$\frac{d}{dt}\imath(v) = \imath(v') = \imath\Big(\Big(Y + \sum_{i=1}^m u_i l_i\Big)v\Big) = \Big(X + \sum_{i=1}^m \mathrm{ad}_{k_i}\Big)\imath(v)$$

almost everywhere. Hence $\imath(v)$ is a solution of the corresponding control system on $\mathfrak{p}$, and we may apply Theorem 2.3.8 to obtain that $\imath(v)^\downarrow$ is a solution of the reduced control system on $\mathfrak{a}$, more explicitly, for almost every $t \in [0,T]$ there is some $K \in \mathbf{K}$ such that $\frac{d}{dt}\imath(v^\downarrow(t)) = X_K(\imath(v^\downarrow(t)))$. Next we show that $v^\downarrow$ solves (2.5). Indeed for the same t we obtain by linearity of $\imath$ that $(v^\downarrow)' = \imath^\star(X_K)(v^\downarrow) = Y_L v^\downarrow$ for some $L \in \mathbf{L}$. This concludes the projection part of the proof.

Conversely, assume that we have a solution $w : [0,T] \to W$ to the reduced system (2.5). Again we find for almost all $t \in [0,T]$ some $K \in \mathbf{K}$ such that $\frac{d}{dt}\imath(w(t)) = \imath(w'(t)) = X_K(\imath(w(t)))$, and so $a := \imath(w)$ solves the corresponding control system on $\mathfrak{a}$. Hence there exists a corresponding control function $K : [0,T] \to \mathbf{K}$ which is measurable. Using Theorem 2.3.14 we find approximate lifted solutions $p_\varepsilon : [0,T] \to \mathfrak{p}$ with $\|p_\varepsilon - \mathrm{Ad}_K(a)\|_\infty \leq \infty$. As above we can compute with $v_\varepsilon := \imath^{-1}(p_\varepsilon)$

$$\frac{d}{dt}v_\varepsilon = \imath^{-1}(p_\varepsilon') = \imath^{-1}\Big(\Big(X + \sum_{i=1}^m u_i \,\mathrm{ad}_{k_i}\Big)p_\varepsilon\Big) = \Big(\imath^\star(X) + \sum_{i=1}^m u_i l_i\Big)v_\varepsilon$$

and see that v_ε is a solution to (2.3). Since $\imath$ is an isometry, for any measurable lift L of Ad_K along $\jmath$ (which exists due to Lemma 1.6.5) we get

$$\|v_\varepsilon - Lw\|_\infty = \|\imath(v_\varepsilon - Lw)\|_\infty = \|\imath(v_\varepsilon) - \jmath(L)\imath(w)\|_\infty = \|p_\varepsilon - \mathrm{Ad}_K(a)\|_\infty \leq \varepsilon,$$

so v_ε is also an ε-approximation. This concludes the proof. □

2.4 Reachability, Stabilizability and Much More

The equivalence results proven above allow us to easily deduce several useful consequences on important control theoretic notions like reachability, stabilizability, controllability, and accessibility. In each case we describe how the notions in the full and reduced control systems are related. Some additional basic properties are collected in Appendix 2.A for reference.

Speed Limit

One of the reasons why the original control system (F) is difficult to work with is the presence of unbounded controls, and the resulting fact that there are points in the state space which are far apart but can be joined in an arbitrarily short amount of time. Since these are exactly the points which are identified in the reduced control system, this cannot occur anymore. Indeed, we can define the *speed limit* $c : \mathfrak{a} \to \mathbb{R}_{\geq 0}$ by

$$c(a) = \max_{K \in \mathbf{K}} \|X_K(a)\|.$$

Then we have the following result:

Proposition 2.4.1. *If the drift X is continuous, then the speed limit c is well-defined and continuous. In particular, c is bounded on bounded subsets of $\mathfrak{a}$. Given any solution $\mathrm{a} : [0,T] \to \mathfrak{a}$ to ($\bar{\mathsf{R}}$) such that $c(\mathrm{a}(t)) \neq 0$, it holds that*

$$T \geq \int_{\mathrm{a}(0)}^{\mathrm{a}(T)} \frac{\|da\|}{c(a)} \geq \frac{\ell(\mathrm{a})}{\max_{t\in[0,T]} c(\mathrm{a}(t))} \geq \frac{\|\mathrm{a}(T) - \mathrm{a}(0)\|}{\max_{t\in[0,T]} c(\mathrm{a}(t))},$$

where $\ell(a)$ denotes the length of a.

Proof. As $\mathbf{K}$ is compact and $K \mapsto X_K(a)$ is continuous, the image is also compact and hence c is well defined. Since all the vector fields X_K are continuous, so is the map $f : \mathbf{K} \times (\overline{B_\varepsilon(a_0)} \cap \mathfrak{a}) \to \mathbb{R}$, $(K, a) \mapsto \|X_K(a)\|$ for all $a_0 \in \mathfrak{a}$, $\varepsilon > 0$. In particular f is uniformly continuous which readily implies continuity of $a \mapsto \max_{K\in\mathbf{K}} \|X_K(a)\| = c(a)$. The lower bound on T follows immediately. □

Reachability

We start with the *reachable set of a_0 at time T* for (R). We denote

$$\mathsf{reach}_{\mathsf{R}}(a_0, T) := \{\mathrm{a}(T) : \mathrm{a} : [0,T] \to \mathfrak{a} \text{ solves } (\mathsf{R}),\ \mathrm{a}(0) = a_0\}$$

for any $T \geq 0$. By $\mathsf{reach}_{\mathsf{R}}(a_0) := \bigcup_{T\geq 0} \mathsf{reach}_{\mathsf{R}}(a_0, T)$ we denote the *(all-time) reachable set of a_0*, and we define the *reachable set of a_0 up to time T* by $\mathsf{reach}_{\mathsf{R}}(a_0, [0,T]) := \bigcup_{t\in[0,T]} \mathsf{reach}_{\mathsf{R}}(a_0, t)$ for any $T \geq 0$. The definitions for the control systems (F), (R'), ($\bar{\mathsf{R}}$), (FL), and (RL) are analogous.

Remark 2.4.2. *Note that, although the reduced control system (R) is symmetric under the Weyl group action, the reachable set in general does not have the same symmetry as it depends on the initial state. However, due to Proposition 2.A.4, if the solution starts in the Weyl chamber $\mathfrak{w}$, then it holds that $\pi(\mathsf{reach}_{\mathsf{R}}(a_0, T)) = \pi(\mathsf{reach}_{\mathsf{R}}(a_0, T) \cap \mathfrak{w})$. Together with Proposition 2.4.3 below this shows that all relevant information concerning reachability is held in the Weyl chamber which contains the initial state.*

The equivalence results of Section 2.3 are formulated at the level of solutions, and they immediately imply the equivalence of reachable sets up to closure and $\mathbf{K}$-orbits.

Proposition 2.4.3. *Assume that X is locally Lipschitz and linearly bounded. Let $T > 0$ and $p_0 \in \mathfrak{p}$, $a_0 \in \mathfrak{a}$ with $\pi(p_0) = \pi(a_0)$ be given. Then it holds that*

$$\mathsf{reach}_{\mathsf{F}}(p_0, T) \subseteq \mathrm{Ad}_{\mathbf{K}}(\mathsf{reach}_{\bar{\mathsf{R}}}(a_0, T)) \subseteq \overline{\mathsf{reach}_{\mathsf{F}}(p_0, T)},$$

where the reachable sets refer the the control-affine system (F) on $\mathfrak{p}$ and the relaxed control system ($\bar{\mathsf{R}}$) on $\mathfrak{a}$. In particular the closures coincide:

$$\overline{\mathsf{reach}_{\mathsf{F}}(p_0, T)} = \mathrm{Ad}_{\mathbf{K}}(\overline{\mathsf{reach}_{\bar{\mathsf{R}}}(a_0, T)})\,.$$

Finally, all statements remain true if we substitute ($\bar{\mathsf{R}}$) with (R).

Proof. We prove the result only for (R) since the proof for ($\bar{\mathsf{R}}$) is analogous. First let $p : [0,T] \to \mathfrak{p}$ be a solution to (F). By Theorem 2.3.8 we obtain a solution $a^{\downarrow} : [0,T] \to \mathfrak{w}$ with $\pi(a(T)) = \pi(p(T))$ to (R). This proves the first inclusion. Conversely, let $a : [0,T] \to \mathfrak{a}$ be a solution to (R) and let $p_1, p_2 \in \mathfrak{p}$ be such that $\pi(p_1) = \pi(\mathrm{a}(0))$ and $\pi(p_2) = \pi(\mathrm{a}(T))$. Due to Theorem 2.3.14 there exists for every $\varepsilon > 0$ a solution $p : [0,T] \to \mathfrak{p}$ to (F) such that $d(\pi(\mathrm{a}(t)), \pi(p(t))) \leq \varepsilon$ where d refers to the quotient metric induced by π. Now let some $\varepsilon > 0$ be given and let $K_1, K_2 \in \mathbf{K}$ be such that $\mathrm{Ad}_{K_1}(p(0))$ is ε-close

to p_1 and such that $\mathrm{Ad}_{K_2}(p(T))$ is ε-close to p_2. By approximately implementing $\mathrm{Ad}_{K_1}^{-1}$ on $[0,\varepsilon]$ and Ad_{K_2} on $[T-\varepsilon,T]$ we can find a solution to (F) which equals p on $[\varepsilon,T-\varepsilon]$ and approximately starts at p_1 and approximately ends at p_2. Using [Son98, Thm. 1] as in the proof of Theorem 2.3.14 one can deduce that $p_2 \in \overline{\mathsf{reach}_{\mathsf{F}}(p_1,T)}$. □

Note that the analogous result is true for the all-time reachable sets.

Now consider X linear. We can also use the operator lift (RL) to understand reachability in the reduced system (R). Indeed, it is clear that $b \in \mathsf{reach}_{\mathsf{R}}(a,T)$ if and only if there is some $L \in \mathsf{reach}_{\mathsf{RL}}(\mathbb{1},T)$ such that $La = b$. In fact it holds that $\overline{\mathsf{reach}_{\mathsf{RL}}(\mathbb{1})}$ is the Lie subsemigroup of $\mathrm{GL}(\mathfrak{a})$ generated by $\mathfrak{X}$, see [Law99, Prop. 6.2]. Proposition 2.4.1 shows that $\mathsf{reach}_{\mathsf{R}}(a_0,[0,T])$ is bounded. If X is Lipschitz, then Proposition 2.A.6 (ii) guarantees compactness of $\mathsf{reach}_{\bar{\mathsf{R}}}(a_0,[0,T])$, cp. [BP04, Thm. 3].

Stabilizability

In practice one often wants to keep the system close to a certain state, i.e. one wants to stabilize the state. We define the *set of stabilizable states* with respect to (R), denoted $\mathsf{stab}_{\mathsf{R}}$, as follows: a point $a_0 \in \mathfrak{a}$ is in $\mathsf{stab}_{\mathsf{R}}$ if for all $T>0$ and all $\varepsilon>0$ there is a solution $a:[0,T]\to\mathfrak{a}$ to (R) with $a(0)=a_0$ and which takes values in $B_\varepsilon(a_0)$. Moreover, we say that a point a_0 is *strongly stabilizable*[12] if the constant path $a\equiv a_0$ is a solution to (R). The definition for the other control systems is analogous. Note that we only consider open-loop controls here and that we are not using feedback.

Lemma 2.4.4. *Assume that X is Lipschitz. Given any point $a_0\in\mathfrak{a}$ the following statements hold:*

(i) *a_0 is strongly stabilizable w.r.t (R) if and only if $0\in\mathsf{derv}(a_0)$;*

(ii) *a_0 is stabilizable w.r.t (R) if and only if $0\in\mathrm{conv}(\mathsf{derv}(a_0))$.*

In fact these statements hold true for all continuous differential inclusions with closed values.

Proof. (i): If $a\equiv a_0$ is a solution to (R), then $0=a'(t)\in\mathsf{derv}(a_0)$ for almost all $t\in[0,T]$. Conversely, if $0\in\mathsf{derv}(a_0)$, then $a\equiv a_0$ is a solution to (R). (ii): If $0\in\mathrm{conv}(\mathsf{derv}(a_0))$ then $a\equiv a_0$ is a solution to ($\bar{\mathsf{R}}$). By the Relaxation Theorem [AC84, Ch. 2.4, Thm. 2] (which requires the Lipschitz property) the constant solution can be approximated in (R) and hence a_0 is stabilizable. If $0\notin\mathrm{conv}(\mathsf{derv}(a_0))$ there is a linear functional α on $\mathfrak{a}$ such that $\alpha\le-\delta$ on $\mathsf{derv}(a_0)$ for some $\delta>0$. By continuity we may assume that this is true for all $b\in\mathfrak{a}$ in some neighborhood $B_\varepsilon(a_0)$ of a_0. Hence there is some time $T>0$ such that every solution to (R) and starting at a_0 must leave $B_\varepsilon(a_0)$ until time T. □

In particular a point is stabilizable for (R) if and only if it is strongly stabilizable for ($\bar{\mathsf{R}}$), and for ($\bar{\mathsf{R}}$) both notions coincide.

We have the following specialization of Proposition 2.3.10 for strongly stabilizable states.

Proposition 2.4.5. *The following statements hold.*

(i) *If there exists $p_0=\mathrm{Ad}_K(a_0)\in\mathfrak{p}$ as well as $k\in\mathfrak{k}$ such that $(X+\mathrm{ad}_k)(p_0)=0$, then a_0 is strongly stabilizable. In fact it holds that $X_K(a_0)=0$.*

(ii) *Conversely, assume that a_0 is regular and strongly stabilizable with $X_K(a_0)=0$. Then setting $k_c=\mathrm{ad}_{p_0}^{-1}\circ\Pi_{p_0}^{\perp}\circ X(p_0)$ it holds that $(X+\mathrm{ad}_{k_c})(p_0)=0$, where $p_0=\mathrm{Ad}_K(a_0)$. We call k_c the* compensating control.

[12] Strongly stabilizable states are also called equilibrium states, see [Son98, p. 124].

Proof. (i): Using (2.1) and Lemma 1.A.24 (i), the assumption $(X + \mathrm{ad}_k)(p_0) = 0$ yields $X_K(a_0) = 0$ after a short computation. (ii): First note that for a_0 regular and $p_0 = \mathrm{Ad}_K(a_0)$ it holds that $X_K(a_0) = \mathrm{Ad}_K^{-1} \circ \Pi_{p_0} \circ X(p_0)$, and in particular $X_K(a_0) = 0$ if and only if $\Pi_{p_0} \circ X(p_0) = 0$. Then it just remains to plug in and compute: $(X + \mathrm{ad}_{k_c})(p_0) = (X - \Pi_{p_0}^{\perp} \circ X)(p_0) = \Pi_{p_0} \circ X(p_0) = 0$. □

Suppose that the control directions $k_1, \ldots, k_m$ in (F) span the entire Lie algebra $\mathfrak{k}$ (Assumption (I')). Then we can rephrase the proposition above as follows: If p_0 is strongly stabilizable, then so is a_0. Conversely, if a_0 is strongly stabilizable and regular, then there is a corresponding strongly stabilizable p_0 in the $\mathbf{K}$-orbit of a_0.

Viability

Let R be a subset of $\mathfrak{a}$. We call R *viable* for (R) if for every $a_0 \in R$, there exists a solution $a : [0, \infty) \to \mathfrak{a}$ to (R) with $a(0) = a_0$ which takes values only in R. For differential inclusions, and hence also for the corresponding control systems, viability of closed subsets can be restated more geometrically using tangent cones, see [Smi02, Thm. 5.2] as well as [CNV07, Thm. 6.5.5] for the time-dependent version. Note that a point a_0 is strongly stabilizable if and only if $\{a_0\}$ is viable for (R).

First some notation: for a set $S \subseteq \mathfrak{p}$ we denote by $S^\flat \subseteq \mathfrak{a}$ the set of all $a \in \mathfrak{a}$ with $\pi(a) \in \pi(S)$. For a set $R \subseteq \mathfrak{a}$ we denote by $R^\sharp \subseteq \mathfrak{p}$ the set of all $p \in \mathfrak{p}$ with $\pi(p) \in \pi(R)$. Note that $S^\flat$ is always $\mathbf{W}$-invariant and $R^\sharp$ is always $\mathbf{K}$-invariant.

Lemma 2.4.6. *Let $S \subset \mathfrak{p}$ be viable for (F), then $S^\flat$ is viable for (R).*

Proof. Let $a_0 \in S^\flat$ and let $p_0 \in S$ be any lift of a_0. By viability of S there is a solution $p : [0, \infty) \to S$ and by Theorem 2.3.8 there is a corresponding solution $a^\downarrow : [0, \infty) \to \mathfrak{a}$ with values in $S^\flat$. □

Due to Example 2.3.12 the converse cannot hold exactly. However, we have the following approximate result. We say that $S \subseteq \mathfrak{p}$ is *approximately viable* for (F) if for every $p_0 \in S$, every $T > 0$ and every ε-neighborhood U of S there is a solution $p : [0, T] \to \mathfrak{p}$ with $p(0) = p_0$ and taking values only in U.

Proposition 2.4.7. *Let $R \subset \mathfrak{a}$ be viable for ($\bar{\text{R}}$). Then $R^\sharp$ is approximately viable for (F).*

Proof. Let $p_0 \in R^\sharp$, some $T > 0$, and an ε-neighborhood U of $R^\sharp$ be given. Let $a_0 \in R$ be such that $\pi(a_0) = \pi(p_0)$. Since R is viable, there exists some solution $a : [0, T] \to R$. By Theorem 2.3.14 there is some ε-approximate lift p of a. As in the proof of Proposition 2.4.3 we may assume that $p(0) = p_0$. Hence p remains in U and $R^\sharp$ is approximately viable. □

Note that even if R consists of regular points, $R^\sharp$ need not be (exactly) viable.

Invariant Subsets

A subset is called *invariant* if no solution can leave it. As above, for differential inclusions invariance of closed subsets can be characterized using a tangent cone condition, cf. [Smi02, Thm. 5.6].

Proposition 2.4.8. *Let $S \subseteq \mathfrak{p}$ be a closed, $\mathbf{K}$-invariant subset. Then S is invariant with respect to (F) if and only if $S^\flat$ is invariant with respect to (R) (equivalently ($\bar{\text{R}}$)).*

Proof. Note that S is invariant if and only if for every $p_0 \in S$ and $T > 0$ it holds that $\mathsf{reach}_{\mathsf{F}}(p_0, T) \subseteq S$. Hence the result follows immediately from Proposition 2.4.3. □

Accessibility

Let $\langle \mathfrak{X} \rangle_{\mathsf{Lie}}$ denote the Lie algebra generated by the induced vector fields, and $(\langle \mathfrak{X} \rangle_{\mathsf{Lie}})_{a_0} = \{Y(a_0) : Y \in \langle \mathfrak{X} \rangle_{\mathsf{Lie}}\}$ the evaluation at $a_0 \in \mathfrak{a}$. Let a closed embedded submanifold $R \subseteq \mathfrak{a}$ be invariant for (R). If $(\langle \mathfrak{X} \rangle_{\mathsf{Lie}})_{a_0} = T_{a_0} R$ for all $a_0 \in R$, we say that $\mathfrak{X}$ satisfies the *accessibility rank condition* [Son98, Def. 4.3.2]. Note that if X is linear, then $\mathfrak{X} \subset \mathfrak{gl}(\mathfrak{a})$. In particular $\langle \mathfrak{X} \rangle_{\mathsf{Lie}}$ is finite dimensional. The system (R) is *accessible at* a_0 *on* R if $\mathsf{reach}_{\mathsf{R}}(a_0, [0, T])$ has non-empty interior in R for all $T > 0$. The accessibility rank condition implies accessibility, see [Son98, Thm. 9].

In the differential inclusion picture we can define a stronger notion. We say that (R) is *directly accessible at* a_0 *on* R if $\operatorname{span}(X_K(a_0) : K \in \mathbf{K}) = T_{a_0} R$. This means linear combinations suffice to generate the entire tangent space without the use of Lie brackets. Conveniently, this property is relatively easy to check by considering the differential inclusion. Note that direct accessibility in (R) is equivalent to direct accessibility in ($\bar{\mathsf{R}}$).

Proposition 2.4.9. *Let X be real analytic and assume that there is some $a_0 \in \mathfrak{a}$ such that (R) is directly accessible at a_0. Then (R) is directly accessible on an open dense subset of $\mathfrak{a}$ whose complement has measure zero. In particular this happens if $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is simple and $X(0) \neq 0$.*

Proof. Choose a set of induced vector fields $X_{K_i} \in \mathfrak{X}$ for $i = 1, \dots, n$ such that the $X_{K_i}(a_0)$ form a basis of $T_{a_0}\mathfrak{a}$. Now consider the determinant of these vector fields $a_0 \mapsto \det(X_{K_1}(a_0), \dots, X_{K_n}(a_0))$ as a function on $\mathfrak{a}$. By assumption this is a real analytic function on $\mathfrak{a}$ which does not vanish at a_0. Hence it is non-zero on an open dense set whose complement has measure zero, and clearly (R) is directly accessible whenever the function is non-zero.

Now assume that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is simple and $X(0) \neq 0$. Let $K \in \mathbf{K}$ be such that $\operatorname{Ad}^{\star}_K(X)(0) \in \mathfrak{a}$ and so $X_K(0) \neq 0$. Since the Weyl group acts irreducibly on $\mathfrak{a}$, and since, as in Example 2.3.12, it holds that $\mathsf{derv}(0) = \operatorname{conv}(\mathbf{W} X_K(0))$, we see that $0 \in \operatorname{int}(\mathsf{derv}(0))$ and so (R) is directly accessible at 0. □

Proposition 2.4.10. *Let $R \subseteq \mathfrak{a}$ be a $\mathbf{W}$-invariant closed embedded submanifold and assume that the reduced system (R) is directly accessible at some regular a_0 on R. Then (F) is accessible on $R^\sharp$ at every p_0 with $\pi(p_0) = \pi(a_0)$.*

Proof. The Lie algebra corresponding to (F) is $\langle X + \operatorname{ad}_{\mathfrak{k}} \rangle_{\mathsf{Lie}} = \langle X, \operatorname{ad}_{\mathfrak{k}} \rangle_{\mathsf{Lie}}$. Since every Lie algebra is invariant under its adjoint action, it holds that $\operatorname{Ad}_K^{-1} \circ X \circ \operatorname{Ad}_K \in \langle X, \operatorname{ad}_{\mathfrak{k}} \rangle_{\mathsf{Lie}}$ for all $K \in \mathbf{K}$. For the same reason we can also assume that $p_0 = a_0$. The tangent space at p_0 takes the form $T_{p_0} R^\sharp = T_{p_0} R \oplus \operatorname{ad}_{\mathfrak{k}}(p_0)$. The assumption means that $\operatorname{span}(X_K(a_0) : K \in \mathbf{K}) = T_{a_0} R$, and so every element in $T_{p_0} R^\sharp$ is a linear combination of some $\operatorname{Ad}_K^{-1} \circ X \circ \operatorname{Ad}_K(p_0)$ with $K \in \mathbf{K}$ and some $\operatorname{ad}_k(p_0)$ with $k \in \mathfrak{k}$. This concludes the proof. □

Controllability

Let R be an invariant subset for (R). Then we say that (R) is *controllable on* R if for every $a_0 \in R$ it holds that $\mathsf{reach}_{\mathsf{R}}(a_0) = R$. We say that (R) is *controllable on* R *in time* T if for every $a_0 \in R$ it holds that $\mathsf{reach}_{\mathsf{R}}(a_0, [0, T]) = R$, see [Son98, Ch. 3]. We define *approximate controllablility* analogously except that we consider the closure of the reachable set. Then the following is an immediate consequence of Proposition 2.4.3.

Proposition 2.4.11. *Assume that X is locally Lipschitz and linearly bounded. Let $S \subseteq \mathfrak{p}$ be $\mathbf{K}$-invariant and invariant for (F). Then the following statements hold.*

(i) If (F) is (approximately) controllable on S, then (R) is (approximately) controllable on $S^\flat$.

(ii) If (R) *is approximately controllable on* S^ϕ*, then* (F) *is approximately controllable on* S*.*

All statements remain true if we consider (approximate) controllability in time T*.*

Let $a_0 \in R$. We say that (R) is *locally controllable at* a_0 *on* R if $\mathsf{reach}_{\mathsf{R}}(a_0, [0,T])$ contains an open neighborhood of a_0 (in the subspace topology of R) for all $T > 0$. Moreover we say that (R) is *locally directly controllable at* a_0 *on* R if $0 \in \operatorname{int}(\mathsf{derv}(a_0))$ (where the interior is taken in the topology of $T_{a_0}R$).

The Weyl group acts on the Lie algebra $\mathfrak{gl}(\mathfrak{a})$ by Lie algebra automorphisms and, by Lemma 2.A.2, the set $\mathfrak{X}$ of induced vector fields is invariant under this action.

Proposition 2.4.12. *Assume that* X *is linear. Let* $\mathbf{T} \subseteq \mathrm{GL}(\mathfrak{a}; R)$ *be a connected Lie subgroup with Lie algebra* $\mathfrak{t}$ *and assume that* $\{0\} \neq \mathfrak{X} \subseteq \mathfrak{t}$*. If* $\mathbf{T}$ *acts transitively on* R*, and* $\mathbf{W}$ *acts irreducibly on* $\mathfrak{t}$*, then* (R) *is controllable on* R*.*

Proof. Due to irreducibility, it holds that $\operatorname{span}(\mathfrak{X}) = \mathfrak{t}$. In particular $\mathsf{reach}_{\mathsf{RL}}(\mathbb{1}) = \mathbf{T}$ which implies the result. □

2.5 The Majorization Theorem

In this section we assume that X is an affine linear vector field on $\mathfrak{p}$. Using the Weyl group action we obtain a preorder on $\mathfrak{a}$ which acts as a kind of resource, allowing one system to simulate another.

Let $a \in \mathfrak{a}$. We define the *Weyl polytope* of a via $P(a) := \operatorname{conv}(\mathbf{W}a)$, that is, $P(a)$ is the convex hull of the Weyl group orbit of a. Since $\mathbf{W} \subseteq \mathfrak{gl}(\mathfrak{a})$, we can consider the convex hull of $\mathbf{W}$ in $\mathfrak{gl}(\mathfrak{a})$. For $a, b \in \mathfrak{a}$, it is clear that $a \in P(b)$ if and only if there is some $\overline{w} \in \operatorname{conv}(\mathbf{W})$ such that $a = \overline{w}b$. It is easy to show that $\operatorname{conv}(\mathbf{W})$ is a semigroup. Hence we can define a preorder, called majorization, on $\mathfrak{a}$ by declaring for $a, b \in \mathfrak{a}$ that $a \preceq b :\iff a \in P(b)$. Indeed, reflexivity is clear and transitivity follows immediately from the fact that $\operatorname{conv}(\mathbf{W})$ is a semigroup.

Lemma 2.5.1. *The set of vertices of* $P(a)$ *is exactly* $\mathbf{W}a$*. In particular if* $P(a) = P(b)$*, then* $\mathbf{W}a = \mathbf{W}b$*.*

Proof. By definition the set of vertices is a subset of the Weyl group orbit of a. However, since $P(a)$ is invariant under $\mathbf{W}$, and since $\mathbf{W}$ acts transitively on the orbit, all elements of the orbit must be vertices. □

Note that if $a \preceq b$ and $b \preceq a$, then $P(a) = P(b)$. Hence a and b belong to the same Weyl group orbit so $\preceq$ induces a partial order on the orbits (or, equivalently, in a closed Weyl chamber).

The following continuity property will be useful later.

Lemma 2.5.2. *The set-valued map* $P : \mathfrak{a} \to \mathcal{P}(\mathfrak{a})$ *(where* $\mathcal{P}(\cdot)$ *denotes the power set) defined by* $a \mapsto \operatorname{conv}(\mathbf{W}a)$ *is Lipschitz continuous with Lipschitz constant* 1*.*

Proof. Let $\mathbf{W} = \{w_i : i = 1, \dots m\}$ with m the order of the Weyl group and let Δ^{m-1} denote the standard simplex. Consider the map $f : \mathfrak{a} \times \Delta^{m-1} \to \mathfrak{a}$ given by $(a, \lambda) \mapsto \sum_{i=1}^m \lambda_i w_i \cdot a$. This map is clearly 1-Lipschitz in a, and by [Smi02, Prop. 2.4] it holds that $a \mapsto P(a) = f(a, \Delta^{m-1})$ is 1-Lipschitz as well. □

The main result of this section is the following:

Theorem 2.5.3 (Majorization Theorem). *Assume that X is affine linear. Let $\mathrm{a} : [0,\infty) \to \mathfrak{a}$ be a solution to the relaxed control system ($\bar{\mathsf{R}}$) and let $b_0 \in \mathfrak{a}$ such that $\mathrm{a}(0) = \mathrm{a}_0 \preceq b_0$. Then there exists a solution $b : [0,\infty) \to \mathfrak{a}$ to ($\bar{\mathsf{R}}$) with $b(0) = b_0$ such that $\mathrm{a}(t) \preceq b(t)$ for all $t \in [0,\infty)$.*

Proof. First we prove the result with the additional assumption that a is differentiable. Then by Proposition 2.A.4, $a^{\downarrow}$ is also a solution. Since it is continuous, Lemma 1.B.5 (iii) shows that $a^{\downarrow}$ is right-differentiable.

Consider the set-valued maps $A, C : [0,\infty) \to \mathcal{P}(\mathfrak{a})$ defined by $A(t) = \{x \in \mathfrak{a} : x \succeq \mathrm{a}(t)\}$ and $C(t) = A(t) \cap \mathfrak{w}$. The main idea is to show that for each $t \in [0,\infty)$ and $x \in C(t)$ there is some $v \in \mathsf{derv}(x)$ such that $v \in T_x\mathfrak{w}$ and $v - \mathrm{a}'(t) \in T_x A(t)$. Intuitively this means that for every point majorizing $\mathrm{a}(t)$, there exists a derivative preserving majorization and the Weyl chamber for an infinitesimal amount of time.

By assumption, $a(t) \in \operatorname{relint}(F)$ for some face F of $\operatorname{conv}(\mathbf{W}x)$. By Result 2.B.10 there is some $\Omega \subset \mathfrak{w}$ such that $F = \operatorname{conv}(\mathbf{W}_\Omega x)$ For some enumeration w_i, with $i = 1,\dots,k$, of $\mathbf{W}_\Omega$ and some $\lambda \in \Delta^{k-1}$ it holds that $a(t) = \sum_{i=1}^k \lambda_i w_i \cdot x$. Hence using affine linearity of X and Lemma 2.A.2 we compute

$$a'(t) = \sum_j \mu_j X_{K_j}\big(\sum_i \lambda_i w_i x\big) = \sum_{i,j} \lambda_i\mu_j w_i X_{K_jN_i}(x) = \sum_i w_i \sum_j \lambda_i\mu_j X_{K_jN_i}(x)$$

where $N_i \in \mathbf{K}$ is any representative of w_i. Now consider the achievable derivative

$$v = \sum_{i,j} \lambda_i\mu_j X_{K_jN_i}(x) \in \operatorname{conv}(\mathsf{derv}(x))\,,$$

then

$$v - \mathrm{a}'(t) = \sum_i (\mathbb{1} - w_i) \sum_j \lambda_i\mu_j X_{K_jN_i} x,$$

which lies in the tangent space $T_{a(t)}F$. This shows that $v - \mathrm{a}'(t) \in T_x A(t)$. Moreover there exists some $w \in \mathbf{W}_x$ such that $\overline{v} = w\cdot v \in T_x\mathfrak{w}$. But then we still have $\overline{v} - \mathrm{a}'(t) = (w\cdot v - v) + (v - \mathrm{a}'(t)) \in T_x A(t)$.

To show that existence of $\bar{v}$ implies existence of the desired solution b, we employ a sequence of rather technical results detailed in Appendix 2.B. It follows from Corollary 2.B.11 that $T_{\mathrm{a}(t)}A(t)$ is the negative dual of $\mathfrak{w}$. Thus we can in fact apply Lemma 2.B.8 and Proposition 2.B.9, showing that $\overline{v} \in \mathrm{D}^- C(t,c)(1)$, and together with Result 2.B.5 this tells us that there exists a solution b to the relaxed control system ($\bar{\mathsf{R}}$) such that $b(0) = b_0$ and such that $b(t) \in C(t)$ for all $t \geq 0$, or equivalently, $b(t) \preceq \mathrm{a}(t)$ and $b(t) \in \mathfrak{w}$. This concludes the proof in the differentiable case.

Now we drop the assumption that a is differentiable. By [Son98, Thm. 1] and [AC84, Ch. 2.4 Thm. 2] there exists a sequence a_n of differentiable solutions to the relaxed control system converging uniformly to a on compact time intervals. By the above, there exist solutions b_n satisfying $\mathrm{a}_n(t) \preceq b_n(t)$ for all t. By compactness of solution set on compact time interval, cf. Proposition 2.A.6 (ii), there is a uniformly converging subsequence b_{k_n} with limit b. Since $\mathrm{a}_n(t) \to \mathrm{a}(t)$ it holds that $\mathrm{a}(t) \preceq b(t)$ by [Smi02, Prop. 2.1] since $b \mapsto \operatorname{conv}(\mathbf{W}b)$ is upper semi-continuous, see Lemma 2.5.2, with closed values. □

Theorem 2.5.3 leads to a powerful maximum principle used in Chapter 6, see [2, 6] for precise statements and proofs.

2.6 Worked Example

We now revisit the motivating example given in the introduction in order to apply to it the theory we have developed. We will consider the following vector field X on $\mathbb{R}^2$, which leaves the disk $D = \{(y,z) \in \mathbb{R}^2 : y^2 + z^2 \leq 1\}$ invariant:

$$X(y,z) = (-\Gamma y, -\gamma(z-1))$$

where $\Gamma, \gamma > 0$. This system is ubiquitous in quantum mechanics since it describes the relaxation of a two-level system under the Bloch equations. The corresponding control system has been studied in [Lap+10] using the Pontryagin Maximum Principle (PMP). In this example we will show how the same control system can be studied using our reduction method.

By rescaling, it suffices to consider $\gamma = 1$. Moreover, to ensure that the flow does not leave the disk we have to require $\Gamma \geq \frac{\gamma}{2} = \frac{1}{2}$. In fact we will consider $\Gamma \geq \frac{3}{2}$ in the following to simplify the exposition. (All figures use the value $\Gamma = 3$.)

We already stated that this problem can be described using the symmetric Lie algebra given in Example 1.2.4. The reduced control system is defined on the set $[-1,1]$, which can be seen as the intersection of the disk with x-axis[13]. Since the map $(a,\phi) \mapsto X_\phi(a)$, where the angle ϕ parametrizes $\mathrm{SO}(2)$, is continuous, the values of the set-valued map derv are compact intervals, see Figure 2.1. In order to understand derv, it suffices to find the upper bound $\mu(a) := \max(\mathsf{derv}(a))$, which we will call the *optimal derivative function*. One can show that (see Lemma 5.A.8)

$$\mu(a) = \begin{cases} -\left(\frac{1}{4(\Gamma-1)a} + \Gamma a\right) & a \leq a_0 := \frac{-1}{2(\Gamma-1)} \\ 1 - a & a \geq a_0\,. \end{cases} \tag{2.6}$$

We consider the optimal control problems of moving from the boundary of the disk to the center and vice-versa. In the reduced control system this is equivalent to moving from -1 to 1. The form of (2.6) shows that this is indeed possible, but it takes infinite time to reach 1. The optimal solution is then defined by the differential equation $a'(t) = \mu(a(t))$ with $a(0) = -1$, which can be solved explicitly, and one obtains

$$a^\star(t) = \begin{cases} -\frac{\sqrt{-1+(1-2\Gamma)^2 e^{-2\Gamma t}}}{2\sqrt{\Gamma(\Gamma-1)}} & t \leq t_0 := \frac{\log((\Gamma-1)(2\Gamma-1))}{2\Gamma} \\ 1 - \frac{2\Gamma-1}{2(\Gamma-1)}((\Gamma-1)(2\Gamma-1))^{\frac{1}{2\Gamma}} e^{-t} & t \geq t_0\,. \end{cases}$$

The next step is to lift the optimal solution $a^\star$ to the original control system to obtain a solution $p^\star$ on the disk. This solution will start on the boundary of the disk, pass through the center, and again approach the boundary of the disk. Once this optimal solution is found, we will determine the corresponding control function $\omega^\star$.

Above we have determined the upper boundary $\mu(a)$ of $\mathsf{derv}(a)$. More precisely, one can show that $\mu(a) = X_{\phi^\star}(a)$ where

$$\phi^\star(a) = \begin{cases} \arccos\left(\frac{-1}{2(\Gamma-1)a}\right) + \frac{\pi}{2} & a \leq a_0 \\ \frac{\pi}{2} & a \geq a_0\,. \end{cases}$$

Thus the optimal path $p^\star(t)$ in the disk (in polar coordinates) is $(a^\star(t), \phi^\star(a^\star(t)))$, cf. Figure 2.3.

Finally, it remains to calculate the control function $\omega^\star$ which generates the optimal solution $p^\star$. There are two components, the direct term ω_0 and the compensation term ω_c, cf. Proposition 2.3.10:

$$\omega^\star = \omega_0 + \omega_c = \frac{d}{dt}\phi^\star(a^\star(t)) + \mathrm{ad}^{-1}_{p^\star(t)} \circ \Pi^\perp_{p^\star(t)} \circ X(p^\star(t))\,.$$

[13] Any axis would do as they are equivalent under rotation.

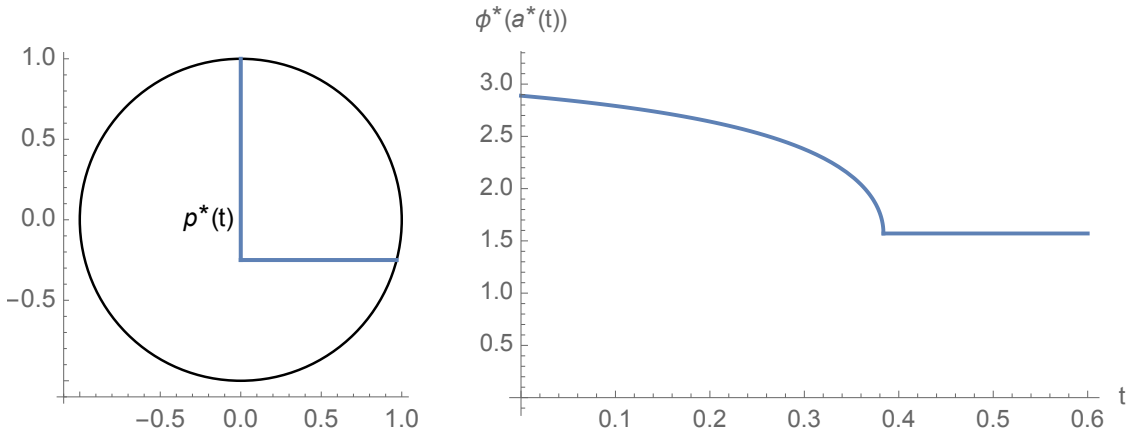

Figure 2.3: (Left) Optimal path $p^\star(t)$ on the unit disk. The horizontal part satisfies $z = a_0 = \frac{-1}{2(\Gamma-1)}$. (Right) Optimal angle $\phi^\star$ as a function of time.

Explicitly we obtain for $t < t_0$:

$$\omega_0(t) = -\Gamma \frac{\delta(t)+1}{\delta(t)} \sqrt{\frac{\eta}{\delta(t)-\eta}}, \quad \omega_c(t) = -\frac{\Gamma}{\delta(t)} \sqrt{\frac{\delta(t)-\eta}{\delta(t)}}$$

where $\eta = \frac{\Gamma}{\Gamma-1}$ and $\delta(t) = (1-2\Gamma)^2 e^{-2\Gamma t} - 1$. For $t \geq t_0$ it holds that $\omega_0(t) = \omega_c(t) = 0$. The optimal controls are plotted in Figure 2.4.

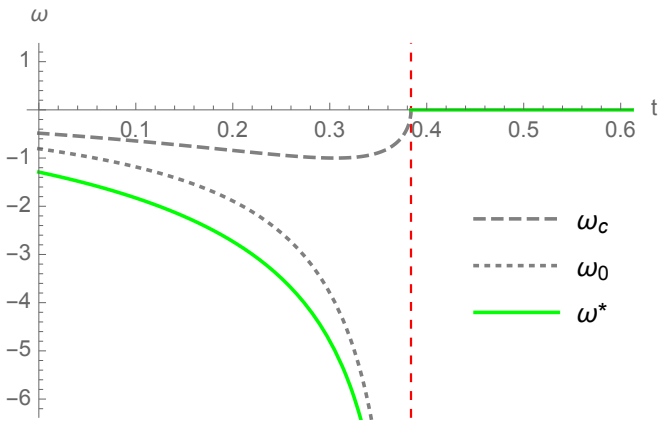

Figure 2.4: Optimal control $\omega^\star$ of the full system for the worked example. The compensating term ω_c and the direct term ω_0 are shown separately.

Since $\delta(t_0) = \eta$, we see that $\lim_{t\to t_0} \omega_0(t) = -\infty$ and $\lim_{t\to t_0} \omega_c(t) = 0$. In particular ω_0 explodes at t_0 whereas ω_c is continuous. Note also that the controls are smooth when $a^\star(t) = 0$, so in this example the orbifold singularity does not pose any problems.

2.A Some Basic Properties of the Reduced Control Systems

Here we give some basic properties of the control systems defined in Section 2.2.

Weyl Symmetry

We start with some symmetry considerations.[14]

Lemma 2.A.1. *If $w = NZ_{\mathbf{K}}(\mathfrak{a}) \in \mathbf{W}$ with $N \in N_{\mathbf{K}}(\mathfrak{a})$, then $\Pi_{\mathfrak{a}} \circ \mathrm{Ad}_N = w \circ \Pi_{\mathfrak{a}}$.*

Proof. First note that for $x \in \mathfrak{p}$ and $K \in \mathbf{K}$ it holds that $\Pi_{\mathrm{Ad}_K(\mathfrak{a})} \circ \mathrm{Ad}_K(x) = \mathrm{Ad}_K \circ \Pi_{\mathfrak{a}}(x)$, see for instance Lemma 1.A.24 (iii). Since $N \in N_{\mathbf{K}}(\mathfrak{a})$ we have that $w \cdot \Pi_{\mathfrak{a}}(x) = \mathrm{Ad}_N \circ \Pi_{\mathfrak{a}}(x) = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}_N(x)$, as desired. □

Lemma 2.A.2. *Let $N \in N_{\mathbf{K}}(\mathfrak{a})$ and $w = NZ_{\mathbf{K}}(\mathfrak{a}) \in \mathbf{W}$ and $K \in \mathbf{K}$, then $X_{KN} = w^{-1} \circ X_K \circ w$, and hence for all $\mathrm{a} \in \mathfrak{a}$ we get $\mathsf{derv}(w \cdot \mathrm{a}) = w \cdot \mathsf{derv}(\mathrm{a})$.*

Proof. We compute using Lemma 2.A.1: $X_{KN} = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_{KN}(X) \circ \iota = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_N(\mathrm{Ad}^{\star}_K(X)) \circ \iota = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{-1}_N \circ \mathrm{Ad}^{\star}_K(X) \circ \mathrm{Ad}_N \circ \iota = w^{-1} \circ \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_K(X) \circ \iota \circ w = w^{-1} \circ X_K \circ w$. In particular $X_K \circ w = w \circ X_{NK}$, which shows the second claim. □

Lemma 2.A.3. *Let $\mathbf{K}_X = \{K \in \mathbf{K} : \mathrm{Ad}^{\star}_K X = X\}$. Then $X_{SK} = X_K$ for all $S \in \mathbf{K}_X$ and $K \in \mathbf{K}$.*

Proof. As $\mathrm{Ad}^{\star}_{SK} = (\mathrm{Ad}_S \circ \mathrm{Ad}_K)^{\star} = \mathrm{Ad}^{\star}_K \circ \mathrm{Ad}^{\star}_S$, we have $X_{SK} = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_{SK}(X) \circ \iota = \Pi_{\mathfrak{a}} \circ \mathrm{Ad}^{\star}_K(X) \circ \iota = X_K$. □

Proposition 2.A.4. *If $\mathrm{a} : [0, \infty) \to \mathfrak{a}$ is a solution to (R′) or ($\bar{\mathsf{R}}$), then the unique $\mathrm{a}^{\downarrow} : [0, \infty) \to \mathfrak{w}$ satisfying $\pi \circ \mathrm{a} = \pi \circ a^{\downarrow}$ is also a solution.*

Proof. By Proposition 1.3.1 (v), $a^{\downarrow}$ is still absolutely continuous. Assume that a and $a^{\downarrow}$ are both differentiable at t. By Lemma 1.B.5 (i) there is some $w \in \mathbf{W}$ such that $a^{\downarrow}(t) = w \cdot \mathrm{a}(t)$ and $(a^{\downarrow})'(t) = w \cdot \mathrm{a}'(t)$. By Lemma 2.A.2 we get, using $N \in N_{\mathbf{K}}(\mathfrak{a})$ with $w = NZ_{\mathbf{K}}(\mathfrak{a})$: $(a^{\downarrow})'(t) = w \cdot \mathrm{a}'(t) = wX_K\mathrm{a}(t) = X_{KN^{-1}}w\mathrm{a}(t) = X_{KN^{-1}}a^{\downarrow}(t)$ and so $a^{\downarrow}$ satisfies the differential inclusion at t, and hence almost everywhere. □

Continuity and Compactness

Lemma 2.A.5. *If X is Lipschitz, then the set-valued function derv is also Lipschitz. This means that for all $x, y \in \mathfrak{a}$, $\mathsf{derv}(x) \subseteq \mathsf{derv}(y) + L\|x - y\|B_1$ for some (global) Lipschitz constant $L > 0$ and where B_1 denotes the closed unit ball in $\mathfrak{a}$.*

This implies some convenient properties of the relaxed control system ($\bar{\mathsf{R}}$), see [Smi02, Ch. 4]. Here we denote the set of solutions $\mathrm{a} : [0, T] \to \mathfrak{a}$ to ($\bar{\mathsf{R}}$) with $\mathrm{a}(0) = a_0$ by $\mathsf{sols}_{\bar{\mathsf{R}}}(a_0, [0, T])$.

Proposition 2.A.6. *Let X be Lipschitz and let $a_0 \in \mathfrak{a}$. The following hold:*

(i) The set of solutions $\mathsf{sols}_{\bar{\mathsf{R}}}(a_0, [0, T])$ is path-connected in the AC-topology[15].

(ii) If derv is bounded, then $\mathsf{sols}_{\bar{\mathsf{R}}}(a_0, [0, T])$ is compact in the topology of uniform convergence.

(iii) If $\mathrm{a} \in \mathsf{sols}_{\bar{\mathsf{R}}}(a_0, [0, T])$ is a solution to ($\bar{\mathsf{R}}$) with $\mathrm{a}(T) \in \partial\mathsf{reach}_{\bar{\mathsf{R}}}(a_0, [0, T])$, then for all $t \in [0, T]$ we have $\mathrm{a}(t) \in \partial\mathsf{reach}_{\bar{\mathsf{R}}}(a_0, [0, t])$.

[14] Recall that the Weyl group is defined as $\mathbf{W} = N_{\mathbf{K}}(\mathfrak{a})/Z_{\mathbf{K}}(\mathfrak{a})$ where $N_{\mathbf{K}}(\mathfrak{a})$ denotes the normalizer of $\mathfrak{a}$ in $\mathbf{K}$ and $Z_{\mathbf{K}}(\mathfrak{a})$ the centralizer.

[15] By $AC([0, T], \mathfrak{a})$ we denote the Banach space of absolutely continuous functions $\mathrm{a} : [0, T] \to \mathfrak{a}$ equipped with the norm $\|\mathrm{a}\|_{AC} = |\mathrm{a}(0)| + \int_0^T |\mathrm{a}'(t)|dt$.

(iv) If derv *is bounded, then there exist time-optimal solutions to* ($\bar{\mathsf{R}}$) *starting in a given compact set and ending in a given closed set, assuming any such solution exists in the first place.*

(v) If X is Lipschitz with Lipschitz constant L, then the map $\mathfrak{a} \to \mathcal{P}(AC([0,T],\mathfrak{a}))$ given by $a_0 \mapsto \mathsf{sols}_{\bar{\mathsf{R}}}(a_0,[0,T])$ is Lipschitz with Lipschitz constant $1 + TLe^{TL}$.

2.B Some Technical Results for the Majorization Theorem

We recall some basic facts from convex analysis and prove some technical results needed for the proof of Theorem 2.5.3. Our main reference is [Smi02]. We start with the concept of a tangent cone to a convex set at a certain point.

Definition 2.B.1 (Tangent cone). *Let X be a normed space and let $C \subseteq X$ be a convex subset. Given any $x \in C$, the* tangent cone to C at x *is defined by*

$$T_x C = \overline{\bigcup_{\lambda>0} \frac{C-x}{\lambda}} = \{v \in X : \lim_{\lambda \to 0^+} d(x+\lambda v, A)/\lambda = 0\}.$$

Intuitively, $T_x C$ is the closure of the set of all directions which lie in C for some small enough distance. As soon as non-convex sets come into play, the situation becomes more complicated.

Definition 2.B.2 (Bouligand contingent cone). *Let X be a normed space and let $A \subseteq X$ be any subset. Given any $x \in A$, the* contingent cone to A at x *is defined by*

$$T_x^- A = \{v \in X : \liminf_{\lambda \to 0^+} d(x+\lambda v, A)/\lambda = 0\}.$$

The contingent cone is indeed a closed cone and for convex sets it coincides with the tangent cone, cf. [Smi02, p. 38].

A common geometric way to think of the derivative of a function f in standard calculus is as a tangent space to the graph Γ_f of the function at a given point. Using the Bouligand contingent cone we can define a derivative for set-valued function in much the same way, cf. [Smi02, p. 41].

Definition 2.B.3 (Contingent derivative). *Let normed spaces X, Y and a set-valued function $F : X \to \mathcal{P}(Y)$ be given. For $(x_0, y_0) \in \Gamma_F$, the set-valued map $D^- F(x_0,y_0) : X \to \mathcal{P}(Y)$ defined by*

$$\Gamma_{D^- F(x_0,y_0)} = T^-_{(x_0,y_0)} \Gamma_F$$

is called the contingent derivative of F at (x_0, y_0).

Example 2.B.4. *Let $f(x) = x\sin(1/x)$ (with $f(0) = 0$). Then f is continuous, but not differentiable at 0. The contingent derivative is $D^- F(0,0)(x) = [-|x|, |x|]$. In particular $D^- F(0,0)(1) = [-1,1]$.*

Result 2.B.5 (Thm. 6.5.5 in [CNV07]). *Assume that the set-valued map $C : \mathbb{R} \to \mathcal{P}(\mathbb{R}^n)$ has a closed graph and the set-valued map $F : \Gamma_C \to \mathcal{P}(\mathbb{R}^n)$ is upper semi-continuous and has closed, convex values. Then the following statements are equivalent.*

(i) For any point $(t_0, x_0) \in \Gamma_C$ there is a solution x on $[t_0, \infty)$ to the differential inclusion $x(t) \in F(t, x(t))$ with $x(t_0) = x_0$.

(ii) For any $(t,x) \in \Gamma_C$ it holds that $F(t,x) \cap D^- C(t,x)(1) \neq \emptyset$.

Lemma 2.B.6. *Let X be a metric space and consider $\mathcal{P}(X)$ with the Hausdorff distance d. Then it holds that $d(x,A) \leq d(x,B) + d(B,A)$.*

Recall that for a cone $C \subseteq \mathbb{R}^n$, the *dual cone* of C is defined as $C^* = \{x' \in \mathbb{R}^n : \langle x', x\rangle \geq 0 \, \forall x \in C\}$. Note that if $C \subseteq D$ for two cones, then $C^* \supseteq D^*$. Moreover, for a convex set A and some $x \in A$, the *normal cone* of A at x is defined by $N_x A := -(T_x A)^*$, i.e. the negative of the dual of the tangent cone. The next result follows for instance from [Cla13, Sec. 11.2].

Lemma 2.B.7. *Let a closed convex set $A \subseteq \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$ be given. If $\bar{x} \in A$ denotes the closest point in A to x, then $x - \bar{x} \in N_{\bar{x}} A$.*

Lemma 2.B.8. *Let $C \subseteq \mathbb{R}^n$ be a closed convex cone and let $x \in C$. Then for any $y \in C$ it holds that $d(y, C^* + x) = d(y, (C^* + x) \cap C)$.*

Proof. Since C^* is closed and convex, there exists unique $\overline{y} \in C^* + x$ such that $d(y, C^* + x) = d(y, \overline{y})$. We will show that $\overline{y} \in C$. Let $v = y - \overline{y}$. By Lemma 2.B.7 it holds that $v \in N_{\overline{y}}(C^* + x)$. Since for any closed convex set A and $a \in A$ we have $T_a A \supseteq A$, we compute

$$v \in N_{\overline{y}}(C^* + x) = N_{\overline{y}-x}(C^*) = -(T_{\overline{y}-x}(C^*))^* \subseteq -(C^*)^* = -C,$$

so $-v \in C$ and hence $\overline{y} = y - v \in C$, as desired. □

Proposition 2.B.9. *Let $I \subseteq \mathbb{R}$ be an open interval and let $\tilde{A}, \tilde{B} \subseteq \mathbb{R}^n$ be subsets. Assume that $\tilde{B}$ is a closed, convex, polyhedral cone. Let $\mathrm{a} : I \to \tilde{B}$ be right differentiable and define the set-valued functions $A(t) = \tilde{A} + \mathrm{a}(t)$, and $C(t) = A(t) \cap \tilde{B}$. Assume that for all $b \in \tilde{B}$ and $t \in I$ it holds that $d(b, A(t)) = d(b, C(t))$. Let $t \in I$ and $c \in C(t)$, and assume that there is v such that $v \in T_c \tilde{B}$ and $v - \mathrm{a}'_+(t) \in T_c A(t)$. Then $v \in \mathrm{D}^- C(t,c)(1)$.*

Proof. We assume that $t = 0$. By definition, $v \in \mathrm{D}^- C(0,c)(1)$ if and only if $(1,v) \in \Gamma_{\mathrm{D}^- C(0,c)} = T^-_{(0,c)} \Gamma_C$. So we have to show that

$$\liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} d((\varepsilon, c + \varepsilon v), \Gamma_C) = 0\,.$$

In fact it is easy to see that $d((\varepsilon, c+\varepsilon v), \Gamma_C) \leq d(c+\varepsilon v, C(\varepsilon))$. For ε small enough, $c + \varepsilon v \in \tilde{B}$ since $\tilde{B}$ is a convex polyhedron and $v \in T_c \tilde{B}$. So, for any $x \in \tilde{B}$ we have by assumption $d(c + \varepsilon v, C(\varepsilon)) = d(c + \varepsilon v, A(\varepsilon))$. Moreover using Lemma 2.B.6 we find

$$\begin{aligned} d(c + \varepsilon v, A(\varepsilon)) &= d(c + \varepsilon v, A(0) + \mathrm{a}(\varepsilon) - \mathrm{a}(0)) \\ &\leq d(c + \varepsilon v, A(0) + \varepsilon \mathrm{a}'_+(0)) + d(A(0) + \mathrm{a}(\varepsilon) - \mathrm{a}(0), A(0) + \varepsilon \mathrm{a}'_+(0)) \\ &\leq d(c + \varepsilon v, A(0) + \varepsilon \mathrm{a}'_+(0)) + |\mathrm{a}(\varepsilon) - \mathrm{a}(0) - \varepsilon \mathrm{a}'_+(0)|\,. \end{aligned}$$

Combining the results above and the assumption that $v - \mathrm{a}'_+(0) \in T_c A(0)$ we see that

$$\liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} d((\varepsilon, c + \varepsilon v), \Gamma_C) \leq \liminf_{\varepsilon \to 0^+} \frac{1}{\varepsilon} d(c + \varepsilon v, A(0) + \varepsilon \mathrm{a}'_+(0)) + \left| \frac{\mathrm{a}(\varepsilon) - \mathrm{a}(0)}{\varepsilon} - \mathrm{a}'_+(0) \right| = 0$$

which concludes the proof. □

Note that although $T_x(A \cap B) \subseteq T_xA \cap T_xB$, the converse need not hold, which complicates the proof above.

The following result is a restatement of [McC+03, Thm. 4.1].

Result 2.B.10. *Let* $\mathbf{W}$ *be a Coxeter group acting on a real,* n*-dimensional vector space* V*, and let* $\mathfrak{w}$ *be a (closed) Weyl chamber. Let* F *be a codimension-*k *face of the orbitope* $\operatorname{conv}(\mathbf{W}x)$ *for some* $x \in \mathfrak{w}$*. Then there exists a set* Ω *of* k *fundamental weights belonging to the same Weyl chamber* $\mathfrak{w}$ *such that* $F = \operatorname{conv}(\mathbf{W}_\Omega x)$*.*

Corollary 2.B.11. *Fix a Weyl chamber* $\mathfrak{w}$*, and let* $x \in \mathfrak{w}$ *be a regular point. Then* $T_x(\operatorname{conv}(\mathbf{W}x)) = -\mathfrak{w}^*$*, i.e. the negative dual cone of* $\mathfrak{w}$*.*

Proof. Applying Result 2.B.10 with $k = 1$ we see that the fundamental weights ω_i of $\mathfrak{w}$ are exactly the outward normals of the facets of $T_x(\operatorname{conv}(\mathbf{W}x))$. Hence $v \in T_x(\operatorname{conv}(\mathbf{W}x))$ if and only if $\langle \omega, v \rangle \leq 0$, and since the fundamental weights generate the Weyl chamber $\mathfrak{w}$, it holds that $T_x(\operatorname{conv}(\mathbf{W}x)) = -\mathfrak{w}^*$. □

Part II

Markovian Systems with Unitary Control

« La lutte elle-même vers les sommets suffit à remplir un cœur d'homme; il faut imaginer Sisyphe heureux. »
— Albert Camus, *Le Mythe de Sisyphe* (1942)

"If you optimize everything, you will always be unhappy."
— Donald Knuth, *on* 𝕏 (2016)

One of the main obstacles towards realizing quantum technologies is uncontrolled or unmitigated noise leading to the loss of quantum coherence. Every quantum system that can be externally controlled and measured must interact with its environment and hence is also subject to decoherence. Control is necessary for reducing deleterious effects of noise and to cool systems into a ground state useful for information processing tasks. Moreover, it is possible to modulate noise in order to exploit it as an additional control resource beyond coherent controls.

In this work, we assume the noise to be Markovian and time-independent in the sense that it is described by a master equation of Lindblad form. Furthermore, in the systems of concern, we assume that unitary control is fast compared to dissipation. Corresponding results can be obtained assuming that the noise itself is switchable. This allows us to define an equivalent reduced control system on the eigenvalues of the quantum state, which quantify the purity of the system. We study general properties such as reachability and stabilizability and algebraically characterize coolable quantum systems. We obtain a thorough understanding of the problem in the qubit case, leading to a general method for deriving optimal controls. Finally we consider the task of optimal cooling where we introduce a powerful method based on majorization which significantly simplifies the search for optimal solution.

Outline Chapter 3 applies the methods of Part I to the setting of Markovian quantum systems subject to fast unitary control in order to obtain a reduced control system on the eigenvalues. Chapter 4 studies general finite dimensional systems and characterizes important control theoretic notions such as reachability and stabilizability using the reduced control system as an essential tool. Chapter 5 gives a thorough treatment of the single qubit case. Chapter 6 addresses the task of cooling quantum mechanical systems in a time-optimal way.

Acknowledgements Part II is based on [3–6, 9]. The works [3, 4] are joint work with Frederik vom Ende, Thomas Schulte-Herbrüggen and Gunther Dirr. The paper [9] has profited from feedback from Robert Zeier and Thomas Schulte-Herbrüggen. Frederik vom Ende and Gunther Dirr have given feedback on [5] and [6] respectively.

CHAPTER 3

Reduction to the Eigenvalues

Introduction

In this chapter we apply the methods developed in Chapter 2 to open Markovian quantum systems subject to fast unitary control. The connection to symmetric Lie algebras is made via the Hermitian EVD, recall Example 1.2.1. We derive the corresponding equivalent reduced control system defined on the eigenvalues of the state density matrix. Its state space is the standard simplex, and its dynamics are given by stochastic transformations. This reduced control system is the main tool used in the subsequent chapters of Part II.

The reduced control system in the present Lindbladian setting has, to the best of my knowledge, first been formulated in [STK04, Yua10] (see also [TM92]), and it has been studied in [RBR18], without however giving a full proof of the equivalence. Our treatment in Chapter 2 removes those assumptions made in [RBR18] whereby certain singularities are essentially ignored. Indeed, these singularities are an inherent feature and present the main complication of the reduction.

A natural way to simplify the reduced control system is to restrict its controls to values in a finite set (in our case the permutations of the eigenvalues of the state). This is also considered in [RBR18], and explored more thoroughly in [3, DES19, End20, SED22] in the context of quantum thermodynamics: there, "thermal operations" come with separate time evolutions for diagonal and off-diagonal elements in the density matrix, which naturally inspires the reduction to "toy models" of diagonal states. Applications to unital systems were already given in [Yua10, SAZ19] and the results will be recovered in the subsequent chapters as special cases.

Full Control System

Throughout Part II, we use n-dimensional Hilbert spaces ($2 \leq n < \infty$) represented by $\mathbb{C}^n$. A *mixed state* ρ is a density matrix, i.e. a positive semi-definite operator with unit trace. The set of all states is denoted $\mathfrak{pos}_1(n)$. The Markovian evolution of a state is described by the *Lindblad equation* [GKS76, Lin76], which has the form

$$\dot{\rho} = -L(\rho) = -\Big(\mathrm{i}\,\mathrm{ad}_{H_0} + \sum_{k=1}^{r} \Gamma_{V_k}\Big)(\rho), \tag{3.1}$$

where $\mathrm{ad}_{H_0}(\rho) := [H_0, \rho]$ and $\Gamma_V(\rho) := \frac{1}{2}(V^*V\rho + \rho V^*V) - V\rho V^*$. The Hamiltonian $H_0 \in \mathfrak{iu}(n)$ is a Hermitian matrix and the Lindblad terms $\{V_k\}_{k=1}^r \subset \mathbb{C}^{n\times n}$ are arbitrary matrices. We call $-L$

a *Kossakowski–Lindblad generator*,[1] and we denote the set of all Kossakowski–Lindblad generators in n-dimensions by $\mathfrak{w}_{\mathsf{KL}}(n)$, called the *Kossakowski–Lindblad Lie wedge*, cf. [Dir+09].

The following definition encapsulates what we mean by an open Markovian quantum system with fast unitary control. Let $\{H_j\}_{j=1}^m$ be a set of Hermitian matrices, called *control Hamiltonians*, $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ a Kossakowski–Lindblad generator representing the uncontrolled drift of the system, and I an interval of the form $[0,T]$ or $[0,\infty)$. A path $\rho : I \to \mathfrak{pos}_1(n)$ of density matrices is a solution to the bilinear control system

$$\dot{\rho}(t) = -\Big(\mathrm{i}\sum_{j=1}^{m} u_j(t)\,\mathrm{ad}_{H_j} + L\Big)(\rho(t)), \quad \rho(0) = \rho_0 \in \mathfrak{pos}_1(n), \tag{$\mathcal{D}$}$$

with locally integrable *control functions* $u_j : I \to \mathbb{R}$ if ρ is absolutely continuous and satisfies ($\mathcal{D}$) almost everywhere. We will always assume that the control Hamiltonians generate at least the special unitary Lie algebra: $\langle \mathrm{i}H_j : j = 1,\dots,m\rangle_{\mathsf{Lie}} \supseteq \mathfrak{su}(n)$.

Reduced Control System

Our main focus is the reduced control system obtained from the full system ($\mathcal{D}$) using the fast controllability over the unitary group. The definitions in this section specialize those of Section 2.2, and the results of Chapter 2 then establish equivalence of these systems in a certain sense. An overview of the notation is given in Table 1.

Since the bilinear control system ($\mathcal{D}$) allows for unbounded control functions, and since the control Hamiltonians generate the entire special unitary Lie algebra — meaning that we have fast unitary control — we can quickly move within the unitary orbits. Thus we may concentrate on the dynamics of the eigenvalues of the state (as two density matrices have the same spectrum if and only if they lie on the same unitary orbit).[2]

The reduced state space will be the standard simplex

$$\Delta^{n-1} = \left\{(x_1,\dots,x_n) \in \mathbb{R}^n : \textstyle\sum_{i=1}^n x_i = 1,\ x_i \geq 0\ \ \forall i\right\},$$

representing the subset of diagonal density matrices. The standard simplex Δ^{n-1} is a convex polytope of dimension $n-1$, and its faces are lower dimensional simplices. An important group action on the simplex, stemming from the action of $\mathrm{SU}(n)$ on $\mathfrak{pos}_1(n)$, is that of the symmetric group S_n acting by coordinate permutations. Indeed, every unitary orbit in $\mathfrak{pos}_1(n)$ intersects Δ^{n-1} a finite number of times and the intersections form a permutation group orbit (this is just the unitary diagonalization of Hermitian matrices). Two faces of the same dimension can always be mapped to each other and can thus be considered equivalent. The points in the (relative) interior of a $(d-1)$-dimensional face correspond to quantum states of rank d. In particular the vertices $e_1,\dots,e_n$ correspond to the pure states and the barycenter $\mathbf{e}/n$ where, here and henceforth, $\mathbf{e} := (1,\dots,1)^\top$ corresponds to the maximally mixed state $\mathbb{1}/n$. Moreover we use the notation $\Delta_{\downarrow}^{n-1} := \{x \in \Delta^{n-1} : x_1 \geq \ldots \geq x_n\}$, which we also call the *ordered Weyl chamber*, as well as $\mathrm{spec}^{\downarrow} : \mathfrak{pos}_1(n) \to \Delta_{\downarrow}^{n-1}$ for the map which arranges the eigenvalues of the input into a vector in non-increasing order. Conversely, $\mathrm{diag} : \Delta^{n-1} \to \mathfrak{pos}_1(n)$ maps a vector to the corresponding diagonal matrix. Finally, for $\lambda \in \Delta^{n-1}$ we write $\lambda^{\downarrow}$ for the non-increasingly ordered version in $\Delta_{\downarrow}^{n-1}$.

[1] The signs are chosen such that the real parts of the eigenvalues of $-L$ are non-positive.

[2] As mentioned in the introduction, if the noise term is switchable, then one can effectively emulate ($\mathcal{D}$) even if the unitary control is not fast, at the expense of working on slower time scales.

The next step is to define the appropriate control system on the standard simplex. To motivate the definition, consider any solution $\rho : I \to \mathfrak{pos}_1(n)$ to the bilinear control system ($\mathcal{D}$) and assume that ρ is regular[3] on I. Moreover let $\rho = \mathrm{Ad}_U(\mathrm{diag}(\lambda)) := U\,\mathrm{diag}(\lambda)U^*$ be a differentiable (in time) diagonalization of ρ. Then by differentiating (cf. Lemma 1.4.1) one can show that $\dot\lambda = -L_U\lambda$, where

$$-L_U := -\Pi_{\mathrm{diag}} \circ \mathrm{Ad}_U^{-1} \circ L \circ \mathrm{Ad}_U \circ \mathrm{diag}(\cdot)\,, \tag{3.2}$$

and where $\Pi_{\mathrm{diag}} : \mathfrak{pos}_1(n) \to \Delta^{n-1}$ maps a matrix to the vector of its diagonal elements. The $-L_U$ are clearly linear and we call them *induced vector fields* on $\mathbb{R}^n$. Note that, by definition, the L_U are independent of the choice of Lindblad terms V_k. A more explicit form of the induced vector fields is given by

$$-L_U = J(U) - \mathrm{diag}(J(U)^\top \mathbf{e}), \qquad J(U) := \sum_{k=1}^r U^*V_kU \circ \overline{U^*V_kU} \tag{3.3}$$

with $\circ$ the Hadamard product, i.e. $J_{ij}(U) = \sum_{k=1}^r |\langle i|U^*V_kU|j\rangle|^2$ for all i, j. Indeed this follows from the following computation:

$$\begin{aligned} -(L_U)_{ij} &= \langle i|(\mathrm{Ad}_U^{-1} \circ L \circ \mathrm{Ad}_U)(|j\rangle\langle j|)|i\rangle = \sum_{k=1}^r |(U^*V_kU)_{ij}|^2 - \delta_{ij}\sum_{k=1}^r (U^*V_k^*V_kU)_{ii} \\ &= \sum_{k=1}^r |(U^*V_kU)_{ij}|^2 - \delta_{ij}\sum_{k=1}^r\sum_{\ell=1}^n |(U^*V_kU)_{\ell i}|^2 = J_{ij}(U) - \delta_{ij}\sum_{\ell=1}^n J_{\ell i}(U)\,, \end{aligned} \tag{3.4}$$

where δ_{ij} is the Kronecker symbol. We denote the set of induced vector fields as

$$\mathfrak{L} := \{-L_U : U \in \mathrm{SU}(n)\}\,.$$

Note that $\mathfrak{L}$ is the image of a compact set under the continuous function $U \mapsto -L_U$, hence compact itself. Also, the elements of $\mathfrak{L}$ are generators of stochastic matrices, i.e. $\mathbf{e}^\top L_U = 0$ and the off-diagonal elements are non-negative. We write this as $\mathfrak{L} \subseteq \mathfrak{stoch}(n)$, where $\mathfrak{stoch}(n)$ denotes the Lie wedge corresponding to $\mathrm{Stoch}(n)$ which is the closed subsemigroup of $\mathrm{GL}(n,\mathbb{R})$ consisting of all invertible stochastic matrices. Now $\mathfrak{L} \subseteq \mathfrak{stoch}(n)$ in particular means that the standard simplex Δ^{n-1} is (forward) invariant under the flow of the induced vector fields $-L_U$.

We define on Δ^{n-1} the set-valued function derv of *achievable derivatives* by

$$\mathsf{derv}(\lambda) := \{-L_U\lambda : U \in \mathrm{SU}(n)\} = \mathfrak{L}\lambda \subset T_\lambda\Delta^{n-1}$$

where $T_\lambda\Delta^{n-1}$ denotes the tangent cone at λ, which can always be identified with a subset of $\mathbb{R}^n_0 := \{x \in \mathbb{R}^n : x_1 + \cdots + x_n = 0\}$. With this we are ready to define the reduced control system (in two equivalent ways):

Definition 3.0.1. *A function $\lambda : I \to \Delta^{n-1}$ is a solution to the control system*

$$\dot\lambda(t) = -L_{U(t)}\lambda(t)\,, \quad \lambda(0) = \lambda_0 \in \Delta^{n-1} \tag{Λ}$$

with measurable control function $U : I \to \mathrm{SU}(n)$, if λ is absolutely continuous and satisfies (Λ) almost everywhere. Equivalently[4], a solution $\lambda : I \to \Delta^{n-1}$ is an absolutely continuous function which satisfies the differential inclusion

$$\dot\lambda(t) \in \mathsf{derv}(\lambda(t)), \quad \lambda(0) = \lambda_0 \in \Delta^{n-1}$$

almost everywhere.

[3] A state ρ is called *regular* if its eigenvalues are all distinct.

[4] This is due to Filippov's Theorem, cf. [Smi02, Thm. 2.3]. Here by "equivalent" we mean that the two systems have exactly the same set of solutions.

A convenient relaxation of the reduced control system can be obtained by allowing convex combinations of achievable derivatives.

Definition 3.0.2. *A function $\lambda : I \to \Delta^{n-1}$ is a solution of the control system*

$$\dot{\lambda}(t) \in \operatorname{conv}(\mathsf{derv}(\lambda(t))), \quad \lambda(0) = \lambda_0, \tag{$\bar{\Lambda}$}$$

if it is absolutely continuous and satisfies the differential inclusion ($\bar{\Lambda}$) almost everywhere.

The relaxation to the convex hull will slightly enlarge the set of solutions; however, every solution of ($\bar{\Lambda}$) can still be approximated uniformly (on compact time intervals) by solutions to (Λ), see [AC84, Ch. 2.4, Thm. 2].

Remark 3.0.3. *Again due to Filippov's Theorem, a solution λ to (Λ) (resp. ($\bar{\Lambda}$)) is an absolutely continuous function for which there exists $(-M_\tau)_{\tau\in I} \subset \mathfrak{L}$ (resp. $\operatorname{conv}\mathfrak{L}$) measurable such that $\lambda(t) = \lambda_0 + \int_0^t (-M_\tau)\lambda(\tau)\, d\tau$ holds for all $t \in I$.*

The main results of Chapter 2, namely Theorems 2.3.8 and 2.3.14, pertain to this setting as follows.

Theorem 3.0.4 (Equivalence Theorem). *Let $\rho : [0,T] \to \mathfrak{pos}_1(n)$ be a solution to the bilinear control system ($\mathcal{D}$) and let $\lambda^\downarrow : [0,T] \to \Delta_\downarrow^{n-1}$ be the unique path which satisfies $\lambda^\downarrow = \operatorname{spec}^\downarrow(\rho)$. Then $\lambda^\downarrow$ is a solution to the reduced control system (Λ). Conversely, let $\lambda : [0,T] \to \Delta^{n-1}$ be a solution to the reduced control system (Λ) with control function $U : [0,T] \to \mathrm{SU}(n)$. Then for every $\varepsilon > 0$ there exists a solution $\rho_\varepsilon : [0,T] \to \mathfrak{pos}_1(n)$ to the bilinear control system ($\mathcal{D}$) such that*

$$\| \operatorname{Ad}_U(\operatorname{diag}(\lambda)) - \rho_\varepsilon \|_\infty \leq \varepsilon\,.$$

Proof. We only show how the bilinear control system ($\mathcal{D}$) can be reinterpreted in the setting of semisimple orthogonal symmetric Lie algebras. The Lie algebra in question is $\mathfrak{sl}(n,\mathbb{C}) = \mathfrak{su}(n) \oplus \mathfrak{herm}_0(n,\mathbb{C})$, where $\mathfrak{herm}_0(n,\mathbb{C})$ denotes the traceless Hermitian matrices, see Example 1.2.1. Since for density matrices ρ it holds that $\operatorname{tr}(\rho) = 1$, we will consider the shifted operator $\tilde{\rho} := \rho - \mathbb{1}/n$ which satisfies $\tilde{\rho} \in \mathfrak{herm}_0(n,\mathbb{C})$. The corresponding shifted Kossakowski–Lindblad generator has the form $\tilde{L}(\bar{\rho}) := L(\rho) = L(\tilde{\rho}) + L(\mathbb{1})/n$, which is affine linear. Since $\operatorname{Ad}_U(\tilde{\rho}) = \widetilde{\operatorname{Ad}_U(\rho)}$ and $\operatorname{ad}_H(\rho) = \operatorname{ad}_H(\tilde{\rho})$, the "shifted" control system $\dot{\tilde{\rho}} = -(\operatorname{ad}_H + \tilde{L})(\tilde{\rho})$ and ($\mathcal{D}$) are state space equivalent[5]. After this transformation, the new system is in the form considered in Chapter 2, and hence we can define the equivalent reduced control system. The reduced state space is $\mathfrak{diag}_0(n,\mathbb{R}) \cong \mathbb{R}_0^n$, and the induced vector fields are $-\tilde{L}_U\tilde{\lambda} = -\Pi_{\mathrm{diag}} \circ \operatorname{Ad}_U^{-1} \circ \tilde{L} \circ \operatorname{Ad}_U \circ \operatorname{diag}(\tilde{\lambda})$. Applying the definitions one finds that $L_U\lambda = \tilde{L}_U\tilde{\lambda}$. This shows that the "shifted" reduced control system is state space equivalent to the reduced control system (Λ). This (linear) state space equivalence shows that show that Theorems 2.3.8 and 2.3.14 imply the desired result. More generally, all results from Chapter 2 presupposing an affine linear drift term can be applied to our system. □

Corollary 3.0.5. *Let $H : [0,T] \to \mathfrak{isu}(n)$ be an integrable Hamiltonian. Consider a solution $\rho : [0,T] \to \mathfrak{pos}_1(n)$ to $\dot{\rho}(t) = -(\mathrm{i}\operatorname{ad}_{H(t)} + L)\rho(t)$. Then $\operatorname{spec}^\downarrow(\rho(t))$ is a solution to (Λ) and a fortiori to ($\bar{\Lambda}$).*

[5] Two control systems are state space equivalent if there is a smooth diffeomorphism between their state spaces which also maps the drift and control vector fields of one system to the other.

Proof. Choose any control Hamiltonians $\{H_j\}_{j=1}^m$ which linearly span $\mathfrak{isu}(n)$. Then we may write $H(t) = \sum_{j=1}^m u_j(t) H_j$ with integrable control functions u_j. Hence ρ can be seen as a solution to ($\mathcal{D}$) (with the chosen controls) and so we may apply Theorem 3.0.4 to obtain that $\mathrm{spec}^{\downarrow}(\rho(t))$ is a solution of (Λ). □

The proof above uses the fact that while "the" bilinear control system ($\mathcal{D}$) requires a choice of control Hamiltonians to be fully defined, the reduced control system (Λ) is independent of this choice so long as the control Hamiltonians generate the entire special unitary Lie algebra.

We emphasize that although the "lifting" part in Theorem 3.0.4 is only approximate in general, in many relevant cases one can obtain stronger results. When the path is regular (recall footnote 3), then the lift can be performed exactly and explicitly using Proposition 2.3.10. Moreover, due to the structure theory of Kossakowski–Lindblad generators, it is often the case that one can choose time-independent Hamiltonians to achieve practical tasks. This is summarized in Appendix 4.A.

To close out we briefly give two useful properties of the matrices $J(U)$ and L_U.

Lemma 3.0.6. *Given arbitrary $\{V_k\}_{k=1}^r \subset \mathbb{C}^{n\times n}$ as well as any $i, j \in \{1, \ldots, n\}$ with $i \neq j$, the following statements hold.*

(i) Given $\alpha_1, \ldots, \alpha_m > 0$ with $\sum_{k=1}^m \alpha_k = 1$ as well as $U_1, \ldots, U_m \in \mathrm{SU}(n)$ define $M := \sum_{k=1} \alpha_k L_{U_k}$. If $M_{ij} = 0$, then $(L_{U_k})_{ij} = 0$ for all $k = 1, \ldots, m$.

(ii) If $(L_U)_{ij} = 0$ for some $U \in \mathrm{SU}(n)$, then $(U^* V_k U)_{ij} = 0$ for all $k = 1, \ldots, r$.

Proof. (i): As $-L_{U_k}$ is the generator of a stochastic matrix we know $(-L_{U_k})_{ij} \geq 0$ because $i \neq j$. Thus $(-M)_{ij} = 0$ implies $(-\alpha_k L_{U_k})_{ij} = 0$ for all k, hence $(L_{U_k})_{ij} = 0$. (ii): Because $i \neq j$ we know that $(-L_U)_{ij} = (J(U))_{ij}$ so the former being zero forces $(U^* V_k U)_{ij} = 0$ for all k. □

Lemma 3.0.7. *Let $\{V_k\}_{k=1}^r$ be a family of n-dimensional Lindblad terms and let $U \in \mathrm{U}(n)$ be arbitrary. The row and column sums of $J(U)$ (as defined in (3.3)) are*

$$(J(U)\mathbf{e})_i = \sum_{k=1}^r \langle u_i | V_k V_k^* | u_i \rangle, \quad (J(U)^\top \mathbf{e})_j = \sum_{k=1}^r \langle u_j | V_k^* V_k | u_j \rangle$$

respectively, where $|u_i\rangle = U|i\rangle$. In particular if $\preceq$ denotes standard majorization, then $J(U)\mathbf{e} \preceq \mathrm{spec}^{\downarrow}(\sum_{k=1}^r V_k V_k^*)$ and $J(U)^\top \mathbf{e} \preceq \mathrm{spec}^{\downarrow}(\sum_{k=1}^r V_k^* V_k)$. It follows that

$$((J(U) - J(U)^\top)\mathbf{e})_i = \sum_{k=1}^r \langle u_i | [V_k, V_k^*] | u_i \rangle \,, \quad ((J(U) + J(U)^\top)\mathbf{e})_i = \sum_{k=1}^r \langle u_i | \{V_k, V_k^*\} | u_i \rangle \,,$$

where $\{\cdot, \cdot\}$ denotes the anti-commutator. So $(J(U) - J(U)^\top)\mathbf{e} \preceq \mathrm{spec}^{\downarrow}(\sum_{k=1}^r [V_k, V_k^*])$ and $(J(U) + J(U)^\top)\mathbf{e} \preceq \mathrm{spec}^{\downarrow}(\sum_{k=1}^r \{V_k, V_k^*\})$.

Proof. We compute the i-th row sum

$$(J(U)\mathbf{e})_i = \sum_{j=1}^n \langle i | J(U) | j \rangle = \sum_{j,k=1}^{n,r} |\langle u_i | V_k | u_j \rangle|^2 = \sum_{k=1}^r \langle u_i | V_k V_k^* | u_i \rangle.$$

The computation for the j-th column is analogous, and the other claims follow immediately using the Schur–Horn Theorem [Sch23, Hor54]. □

Operator Lifts and Lie Wedges

For completeness, we briefly discuss the operator lifts of the full bilinear system ($\mathcal{D}$) and of the reduced system (Λ). The definitions, in a more general context, are recalled in Section 2.2. Considering the operator lifts immediately leads to the study of Lie semigroups and Lie wedges, cf. [Dir+09, HHL89, Law99]. In particular, the reachable sets of the operator lifts are determined by the Lie saturate $\langle\cdot\rangle_{\mathsf{LS}}$ of their respective generators, see [Law99, Sec. 6]. The generators of ($\mathcal{D}$) are given by $\Omega = \{X + \mathrm{i}\,\mathrm{ad}_{H_j} : j = 1, \dots, m\}$, and the generators of (Λ) are exactly the induced vector fields $\mathfrak{L}$.

Proposition 3.0.8. *The Lie saturate of the full control system ($\mathcal{D}$) is given by*

$$\langle\Omega\rangle_{\mathsf{LS}} = \mathrm{ad}_{\mathfrak{su}(n)} \oplus \langle \mathrm{Ad}_U^{-1} \circ L \circ \mathrm{Ad}_U : U \in \mathrm{SU}(n)\rangle_{\mathsf{wedge}}.$$

Proof. If $-L = \mathrm{i}\,\mathrm{ad}_H$ for some $\mathrm{i}H \in \mathfrak{su}(n)$, then it is clear that the Lie wedge is $\mathrm{ad}_{\mathfrak{su}(n)}$ since it is a compact Lie algebra. Hence we can focus on the case where $-L \notin \mathrm{ad}_{\mathfrak{su}(n)}$. Since $\mathfrak{w}_{\mathsf{KL}}(n)$ is a global Lie wedge (cf. [Dir+09, Thm. 3.3]) and since the edge of $\mathfrak{w}_{\mathsf{KL}}(n)$ equals $\mathrm{ad}_{\mathfrak{su}(n)}$ (e.g. by Lemma 4.4.8 (ii)), the result follows from Proposition 2.2.1 (iii) (which itself is based on [HN12, Prop. 1.37]). □

Furthermore, the Lie saturate of the reduced control system (Λ) is $\langle\mathfrak{L}\rangle_{\mathsf{LS}} = \langle\mathfrak{L}\rangle_{\mathsf{wedge}}$. This follows immediately from the fact that $\mathfrak{stoch}(n)$ is a pointed (hence global) Lie wedge by use of [HN12, Prop. 1.37]. As a consequence we obtain that the Lie saturates are related by $\Pi_{\mathrm{diag}} \circ \langle\Omega\rangle_{\mathsf{LS}} \circ \mathrm{diag} = \langle\mathfrak{L}\rangle_{\mathsf{LS}}$, see also Lemma 2.2.3.

CHAPTER 4

Reachability and Stabilizability

4.1 Introduction

Before one can find optimal controls for concrete problems, it is important to understand some fundamental properties of the control system in question. For example: Which states and subspaces can be stabilized? Which states can be reached from a given initial state? Is the system coolable, controllable, or accessible? We will give quite general answers to these questions for Markovian systems with fast unitary control using the reduced control system of Chapter 3. We emphasize that certain stabilization and reachability tasks of interest can be implemented using only time-independent control Hamiltonians. Moreover, in a toy-model setting, we obtain some explicit reachable and stabilizable sets exhibiting non-trivial geometries. The results of this chapter will prove to be very useful in studying concrete systems in Chapters 5 and 6.

Previous results for time-independent Hamiltonian control are given in [Kra+08, TV09, SW10], and our results show what improvements can (and cannot) be obtained from using time-dependent Hamiltonian control. The toy model for quantum thermodynamics was thoroughly explored in [DES19, End20, SED22], see also [RBR18], and we will improve on these results by deriving explicit solutions. A particular section is devoted to the unital case, where our general answers can be further specialized, and where we recover and generalize a number of known results [Yua10, SAZ19].

Outline and Main Results

At the end of Section 4.1 we introduce two matrix algebras whose invariant subspaces are important in the structure theory of Kossakowski–Lindblad generators, as well as in the study of time-independent Hamiltonian control and of the reduced control system. When proceeding to study the control-theoretic properties of our open Markovian quantum system, we use (i) the reduced control system to find algebraic characterizations and (ii) the Equivalence Theorem 3.0.4 to lift the results to the full control system. A compact overview of the main results is provided in Table 4.1.

In Section 4.2 we establish stabilizability of individual points and the entire system (Thm. 4.2.7), the viability of faces of the simplex, and the accessibility of the system (Prop. 4.2.12, and for unital systems Prop. 4.4.11).

Section 4.3 is devoted to reachability, in particular (asymptotic) coolability (Thm. 4.3.7), reverse coolability, and the reachability of faces, with the conditions for (approximate) controllability of the system settled by Prop. 4.3.20.

Section 4.4 treats the special case of unital quantum systems, where the algebraic structure simplifies considerably and allows to derive stronger results on stabilizability and reachability (Thm. 4.4.9).

In Appendix 4.A we explore the structure of Kossakowski–Lindblad generators via associated matrix algebras. In Appendix 4.B we compute explicit stabilizable and reachable sets in a simplified toy model setting. Appendix 4.C shows how common eigenvectors and simultaneous triangulations of matrices can be efficiently computed, thus giving a method for checking coolability and stabilizability of systems.

Kossakowski–Lindblad Generators and Their Relaxation Algebras

Throughout the chapter we assume that $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ is a Kossakowski–Lindblad generator on the n-dimensional Hilbert space $\mathbb{C}^n$, given by Lindblad terms $\{V_k\}_{k=1}^r$ and Hamiltonian H_0. We say that a given set $\{V'_k\}_{k=1}^s \subset \mathbb{C}^{n,n}$ is a choice of Lindblad terms of $-L$ if there exists a Hermitian $H'_0 \in \mathbb{C}^{n,n}$ such that $-L = -(\mathrm{i}\,\mathrm{ad}_{H'_0} + \sum_{k=1}^s \Gamma_{V'_k})$. The freedom of representation of $-L$ is summarized in Lemma 4.A.3. Recall that $-L$ is called *unital* if $L(\mathbb{1}) = 0$ and *purely Hamiltonian* if $-L = -\mathrm{i}\,\mathrm{ad}_{H_0}$.

We will briefly introduce the main concepts here, referring to Appendix 4.A for precise statements. To the generator $-L$ we define an associated (complex) matrix algebra $\mathcal{V}$, called its *relaxation algebra*, generated by "the" Lindblad terms $\{V_k\}_{k=1}^r$ and the identity matrix, as well as its *extended relaxation algebra* $\mathcal{V}^+$ additionally generated by $K = \mathrm{i}H_0 + \frac{1}{2}\sum_{k=1}^r V_k^* V_k$. Importantly, these matrix algebras are well-defined and their invariant subspaces encode important information about the generator $-L$. Indeed, many results about the structure of Kossakowski–Lindblad generators presented in [BN08] can be formulated succinctly using the algebras $\mathcal{V}$ and $\mathcal{V}^+$. Moreover, $\mathcal{V}$ was used to give a sufficient condition for the evolution to be relaxing in the sense of having a unique attractive fixed point, cf. [Dav70, Spo77]. The invariant subspaces of $\mathcal{V}$ are called *lazy subspaces*, and invariant subspaces of $\mathcal{V}^+$ are called *collecting subspaces*. Under the evolution of $-L$, a state supported on a lazy subspace will only leave the subspace "slowly", whereas a state supported on a collecting subspace will not leave the subspace at all. Importantly, using time-independent controls, any lazy subspace can be turned into a collecting one. A collecting subspace whose orthocomplement is also collecting is called an *enclosure*, and it corresponds to a symmetry of the generator $-L$, cf. [BN08, AJ14]. Interestingly, in the unital case, $\mathcal{V}$ and $\mathcal{V}^+$ turn out to be $*$-algebras, which are highly structured. This allows us to derive strong results, beyond the known fact that a unital system is relaxing if and only if $\mathcal{V} = \{V_k : k = 1, \ldots, r\}'' = \mathbb{C}^{n,n}$ (with $\{\cdot\}''$ denoting the double-commutant in $\mathbb{C}^{n,n}$), which is similar to the result in [Spo77].

These relaxation algebras are useful in understanding what can be achieved using time-independent Hamiltonian control, cf. [Kra+08, TV09, SW10] as well as Appendix 4.A, and they also turn out to be crucial in the study of the reduced control system.

4.2 Stabilizability, Viability and Accessibility

An important task in control theory is that of keeping the state in a certain region of the state space, called viability, or close to some desired state, called stabilizability. In this section, we characterize viability and stabilizability in the reduced control system and deduce the implications for the full bilinear system. Moreover, we study accessibility in the reduced system and show that non-unital systems are generically directly accessible.

Stabilizable and Strongly Stabilizable Points

We begin with stabilizable points, emphasizing that our systems do not have feedback, hence why we talk about open-loop stabilizability only.

Definition 4.2.1. *A point $\lambda \in \Delta^{n-1}$ is called* stabilizable *for (Λ) if it holds that $0 \in \operatorname{conv}(\mathsf{derv}(\lambda))$. The set of all stabilizable points is denoted stab_Λ. We say that λ is* strongly stabilizable *for (Λ) if $0 \in \mathsf{derv}(\lambda)$.*

Using previously established notation, λ is strongly stabilizable if and only if $0 \in \mathfrak{L}\lambda$, and λ is stabilizable if and only if $0 \in \operatorname{conv}(\mathfrak{L}\lambda) = (\operatorname{conv}\mathfrak{L})\lambda$.

It is clear that strong stabilizability implies stabilizability. Intuitively, λ is stabilizable if any solution to (Λ) starting at λ can remain close to λ for an arbitrarily long amount of time, and λ is strongly stabilizable if the constant path at λ is a solution to (Λ). For a more precise statement, see Section 2.4. The following result is a direct specialization of Proposition 2.4.5.

Proposition 4.2.2. *Let $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ be any Kossakowski–Lindblad generator, and let $\rho_0 \in \mathfrak{pos}_1(n)$, $\lambda_0 \in \Delta^{n-1}$ and $U \in \mathrm{SU}(n)$ be given such that $\rho_0 = \mathrm{Ad}_U(\mathrm{diag}(\lambda_0))$. Then the following hold:*

(i) If there is some Hamiltonian H such that $-(\mathrm{i}\,\mathrm{ad}_H + L)(\rho_0) = 0$, then $-L_U\lambda_0 = 0$ and hence λ_0 is strongly stabilizable for (Λ).

(ii) Conversely, if λ_0 is regular and strongly stabilizable for (Λ) with $-L_U\lambda_0 = 0$, then the compensating Hamiltonian *$H_c := \mathrm{ad}_{\rho_0}^{-1} \circ \Pi_{\rho_0}^{\perp} \circ L(\rho_0)$ satisfies*

$$-(\mathrm{i}\,\mathrm{ad}_{H_c} + L)(\rho_0) = 0\,.$$

Note that the assumption on regularity is generally necessary, as exemplified in Example 2.3.12.

Denoting $\tilde{V}_k = U^* V_k U$, and analogously $\tilde{H}_c$ and $\tilde{H}_0$, the compensating Hamiltonian H_c for a regular, strongly stabilizable state $\rho = \mathrm{Ad}_U(\mathrm{diag}(\lambda))$ can be written more explicitly as follows. Using equivariance and the fact that, by regularity of ρ, we have $\Pi_{\mathrm{diag}(\lambda)} = \Pi_{\mathrm{diag}}$, we find $-\mathrm{i}H_c = \mathrm{Ad}_U \circ \mathrm{ad}_\lambda^{-1} \circ \Pi_{\mathrm{diag}}^{\perp} \circ \mathrm{Ad}_U^{-1} \circ L \circ \mathrm{Ad}_U(\mathrm{diag}(\lambda))$. Hence $-\mathrm{i}\langle i|\tilde{H}_c + \tilde{H}_0|j\rangle = \sum_{k=1}^r \langle i|\Gamma_{\tilde{V}_k}(\lambda)|j\rangle/(\lambda_i - \lambda_j)$ which expands to

$$-\mathrm{i}\langle i|\tilde{H}_c|j\rangle = \mathrm{i}\langle i|\tilde{H}_0|j\rangle + \sum_{k=1}^{r} \frac{\frac{\lambda_i+\lambda_j}{2}\langle i|\tilde{V}_k^*\tilde{V}_k|j\rangle - \sum_{\ell=1}^n \lambda_\ell \langle i|\tilde{V}_k|\ell\rangle\langle\ell|\tilde{V}_k^*|j\rangle}{\lambda_i - \lambda_j}\,,$$

as desired.

Note that more generally, for regular $\rho_0 = \mathrm{Ad}_U(\mathrm{diag}(\lambda_0))$ the compensating Hamiltonian H_c exactly cancels out the part of $-L(\rho_0)$ which is tangent to the orbit.

Computing the set of all (strongly) stabilizable states is difficult in general. Even in the three-dimensional toy model case considered in Appendix 4.B this is non-trivial, see also [RBR18, Sec. IV].

Viable Faces

Viable subsets of the state space are those in which one can remain for an arbitrary amount of time. Note that the singleton set $\{\lambda\}$ is viable for ($\bar{\Lambda}$) if and only if λ is stabilizable for (Λ), cf. Section 2.4.

We are primarily interested in the case where the subset in question is a face of the simplex Δ^{n-1}. Due to the permutation symmetry of the reduced system, as well as the geometry of the simplex Δ^{n-1}, all $(d-1)$-dimensional faces of the simplex are equivalent. The main result of this section will relate viable faces of dimension $d-1$ to lazy subspaces (i.e. common invariant subspaces of the relaxation algebra $\mathcal{V}$, or equivalently, of any choice of Lindblad terms V_k which define $-L$, cf. Appendix 4.A) of dimension d. We begin with a simple lemma:

Lemma 4.2.3. *Let F be a face of Δ^{n-1} of dimension $d-1$ with $1 \leq d \leq n$. Let λ be a point in the relative interior[1] of F and let $-M \in \mathrm{conv}(\mathfrak{L})$ be such that $-M\lambda$ is tangent to F. Then there exists a lazy subspace of dimension d.*

Proof. Using permutations we can assume that F consist of all vectors in Δ^{n-1} whose last $n-d$ entries are zero. Since λ lies in the relative interior of F, its first d entries must be strictly positive. By assumption, the last $n-d$ entries of $M\lambda$ must be zero. But this can only be if the block $(d+1,\ldots,n)\times(1,\ldots,d)$ of M is identically 0. By Lemma 3.0.6 there exists $U \in \mathrm{U}(n)$ such that all U^*V_kU have this property, hence the first d columns of any such U^* span the desired lazy subspace. □

Proposition 4.2.4. *Let F be a face of Δ^{n-1} of dimension $d-1$ (where $1 \leq d \leq n$). Given any Kossakowski–Lindblad generator $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$, the following are equivalent:*

(i) *F is viable for ($\bar{\Lambda}$).*

(ii) *There exists some $-M \in \mathrm{conv}(\mathfrak{L})$ whose flow leaves F invariant.*

(iii) *There exists a lazy subspace of dimension d.*

(iv) *There exists a stabilizable point for (Λ) in the relative interior of F.*

(v) *There exists a strongly stabilizable point for (Λ) in the relative interior of F.*

(vi) *There exists some $-M \in \mathrm{conv}(\mathfrak{L})$ and some λ in the relative interior of F such that $-M\lambda$ is tangent to F.*

(vii) *There exists some $U \in \mathrm{U}(n)$ and some λ in the relative interior of F such that $-L_U\lambda$ is tangent to F.*

(viii) *There exists a* stabilizing Hamiltonian *H_S and a state ρ of rank d such that ρ is the unique fixed point of $-(\mathrm{i}\,\mathrm{ad}_{H_S}+L)$ restricted to the support of ρ (which is automatically a collecting subspace).*

Proof. The implications (iii) ⇒ (ii) ⇒ (i) ⇒ (vi) are easy. By Lemma 4.2.3 and its proof it holds that (vi) ⇒ (vii) ⇒ (iii). That (iii) ⇒ (viii) follows from Lemma 4.A.19. Finally, (viii) ⇒ (v) follows from Proposition 4.2.2, and (v) ⇒ (iv) ⇒ (vi) are clear. □

Recall from Corollary 4.A.11 that for a lazy subspace S, any H_S satisfying $P_S^{\perp}H_SP_S = -P_S^{\perp}\big(H_0+\frac{1}{2\mathrm{i}}\sum_{k=1}^r V_k^*V_k\big)P_S$ is a stabilizing Hamiltonian, where P_S is the orthogonal projection onto S. This is a well-known result, cf. [BN08, LCW98], and Proposition 4.2.4 shows that such time-independent control is always sufficient for making a face viable.

For the vertices e_i of Δ^{n-1}, which correspond to pure states, Proposition 4.2.4 specializes as follows:

Corollary 4.2.5. *Given any Kossakowski–Lindblad generator $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$, the following are equivalent:*

(i) *Some (equivalently: each) e_i is stabilizable for (Λ).*

(ii) *Some (equivalently: each) e_i is strongly stabilizable for (Λ).*

[1] The relative interior of a set S, denoted $\mathrm{relint}(S)$, is the interior of S within its affine hull. Note that for a singleton set $\{s\}$ the relative interior is $\{s\}$ itself.

(iii) For some (equivalently: each) choice of Lindblad terms $\{V_k\}_{k=1}^r$ *of* $-L$*, the* V_k *have a common eigenvector.*

(iv) There is a stabilizing Hamiltonian H_ψ *for a pure state*[2]*, that is, there exists* $\|\psi\| = 1$ *such that* $-(\mathrm{i}\,\mathrm{ad}_{H_\psi} + L)|(|\psi\rangle\langle\psi|) = 0$.

Moreover, if $|\psi\rangle$ *is a common eigenvector of the* V_k *with* $V_k|\psi\rangle = \lambda_k|\psi\rangle$*, then any Hermitian* H_ψ *satisfying* $H_\psi|\psi\rangle = -\big(H_0 + (\lambda - \frac{1}{2\mathrm{i}}\sum_{k=1}^r |\lambda_k|^2)\mathbb{1} + \frac{1}{2\mathrm{i}}\sum_{k=1}^r V_k^* V_k\big)|\psi\rangle$ *where* $\lambda \in \mathbb{R}$ *is arbitrary is an admissible choice for the stabilizing Hamiltonian.*

Later we will see how certain systems can be asymptotically cooled into a pure state. The corollary above then shows how such a pure state can be stabilized.

Remark 4.2.6. *The permutation symmetry of the reduced control system (cf. Lemma 2.A.2) yields some further viable subsets. Indeed every degeneracy plane (i.e. the fixed point set of some permutation group) is viable. Similarly one can show that the ordered Weyl chamber* $\Delta_\downarrow^{n-1}$ *as well as each of its permutations are viable, see Proposition 2.A.4.*

Stabilizable Systems

If every point in Δ^{n-1} is stabilizable for (Λ), we say that the system is *stabilizable* for (Λ). It turns out that stabilizable systems can be characterized in simple algebraic terms.

Theorem 4.2.7 (Stabilizable Systems). *Given any* $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$*, the following are equivalent:*

(i) The system is stabilizable for (Λ).

(ii) For some (equivalently: each) choice of Lindblad terms $\{V_k\}_{k=1}^r$ *of* $-L$*, all* V_k *are simultaneously triangulable.*

(iii) For some (equivalently: each) choice of Lindblad terms $\{V_k\}_{k=1}^r$ *of* $-L$*, the* V_k *generate a solvable Lie algebra.*

In particular if all Lindblad terms V_k commute (e.g., if there is only one V_k), then the system is stabilizable for (Λ). If the system is stabilizable for (Λ), then one can show that in the full system ($\mathcal{D}$), every $\mathrm{SU}(n)$-orbit is approximately viable, i.e. one can stay arbitrarily close to any orbit, see Proposition 2.4.7.

We will prove the theorem as a sequence of lemmas:

Lemma 4.2.8. *Let* $A \in \mathbb{R}^{n,n}$ *be a matrix with non-positive values on the diagonal and non-negative values on the off-diagonal. Let* $v \in \mathbb{R}^n$ *be a vector with non-negative elements such that* $Av = 0$*. For all* $i \neq j$ *we have that* $A_{ij}\, v_j \leq \|A\|_{\max} v_i$*, where* $\|A\|_{\max} := \max_{i,j} |A_{ij}|$.

Proof. Note that $(Av)_i = 0$ is equivalent to $\sum_{k\neq i} A_{ik}v_k = -A_{ii}v_i = |A_{ii}|v_i$. This yields $A_{ij}v_j \leq \sum_{k\neq i} A_{ik}v_k = |A_{ii}|v_i \leq \|A\|_{\max}v_i$, for all $i \neq j$ as desired. □

Lemma 4.2.9. *If the system is stabilizable, then the Lindblad terms* $\{V_k\}_{k=1}^r$ *are simultaneously triangulable.*

[2]Note that this does *not* follow from Proposition 4.2.2 since e_i is not regular.

Proof. For $\varepsilon > 0$ small enough we consider the curve $\lambda : [0, \varepsilon] \to \Delta^{n-1}$ defined as $\lambda(t) = (1 - (t + t^2 + t^3 + \ldots + t^{n-1}), t, t^2, t^3, \ldots, t^{n-1})$. By assumption, for all $\varepsilon > 0$, there exists $-M_t \in \mathrm{conv}(\mathfrak{L})$ such that $-M_t\lambda(t) = 0$. Applying Lemma 4.2.8 to $A = -M_t$ and $v = \lambda(t)$ we have that if $i > j$,

$$|(M_t)_{ij}| \leq m^\star \frac{\lambda_i(t)}{\lambda_j(t)} \to 0 \tag{4.1}$$

as $t \to 0$ where $m^\star = \sup_t \|M_t\|_{\max} \leq \sup_{U \in \mathrm{SU}(n)} \|L_U\|_{\max}$ which is finite by compactness of $\mathrm{SU}(n)$ and continuity of $-L$. Moreover, by compactness of $\mathrm{conv}(\mathfrak{L})$ we can pass to a subsequence of the $-M_t$ which converges to some $-M \in \mathrm{conv}(\mathfrak{L})$. Eq. (4.1) then implies that M is upper triangular. With this, Lemma 3.0.6 shows that all Lindblad terms can be simultaneously triangulated. □

Lemma 4.2.10. *If all Lindblad terms are simultaneously triangulable, then the system is stabilizable.*

Proof. By assumption there exists $U \in \mathrm{SU}(n)$ such that U^*V_kU is upper triangular for all $k = 1, \ldots, r$. Given any $m = 1, \ldots, n$ we define $J^{(m)}(U)$ as the $m \times m$ block in the lower right corner of $J(U)$, as well as $-L_U^{(m)} := J^{(m)}(U) - \mathrm{diag}((J^{(m)}(U))^\top \mathbf{e})$ (note that $-L_U^{(m)}$ is a stochastic generator, but it is generally not a submatrix of L_U). This definition yields an inductive structure: because all U^*V_kU are upper triangular, so is $J^{(m)}(U)$, meaning for all $m = 1, \ldots, n-1$ one has

$$-L_U^{(m+1)} = \begin{pmatrix} 0 & (J^{(m+1)}(U))_{1,2} & \cdots & \cdots & (J^{(m+1)}(U))_{1,m} \\ 0 & -(J^{(m+1)}(U))_{1,2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & -(J^{(m+1)}(U))_{1,m} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -L_U^{(m)} \end{pmatrix}$$

as is readily verified. Our goal is to show that for every $1 \leq m \leq n$ and for every $\lambda \in \Delta^{m-1}$ there exist $\ell \in \mathbb{N}$, $\mu \in \Delta^{\ell-1}$, and permutation matrices $P_1, \ldots, P_\ell \in \mathbb{R}^{m,m}$ such that $-\sum_{i=1}^{\ell} \mu_i P_i^\top L_U^{(m)} P_i \lambda = 0$. This would conclude the proof as then $-\sum_{i=1}^{\ell} \mu_i P_i^\top L_U^{(n)} P_i = -\sum_{i=1}^{\ell} \mu_i L_{UP_i} \in \mathrm{conv}(\mathfrak{L})$ maps λ to zero.

We proceed by induction on m. The case $m = 1$ is trivial as $-L_U^{(1)} = 0$. For the induction step, let any $1 \leq m < n$ and any $\lambda \in \Delta^m$ be given. We distinguish two cases: if there exists a permutation matrix $P \in \mathbb{R}^{(m+1),(m+1)}$ such that $P\lambda = e_1$, then $-P^\top L_U^{(m+1)} P\lambda = 0$ by the triangular structure of L_U. If this is not the case, then the vector $\lambda^{(q)} := (\lambda_1, \ldots, \lambda_{q-1}, \lambda_{q+1}, \ldots, \lambda_{m+1})/(1 - \lambda_q) \in \Delta^{m-1}$ is well-defined for all $q = 1, \ldots, m+1$. By the induction hypothesis, for all $q = 1, \ldots, m+1$ there exist $\ell_q \in \mathbb{N}$, $\mu_q \in \Delta^{m-1}$, as well as permutation matrices $P_{q,1}, \ldots, P_{q,\ell_q} \in \mathbb{R}^{m,m}$ such that $-\sum_{i=1}^{\ell_q} (\mu_q)_i P_{q,i}^\top L_U^{(m)} P_{q,i} \lambda^{(q)} = 0$. Defining $P'_{q,i} := 1 \oplus P_{q,i} \in \mathbb{R}^{(m+1),(m+1)}$ for all $i = 1, \ldots, \ell_q$, and $q = 1, \ldots, m+1$, the block structure of $-L_U^{(m+1)}$ from above readily implies that $-\sum_{i=1}^{\ell_q} (\mu_q)_i (P'_{q,i})^\top L_U^{(m+1)} P'_{q,i} (\lambda_q, \lambda^{(q)}(1 - \lambda_q))^\top$ has a non-negative first element and all other elements are non-positive. Therefore

$$x_q := -\sum_{i=1}^{\ell_q} (\mu_q)_i (P'_{k,i}\pi_q)^\top L_U^{(m+1)} P'_{q,i}\pi_q\lambda = -\pi_q \Big(\sum_{i=1}^{\ell_q} (\mu_q)_i (P'_{q,i})^\top L_U^{(m+1)} P'_{q,i} \begin{pmatrix} \lambda_q \\ \lambda^{(q)}(1 - \lambda_q) \end{pmatrix} \Big)$$

for all $q = 1, \ldots, m+1$ has a non-negative element in the q-th component while all others are non-positive; here $\pi_q \in \mathbb{R}^{(m+1),(m+1)}$ which only swaps the first and the q-th component. It remains to find $\xi \in \Delta^m$ such that $\sum_{q=1}^{m+1} \xi_q x_q = 0$ as then $0 = -\sum_{q=1}^{m+1} \sum_{i=1}^{\ell_q} \xi_q \mu_i \pi_q^\top (P'_{q,i})^\top L_U^{(m+1)} P'_{q,i} \pi_q \lambda$ so $(\xi_k \mu_i)_{q,i}$ and $\{P'_{q,i}\pi_q : q, i\}$ satisfy the property we are aiming to verify.

For the final step the key is that $X := \begin{pmatrix} x_1 & \cdots & x_{m+1} \end{pmatrix} \in \mathbb{R}^{(m+1),(m+1)}$ satisfies $\mathbf{e}^\top X = 0$ and $X_{ij} \geq 0$ for all $i \neq j$ by construction of the x_q. Therefore there exists $\varepsilon > 0$ such that $\mathbb{1} + \varepsilon X$ is a stochastic matrix (if $X = 0$ this is trivial, else choose $\varepsilon := (\max_j |X_{jj}|)^{-1}$). In particular $(\mathbb{1} + \varepsilon X)\Delta^m \subseteq \Delta^m$ so by the Brouwer fixed-point theorem [Bro11] this matrix has a fixed point $\xi \in \Delta^m$. But this means $\xi + \varepsilon X\xi = \xi$ which, due to $\varepsilon > 0$, is equivalent to $X\xi = 0$, as desired. □

Note that if the V_k are upper triangular in the standard basis, only permutations are used in the proof of Lemma 4.2.10 to stabilize a given state. Hence the proof also works for the toy model discussed in Appendix 4.B.

Lemma 4.2.11. *A set of matrices in $\mathbb{C}^{n,n}$ is simultaneously triangulable if and only if they generate a solvable Lie algebra $\mathfrak{g}$.*

Proof. By Lie's Theorem [Kna02, Thm. 1.25], if $\mathfrak{g}$ is solvable, then it is triangulable. Conversely, it is clear that a Lie algebra generated by simultaneously triangulable matrices is solvable. □

This completes the proof of Theorem 4.2.7.

The theorem shows that determining whether a collection of square matrices over $\mathbb{C}$ is simultaneously (unitarily)[3] triangulable is a relevant task for studying our (reduced) control system. Simultaneous triangulation is certainly possible if all matrices commute, but we can say more. Many equivalent conditions are listed in [RR00, Ch. 1]. Also recall that for a Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{C})$ it holds that $\mathfrak{g}$ is strictly triangulable if and only if every element of $\mathfrak{g}$ is nilpotent (Engel's Theorem [Kna02, Thm. 1.35]). An efficient algorithm for deciding if a set of matrices is simultaneously triangulable and for finding a corresponding basis is presented in Appendix 4.C.

An interesting follow-up question would be to understand when every state in $\lambda \in \Delta^{n-1}$ is strongly stabilizable and how to find U such that $-L_U\lambda = 0$.

Accessibility

As we will see below, the systems we consider are never (exactly) controllable. This is due to the dissipative nature of the dynamics. In this case, relevant notions are approximate controllability or accessibility. Indeed, we will show that our systems are almost everywhere directly accessible. The reduced system (Λ) is *directly accessible* at some $\lambda \in \Delta^{n-1}$ if $\operatorname{span}(\mathsf{derv}(\lambda)) = \mathbb{R}^n_0$. Intuitively this means that one can move directly in all directions of some cone with non-empty interior. Using Propositions 2.4.9 and 2.4.10, we immediately obtain:

Proposition 4.2.12. *If $-L$ is non-unital, then the reduced control system (Λ) is generically[4] directly accessible, and hence the full bilinear system ($\mathcal{D}$) is generically accessible.*

Direct accessibility in the unital case will be considered in Section 4.4. The accessibility of the full system ($\mathcal{D}$) and its operator lift have been addressed in greater generality in [KDH12] using Lie-theoretic methods.

[3] A family of complex matrices is triangulable if and only if it is unitarily triangulable. This can be seen, for instance, using the QR decomposition.

[4] We say that a property holds generically on the state space if it holds on an open, dense subset which has full measure.

4.3 Reachability, Coolability and Controllability

One of the main questions in control theory is that of reachability, i.e. given an initial state, what is the set of all states that can be generated within the control system? Reaching pure states is of particular importance since it corresponds to coolability of the system.[5] Furthermore, we will characterize which faces of the simplex Δ^{n-1} can be reached and when the system is controllable. The reachability of faces of the simplex is related to the reachability of subspaces and to the cooling of subsystems.

First we define some important notions. The *reachable set* of λ_0 at time $T \geq 0$ of the reduced control system (Λ), denoted $\mathsf{reach}_\Lambda(\lambda_0, T)$, is the set of all $\lambda(T)$ where $\lambda : [0, T] \to \Delta^{n-1}$ is a solution to (Λ) with $\lambda(0) = \lambda_0$. The *(all-time) reachable set* is $\mathsf{reach}_\Lambda(\lambda_0) = \bigcup_{T\geq 0} \mathsf{reach}_\Lambda(\lambda_0, T)$. Moreover we say that λ_f is *approximately reachable* from λ_0 if $\lambda_f \in \overline{\mathsf{reach}_\Lambda(\lambda_0)}$ and it is *asymptotically reachable* if there is a solution $\lambda : [0, \infty) \to \Delta^{n-1}$ with $\lambda(0) = \lambda_0$ and $\lambda(t) \to \lambda_f$ as $t \to \infty$. The definitions for other control systems are entirely analogous.

The Equivalence Theorem 3.0.4 implies that the reachable sets of (Λ) and ($\mathcal{D}$) are closely related, see Proposition 2.4.3 for a proof.

Proposition 4.3.1. *Given $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$, let $\rho_0 \in \mathfrak{pos}_1(n)$ and $\lambda_0 \in \Delta^{n-1}$ be such that $\mathrm{spec}^{\downarrow}(\rho_0) = \lambda_0^{\downarrow}$. Then it holds that*

$$\mathsf{reach}_{\mathcal{D}}(\rho_0, T) \subseteq \{U \operatorname{diag}(\lambda) U^* : \lambda \in \mathsf{reach}_\Lambda(\lambda_0, T), U \in \mathrm{SU}(n)\} \subseteq \overline{\mathsf{reach}_{\mathcal{D}}(\rho_0, T)}.$$

Let us continue with some general facts about reachable states. The following result extends some previous results derived in the toy model, see Lemmas 4.B.2 and 4.B.7. We omit the proof, since it is essentially the same.

Proposition 4.3.2. *Let $\lambda \in \Delta^{n-1}$ and assume that $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ is not purely Hamiltonian. Then $\overline{\mathsf{reach}_{\bar\Lambda}(\lambda)}$ contains $\mathbf{e}/n$ and is contractible.*

If a state can be reached asymptotically in the relaxed control system ($\bar\Lambda$), then it is stabilizable. The proof is similar to the one of Lemma 2.4.4 (ii), and hence omitted.

Lemma 4.3.3. *Let $\lambda : [0, \infty) \to \Delta^{n-1}$ be a solution to ($\bar\Lambda$) such that $\lambda(t) \to \mu$ as $t \to \infty$ for some $\mu \in \Delta^{n-1}$. Then $\mu \in \mathsf{stab}_\Lambda$.*

A similar result also holds for faces.

Lemma 4.3.4. *Let F be a face of Δ^{n-1} which is not a vertex. Assume that there is some $\mu \in \mathrm{relint}(F)$ and some $\lambda \neq \mu$ such that $\mu \in \overline{\mathsf{reach}_{\bar\Lambda}(\lambda)}$. Then F is viable for ($\bar\Lambda$).*

Proof. We show the contrapositive, so assume that F is not viable for ($\bar\Lambda$). By Proposition 4.2.4 it holds for every $\mu \in \mathrm{relint}(F)$ that no element of $\mathsf{derv}(\mu)$ is tangent to F. In particular, if α is any linear functional on $T_\mu \mathbb{R}^n_0$ which vanishes on tangent vectors along F and is non-positive on the tangent cone $T_\mu \Delta^{n-1}$ (which is a subset of $T_\mu \mathbb{R}^n_0$), then $\alpha(\mathsf{derv}(\mu)) \subset [-\infty, -\varepsilon]$ for some $\varepsilon > 0$. By continuity, and if necessary decreasing ε, we may assume that $\alpha(\mathsf{derv}(\nu)) \subset [-\infty, -\varepsilon]$ for every ν in some ball $B(R, \mu)$ about μ with radius $R > 0$. Moreover, by compactness there is some $C > 0$ such that $\mathsf{derv}(\nu) \subseteq B(C, 0)$ for every ν in $B(R, \mu)$. Thus there is $r > 0$ small enough such that any solution to ($\bar\Lambda$) starting outside of $B(R, \mu)$ cannot enter $B(r, \mu)$. In particular, for any λ we can choose $R < \|\lambda - \mu\|$ and hence μ is not approximately reachable from λ. □

[5] Due to the assumption of fast unitary control, reaching a pure state, i.e. a vertex of Δ^{n-1}, is sufficient for reaching an energy minimizing state.

Note, however, that there are reachable states that are not stabilizable. This property and more can be shown explicitly in the toy model, see Appendix 4.B, and in particular Figure 4.4. In general it is very difficult to compute the reachable set, hence why we will focus on the reachability of faces of the simplex (including coolability), and we will characterize approximate controllability, which is the situation where all states are approximately reachable from all other states.

Asymptotically Coolable Systems

One of the DiVincenzo criteria [DiV00] necessary for quantum computation requires the ability to initialize the system in a simple reference state. In practice one often starts in a thermal state and cools the system to a temperature near absolute zero. Since we assume fast unitary controllability, we will use the term "cooling" to refer to the preparation of any pure state. In this section we characterize (asymptotic) coolability and show how it can be implemented in the bilinear control system ($\mathcal{D}$).

We start with some technical results. If we think of the dynamics on the simplex Δ^{n-1} as a (controlled) Markov chain on n states, the task of cooling corresponds to moving the entire population into a single state. To achieve this one has to minimize the outflow of this state and maximize the inflow. The next result yields some bounds on the ratio of inflow to outflow for certain systems.

Lemma 4.3.5. *Let $\{V_k\}_{k=1}^r$ denote a finite set of complex $n \times n$ matrices such that every common eigenvector of all V_k is also a common eigenvector of all V_k^*. With $J(U)$ as in (3.3), there is a constant $C > 0$ depending on the V_k such that the flow ratio satisfies*

$$\sum_{i=2}^{n} J_{1i}(U) \Big/ \sum_{i=2}^{n} J_{i1}(U) \leq C \tag{4.2}$$

for all U where the expression is defined.

Proof. Our key object will be the subspace $S \subseteq \mathbb{C}^n$ defined as follows. If there is no eigenvector which all the V_k have in common set $S := \{0\}$. Else let v_1 denote such a common (unit) eigenvector of V_k and set $S_1 := \operatorname{span}(v_1)$. From this we proceed inductively over $i = 1, \ldots, n$: if $S_i^\perp$ contains a common (unit) eigenvector v_{i+1} we set $S_{i+1} = S_i \oplus \operatorname{span}(v_{i+1})$. Else $S := S_i$. The advantage of this explicit construction is that, if $S \neq \{0\}$, we now have an orthonormal basis $\{v_i\}_{i=1}^s$ ($s \geq 1$) of S consisting of common eigenvectors of all V_k.

Assume w.l.o.g. that $S \neq \mathbb{C}^n$ (else the V_k are all normal and commute with each other so the flow ratio equals 1 where it is defined). This assumption is equivalent to $S^\perp \neq \{0\}$ which means the domain of the following map is well defined.

$$b : \mathrm{SU}(n) \times \{\psi \in S^\perp \,:\, \|\psi\| = 1\} \to [0, \infty), \quad (U, \psi) \mapsto \sum_{k=1}^{r} \left\| (V_k - \langle e_1 | U^* V_k U | e_1 \rangle \mathbb{1}) \psi \right\|^2.$$

As b is a continuous function on a compact domain it attains its minimum, that is,

$$C_D := \min_{U \in \mathrm{SU}(n), \psi \in S^\perp, \|\psi\| = 1} b(U, \psi) \geq 0$$

exists. Most importantly, $C_D > 0$: if C_D were zero, then there exist $U \in \mathrm{SU}(n)$ and a normalized vector $\psi \in S^\perp$ such that $V_k \psi = \langle e_1 | U^* V_k U | e_1 \rangle \psi$ for all k. Thus ψ is a common eigenvector for all V_k which contradicts $\psi \in S^\perp$.

With this let $U \in \mathrm{SU}(n)$ be given such that the flow ratio is well-defined (that is, $\sum_{i=2}^{n} J_{i1}(U) > 0$). Our goal is to show that[6]

$$R(U) \leq \max\left\{1, \frac{4\sum_{k=1}^{r} \|V_k\|_\infty^2}{C_D}\right\} \tag{4.3}$$

which would conclude the proof because the right-hand side of Eq. (4.3) is finite as we just saw. Defining $w := Ue_1$ and using Lemma 3.0.7, a straightforward computation shows

$$R(U) = \frac{\sum_{k=1}^{r} \langle w|V_k V_k^*|w\rangle - |\langle w|V_k|w\rangle|^2}{\sum_{k=1}^{r} \langle w|V_k^* V_k|w\rangle - |\langle w|V_k|w\rangle|^2} = \frac{\sum_{k=1}^{r} \|(V_k^* - \langle w|V_k^*|w\rangle \mathbb{1})w\|^2}{\sum_{k=1}^{r} \|(V_k - \langle w|V_k|w\rangle \mathbb{1})w\|^2}.$$

Next we use the orthogonal projection onto S, denoted Π_S, to split up the above norms via $\|\psi\|^2 = \|\Pi_S\psi\|^2 + \|(\mathbb{1} - \Pi_S)\psi\|^2 = \|\Pi_S\psi\|^2 + \|\Pi_{S^\perp}\psi\|^2$. Importantly, by assumption on V_k, both V_k and V_k^* commute with Π_S (and thus with $\Pi_{S^\perp}$). Thus we find that $R(U)$ is equal to

$$\begin{aligned}
&\frac{\sum_{k=1}^{r} \|\Pi_S(V_k^* - \langle w|V_k^*|w\rangle \mathbb{1})w\|^2 + \|\Pi_{S^\perp}(V_k^* - \langle w|V_k^*|w\rangle \mathbb{1})w\|^2}{\sum_{k=1}^{r} \|\Pi_S(V_k - \langle w|V_k|w\rangle \mathbb{1})w\|^2 + \|\Pi_{S^\perp}(V_k - \langle w|V_k|w\rangle \mathbb{1})w\|^2} \\
&= \frac{\sum_{k=1}^{r} \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\Pi_S Ue_1\|^2 + \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2}{\sum_{k=1}^{r} \|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\Pi_S Ue_1\|^2 + \|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2}.
\end{aligned}$$

Observe that $\|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\psi\|^2$ and $\|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\psi\|^2$ are equal for all k, all $U \in \mathrm{SU}(n)$, and all $\psi \in S$. This can be seen, e.g., by expanding $\|\cdot\|^2$ into $\sum_{i=1}^{s} |\langle v_i|\cdot\rangle|^2$ (Parseval's identity) and using that if $\lambda_{k,i}$ is the eigenvalue of V_k w.r.t. the eigenvector v_i, then[7] $V_k^* v_i = \overline{\lambda_{k,i}} v_i$. Therefore the first summand of the numerator and the denominator of $R(U)$ coincide. At this point we have to distinguish two cases: if $\Pi_{S^\perp} Ue_1 = 0$, then $R(U) = 1$. Thus (4.3) holds and we are done. Else we use the mediant inequality to obtain the (well-defined) expression

$$R(U) \leq \max\left\{1, \frac{\sum_{k=1}^{r} \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2}{\sum_{k=1}^{r} \|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2}\right\}.$$

Finally we upper bound the second argument of this maximum as follows:

$$\begin{aligned}
&\frac{\sum_{k=1}^{r} \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2}{\sum_{k=1}^{r} \|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\Pi_{S^\perp} Ue_1\|^2} \\
&\leq \frac{\sum_{k=1}^{r} \max_{\psi \in \mathbb{C}^n, \|\psi\|=1} \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\psi\|^2}{\min_{\psi \in S^\perp, \|\psi\|=1} \sum_{k=1}^{r} \|(V_k - \langle e_1|U^*V_kU|e_1\rangle \mathbb{1})\psi\|^2} \\
&\leq \frac{\sum_{k=1}^{r} \|(V_k^* - \langle e_1|U^*V_k^*U|e_1\rangle \mathbb{1})\|_\infty^2}{\min_{\psi \in S^\perp, \|\psi\|=1} b(U,\psi)} \leq \frac{\sum_{k=1}^{r} (2\|V_k\|_\infty)^2}{C_D}.
\end{aligned}$$

In either case (4.3) holds so we are done. □

The lemma above shows that under certain circumstances, the "flow ratio" of a set of matrices, as defined in (4.2), remains bounded, even if it is undefined at some points. This limits the ability of solutions to ($\bar{\Lambda}$) to converge to a vertex of Δ^{n-1}:

[6] Here $\|\cdot\|_\infty$ denotes the usual operator norm, i.e. the largest singular value of the input.

[7] By assumption v_i is a normalized eigenvector of V_k^* (to an eigenvalue $\mu_{k,i}$) which implies $\overline{\lambda_{k,i}} = \langle V_k v_i|v_i\rangle = \langle v_i|V_k^* v_i\rangle = \mu_{k,i}$.

Corollary 4.3.6. *Let $\lambda : [0,\infty) \to \Delta^{n-1}$ be a solution to ($\bar{\Lambda}$). If every common eigenvector of the Lindblad terms $\{V_k\}_k$ is also a common eigenvector of $\{V_k^*\}_k$, then the first component $\lambda_1(t) \leq \max\{\lambda_1(0), \frac{C}{1+C}\}$ for all $t \geq 0$, where C is as in (4.2). Therefore, if $\lambda(0) \neq e_1$ and if there is a sequence $t_n \to \infty$ with $\lambda(t_n)$ converging to e_1, then there exists a common eigenvector of all Lindblad terms which is not a common left eigenvector.*

Proof. It suffices to prove the first statement as the second one is an immediate consequence. For this assume that every common eigenvector is also a left eigenvector. Indeed, a direct consequence of Lemma 4.3.5 and the mediant inequality is that $\sum_{i=2}^n(-M_{1i}) \leq C\sum_{i=2}^n(-M_{i1})$ for all $-M \in \operatorname{conv}(\mathfrak{L})$. Note that this also holds when $\sum_{i=2}^n(-M_{i1}) = 0$. Now let $\tilde{\lambda} \in \Delta^{n-1}$ and $-M \in \operatorname{conv}(\mathfrak{L})$ be given such that $(-M\tilde{\lambda})_1 > 0$. Then, because $(-M_{i1}) \geq 0$ for all $i \neq 1$ we compute

$$\begin{aligned} 0 &< (-M_{11})\tilde{\lambda}_1 + \sum_{i=2}^n(-M_{1i})\tilde{\lambda}_i = -\tilde{\lambda}_1\sum_{i=2}^n(-M_{i1}) + \sum_{i=2}^n(-M_{1i})\tilde{\lambda}_i \\ &\leq -\tilde{\lambda}_1\sum_{i=2}^n(-M_{i1}) + (1-\tilde{\lambda}_1)C\sum_{i=2}^n(-M_{i1}) = \Big(\sum_{i=2}^n(-M_{i1})\Big)(C - \tilde{\lambda}_1(1+C))\,. \end{aligned}$$

Rearranging this inequality yields $\tilde{\lambda}_1 < \frac{C}{1+C}$. Now let λ be any solution to ($\bar{\Lambda}$) and denote $C' := \max\{\lambda_1(0), \frac{C}{1+C}\}$. Towards a contradiction assume that there is some $t_1 > 0$ such that $\lambda_1(t_1) > C'$. Let $t_0 = \max(\{t \in [0,t_1] : \lambda_1(t) = C'\})$. Then $t_0 < t_1$ and $\lambda_1(t) \geq C'$ for all $t \in [t_0, t_1]$, and by the above $(-M\lambda(t))_1 \leq 0$ for every $-M \in \operatorname{conv}(\mathfrak{L})$ on the same interval. By Remark 3.0.3 there is $(M_t)_{t\in[t_0,t_1]}$ corresponding to λ. Then it holds that $\lambda_1(t_1) = \lambda_1(t_0) + \int_{t_0}^{t_1}(-M_t\lambda(t))_1\,dt \leq \lambda_1(t_0) = C'$, which yields the desired contradiction. □

This shows that determining upper bounds for the flow ratio (4.2) of a system allows to find bounds on the purest reachable state.

Now we are ready to characterize asymptotically coolable systems.

Theorem 4.3.7 (Asymptotically Coolable Systems). *Given any $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$, the following are equivalent.*

(i) For each choice of Lindblad terms $\{V_k\}_{k=1}^r$ of $-L$, there exists a common eigenvector of all V_k which is not a common left eigenvector.

(ii) There exists a (time-independent) Hamiltonian H such that $-(\mathrm{i}\,\mathrm{ad}_H + L)$ has a (unique) attractive fixed point[a], and this fixed point is pure.

(iii) For every initial state, there exists some solution λ converging to e_1.

(iv) $e_1 \in \overline{\mathsf{reach}(\lambda)}$ for some $\lambda \in \Delta^{n-1} \setminus \{e_1\}$.

[a] We say that ρ is an attractive fixed point if every solution converges to ρ. If such an attractive fixed point exists, it is clearly unique.

Proof. (i) ⇒ (ii) follows from Lemma 4.A.19, (ii) ⇒ (iii) follows from Corollary 3.0.5, and (iii) ⇒ (iv) is trivial. Finally (iv) ⇒ (i) is Corollary 4.3.6. □

Most of our effort went into proving that if the system is coolable, then the Lindblad terms must have a common eigenvector which is not a common left eigenvector. The other implications are mostly known and have been rediscovered several times, see for instance [Kra+08, TV09, BN08]. The task of

efficiently determining if a common eigenvector exists, and of computing such a common eigenvector, is not trivial. In Appendix 4.C we present an efficient algorithm for this problem.

Theorem 4.3.7 shows that for the purpose of cooling, time-independent controls are sufficient. In the following section we will consider the reachability of faces, where time-dependent controls will offer new possibilities.

Directly and Indirectly Reachable Faces

In the previous section we characterized coolability by studying reachability of vertices of Δ^{n-1} in the reduced system. One consequence of Theorem 4.3.7 is that for vertices, asymptotic and approximate reachability coincide. For interior points of higher-dimensional faces of Δ^{n-1} this need not hold anymore. The reachability of faces of the simplex corresponds to the reachability of certain subspaces of Hilbert space. Of particular interest might be the reachability of decoherence free subspaces [LCW98]. Moreover, the reachability of certain faces of low enough dimension corresponds to the cooling of subsystems.

Faces of Δ^{n-1} whose interior is reachable in an approximate sense satisfy the following dichotomy: Either the interior can be reached directly, or one first has to approach the boundary and then move parallel to the face, which we call indirect reachability. The precise result is given in Proposition 4.3.19 at the end of the section. Note that indirectly reachable faces provide a concrete problem for which time-independent controls are insufficient.

We start by characterizing direct reachability of faces. Like in the previous section, flow ratios play an important role here, so we begin by formalizing the concept.

Definition 4.3.8. *Let J be an $n \times n$ matrix with non-negative off-diagonal entries. Given any $d \in \{1, \dots, n-1\}$, the* d-dimensional inflow f^d_{in}*, the* d-dimensional outflow f^d_{out}*, and the* d-dimensional flow ratio R_d *of J are defined by*

$$f^d_{\mathsf{in}} := \sum_{i=1}^{d} \sum_{j=d+1}^{n} J_{ij}\,, \quad f^d_{\mathsf{out}} := \sum_{i=d+1}^{n} \sum_{j=1}^{d} J_{ij}\,, \quad R_d := \frac{f^d_{\mathsf{in}}}{f^d_{\mathsf{out}}}\,,$$

respectively. If the matrix J is of the form $J(U)$ as in (3.3), then we denote the objects above by $f^d_{\mathsf{in}}(U)$, $f^d_{\mathsf{out}}(U)$ and $R_d(U)$[8]*. We allow $R_d(U)$ to take values in $[0, +\infty]$, where expressions of the form $\frac{c}{0}$ with $c > 0$ are interpreted as $+\infty$. Only expressions of the form $\frac{0}{0}$ are considered undefined. We will say that the system has* bounded d-dimensional flow ratio *if $\sup\{R_d(U) : U \in \mathrm{SU}(n)\} < \infty$, otherwise we say it is* unbounded*. Here and henceforth, suprema and infima of this form always implicitly ignore undefined values.*

Note that the flow ratio $R_d(U)$ is infinite or undefined whenever the first d columns of U^* span a lazy subspace, cf. Lemma 4.A.8 (v). Outside of these points, however, R_d is a continuous function. The behavior of $R_d(U)$ near these singularities has important consequences for the reachability of faces, see Remark 4.3.14.

In the proof of Corollary 4.3.6 we showed that if the 1-dimensional flow ratio of the system is bounded, then it is impossible to reach a vertex of Δ^{n-1}. In higher dimensions the situation becomes more nuanced, as a bounded flow ratio only prohibits approaching the interior of a face directly, but it does not prohibit approaching the boundary of the face:

Lemma 4.3.9. *Let $1 \leq d \leq n-1$ and set $p_d : \Delta^{n-1} \to \mathbb{R}$, $p_d(\lambda) := \sum_{i=1}^{d} \lambda_i$. Given $\lambda \in \Delta^{n-1}$, $\varepsilon > 0$ assume that $\lambda_i \geq \varepsilon p_d(\lambda)$ for all $i = 1, \dots, d$, and that the system has bounded flow ratio $R_d(U)$ with R*

[8] Since $-L_U$ is well-defined, so are the off diagonal elements of $J(U)$, and hence also all the quantities defined here.

denoting the supremum. Then if $p_d(\lambda) \geq \frac{R}{R+\varepsilon}$, it holds that $p_d(\mathsf{derv}(\lambda)) \subseteq (-\infty, 0]$. In particular, no solution to ($\bar{\Lambda}$) can converge to the interior of a $(d-1)$-dimensional face if it starts outside of the face.

Proof. Using (3.4) we for any $U \in \mathrm{SU}(n)$ compute

$$p_d(-L_U\lambda) = \sum_{i=1}^{d}\sum_{j=d+1}^{n} J_{ij}(U)\lambda_j - \sum_{i=d+1}^{n}\sum_{j=1}^{d} J_{ij}(U)\lambda_j \leq (1-p_d(\lambda))f^d_{\mathsf{in}}(U) - \varepsilon p_d(\lambda) f^d_{\mathsf{out}}(U)\,,$$

where we used $J_{ij}(U) \geq 0$ for all $i \neq j$, as well as $\lambda_i \geq \varepsilon p_d(\lambda)$ for all $i = 1, \dots, d$. We distinguish two cases: If $f^d_{\mathsf{out}}(U) = 0$, then $f^d_{\mathsf{in}}(U) = 0$. The reason for this is that R is a continuous, and by assumption bounded, function whose zero set is either all of $\mathrm{U}(n)$ or nowhere dense (as in the proof of Lemma 4.A.17) so if $f^d_{\mathsf{in}}(U)$ were not zero for some U (while $f^d_{\mathsf{out}}(U)$ is), then R cannot be bounded. Thus $p_d(\mathsf{derv}(\lambda)) \subseteq (-\infty, 0]$ either way. Now if $f^d_{\mathsf{out}}(U) \neq 0$, then the above estimate is non-positive if and only if $p_d(\lambda) \geq 1 - \frac{\varepsilon}{R_d(U)+\varepsilon}$. Hence if $p_d(\lambda) \geq \frac{R}{R+\varepsilon} = 1 - \frac{\varepsilon}{R+\varepsilon}$, then $-p_d(L_U\lambda) \leq 0$ for all $U \in \mathrm{SU}(n)$. Now consider any solution $\lambda : [0,\infty) \to \Delta^{n-1}$ to ($\bar{\Lambda}$) and let F be the convex hull of the d first standard basis vectors. If λ converges to some $\mu \in \mathrm{relint}(F)$, then for t large enough and some $\varepsilon > 0$ it holds that $\lambda_i(t) \geq \varepsilon p_d(\lambda(t))$ for all $i = 1, \dots, d$. At the same time $p_d(\lambda(t))$ converges to 1. However, this contradicts the first part of this lemma as in the proof of Corollary 4.3.6. □

Remark 4.3.10. *We have shown that if $p_d(\lambda) \geq \frac{R}{R+\varepsilon}$ then $p_d(\lambda') \leq 0$. This however does not necessarily imply that for $p_d(\lambda) > \frac{R}{R+\varepsilon}$ we have $p_d(\lambda') < 0$. For instance one might consider a system with a single normal V which is not a multiple of the identity (see Corollary 4.4.7). In this case every state is stabilizable but the one-dimensional flow ratio is equal to 1 whenever it is defined.*

With this result at our disposal, we can start characterizing the direct reachability of faces in the simplex. We begin by considering a stronger notion, i.e. which faces of Δ^{n-1} can be reached using a time-independent Hamiltonian, as shown in [TV09].

Lemma 4.3.11. *The following statements are equivalent.*

(i) There exists a d-dimensional lazy subspace which is not an enclosure (recall Lemma 4.A.16).

(ii) There exists a state ρ of rank d and a Hamiltonian H such that ρ is the unique fixed point of $-(\mathrm{i}\,\mathrm{ad}_H + L) \in \mathfrak{w}_{\mathsf{KL}}(n)$, so in particular ρ is attractive.

Proof. (i) ⇒ (ii): Follows from Lemma 4.A.19. (ii) ⇒ (i): Let $S = \mathrm{supp}(\rho)$. Since ρ is a fixed point, S is collecting by Lemma 4.A.10, and hence lazy. If S was an enclosure, there would be at least two fixed points (one supported on S, and one on $S^\perp$). Since $-(\mathrm{i}\,\mathrm{ad}_H + L)$ has a unique fixed point ρ, by [BN08, Thm. 18] the generator has only one eigenvalue equal to 0 corresponding to the fixed point, and all other eigenvalues have strictly negative real part. This shows that ρ is attractive. □

Finally we can give the proper notion of directly reachable faces and characterize them via unbounded flow ratios:

Proposition 4.3.12. *Let $1 \leq d \leq n-1$ and let F be a $(d-1)$-dimensional face of Δ^{n-1}. Then the following are equivalent:*

(i) The d-dimensional flow ratios $R_d(U)$ are unbounded.

(ii) There exists a solution $\lambda : [0,\infty) \to \Delta^{n-1}$ to ($\bar{\Lambda}$) with initial state $\lambda(0) \notin F$ such that $\lim_{t\to\infty} \lambda(t) =: \lambda_F \in \mathrm{relint}(F)$.

In this case we say that F is directly reachable.

Proof. (ii) ⇒ (i): Immediate from Lemma 4.3.9. (i) ⇒ (ii): Due to the permutation symmetry of ($\bar{\Lambda}$) we may assume that $F = \mathrm{conv}(e_1, \ldots, e_d)$. By assumption there is a sequence $(U_i)_{i=1}^{\infty}$ such that $R_d(U_i) \to \infty$ as $i \to \infty$. Let S_n^F be the subgroup of permutation matrices which map F to itself and let $A \subseteq \Delta^{n-1}$ be the line segment consisting of all points fixed by S_n^F. We can parametrize A via $\iota : [0,1] \to A$, $a \mapsto \iota(a) := (\frac{a}{d}, \ldots, \frac{a}{d}, \frac{1-a}{n-d}, \ldots \frac{1-a}{n-d})$. Note that $R_d(U_i) = R_d(U_iP)$ for all $P \in S_n^F$. If we set $M_i = \frac{1}{|S_n^F|} \sum_{P \in S_n^F} P^\top L_{U_i} P = \frac{1}{|S_n^F|} \sum_{P \in S_n^F} L_{U_i P}$, then all $-M_i \in \mathrm{conv}(\mathfrak{L})$ and they leave A invariant. Using $p_d : \Delta^{n-1} \to \mathbb{R} : \lambda \mapsto \sum_{i=1}^{d} \lambda_i$ we compute the derivative along A as $-p_d(M_i \iota(a)) = -p_d(L_{U_i} \iota(a)) = -\frac{a}{d} f_{\mathsf{out}}^d(U_i) + \frac{1-a}{n-d} f_{\mathsf{in}}^d(U_i)$. Hence the unique (attractive) fixed point on A is located at $a = \frac{dR_d(U_i)}{dR_d(U_i)+n-d}$ which can be made arbitrarily close to 1 for i large enough. This shows that there exists a solution to ($\bar{\Lambda}$) converging to $\mu := \iota(1) \in A \cap F$. □

As a consequence of Lemma 4.3.11 together with Corollary 3.0.5 we obtain the following:

Corollary 4.3.13. *Let F be a face of Δ^{n-1} of dimension $d-1$. If there exists a lazy subspace of dimension d which is not an enclosure, then F is directly reachable.*

Remark 4.3.14. *Currently, we do not know whether the converse to Corollary 4.3.13 also holds. A possible generalization of Lemma 4.3.5 to invariant subspaces of dimension higher than $d = 1$ could be used to prove this, but the proof of said lemma does not seem to generalize in a straightforward manner. If the converse does not hold, this would imply that time-dependent Hamiltonians allow one to directly reach faces not directly reachable using time-independent Hamiltonians.*

The following corollaries yield special cases where the converse of Corollary 4.3.13 does hold.

Corollary 4.3.15. *Let F be a face of Δ^{n-1} of dimension $d-1$. If the commutant of $\{H_0, V_1, \ldots, V_r\}$ contains only multiples of the identity and F is directly reachable, then there exists a lazy subspace of dimension d which is not an enclosure.*

Proof. By Lemma 4.A.16 (iii) there are no proper non-trivial enclosures, and since F is directly reachable, by Lemma 4.3.3 and Proposition 4.2.4 there is a lazy subspace of dimension d. □

For the second corollary we observe the following duality relations for flow ratios:

Lemma 4.3.16. *Let $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ and some choice of Lindblad terms $\{V_k\}_{k=1}^r$ of $-L$ be given, and let $-L' \in \mathfrak{w}_{\mathsf{KL}}(n)$ be the Kossakowski–Lindblad generator represented by the Lindblad terms $\{V_k^*\}_{k=1}^r$. Then the following hold (here we use $1/0 = +\infty$ and the suprema and infima are defined as in Definition 4.3.8):*

(i) $\sup_{U \in \mathrm{SU}(n)} R_{n-d}(U) = (\inf_{U \in \mathrm{SU}(n)} R_d(U))^{-1}$

(ii) $\sup_{U \in \mathrm{SU}(n)} R'_d(U) = (\inf_{U \in \mathrm{SU}(n)} R_d(U))^{-1}$,

where $R_d(U)$ is the flow ratio of $-L$ and $R'_d(U)$ is the flow ratio of $-L'$.

Proof. (i): Follows directly from the observation that $f_{\mathsf{in}}^d(U) = f_{\mathsf{out}}^{n-d}(\pi U \pi^\top)$ and vice-versa where $\pi = \sum_{i=1}^{n} e_{n+1-i} e_i^\top$. (ii): Replacing all V_k by V_k^* transforms the d-dimensional inflow into the d-dimensional outflow. □

Corollary 4.3.17. *If F is a facet*[9] *of Δ^{n-1} and if F is directly reachable, then exists a lazy subspace of dimension $n-1$ which is not an enclosure.*

Proof. Proposition 4.3.12 shows that the $(n-1)$-dimensional flow ratio is unbounded. Towards a contradiction assume that all $n-1$ dimensional lazy subspaces are enclosures. Consider the system $-L'$ with Lindblad terms $\{V_k^*\}_{k=1}^r$. Then all common eigenvectors of the V_k^* are also common eigenvectors of all V_k. By Lemma 4.3.5 it holds that $\sup_{U\in\mathrm{SU}(n)} R_1'(U) < C$ for some $C < \infty$. Then by Lemma 4.3.16 (i) and (ii) (with $d=1$) we have $\sup_{U\in\mathrm{SU}(n)} R_{n-1}(U) = \sup_{U\in\mathrm{SU}(n)} R_1'(U)$, i.e. the $(n-1)$-dimensional flow ratio is bounded, yielding the desired contradiction. □

Corollary 4.3.18. *If $n=3$, then the converse of Corollary 4.3.13 holds for all faces.*

Finally we can prove the promised dichotomy alluded to in the beginning. We say that a face F of Δ^{n-1} is approximately reachable if there exists $\lambda \notin F$ such that $\overline{\mathsf{reach}(\lambda)} \cap \mathrm{relint}(F) \neq \emptyset$. A face that is approximately reachable but not directly reachable is called *indirectly reachable face*.

Proposition 4.3.19. *Let F be a face of Δ^{n-1}. Consider the following statements:*

(i) F is directly reachable, otherwise F is viable but not purely Hamiltonian[10] *and some lower dimensional face is approximately reachable.*

(ii) F is approximately reachable.

(iii) F is directly reachable, otherwise F is viable and some lower dimensional face is approximately reachable.

Then we have the implications: (i) ⇒ (ii) ⇒ (iii).

Proof. (i) ⇒ (ii): If F is directly reachable then (ii) is clearly true. Otherwise, some point on the boundary is asymptotically reachable, and since the face is viable and not purely Hamiltonian we can reach, for instance, the center of F as in Proposition 4.3.2. (ii) ⇒ (iii): If F is not directly reachable, then the flow ratio is bounded by Proposition 4.3.12. So in order to approach an interior point of F, one must increase the value of $p_s(\lambda)$ (as defined in Lemma 4.3.9), which implies by compactness that some point on the boundary of F is approximately reachable, and hence some face of lower dimension is approximately reachable. If F is not viable, then no interior point of F is approximately reachable from any other point, cf. Lemma 4.3.4. □

The reason we do not obtain equivalence is that even if every lazy subspace of appropriate dimension has purely Hamiltonian dynamics, it might still be possible to move along the face with arbitrarily small outflow. This can be made rigorous using flow ratio arguments as above, but we will not do so here.

Reverse Coolable and Controllable Systems

So far we have studied under which conditions faces of the simplex can be reached, with special emphasis on reachability of vertices. In this section we try to understand under which conditions all states in Δ^{n-1} can be reached. More precisely, if for all $\lambda, \mu \in \Delta^{n-1}$ it holds that $\mu \in \overline{\mathsf{reach}_\Lambda(\lambda)}$, then we say that the system (Λ) is *approximately controllable*. (Recall that (Λ) and $(\bar{\Lambda})$ have the same reachable sets after

[9] A facet of Δ^{n-1} is a face of dimension $n-2$.

[10] For a viable face F there exists at least some $-L_U \in \mathfrak{L}$ whose restriction to F is tangent to F. If every such vector field vanishes on F, we say that F is purely Hamiltonian.

taking the closure.) Note that since one can never exactly reach the boundary of the simplex from the interior in finite time, the system is never controllable in the usual sense [DH08, Thm. 3.10].

It turns out to be useful to consider time-reversed dynamics on the simplex. Note that under such dynamics Δ^{n-1} ceases to be forward invariant. We say that the reduced control system is *reverse coolable* if $\overline{\mathsf{reach}_\Lambda(e_1)} \supseteq \Delta^{n-1}_\downarrow$. The reason for introducing this artificial concept is that asymptotic coolability together with reverse coolability characterizes controllable systems. The results of this section generalize results on approximate controllability of a quantum system coupled to a heat bath of temperature zero, cf. [DES19, Thm. 1 & 2] and [BWS16].

Proposition 4.3.20. *The following are equivalent:*

(i) The system (Λ) *is approximately controllable.*

(ii) The system (Λ) *is asymptotically coolable and reverse coolable.*

(iii) The system ($\mathcal{D}$) *is approximately controllable.*

Proof. Consider the reachability relation defined by $\lambda \rightsquigarrow \mu$ if $\mu \in \overline{\mathsf{reach}_\Lambda(\lambda)}$. One can show that this is a preorder, in particular, that it is transitive. (i) $\Rightarrow$ (ii): Clearly $\lambda \rightsquigarrow e_1$ for all $\lambda \in \Delta^{n-1}$ and hence by Theorem 4.3.7 the system is asymptotically coolable. Reverse coolability is clear. (ii) $\Rightarrow$ (i): Again by Theorem 4.3.7 $\lambda \rightsquigarrow e_i$ for all $\lambda \in \Delta^{n-1}$ and $i = 1, \dots, n$. By reverse coolability $e_1 \rightsquigarrow \lambda$ for all $\lambda \in \Delta^{n-1}_\downarrow$. By permutation symmetry of the system every point in Δ^{n-1} is reachable from some vertex and by transitivity the system is approximately controllable. (i) $\Leftrightarrow$ (iii): Direct consequence of Proposition 2.4.11. □

This result motivates us to characterize reverse coolability, generalizing a toy model result [DES19, Lem. 3]:

Proposition 4.3.21. *Consider the following statements:*

(i) All faces (except possibly vertices) of Δ^{n-1} *are viable for* ($\bar{\Lambda}$) *but not purely Hamiltonian (cf. footnote 10).*

(ii) The system (Λ) *is reverse coolable.*

(iii) All faces (except possibly vertices) of Δ^{n-1} *are viable for* ($\bar{\Lambda}$).

Then we have the following implications: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

Proof. (i) $\Rightarrow$ (ii): Let F be any face of dimension $d-1$ of the simplex Δ^{n-1} which is not a vertex ($d > 1$), and let $\lambda_d \in F$ be arbitrary. First we show that there is some $\lambda_{d-1} \in \partial F$ on the boundary such that $\lambda_{d-1} \rightsquigarrow \lambda_d$. If $\lambda_d \in \partial F$ we set $\lambda_{d-1} = \lambda_d$.

By assumption there is some L_U such that F is invariant but not fixed. If S_n^F denotes the permutation subgroup which leaves F invariant, then $M = \frac{1}{|S_n^F|} \sum_{P \in S_n^F} L_{UP}$ still leaves F invariant with the center being the unique attractive fixed point (cf. Lemma 4.B.7). Since the entire boundary of F converges to its center, it passes through every point[11] of F, and hence every point is approximately reachable from the boundary. To be precise we have proven the claim in ($\bar{\Lambda}$), but by the Relaxation Theorem (cf. [AC84, Ch. 2.4, Thm. 2]), it still holds in (Λ). Putting everything together, if we start with any $\lambda \in \Delta^{n-1}_\downarrow$, we set $\lambda_n = \lambda$ and find a sequence of λ_d for $d = n-1, \dots, 1$ where necessarily $\lambda_1 = e_i$ for some $i = 1, \dots, n$.

[11] This can be shown rigorously using the fact that $\pi_n(S^n) = \mathbb{Z}$, where π_n denotes the n-th homotopy group [Hat02, Sec. 4.1].

By transitivity we get that $e_i \rightsquigarrow \ldots \rightsquigarrow \lambda$. Using the permutation symmetry (Lemma 2.A.2) and forcing the solution to stay in the ordered Weyl chamber (Proposition 2.A.4) we find that $\lambda = \lambda^{\downarrow} \in \overline{\mathsf{reach}_{\Lambda}(e_1)}$.

(ii) $\Rightarrow$ (iii): Let $\lambda \in \mathrm{relint}(F)$ and assume that F is not viable. Then by Lemma 4.3.4, λ is not approximately reachable from any other point. □

For stabilizable systems we can strengthen the result above:

Corollary 4.3.22. *If (Λ) is stabilizable and has a two-dimensional lazy subspace which is not purely Hamiltonian, then it is reverse coolable.*

Proof. By Theorem 4.2.7 the Lindblad terms are simultaneously triangulable. From the assumption and [RR00, Lem. 1.5.2] it follows that we can choose a triangulation such that the two-dimensional subspace in the chain is not purely Hamiltonian. Hence we satisfy Proposition 4.3.21 (i). □

4.4 Special Structure for Unital Systems

In this section we focus on unital systems, which are precisely those systems for which the identity is a fixed point or, equivalently, on the level of generators, those which satisfy $L(\mathbb{1}) = 0$. Such systems stand in contrast to coolable and controllable systems studied above, since all reachable states are majorized by the initial state. In particular the purity of a state can never increase.

Lemma 4.4.1. *Given any $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$, the following statements are equivalent.*

(i) $-L$ *is unital, i.e.* $L(\mathbb{1}) = 0$.

(ii) $\sum_{k=1}^{r} [V_k, V_k^*] = 0$ *for some (equivalently: each) choice of Lindblad terms* $\{V_k\}_{k=1}^r$ *of* $-L$.

(iii) $L_U \mathbf{e} = 0$ *for all* $U \in \mathrm{SU}(n)$.

(iv) $J_U \mathbf{e} = J_U^{\top} \mathbf{e}$ *for all* $U \in \mathrm{SU}(n)$.

(v) $\mathsf{derv}(\mathbf{e}/n) = \{0\}$.

(vi) $\overline{\mathsf{reach}_{\bar{\Lambda}}(\lambda)} \subseteq \{\mu : \mu \preceq \lambda\}$ *for all* $\lambda \in \Delta^{n-1}$.

Proof. Direct computation shows $-L(\mathbb{1}) = \frac{1}{2} \sum_{k=1}^{r} [V_k, V_k^*]$ for any choice of Lindblad terms $\{V_k\}_{k=1}^r$ of $-L$; hence (i) $\Leftrightarrow$ (ii). For (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v) note that by definition $-L_U = J(U) - \mathrm{diag}(J(U)^{\top}\mathbf{e})$, which shows that $L_U \mathbf{e} = J(U)\mathbf{e} - J(U)^{\top}\mathbf{e}$. Next, Lemma 3.0.7 together with the Schur–Horn Theorem [Sch23, Hor54] shows that $\mathsf{derv}(\mathbf{e}/n)$ is equal to the majorization polytope spanned by the vector of eigenvalues of $\sum_{k=1}^{r} [V_k, V_k^*]$; this shows the equivalence of (ii) and (v). Condition (iii) shows that the time evolution of a unital system is doubly stochastic, which implies (vi) see [MOA11, Thm. A.4]. Conversely, (vi) directly implies (v). (The equivalence of (i) and (vi) was also shown in [Yua10].) □

Importantly, unitality is independent of the Hamiltonian part of the generator, and hence of the control. Note also that property (vi) continues to hold in the infinite-dimensional setting under appropriate assumptions [End+19].

Recall that $\mathcal{V} = \langle \mathbb{1}, V_1, \ldots, V_r \rangle_{\mathsf{alg}}$, i.e. the complex matrix algebra generated by the Lindblad terms and the identity, is called the relaxation algebra of $-L$, and it is independent of the choice of Lindblad terms (Lemma 4.A.6). Moreover, we write $\mathrm{Lat}(\mathcal{V})$ for the lattice of invariant subspaces for $\mathcal{V}$ (i.e. its lazy subspaces), cf. Appendix 4.A. For unital systems, stabilizability and reachability are to a large extent characterized by the structure of the relaxation algebra, which turns out to be particularly nice as the following result shows:

Lemma 4.4.2. *If $-L$ is unital, then the corresponding lattice $\mathrm{Lat}(\mathcal{V})$ of lazy subspaces is orthocomplemented*[12] *and $\mathcal{V}$ is a $*$-algebra*[13].

Proof. Let $S \subseteq \mathbb{C}^n$ be any lazy subspace of $-L$; without loss of generality $S \neq \{0\}$, and $S \neq \mathbb{C}^n$. Choose a unitary U such that the first k columns of U^* span S. Then by Lemma 4.A.8 (v) the matrix $-L_U$ (and thus J_U) is block triangular. Using Lemma 4.4.1 (iv) we compute

$$0 = \sum_{j=1}^{\dim S} \left((J_U \mathbf{e})_j - (J_U^\top \mathbf{e})_j \right) = \sum_{j=1}^{\dim S} \sum_{k=\dim S+1}^{n} J_{jk}(U) - \sum_{j=1}^{\dim S} \sum_{k=\dim S+1}^{n} J_{kj}(U)\,;$$

but the second term vanishes due to J_U being block triangular. Thus, because $J_{jk}(U) \geq 0$ for all j, k, $J_{jk}(U) = 0$ for all $1 \leq j \leq \dim S < k \leq n$, as well. Thus J_U must be block diagonal, hence the orthogonal complement of S is also invariant. This in turn implies that $\mathcal{V}$ is a $*$-algebra [GLR06, Thm. 11.5.1], as claimed. Note that this uses that $\mathcal{V}$ contains the identity. □

Note that the converse is not true, as can be seen by considering the following example of the Bloch equations in Lindblad form: Choosing the Lindblad terms σ_+, $2\sigma_-$, and σ_z, the system is clearly not unital but as the operators do not have a common eigenvector, the lattice of lazy subspaces is trivial and hence orthocomplemented.

The structure theorem for $*$-algebras [Far01, Thm. 5.6] shows that up to a change of orthonormal basis, $*$-algebras have a particularly simple form: Let $\mathcal{A} \subseteq \mathbb{C}^{n,n}$ be a $*$-algebra and let m be the dimension of the center of $\mathcal{A}$. Then there exists $U \in \mathrm{U}(n)$ as well as positive integers $\{q_i\}_{i=1}^m$ and $\{k_i\}_{i=1}^m$ such that

$$U \mathcal{A} U^* = \bigoplus_{i=1}^{m} \mathbb{C}^{q_i, q_i} \otimes \mathbb{1}_{k_i}\,. \tag{4.4}$$

Clearly it holds that $\sum_{i=1}^m q_i k_i = n$. The vector of block-sizes in (4.4)

$$\tau = (\underbrace{q_1, \ldots, q_1}_{k_1}, \ldots, \underbrace{q_m, \ldots, q_m}_{k_m})$$

is of special importance as shown in the following lemma. Either way in this case we say that $\mathcal{A}$ is *of type* τ; moreover, given $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ unital we say that $-L$ is of type τ if the relaxation algebra $\mathcal{V}$ is of type τ. To properly state the results in this section it is convenient to define the notion of refinement:

Definition 4.4.3 (Refinement). *Let n, k, l be positive integers and let $v \in \mathbb{N}^k$ and $w \in \mathbb{N}^l$ be vectors of positive integers such that their elements sum to n respectively. We say that v is a* refinement *of w if $Av = w$ for some $A \in \{0,1\}^{l \times k}$ with $\mathbf{e}^\top A = \mathbf{e}^\top$.*

Such matrices A form a finite semigroup, and hence refinement is a preorder. Vectors differing only up to permutation are equivalent and we will not distinguish between them. Hence, in the following, we will think of τ as a multiset, usually represented in non-increasing order.

The following result is a direct consequence of the Krull–Schmidt Theorem [HGK04, Prop. 3.2.5].

Lemma 4.4.4. *Let $\mathcal{A}$ be a $*$-algebra of type τ. Then for any block diagonalization of $\mathcal{A}$ the vector of block sizes is refined by the vector τ.*

We will see below that the type of a unital system determines many of its control-theoretic properties.

[12] This means that for every invariant subspace $S \in \mathrm{Lat}(\mathcal{V})$, the orthocomplement $S^\perp$ is also invariant.

[13] $\mathcal{V}$ is a $*$-algebra if for every $V \in \mathcal{V}$ it holds that $V^* \in \mathcal{V}$.

Stabilizability

In contrast to general systems, where the exact shape of the stabilizable set stab_Λ is very difficult to describe, unital systems allow for a complete characterization in algebraic terms using the relaxation algebra $\mathcal{V}$, as we will show in Theorem 4.4.6. We begin with a simple result about stabilizable states in unital systems:

Lemma 4.4.5. *For unital systems, the set of stabilizable states equals the set of strongly stabilizable states.*

Proof. Let $\lambda \in \Delta^{n-1}$ be stabilizable, i.e. there exist $\ell \in \mathbb{N}$, $\mu_1, \dots, \mu_\ell > 0$, and $U_1, \dots, U_\ell \in \mathrm{SU}(n)$ such that $-\sum_{k=1}^{\ell} \mu_k L_{U_k} \lambda = 0$ and $\sum_{k=1}^{\ell} \mu_k = 1$. W.l.o.g. there exists k such that $L_{U_k} \neq 0$. Then $\varepsilon := (\max_{j,k} |(L_{U_k})_{jj}|)^{-1} > 0$ is well-defined and, similar to the proof of Lemma 4.2.10, the matrix $\mathbb{1} - \varepsilon L_{U_k}$ is doubly stochastic for all k; here we used unitality of the system together with Lemma 4.4.1 (iii). This lets us compute $\lambda = \lambda - \varepsilon(\sum_{k=1}^{\ell} \mu_k L_{U_k} \lambda) = \sum_{k=1}^{\ell} \mu_k (\mathbb{1} - \varepsilon L_{U_k})\lambda$, meaning we expressed λ, which is an extreme point of the majorization polytope $\{\lambda' \in \Delta^{n-1} : \lambda' \preceq \lambda\}$ [Dah10, Thm. 1], as a non-trivial convex combination of elements of $\{\lambda' \in \Delta^{n-1} : \lambda' \preceq \lambda\}$ [MOA11, Ch. 2, Thm. B.2]. By definition of an extreme point this is only possible if $(\mathbb{1} - \varepsilon L_{U_k})\lambda = \lambda$ for all k; hence $-L_{U_k}\lambda = 0$ so λ is strongly stabilizable. □

Theorem 4.4.6 (Stabilizable Set for Unital Systems). *Assume that $-L \in \mathfrak{vo}_{\mathsf{KL}}(n)$ is unital and let $\lambda \in \Delta^{n-1}$. Then $\lambda \in \mathsf{stab}_\Lambda$ if and only if the type of $-L$ is a refinement of the vector of multiplicities of λ.*

Proof. "⇐": First assume that all Lindblad terms are in block-diagonal form such that the block sizes give a refinement of the multiplicities of λ. In particular L_U will have the same block structure. Then there exists a permutation λ' of λ such that in the expression $L_U\lambda'$, all elements of λ' which are multiplied with a given block of L_U must have the same value. Thus Lemma 4.4.1 (iii) shows $L_U\lambda' = 0$ and hence $\lambda \in \mathsf{stab}_\Lambda$.

"⇒": Assume that no such block diagonal structure is achievable. Then the product $L_U\lambda$ will always have at least two distinct elements of λ multiplied with a single block of L_U. Now consider a block where this happens. Then we argue that (at least one copy of) the smallest value of λ falling into this block must have a strictly positive derivative. Without loss of generality we assume that L_U consists of a single block, and all copies of the smallest element of λ are exactly the first s elements of λ. Towards a contradiction assume that $\lambda'_i = 0$ for all $1 \leq i \leq s$ (note that by unitality the smallest element of λ cannot strictly decrease). This means that $(L_U)_{ij} = 0$ for all $i \leq s < j$. By Lemma 4.4.1 (v) this implies that $(L_U)_{ji} = 0$, contradicting the assumption that L_U consists of a single block. Hence λ is not strongly stabilizable and by Lemma 4.4.5 it is not stabilizable at all. □

As a special case we can characterize unital systems which are stabilizable (that is, every state is stabilizable).

Corollary 4.4.7. *The following statements are equivalent.*

(i) *$-L$ is unital and $\mathsf{stab}_\Lambda = \Delta^{n-1}$.*

(ii) *The U^*V_kU are simultaneously diagonal for some $U \in \mathrm{SU}(n)$ and for some (equivalently: every) choice of $\{V_k\}_{k=1}^r$.*

(iii) *The V_k are normal and commute for some (equivalently: every) choice of $\{V_k\}_{k=1}^r$.*

(iv) $J(U)$ *is diagonal (equivalently:* $L_U = 0$*) for some* $U \in \mathrm{SU}(n)$.

If $-L$ *satisfies any, and hence all, of the above, we call it* unital stabilizable.

Proof. (i) ⇔ (ii) is a consequence of Theorem 4.4.6. (ii) ⇔ (iii): [HJ87, Thm. 2.5.5]. (ii) ⇔ (iv) is trivial. □

Note that one could also use Theorem 4.2.7 instead of Theorem 4.4.6 in the proof above.

Due to Lemma 4.4.5, in a unital stabilizable system every state is also strongly stabilizable, and any unitary U which diagonalizes all Lindblad operators satisfies $L_U\lambda = 0$ due to property (iv) above. Hence using Proposition 4.2.2 one can find a corresponding compensating Hamiltonian for each regular $\lambda \in \Delta^{n-1}$. Finally let us characterize the following (trivial) case:

Lemma 4.4.8. *The following are equivalent.*

(i) $L = \mathrm{i}\,\mathrm{ad}_H$ *for some Hermitian matrix* H.

(ii) All eigenvalues of $-L$ *are on the imaginary axis.*

(iii) For some (equivalently: each) choice of Lindblad terms $\{V_k\}_{k=1}^r$ *of* $-L$*, all* V_k *are multiples of the identity.*

(iv) Every subspace is lazy for $-L$.

(v) $J(U)$ *is diagonal (equivalently:* $L_U = 0$*) for all* $U \in \mathrm{SU}(n)$.

We call such systems purely Hamiltonian.

Proof. (i) ⇒ (ii) is obvious and the converse follows from [BN08, Thm. 18]. (i) ⇔ (iii): This is clear from the fact that the relaxation algebra is well-defined, cf. Lemma 4.A.6. (iii) ⇔ (iv): Elementary. (iii) ⇒ (v): Immediate from the definitions. (v) ⇒ (iv): Follows from Lemma 3.0.6. □

It is clear that purely Hamiltonian systems are always unital stabilizable, which in turn are always unital. Moreover, these inclusions are strict.

Reachability

Similar to stabilizability, reachability properties are also highly dependent on the structure of the relaxation algebra $\mathcal{V}$. From Lemma 4.4.1 (vi) we know that the reachable set is contained in the majorization polytope of the initial state. Hence it is natural to ask when the reachable set is as large as it could be. Due to the continuity of solutions, it is clear that in the reduced control system not every point in the majorization polytope can be reached. This however is an artifact of the Weyl symmetry, and one should really ask which states in the Weyl chamber of the initial state can be reached.

Theorem 4.4.9 (Reachable Set for Unital Systems). *Assume that* $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ *is unital and consider an initial state* $\lambda \in \Delta_{\downarrow}^{n-1}$ *in the ordered Weyl chamber. Then exactly one of the following is the case:*

(i) If $-L$ *is purely Hamiltonian, then* $\overline{\mathsf{reach}_{\bar{\Lambda}}(\lambda)} = \{\lambda\}$.

(ii) If $-L$ *is of type* $(2, 1, \ldots, 1)$ *or type* $(1, \ldots, 1)$ *(i.e. unital stabilizable) but not purely Hamiltonian, then* $\overline{\mathsf{reach}_{\bar{\Lambda}}(\lambda)} \cap \Delta_{\downarrow}^{n-1} = \{\mu \in \Delta_{\downarrow}^{n-1} : \mu \preceq \lambda\}$.

(iii) If $-L$ is of type greater than $(2,1,\dots,1)$, then

$$\overline{\mathsf{reach}_{\bar{\Lambda}}(\lambda)} \cap \Delta_{\downarrow}^{n-1} \begin{cases} = \{\mathbf{e}/n\} & \text{if } \lambda = \mathbf{e}/n \\ \subsetneq \{\mu \in \Delta_{\downarrow}^{n-1} : \mu \preceq \lambda\} & \text{else}\,. \end{cases}$$

Proof. It is clear that the three cases are mutually exclusive and cover all possible $-L$. (i): This follows most easily from Lemma 4.4.8 (v). (ii): Let U be a unitary such that L_U is block diagonal with blocks of sizes $(2,1,1,\dots,1)$ (this includes the $(1,1,\dots,1)$ case). Because L is not purely Hamiltonian, we can choose L_U such that it is not diagonal, and thus the off-diagonal elements in the block of size 2 are non-zero. In addition, we have access to all $L_{UP} = P^\top L_U P$ where P is any permutation matrix meaning we have (approximate) access to arbitrary two-level weight shifts (also called "T-transforms"). By [MOA11, Ch. 2, Thm. B.6] this implies the claim.

(iii): Without loss of generality $\lambda \neq \mathbf{e}/n$. By assumption, M_U must always have either one block of size at least 3 or at least two blocks of size at least 2. This means that at least 3 elements of λ will be acted on by M_U. In particular, as in the proof of Theorem 4.4.6, the largest (smallest) of these elements will have a strictly negative (positive) derivative. Since the edges of the majorization polytope incident to λ correspond to the mixing of two neighboring elements in λ, no derivative pointing along an edge can be achieved. Let v be a unit vector pointing along one of the edges. Let α be a linear functional with $\alpha(v) = 0$ such that α is strictly positive on any point in the majorization polytope that does not lie on the edge defined by v. Consider the function $f_{U,\alpha}(\lambda) = -\alpha(L_U\lambda)$ on a compact neighborhood V of λ which is disjoint from stab_Λ. One can show that $f^-(\lambda) = \min_U f_{U,\alpha}(\lambda)$ is continuous on V. By shrinking V (while keeping λ inside) we can assume that $f^-(\lambda) \geq \varepsilon$ on V for some $\varepsilon > 0$. Let $\mu(t)$ be any solution to ($\bar{\Lambda}$) starting at λ, and let $T > 0$ be such that $\mu(t) \in V$ for all $t \in [0,T]$. Then

$$\alpha(\mu(T)) = \int_0^T \partial_t \alpha(\mu(t))dt = \int_0^T \alpha(\partial_t \mu(t))dt \geq T\varepsilon > 0\,.$$

Thus no point on an edge incident to λ which is not in V can be in $\overline{\mathsf{reach}_{\bar{\Lambda}}(\lambda)}$. □

This recovers the main result of [SAZ19] where it was shown, using similar arguments, and based on [RBR18], that a generator $-L$ can effect all state transfers respecting majorization if and only if we are in case (ii). Moreover, the case of a single Lindblad term V_k which is not a multiple of the identity (again covered by case (ii)) continues to hold in the infinite dimensional case [End+19].

Note that the results of Theorem 4.4.9 immediately lift to the full control system ($\mathcal{D}$) via Proposition 4.3.1 and using the fact that $\rho \preceq \sigma$ if and only if $\mathrm{spec}^{\downarrow}(\rho) \preceq \mathrm{spec}^{\downarrow}(\sigma)$.

Corollary 4.4.10. *For $-L$ unital, the reduced control system is reverse coolable if and only if case (ii) of Theorem 4.4.9 is satisfied.*

Accessibility

In our discussion of accessibility in Section 4.2 we forwent the unital case, which we will treat now.

Proposition 4.4.11. *Let $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ unital be given. If $n = 2$, the reduced control system (Λ) is generically (cf. footnote 4) directly accessible if and only if $-L$ is not purely Hamiltonian. If $n > 2$, then exactly one of the following holds.*

(i) The reduced control system (Λ) is generically directly accessible.

(ii) The reduced control system (Λ) is nowhere directly accessible and each $J(U)$ has identical off-diagonal elements.

Proof. If $n = 2$ and $-L$ is not purely Hamiltonian, then by Lemma 4.4.8 there is some $U \in \mathrm{SU}(2)$ such that L_U is not identically 0. But then L_U vanishes at only one point on Δ^1, and hence reduced control system is generically directly accessible.

Now let $n > 2$. We consider accessibility at the vertices e_i of the simplex. If there is some $i \in \{1, \dots, n\}$ and some $U \in \mathrm{SU}(n)$ such that $-L_U e_i$ is not invariant under all permutations stabilizing e_i, then the system is accessible at e_i and hence, using Proposition 2.4.9, it is generically directly accessible. If no such e_i and U exist, then all columns of all $J(U)$ are constant on their off-diagonal elements. It is easy to see that this implies that all off-diagonal elements of all $J(U)$ are the same. In this case $\mathfrak{L}$ is one dimensional, and hence the system is nowhere directly accessible. □

As in the non-unital case, if the reduced system (Λ) is generically directly accessible, then by Proposition 2.4.10, the full bilinear system ($\mathcal{D}$) is generically accessible.

A sufficient condition for being nowhere directly accessible is the following:

Corollary 4.4.12. *If the non-Hamiltonian part of $-L$ is unitarily invariant, then $\mathfrak{L}$ contains a single vector field, which is necessarily permutation invariant. If, in addition, $n > 2$, then the system is nowhere directly accessible.*

4.A Relaxation Algebras

In this section we recall some results pertaining to the structure theory of the Lindblad equation based primarily on [BN08], taking care to present them in their proper mathematical context. Since many of these results can also be found in other sources, such as [Kra+08, TV09, SW10], often with significantly differing terminology, we give precise definitions of all concepts used in this chapter. See also [Des+16] and references therein for related results (obtained in the Heisenberg picture) which extend into the infinite-dimensional setting. We introduce several matrix algebras generated by the Lindblad terms and consider their invariant subspace lattices. Once this structure is understood, we can modify the Lindblad equation to achieve certain dynamical properties.

The celebrated result by Gorini, Kossakowski, Sudarshan [GKS76] and Lindblad [Lin76] establishes the form of the generators of so-called *quantum-dynamical semigroups* (i.e. of continuous maps $t \mapsto \Phi_t$ on $\mathbb{R}_+$ into the completely positive trace-preserving linear maps, which satisfy $\Phi_0 = \mathbb{1}$ and $\Phi_t \circ \Phi_s = \Phi_{t+s}$ for all $t, s \geq 0$).

Remark 4.A.1. *For our purposes* completely positive dynamical semigroups *(i.e. quantum-dynamical semigroups without the trace preservation condition) and their generators turn out to be more convenient. Hence we will formulate many results for these more general objects.*

First, generators of completely positive dynamical semigroups can be characterized as follows, cf. [Lin76, Thm. 3]:

Lemma 4.A.2. *A linear map $-L \in \mathcal{L}(\mathbb{C}^{n,n})$ is the generator of a completely positive dynamical semigroup if and only if there is some completely positive*[14] *linear map $\phi = \sum_{k=1}^{r} V_k(\cdot)V_k^*$ and some $K \in \mathbb{C}^{n,n}$ such that*

$$-L(\rho) = \phi(\rho) - K\rho - \rho K^*. \tag{4.5}$$

[14] Recall that a map ϕ is completely positive if and only if it can be written in the form $\sum_{k=1}^{r} V_k(\cdot)V_k^*$, and any such operators V_k are called Kraus operators.

Table 4.1: The Equivalence Theorem 3.0.4 shows how the full control system ($\mathcal{D}$) and the reduced control system (Λ) are equivalent. The table summarizes the equivalence for several important control-theoretic notions giving the corresponding algebraic conditions.

Reduced control system (Λ)	Algebraic condition	Full control system ($\mathcal{D}$)	Reference
(regular) strongly stabilizable point	kernel of some L_U	(regular) strongly stabilizable point in orbit	Prop. 4.2.2
viable face of dimension $d-1$	lazy subspace of dimension d	collecting subspace via time-independent Hamiltonian control	Prop. 4.2.4
stabilizable system	simultaneously triangulable	$\mathrm{SU}(n)$-orbits approximately viable	Thm. 4.2.7
generically directly accessible	non-unital or accessible at vertex	generically accessible	Prop. 4.2.12 & Prop. 4.4.11
asymptotically coolable	common right but not left eigenvector	asymptotically coolable with time-independent Hamiltonian	Thm. 4.3.7
directly reachable face	flow ratios, collecting but not enclosing subspace	time-independent Hamiltonian	Lem. 4.3.11 & Prop. 4.3.12
reverse coolable	lazy subspaces of each dimension	approx. reachability from pure state	Prop. 4.3.21
approx. controllable	—	approx. controllable	Prop. 4.3.20
unital stabilizable point	relaxation $*$-algebra and refinement	(regular) strongly stabilizable point in orbit	Thm. 4.4.6
unital reachability of majorization polytope	type of relaxation $*$-algebra	approx. reachability of majorized states	Thm. 4.4.9

NB: the entries in a row are not always exactly equivalent, but some conditions might only be necessary or sufficient. For precise statements we refer to the relevant results.

Moreover, $t \mapsto e^{-tL}$ is a quantum-dynamical semigroup if and only if $-L$ is also trace-annihilating, that is, (4.5) holds and there exists $H_0 \in \mathbb{C}^{n,n}$ Hermitian such that

$$K = \mathrm{i}H_0 + \frac{1}{2}\sum_{k=1}^{r} V_k^* V_k \,. \tag{4.6}$$

Note that condition (4.5) is also known as *conditional complete positivity* of $-L$ [EL77, Thm. 14.7]. Plugging (4.6) into (4.5) recovers the Lindblad equation given in (3.1).

It is important to note that the choice of Hamiltonian H_0 and of Lindblad terms V_k is not unique. The following well-known result characterizes the freedom of representation of a given Kossakowski–Lindblad generator, cf. [BP02, Eq. (3.72) & (3.73)].

Lemma 4.A.3. *Let $\{K, V_k : k = 1, \dots, r\}$ be a representation of some generator of a dynamical semigroup of completely positive maps. Given $\{V'_k\}_{k=1}^{s}$, if it holds that*

$$V'_k = \sum_{j=1}^{r} u_{kj} V_j \quad \text{for all } k = 1, \dots, s \tag{4.7}$$

for some $U := (u_{kj})_{k,j=1}^{s,r} \in \mathbb{C}^{s,r}$ which satisfies either $UU^* = \mathbb{1}_s$ (if $r \geq s$) or $U^*U = \mathbb{1}_r$ (if $r \leq s$), then $\{K, V_k : k = 1, \dots, r\}$ and $\{K, V'_k : k = 1, \dots, s\}$ are equivalent. Similarly, if $\{V'_k\}_{k=1}^{r}$ and K' satisfy, for some $c_k \in \mathbb{C}$ and $\lambda \in \mathbb{R}$, that

$$\begin{aligned} V'_k &= V_k + c_k \mathbb{1} \\ K' &= K + \sum_{k=1}^{r} \overline{c}_k V_k + \Big(\mathrm{i}\lambda + \frac{1}{2}\sum_{k=1}^{r} |c_k|^2\Big)\mathbb{1} \,, \end{aligned} \tag{4.8}$$

then $\{K, V_k : k = 1, \dots, r\}$ and $\{K', V'_k : k = 1, \dots, r\}$ are equivalent. Conversely, if two representations $\{K, V_k : k = 1, \dots V_r\}$ and $\{K', V'_k : k = 1, \dots V_s\}$ define the same generator, then they are related by some sequence of the transformations above.

Proof. For the reader's convenience we provide a short proof. The first statement follows from (4.5) together with the characterization of "uniqueness" of Kraus operators [Wat18, Coro. 2.23]. Next, (4.8) is a straightforward computation. Finally, if $\{K, V_k : k = 1, \dots V_r\}$ and $\{K', V'_k : k = 1, \dots V_s\}$ define the same generator, then shifting $V_k \to V_k - \frac{\mathrm{tr}(V_k)}{n}\mathbb{1}$, $V'_k \to V'_k - \frac{\mathrm{tr}(V'_k)}{n}\mathbb{1}$ turns K, K' into K_1, K'_1 (according to (4.8)), respectively. Thus we may assume that all V_k, V'_k are traceless. In this case, vectorizing [MN07, Ch. 2.4] $\sum_{k=1}^{r} V_k(\cdot)V_k^* - K_1(\cdot) - (\cdot)K_1^* = \sum_{k=1}^{s} V'_k(\cdot)(V'_k)^* - K'_1(\cdot) - (\cdot)(K'_1)^*$ yields $\sum_{k=1}^{r} \overline{V_k} \otimes V_k - \mathbb{1} \otimes K_1 - \overline{K_1} \otimes \mathbb{1} = \sum_{k=1}^{s} \overline{V'_k} \otimes V'_k - \mathbb{1} \otimes K'_1 - \overline{K'_1} \otimes \mathbb{1}$. Using $\mathrm{tr}(V_k) = \mathrm{tr}(V'_k) = 0$ for all k, taking the partial trace over the first system shows $\overline{\mathrm{tr}(K_1 - K'_1)}\frac{\mathbb{1}}{n} = -(K_1 - K'_1)$. Taking the trace again yields $\mathrm{tr}(K_1 - K'_1) \in \mathrm{i}\mathbb{R}$, meaning there exists $\lambda \in \mathbb{R}$ such that $K'_1 = K_1 + \mathrm{i}\lambda\mathbb{1}$ (in accordance with (4.8)). This also means that the shifted V_k, V'_k satisfy $\sum_{k=1}^{r} V_k(\cdot)V_k^* = \sum_{k=1}^{s} V'_k(\cdot)(V'_k)^*$ which, again by [Wat18, Coro. 2.23], shows that there exists an isometry U relating the two sets via (4.7). This concludes the proof. □

In particular, this result shows that the Lindblad terms can always be chosen traceless. Indeed, this is one way to ensure that Kossakowski–Lindblad generators decompose uniquely: if $\mathrm{tr}(V_j) = 0$ for all j, then there exists a unique $H_0 \in \mathfrak{isu}(n)$ such that $-L = -\mathrm{i}\,\mathrm{ad}_{H_0} - \sum_{k=1}^{r} \Gamma_{V_k}$, cf. [GKS76, Thm. 2.2], [Dav80].

Remark 4.A.4. *For the Lindblad equation, the transformation of H_0 resulting from the shift (4.8) is given by $H' = H_0 + \frac{\mathrm{i}}{2}\sum_{k=1}^{r}(c_k^* V_k - c_k V_k^*) + \lambda \mathbb{1}$.*

The central objects considered in studying the structure of the Lindblad equation are the following matrix algebras:

Definition 4.A.5. *Let $-L \in \mathcal{L}(\mathbb{C}^{n,n})$ be the generator of a completely positive dynamical semigroup represented by $\{K, V_k : k = 1, \ldots, r\}$. We define the following unital subalgebras[15] of $\mathbb{C}^{n,n}$:*

$$\mathcal{V} = \langle \mathbb{1}, V_k : k = 1, \ldots, r \rangle_{\mathsf{alg}}, \qquad \mathcal{V}^+ = \langle \mathbb{1}, K, V_k : k = 1, \ldots, r \rangle_{\mathsf{alg}},$$

called the relaxation algebra *and* extended relaxation algebra*, respectively.*

Crucially, as a direct consequence of Lemma 4.A.3, these algebras are well-defined:

Lemma 4.A.6. *The algebras $\mathcal{V}$ and $\mathcal{V}^+$ are well-defined, meaning that they only depend on $-L$, and not on the chosen representation.*

Proof. Let $\{K, V_k' : k = 1, \ldots, s\}$ be another representation of L satisfying (4.7) and let $\tilde{\mathcal{V}}$ and $\tilde{\mathcal{V}}^+$ denote the corresponding algebras. Clearly each V_k' is a complex linear combination of the V_k and hence contained in $\mathcal{V}$ and $\mathcal{V}^+$. Trivially it holds that $K \in \mathcal{V}^+$. This shows that $\tilde{\mathcal{V}} \subseteq \mathcal{V}$ and $\tilde{\mathcal{V}}^+ \subseteq \mathcal{V}^+$. The reversed inclusions are analogous. Now consider another representation $\{K', V_k' : k = 1, \ldots, s\}$ of L satisfying (4.8). Each V_k' is a linear combination of V_k and $\mathbb{1}$ and K' is a linear combination of K, all V_k, and $\mathbb{1}$. Hence again $\tilde{\mathcal{V}} \subseteq \mathcal{V}$ and $\tilde{\mathcal{V}}^+ \subseteq \mathcal{V}^+$, and again the reverse inclusion is analogous. Since by Lemma 4.A.3 any two equivalent representations of L can be transformed into each other using a composition of these transformation, they must generate the same algebras by the above argument. Hence the algebras only depend on L, and so they are well defined. □

Note, however, that

- the algebras are not sufficient to determine the generator;
- for Lemma 4.A.6 to hold it is necessary to include the identity in the definition of the (extended) relaxation algebra.

Below we will be interested in invariant subspaces of the matrix algebras defined above, so let us take a brief moment to recall some basic facts. A subspace $S \subseteq \mathbb{C}^n$ is *invariant* for a set of matrices $\mathcal{A}$ if $AS \subseteq S$ for all $A \in \mathcal{A}$. If P_S is a projection onto S, then S is invariant for $\mathcal{A}$ if and only if $(\mathbb{1} - P_S)AP_S = 0$ for all $A \in \mathcal{A}$. Moreover, the set of all invariant subspaces of $\mathcal{A}$ forms a *lattice*, meaning that if S and T are invariant, then so are $S \cap T$ and $S + T$. Since the invariant subspaces of $\mathcal{A}$ and $\langle \mathcal{A} \rangle_{\mathsf{alg}}$ coincide, we will work with the (well-defined) algebras $\mathcal{V}$ and $\mathcal{V}^+$. For some mathematical background we refer to [RR00, Far01] for matrix algebras, to [GLR06] for invariant subspaces, and to [Bir67] for lattice theory.

The following lemma, which is an immediate consequence of (4.5), is quite useful in relating the algebras defined above to the structure of dynamical semigroups.

Lemma 4.A.7. *Let $-L \in \mathcal{L}(\mathbb{C}^{n,n})$ be the generator of a completely positive dynamical semigroup represented by $\{K, V_k : k = 1, \ldots r\}$. Let $O, P, Q, R \in \mathbb{C}^{n,n}$ be orthogonal projections. Then for all $A \in \mathbb{C}^{n,n}$ it holds that*

$$O\, L(PAR)\, Q = \sum_{k=1}^{r} (OV_kP)A(QV_kR)^* - (OKP)A(RQ) - (OP)A(QKR)^*.$$

[15] A unital matrix algebra is one which contains the identity matrix. Since all our matrix algebras will be unital, we will omit this term in the following to avoid confusion.

Lemma 4.A.8. *Let $-L \in \mathcal{L}(\mathbb{C}^{n,n})$ be the generator of a completely positive dynamical semigroup represented by $\{K, V_k : k = 1, \dots, r\}$. Given any $U \in \mathrm{U}(n)$ and any subspace $\{0\} \neq S \subseteq \mathbb{C}^n$, where P_S denotes the orthogonal projection onto S, the following are equivalent:*

(i) *S is a common invariant subspace of all V_k.*

(ii) *S is an invariant subspace of $\mathcal{V}$.*

(iii) *$P_S^\perp L(P_S A P_S) P_S^\perp = 0$ for all $A \in \mathbb{C}^{n,n}$.*

(iv) *$P_S^\perp L(\rho) P_S^\perp = 0$ for some state ρ with $\operatorname{supp}(\rho) = S$.*

(v) *If the first $\dim S$ columns of U^* span S, then $-L_U$ is block triangular, i.e. for all $1 \leq j \leq \dim S < i \leq n$ one has $(-L_U)_{ij} = 0$.*

If one and thus all of these conditions hold, we say that S is a lazy subspace *of $-L$.*

Proof. (i) $\Leftrightarrow$ (ii): Note that by Lemma 4.A.6 the invariant subspaces of the V_k are well-defined and exactly the invariant subspaces of $\mathcal{V}$. (i) $\Rightarrow$ (iii): By Lemma 4.A.7,

$$P_S^\perp L(P_S A P_S) P_S^\perp = \sum_{k=1}^{r} (P_S^\perp V_k P_S) A (P_S^\perp V_k P_S)^*.$$

(iii) $\Rightarrow$ (iv): Trivial. (iv) $\Rightarrow$ (i): By $0 = P_S^\perp L(P_S \rho P_S) P_S^\perp = \sum_{k=1}^r (P_S^\perp V_k P_S \sqrt{\rho})(P_S^\perp V_k P_S \sqrt{\rho})^*$, we have $P_S^\perp V_k P_S \sqrt{\rho} = 0$ for all k. Since $\operatorname{supp}(\rho) = \operatorname{supp}(\sqrt{\rho}) = S$, this forces $P_S^\perp V_k P_S = 0$ for all k. (i) $\Leftrightarrow$ (v): This is due to $(-L_U)_{ij} = \sum_{k=1}^r |\langle i | U^* V_k U | j \rangle|^2$ as follows from (3.4). □

Intuitively, a lazy subspace is one which is invariant under $-L$ to first order, hence the name. Importantly, they do not depend on K (or H_0), and thus they are independent of the control. A stronger notion of invariance is the following (cf. [BN08, SL05]), which is an immediate consequence of [BN08, Lem. 11]:

Lemma 4.A.9. *Let $-L \in \mathcal{L}(\mathbb{C}^{n,n})$ be the generator of a completely positive dynamical semigroup represented by $\{K, V_k : k = 1, \dots, r\}$. Given any subspace $\{0\} \neq S \subseteq \mathbb{C}^n$, where P_S denotes the orthogonal projection onto S, the following are equivalent:*

(i) *$P_S \,\mathfrak{pos}_1(n) P_S$ is invariant under e^{-tL} for all $t \geq 0$, that is, for all $\rho \in \mathfrak{pos}_1(n)$ it holds that $P_S e^{-tL}(P_S \rho P_S) P_S = e^{-tL}(P_S \rho P_S)$.*

(ii) *$P_S L(P_S \rho P_S) P_S = L(P_S \rho P_S)$.*

(iii) *S is an invariant subspace for $\mathcal{V}^+$.*

If one and thus all of these conditions hold, we say that S is a collecting subspace.

Collecting subspaces are closely related to fixed points, see [BN08, Prop. 5 & Lem. 12].

Lemma 4.A.10. *The following hold:*

(i) *If $\rho \in \mathfrak{pos}_1(n)$ is a fixed point, then $\operatorname{supp}(\rho)$ is collecting.*

(ii) *If $S \subseteq \mathbb{C}^n$ is collecting, then there is a fixed point ρ with $\operatorname{supp}(\rho) \subseteq S$.*

Lemmas 4.A.8 and 4.A.9, together with the preceding discussion, show that the lazy subspaces form a lattice of invariant subspaces, and that the collecting subspaces form a sublattice thereof.[16] A simple but important consequence is that lazy subspaces are exactly those which can be made collecting using a time-independent Hamiltonian:

Corollary 4.A.11. *Let $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ be a Kossakowski–Lindblad generator. If $S \subseteq \mathbb{C}^n$ is a lazy subspace, then there exists a Hamiltonian H_S such that S is collecting for $-L - \mathrm{i}\,\mathrm{ad}_{H_S}$. Indeed, any Hamiltonian satisfying*

$$P_S^\perp H_S P_S = -P_S^\perp \Big(H_0 + \tfrac{1}{2\mathrm{i}} \sum_{k=1}^{r} V_k^* V_k \Big) P_S$$

will do the job. We call H_S a stabilizing Hamiltonian *for S.*

It is clear that such H_S always exists since, in an appropriate basis, the condition only determines matrix elements below the diagonal.

As a consequence of Corollary 4.A.11 and [BN08, Thm. 13] we have the following relation between minimal collecting subspaces and extremal fixed points.

Lemma 4.A.12. *Let $-L \in \mathfrak{w}_{\mathsf{KL}}(n)$ be a Kossakowski–Lindblad generator and let $S \subseteq \mathcal{H}$ be a subspace. The following are equivalent:*

(i) S is a minimal[17] *collecting subspace of S.*

(ii) There is an extremal[18] *fixed point with support equal to S.*

(iii) $-L$ has a unique fixed point on S, and it has full rank (on S).

This result recovers [Kra+08, Thm. 2] and [SW10, Prop. 2].

We say that a subspace S is *decaying* if $\mathrm{tr}(P_S e^{-Lt}\rho) \to 0$ as $t \to \infty$ for all states $\rho \in \mathfrak{pos}_1(n)$. The orthocomplement of a decaying subspace is called an *asymptotic* subspace. Note that the property of being decaying is preserved under taking sums[19] and subspaces. Hence there exists a unique *maximal decaying subspace*, and the decaying subspaces are exactly all of its subspaces. Its complement is then the unique *minimal asymptotic subspace*. Note that this coincides with the so-called "four-corners decomposition" introduced in [Alb+16].

Lemma 4.A.13. *Let $S \subseteq \mathbb{C}^n$ be a subspace. The following are equivalent:*

(i) S is the minimal asymptotic subspace.

(ii) S is the smallest subspace containing the support of all fixed points $\rho \in \ker L$.

(iii) S is the span of all minimal invariant subspaces.

Proof. The equivalence of (ii) and (iii) follows from Lemma 4.A.12. It is easy to see that the support of a fixed point is contained in the minimal asymptotic subspace. The converse is shown in [BN08, Prop. 15]. This proves the result. □

[16] To be precise, we only obtain lattices after including the trivial subspace $\{0\}$ which we excluded from both definitions.

[17] A *minimal collecting subspace* is one which does not contain a smaller collecting subspace. It is an atom in the lattice of collecting subspaces.

[18] An *extremal fixed point* is one which is not a non-trivial convex combination of two distinct fixed points, that is, it is an extreme point of the convex set of fixed points.

[19] This follows from the fact that if $\mathrm{tr}(P_S\rho) = 0$ and $\mathrm{tr}(P_T\rho) = 0$, then $\mathrm{supp}(\rho) \subseteq S^\perp \cap T^\perp = (S+T)^\perp$ and hence $\mathrm{tr}(P_{S+T}\rho) = 0$.

An algebra of complex matrices which is closed under taking the adjoint is called a $*$-algebra. If $\mathcal{A}$ is a $*$-algebra, then S is invariant under $\mathcal{A}$ if and only if $P_S A = A P_S$ for all $A \in \mathcal{A}$, where P_S is the orthogonal projection onto S.

Definition 4.A.14. *We define the* relaxation $*$-algebra $\mathcal{V}_*$ *and the* extended relaxation $*$-algebra $\mathcal{V}_*^+$ *as*

$$\mathcal{V}_* = \langle \mathbb{1}, V_k, V_k^* : k = 1, \dots, r \rangle_{\mathsf{alg}}\,, \qquad \mathcal{V}_*^+ = \langle \mathbb{1}, H, V_k, V_k^* : k = 1, \dots, r \rangle_{\mathsf{alg}}\,.$$

Clearly, $\mathcal{V}_*$ and $\mathcal{V}_*^+$ are the $*$-algebras generated by $\mathcal{V}$ and $\mathcal{V}^+$ respectively, and hence they are well-defined.

Lemma 4.A.15. *The following are equivalent:*

(i) There is no decay, i.e. $\mathbb{C}^n$ is the minimal asymptotic subspace.

(ii) The lattice of collecting subspaces is orthocomplemented (cf. footnote 12) and so $\mathcal{V}^+$ is a $*$-algebra.

Proof. The proof of [BN08, Prop. 14] shows that if there is no decay, the lattice of collecting subspaces is orthocomplemented. Thus, by [GLR06, Thm. 11.5.1] $\mathcal{V}^+$ is a $*$-algebra. Conversely, if the lattice of collecting subspaces is orthocomplemented, it is atomistic[20], and hence by Lemma 4.A.13 there is no decay. □

Lemma 4.A.16. *Let $S \subseteq \mathbb{C}^n$ be a subset, and P_S the orthogonal projection onto S. Then the following are equivalent:*

(i) Both S and $S^\perp$ are collecting.

(ii) S is an invariant subspace of $\mathcal{V}_*^+$.

(iii) $[P_S, V_k] = [P_S, H] = 0$ for $k = 1, \dots, r$.

We call such subspaces enclosing *or simply* enclosures.

Proof. (i) ⇔ (ii): This is clear since $\mathcal{V}_*^+$ is the $*$-algebra generated by $\mathcal{V}^+$. (ii) ⇔ (iii): Elementary. □

The existence of enclosures implies the existence of conserved quantities and (dynamical) symmetries of the Kossakowski–Lindblad generator, but they are not necessary [BN08, AJ14].

Lemma 4.A.17. *Assume that $A \in \mathbb{C}^{n,n}$ is not a scalar multiple of the identity. Then there exists a Hermitian matrix $B \in \mathbb{C}^{n,n}$ such that they generate the entire matrix algebra, that is, $\langle A, B \rangle_{\mathsf{alg}} = \mathbb{C}^{n,n}$.*

Proof. Let $B \in \mathbb{R}^{n,n}$ be any diagonal matrix with distinct non-zero diagonal elements. If we can find $U \in \mathrm{U}(n)$ unitary such that $(U^* A U)_{jk} \neq 0$ for all $j \neq k$, then [Laf92, Lem. 2] implies that the algebra generated by $\{U^* A U, B\}$ is all of $\mathbb{C}^{n,n}$; in particular $\langle A, U B U^* \rangle_{\mathsf{alg}} = \mathbb{C}^{n,n}$ which would conclude the proof. Consider the function $U \mapsto (U^* A U)_{ij}$ defined on the unitary group and let Z_{ij} denote the corresponding zero set. Since the unitary group is a real analytic manifold and since the function is real analytic, Z_{ij} is either all of $\mathrm{U}(n)$ or it has open dense complement in $\mathrm{U}(n)$, cf. [KP02, Thm. 6.3.3]. If A is not a multiple of the identity, none of the Z_{ij} equal $\mathrm{U}(n)$. Hence the union of all Z_{ij} has open dense complement, and hence there is U such that $(U^* A U)_{jk} \neq 0$ for all $j \neq k$, as desired. □

[20] A lattice with least element 0 is atomistic if every element is a least upper bound of a set of atoms, which are the minimal non-zero elements.

Corollary 4.A.18. *Let $A, C \in \mathbb{C}^{n,n}$ such that A is not a multiple of the identity. Then there exists $B \in \mathbb{C}^{n,n}$ Hermitian such that $\langle A, B + C\rangle_{\mathsf{alg}} = \mathbb{C}^{n,n}$.*

Proof. Let B, U be as in the proof of Lemma 4.A.17. By [Kat80, Ch. 2 §1], for all $\varepsilon > 0$ small enough, there exist analytic curves $S(\varepsilon)$ and $B(\varepsilon)$ such that $B + \varepsilon U^* C U = S(\varepsilon)B(\varepsilon)S(\varepsilon)^{-1}$, with $S(0) = \mathbb{1}$ and $B(\varepsilon)$ diagonal with distinct non-zero diagonal elements. Then

$$\langle A, \tfrac{UBU^*}{\varepsilon} + C\rangle_{\mathsf{alg}} = U\langle U^*AU, S(\varepsilon)B(\varepsilon)S(\varepsilon)^{-1}\rangle_{\mathsf{alg}}U^* = US(\varepsilon)\,\langle S(\varepsilon)^{-1}AS(\varepsilon), B(\varepsilon)\rangle_{\mathsf{alg}}\,S(\varepsilon)^{-1}U^*.$$

Since for ε small enough it still holds that $(S(\varepsilon)^{-1}U^*AUS(\varepsilon))_{jk} \neq 0$ for all $j \neq k$, the result follows again from [Laf92, Lem. 2]. □

Using Corollary 4.A.18 together with [TV09, Thm. 12] we obtain a useful extension of Corollary 4.A.11.

Lemma 4.A.19. *Let $S \subseteq \mathbb{C}^n$ be a lazy subspace for $-L$. Then there exists a Hamiltonian H such that $-(\mathrm{i}\,\mathrm{ad}_H + L)$ has a unique (attractive) fixed point with support S if and only if S is not an enclosure.*

Proof. In [TV09, Thm. 12] it is shown how to choose H such that S is attractive. We are still free to choose H on S itself. So using Corollary 4.A.18 we can make sure that S is minimal collecting and hence by Lemma 4.A.12 it supports a unique fixed point. □

4.B Toy Model

In [DES19] the authors introduced a toy model for studying control systems with switchable[21] Markovian noise and unitary control. In this toy model, the states remain at all times diagonal in the energy eigenbasis and the controls are restricted to level permutations. These stipulations suggest the following hybrid (or impulsive) scenario to define the toy model on $\Delta^{n-1} \subset \mathbb{R}^n$ by

$$\begin{aligned} &\dot{x}(t) = -Bx(t)\,, \quad x(t_k) = \pi_k x_k\,, \quad t \in [t_k, t_{k+1})\,, \\ &x_0 \in \Delta^{n-1}\,, \quad x_{k+1} = e^{-(t_{k+1}-t_k)B}x(t_k)\,, \quad k \geq 0\,, \end{aligned} \tag{T′}$$

where $-B$ is a stochastic generator matrix.[22] Furthermore, $0 = t_0 \leq t_1 \leq t_2 \leq \ldots$ is an arbitrary switching sequence and $\pi_k \in S_n$ are arbitrary permutation matrices. Both the switching points and the permutation matrices are regarded as controls for (T′). For simplicity, we assume that the switching points do not accumulate on finite intervals. For more details on hybrid (or impulsive) control systems see, e.g., [LBS89, LMS93, AHS96].

In this section we specify control problems (T′) "compatible with quantum thermodynamics". Given a Hamiltonian $H_0 \in \mathrm{i}\mathfrak{u}(n)$, by unitary controllability it may be chosen diagonal with increasing eigenvalues E_k without loss of generality. The corresponding equilibrium state resulting from coupling to a bath of temperature T is the *Gibbs vector*

$$d = \frac{(e^{-E_k/T})_{k=1}^n}{\sum_{k=1}^n e^{-E_k/T}} \in \Delta^{n-1} \tag{4.9}$$

[21] This is motivated by recent experimental progress [Hof+09, Yin+13, Che+14, Won+19].

[22] Meaning that $\{e^{-Bt}\}_{t\geq 0}$ is a one-parameter semigroup of stochastic matrices. Such matrices define continuous time Markov chains, and are also called transition rate matrices or intensity matrices [Nor97].

with $\rho_{\mathsf{Gibbs}} = \operatorname{diag}(d) \in \mathfrak{pos}_1(n)$. As shown in [DES19], $\operatorname{diag}(d)$ can then be obtained as the unique *fixed point* of the Lindblad equation when choosing two Lindblad terms as

$$V_1 = \sigma_+^d := \sum_{k=1}^{n-1} \sqrt{k(n-k)} \cos(\theta_k) \, |k\rangle\langle k+1| \tag{4.10}$$

$$V_2 = \sigma_-^d := \sum_{k=1}^{n-1} \sqrt{k(n-k)} \sin(\theta_k) \, |k+1\rangle\langle k| \,, \tag{4.11}$$

where

$$\theta_k := \arccos\left(\left(1 + \tfrac{d_{k+1}}{d_k}\right)^{-\frac{1}{2}}\right) \in (0, \tfrac{\pi}{4}]. \tag{4.12}$$

Confining ourselves to σ_+^d and σ_-^d with their non-zero entries on the first off-diagonals is in accordance with the common dipolar selection rules allowing for "one-quantum transitions" (as governed by Wigner's $3j$-symbol) [Zar88, p. 185 ff.].

In [DES19] it was shown that in the $T = 0$ case it holds that $\overline{\mathsf{reach}_{\mathsf{T}'}(x)} = \Delta^{n-1}$ and that for equidistant energies the majorization polytope $M_{\mathbf{e}}(d)$ is invariant. A generalization of this result can be found in [3] and it is based on the concept of d-majorization, which has been explored in [10, ED22], see also Appendix A.

Analytic Results for Qutrits

In this section we will explicitly determine the shape of the reachable set and of the set of stabilizable states for the three-dimensional case $d \in \mathbb{R}^3$, $d > 0$. For this we first introduce some general notions.

It pays off to approach the toy model (T') from a different, but equivalent,[23] perspective: instead of letting the permutations act on the states, leading to discontinuous paths, we let the permutations act on the drift vector field, leading to the following differential inclusion, where, in analogy to $\mathsf{reach}_{\mathsf{T}'}(x)$, we write $\mathsf{derv}(x)$ for the *set of achievable derivatives at* x:

$$\dot{x}(t) \in \operatorname{conv}(\mathsf{derv}(x(t))), \quad \mathsf{derv}(x) := \{-\pi B \pi^{-1} x : \pi \in S_n\}, \tag{T}$$

cf. [Smi02].[24] Many ideas work for any matrix $-B$ which generates a one-parameter semigroup of stochastic matrices and has unique fixed point d, but for some results we will restrict to the following:

(A) $-B$ is of the form resulting from Eqs. (4.10) & (4.11) and the corresponding Hamiltonian has equidistant energies.

Assumption (A) ensures that we obtain sensible formulas, and it is physically motivated, see [3, Remark 5].

Stabilizable States

As before, the set of *stabilizable states* $\mathsf{stab}_{\mathsf{T}}$ is defined to be all $x \in \Delta^{n-1}$ such that $0 \in \operatorname{conv}(\mathsf{derv}(x))$. Intuitively, these are the points in Δ^{n-1} that, when taken as starting point, one can remain arbitrarily close to. More precisely we have the following result:

[23] The systems are equivalent in the sense that every solution of one system has a corresponding solution in the other system differing only by some (time-dependent) permutation. Note however that we allow more general controls in the differential inclusion, so that this equivalence is only approximate in general.

[24] In particular Thm. 2.3 therein shows the equivalence of control systems and the corresponding differential inclusions. Note that taking the convex hull leads to a relaxation of the differential inclusion, which is still approximately equivalent to the original control system, see [AC84, Ch. 2.4, Thm. 2].

Lemma 4.B.1. *A state $x_0 \in \Delta^{n-1}$ is stabilizable if and only if for every $\varepsilon > 0$ and $\tau > 0$ there is a solution $x : [0, \tau] \to \Delta^{n-1}$ to (T) with $x(0) = x_0$ which remains inside of the ε-ball $B_\varepsilon(x_0) \cap \Delta^{n-1}$.*

Proof. If x_0 is stabilizable, then the constant path $x \equiv x_0$ is a solution to (T). Conversely, assume that x_0 is not stabilizable. Then, by continuity, there is some $\delta > 0$ and some linear functional β on $\mathbb{R}^n$ such that β is less than $-\delta$ on $\mathsf{derv}(y)$ for all y in some neighborhood of x_0. Hence there is some time $\tau > 0$ where any solution must leave $B_\varepsilon(x_0)$ for some ε small enough. □

If 0 is not contained in the convex hull of the achievable derivatives at x, then there must exist some linear functional α on $\mathbb{R}^n$ which is negative on $\mathsf{derv}(x)$. Note that while α lives on $\mathbb{R}^n$, only the part parallel to the simplex Δ^{n-1} matters. Based on this observation, the idea is to consider the "permuted" functionals $\alpha_\pi(x) := -\alpha(\pi B \pi^{-1} x)$ because, given any $x \in \Delta^{n-1}$, if there exists α such that $\alpha_\pi(x) < 0$ for all $\pi \in S_n$, then x cannot be stabilisable. Conversely, if x is not stabilisable, then there exists some α for which $\alpha_\pi(x) < 0$ for all $\pi \in S_n$. Obviously, d as well as all permutations of d are stabilizable.

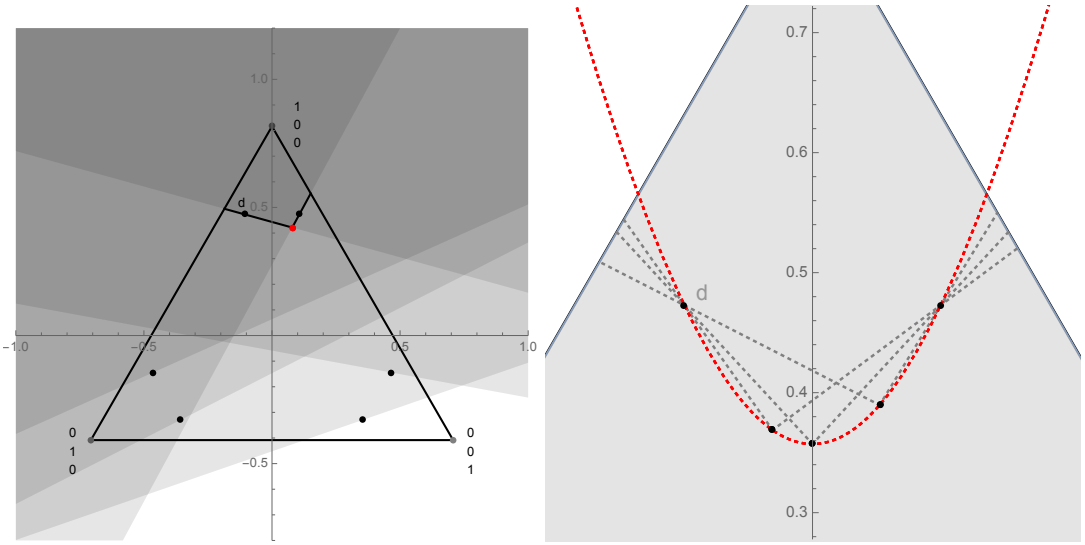

Figure 4.1: Illustration of how to construct the boundary curves of the set of stabilisable points in the case of $a = 0.3$. Left: The shaded regions comprise the points where the functionals α_π are negative; highlighted is the intersection of all the negative regions with the simplex. The points in this region are certainly not stabilizable. In particular the intersection point of $\ker(\alpha_{\mathrm{id}})$ and $\ker(\alpha_{\tau_{23}})$ is marked in red. Right: For three different values of α, parts of $\ker(\alpha_{\mathrm{id}})$ and $\ker(\alpha_{\tau_{23}})$ and their intersections are shown. Taken together, these intersections form the curve given in red, which constitutes a part of the boundary of the set of stabilizable points.

Let us now focus on the three-dimensional case. We will compute a closed curve connecting all these points, which will turn out to be the boundary of the set of stabilizable states: everything (on or) inside the curve will be stabilizable and everything outside will be non-stabilizable, refer to Figure 4.2 below for two examples. Let us, e.g., focus on the part of the boundary curve between d and $\tau_{23}\, d$, where τ_{23} is the transposition acting on the second and third element. Note that d and $\tau_{23}\, d$ are located in neighbouring Weyl chambers since the elements in d are always increasing or decreasing. The idea for determining its shape is: for every functional α (in a certain range) one can compute a point $\ker(\alpha_{\mathrm{id}}) \cap \ker(\alpha_{\tau_{23}}) \cap \Delta^2$ with the property that all points in the simplex "above" it cannot be stabilizable as shown in Figure 4.1. Moreover, due to Assumption (A), the curve will always be part of a conic section. To motivate this

approach, note that any point x with $\alpha_\pi(x) < 0$ for some α and for all $\pi \in S_n$ in contained in an open neighborhood of non-stabilizable points and hence cannot lie on the boundary. Thus we are looking for points lying in the kernel of at least one of the $\alpha_\pi(x)$. Moreover, we really need to find points lying in the intersection of two such kernels, since otherwise a small perturbation applied to α shows that the point has a non-stabilizable neighborhood.

Let us now invoke Assumption (A) so, without loss of generality, $H_0 := \mathrm{diag}(-1, 0, 1) \cdot \Delta E$ for some $\Delta E \in \mathbb{R}$, and thus $d = (1, a, a^2)/(1 + a + a^2)$ with $a = e^{-\Delta E/T}$. The generators of our dissipative dynamics (4.10) & (4.11) are fully characterized by the (constant) angle $\theta = \arccos(\frac{1}{\sqrt{1+a}})$ in (4.12). With this, the generator of the toy model takes the form (cf. also [DES19])

$$-B = \frac{2}{1+a}\begin{pmatrix} -a & 1 & 0 \\ a & -1-a & 1 \\ 0 & a & -1 \end{pmatrix}.$$

Let us go through the construction of the curve for the special (parabolic) case[25] where $a = \frac{1}{4}$. It will turn out that the boundary curve between d and $\tau_{23}d$ is fully determined by the family of functionals[26] α^λ, $\lambda \in [-\frac{1}{7}, \frac{1}{7}]$ where

$$\alpha^\lambda := -(\tfrac{1}{2} + \lambda)\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} - (\tfrac{1}{2} - \lambda)\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \tag{4.13}$$

so $\alpha^\lambda(x) = \lambda(x_2 - x_3) - \frac{1}{2}(x_2 + x_3)$ for all $x \in \mathbb{R}^3$. In order to compute $\ker(\alpha^\lambda_{\mathrm{id}}) \cap \ker(\alpha^\lambda_{\tau_{23}}) \cap \Delta^2$ we find that $\alpha^\lambda_{\mathrm{id}} = \alpha^\lambda \circ (-B)$ (up to a global factor, which we may omit because we have to normalise later on anyway) equals

$$-(\tfrac{1}{2} + \lambda)\begin{pmatrix} 0 & \frac{1}{2} & -2 \end{pmatrix} - (\tfrac{1}{2} - \lambda)\begin{pmatrix} \frac{1}{2} & -\frac{5}{2} & 2 \end{pmatrix} = \begin{pmatrix} \frac{\lambda}{2} - \frac{1}{4} & 1 - 3\lambda & 4\lambda \end{pmatrix}.$$

Also $\alpha^\lambda_{\tau_{23}} = \alpha^{-\lambda}_{\mathrm{id}} \circ \tau_{23}$ is generated by $\begin{pmatrix} -\frac{\lambda}{2} - \frac{1}{4} & -4\lambda & 3\lambda + 1 \end{pmatrix}$. With this we compute $\ker(\alpha^\lambda_{\mathrm{id}}) \cap \ker(\alpha^\lambda_{\tau_{23}})$ to be spanned by "the" vector which is orthogonal to the normal vector of both $\alpha^\lambda_{\mathrm{id}}$ and $\alpha^\lambda_{\tau_{23}}$, that is,

$$\begin{pmatrix} \frac{\lambda}{2} - \frac{1}{4} \\ 1 - 3\lambda \\ 4\lambda \end{pmatrix} \times \begin{pmatrix} -\frac{\lambda}{2} - \frac{1}{4} \\ -4\lambda \\ 3\lambda + 1 \end{pmatrix} = \begin{pmatrix} 1 + 7\lambda^2 \\ -\frac{7}{2}\lambda^2 - \frac{3}{4}\lambda + \frac{1}{4} \\ -\frac{7}{2}\lambda^2 + \frac{3}{4}\lambda + \frac{1}{4} \end{pmatrix}$$

Intersecting the line generated by this vector with the standard simplex only introduces a normalising factor since we have: $\ker(\alpha^\lambda_{\mathrm{id}}) \cap \ker(\alpha^\lambda_{\tau_{23}}) \cap \Delta^2 = \frac{1}{6}(4 + 28\lambda^2, -14\lambda^2 - 3\lambda + 1, -14\lambda^2 + 3\lambda + 1)^\top$. Finally, we reduce the dimensionality of the problem by isometrically embedding[27] the simplex Δ^2 in $\mathbb{R}^2$; this leads to the (parabolic) boundary curve $(\frac{\lambda}{\sqrt{2}}, \frac{1+14\lambda^2}{\sqrt{6}})$ where $\lambda \in [-\frac{1}{7}, \frac{1}{7}]$.

If $a \neq \frac{1}{4}$ we modify the family of functionals α^λ introduced previously by multiplying λ in (4.13) by $\frac{1}{2}\sqrt{1 + 2a}|(3 + 2a)(1 - 4a)|^{-1/2}$; however, the idea and the calculations are analogous. In the hyperbolic[28] case $a > \frac{1}{4}$ the boundary curve can be parametrized via

$$\left(w\frac{-2\lambda}{\lambda^2 - 1}, u\frac{\lambda^2 + 1}{\lambda^2 - 1} + v\right)$$

[25] This is the case where the energy gap $|\Delta E| = \ln(4)k_B T$, where we explicitly write the Boltzmann constant k_B.

[26] Since we only care about the component of the functional parallel to the simplex and since the normalisation does not matter, it suffices to consider a one-parameter family of functionals. The exact parametrisation and parameter range are chosen for ease of computation.

[27] This is done using the partial isometry $P = \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \sqrt{\frac{2}{3}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \end{pmatrix}$.

[28] The unital scenario $a = 1$ is a special case because then $d = \frac{\mathbf{e}}{3}$, so the set of stabilizable points collapses to $\{\frac{\mathbf{e}}{3}\}$.

where

$$v-u=\sqrt{\frac{2}{3}}\frac{1-a}{1+2a}\,,\ u+v=\sqrt{\frac{2}{3}}\frac{1-a}{1-4a}\,,\ w=\frac{\sqrt{2}(1-a)a}{\sqrt{|(1+2a)(3+2a)(1-4a)|}}\,.$$

For the elliptic case $a\in(0,\frac{1}{4})$ one finds

$$\Big(w\frac{2\lambda}{\lambda^2+1},u\frac{\lambda^2-1}{\lambda^2+1}+v\Big).$$

This covers the segment of the curve which connects d and $\tau_{23}d$. For the rest of the boundary curve note that, due to the permutation symmetry, there are only two different curve segments, cf. Figure 4.2. We have just computed one of them. The other one is obtained by re-arranging the elements of d in reverse order and repeating the calculation. One obtains the same formulas with a replaced by a^{-1}.

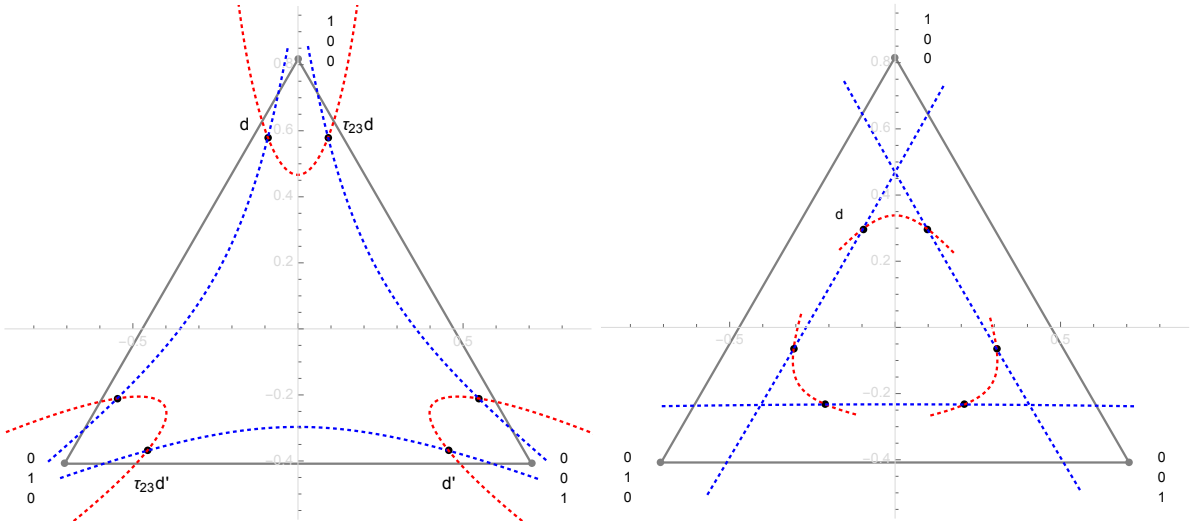

Figure 4.2: Left: The set of stabilisable states for the equidistant energy case with $a=\frac{1}{5}$. The set is bounded by six conics and contains the permutations of d. This is the elliptic case, and part of the ellipse is drawn in red, together with its permuted copies. The blue curve is obtained analogously by taking the hyperbolic case $a=5$ whose fixed point we denote d'. Right: The same approach gives the boundary of the set of stabilisable states for a random generator B numerically. Note that in general the bounding curves need not be conic sections, and one may obtain a convex shape.

We have seen that for each α we obtain an open (convex) region which is certainly not stabilizable. Parametrizing α in a circular fashion, i.e. $\alpha\in S^2\cap\{\mathbf{e}\}^\perp$ in accordance with footnote 26, shows that this region moves continuously around the simplex, and its closure always touches our closed curve in such a way that each point outside of the curve is part of this region at some point, implying that all these points outside are non-stabilizable.

It remains to be shown that every point on the boundary or enclosed within the boundary curve we just computed can in fact be stabilized. We will only give a hand-wavy explanation; again, each α yields a convex region which is not stabilizable, and which touches our curve in some point. Two cases may occur: Either one of the halfplanes on which some α_π is negative lies outside of the majorization polytope of d, in which case no point inside our curve is in this halfplane. Otherwise, we are in the case illustrated in Figure 4.1. Here the convex region of non-stabilisable points given by α, when intersected with the majorisation polytope of d, is a triangle with vertices given by two permutations of d and some

point on our curve. Since the curve is always concave, again no point enclosed by the curve lies in the triangle of non-stabilisable points. The case for arbitrary admissible generators B in the qutrit case is analogous, but the large number of parameters makes the formulas unwieldy. In higher dimensions, the idea of using the functionals α_π to determine non-stabilisable points still applies, but it is unclear how to analytically compute the resulting shapes of the stabilisable set.

Reachable States

Let us now turn towards the set of reachable states (or, more precisely, its closure) for some given initial state and any stochastic generator matrix $-B$ with unique fixed point $d \in \Delta^{n-1}$, $d > 0$. Let us use the notation $y \rightsquigarrow x$ to denote that $y \in \overline{\mathsf{reach}_{\mathsf{T}}(x)}$. Then $\rightsquigarrow$ is a preorder and so it induces an equivalence relation $\sim$ on which it becomes a partial order. In other words, $x \sim y$ if and only if $x \in \overline{\mathsf{reach}_{\mathsf{T}}(y)}$ and $y \in \overline{\mathsf{reach}_{\mathsf{T}}(x)}$ meaning there exists an approximately periodic solution through x and y. Note that up to a viability condition, the equivalence classes $[x]$ of this equivalence relation correspond to *control sets* as defined in [CK00, Def. 3.1.2], and the induced partial order corresponds to the *reachability order* [CK00, Def. 3.1.7].

First we observe that the maximally mixed state can always be reached:

Lemma 4.B.2. *For all $x \in \Delta^{n-1}$, the vectors d and $\frac{1}{n}\mathbf{e}$ are in $\overline{\mathsf{reach}_{\mathsf{T}}(x)}$.*

Proof. Since d is the unique fixed point of e^{-tB} for $t > 0$, and since it is attractive,[29] $d \rightsquigarrow x$ for all $x \in \Delta^{n-1}$. Similarly, consider $\hat{B} = \frac{1}{n!}\sum_{\pi \in S_n} \pi B \pi^{-1}$. Then $\hat{B}$ is invariant under permutations, which implies that $\hat{B}\mathbf{e} = 0$. Moreover $\frac{1}{n}\mathbf{e}$ is the unique fixed point in Δ^{n-1} since otherwise, by permutation symmetry there would be an open set of fixed points in Δ^{n-1}, and hence $\hat{B} \equiv 0$. This would imply that $B \equiv 0$ as one can check by considering the value of $\hat{B}$ at the vertices of Δ^{n-1}. As before, the fixed point $\frac{1}{n}\mathbf{e}$ of $\hat{B}$ is attractive. □

This lemma shows $d \sim \frac{1}{n}\mathbf{e}$, and that the equivalence class $[d] = [\frac{1}{n}\mathbf{e}]$ is an invariant control set as defined in [CK00, Def. 3.1.3].

Let us now, again, restrict to the three-dimensional case. It turns out that this equivalence class is the only one that contains more than a single point: the idea is that equivalence classes with at least two points lead to (approximately) periodic solutions which must enclose a stabilisable point. Since the set of non-stabilsable points is simply connected, when restricting to a Weyl chamber the periodic solution intersects the set of stabilisable states, which are all equivalent to $\frac{1}{3}\mathbf{e}$. We start with an abstract result about differential inclusions.

Lemma 4.B.3. *Let $x : S^1 \to \Delta^2$ be a smooth, periodic, injective solution of the differential inclusion (T) with non-vanishing derivative. Then the region enclosed by x contains a stabilisable point.*

Proof. This is a direct generalisation of [AC84, Ch. 5.2, Thm. 1] which states that if an upper semi-continuous differential inclusion with non-empty, closed, convex values is defined on a compact convex set and satisfies a viability condition, then it has a stabilisable point. By the Schoenfliess Theorem, see [Moi77, Ch. 9, Thm. 6], the interior region of $x(S^1)$ is homeomorphic to an open disk, and hence by the Riemann Mapping Theorem, there is even a biholomorphism. Now note that since x is an injective immersion and S^1 is compact, it is an embedding, and hence the image is a smooth curve. Thus, by [Bel90, Thm. 3.1], the Riemann mapping extends to a diffeomorphism of the closure of the interior

[29] This follows from a basic result on continuous-time Markov chains. Here $-B$ is the transition rate matrix. It is irreducible (in the sense of [Nor97, p. 111]) since $d > 0$ is the unique fixed point. Then [Nor97, Thm. 3.6.2] shows that the corresponding Markov chain is ergodic, i.e. the unique fixed point is attractive.

region to the closed disk. Finally we can pull back the differential inclusion to the disk and apply the aforementioned theorem to find a stabilisable point. □

The idea of the result we want to prove is that if two points are equivalent but distinct, then they must be equivalent to some stabilizable point. To prove this in general we need the following approximation result.

Lemma 4.B.4. *Let $x \neq y$ and $y \rightsquigarrow x$, and assume that y is not stabilizable. Let a solution $\tilde{y}$ starting at y be given. Then for every $\varepsilon > 0$ small enough we can modify the differential inclusion in the region $B_\varepsilon(y)$ without creating new stabilizable points and such that there is a smooth solution $\tilde{x}$ starting at x and ending at y such that the concatenation of $\tilde{x}$ and $\tilde{y}$ is smooth.*

Sketch of proof. Using translations and rotations we may assume that $y = (0, 0)$ and the cone generated by $\mathsf{derv}(y)$ is contained in the upper halfplane and symmetric about the vertical axis. By continuity and assuming that δ is small enough, there are inner and outer approximations of this cone in $B_\delta(y)$ which are both pointed. We may assume (e.g., by extending x backwards) that $x'(0)$ lies in the inner approximating cone. We will only modify the differential inclusion within the lower half of this disk. Now assume that for some small enough $0 < \varepsilon \ll 1$ we have a smooth solution $\tilde{x}$ starting at x that ends ε-close to y. Then by slightly enlarging the outer cone we may assume that $\tilde{x}$ enters the unit disk within the negative of the outer cone. One can see that it is possible to modify the differential inclusion inside $B_R \setminus B_r$ for some $0 < r < R < 1$ such that there is a smooth solution entering B_r inside of the inner approximating cone, while making sure that the cone always lies in the upper halfplane, so that no stabilisable points are created. □

Proposition 4.B.5. *If $x \neq y$ and $x \sim y$, then $x \sim \frac{1}{3}\mathbf{e} \sim d$.*

Proof. If x or y is stabilisable, then by Cor. 4.B.10 it is equivalent to d and we are done. Hence we assume that neither x nor y is stabilisable. Let $\varepsilon > 0$ small enough be given. Since $x \sim y$, we may apply Lemma 4.B.4 twice to obtain a smooth, periodic solution passing through x and y for a slightly modified differential inclusion, which does not introduce new stabilisable points. Without loss of generality we may assume that this solution is injective and has non-vanishing derivative. By Lemma 4.B.3 it encloses a stabilisable point. However, if we work in a Weyl chamber, the non-stabilisable set is simply-connected, and so the periodic solution intersects the stabilisable region in some point s. Hence there is a point ε-close to x which is reachable from s (and by Cor. 4.B.10 also from $\frac{1}{3}\mathbf{e}$). Letting ε go to 0 this shows that $x \rightsquigarrow \frac{1}{3}\mathbf{e}$. □

For any non-stabilizable state x, it holds that the convex cone generated by $\mathsf{derv}(x)$ is pointed (i.e. its edge is a point). Hence there are two extremal derivatives at the boundary of the cone, which we will call the left and right extremal derivatives, as seen from x. The resulting extremal vector fields are depicted in Figure 4.3. More precisely we have the following result.

Lemma 4.B.6. *On the set $\Delta^2 \setminus \mathrm{int}(\mathsf{stab}_\mathsf{T})$ there exist left and right extremal vector fields. The norm of these vector fields might not be continuous, but the direction field is locally Lipschitz continuous, except possibly at d (and its permutations).*

Proof. As already mentioned, for any non-stabilisable point x, the convex cone generated by $\mathsf{derv}(x)$ is pointed. On the other hand, if for some x the convex cone is the plane, then x is in the interior of the stabilisable set. Hence on the boundary of the stabilisable set, the convex cone is either pointed or a half space. Either way there is a well defined left and right extremal derivative, and so the corresponding vector fields are well-defined. Locally, for $x \neq \pi d$ the direction field can be seen as a maximum of finitely many smooth functions, and hence it is locally Lipschitz continuous. □

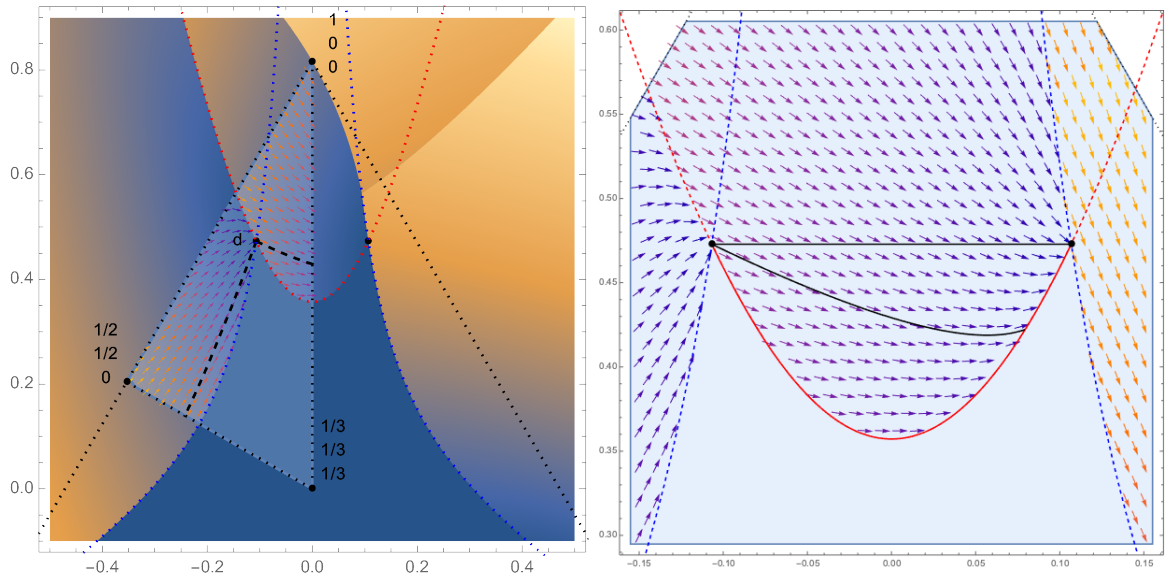

Figure 4.3: Left: The left extremal vector field in the case $a = 0.3$ depicted in the indicated Weyl chamber. The vector field is undefined on the stabilizable set, and hence this yields another way to compute the set of stabilizable points. We plotted again the bounding conics of the stabilizable set and the trajectories bounding the reachable set $\overline{\mathsf{reach}_{\mathsf{T}}(d)}$. In this case, the boundary trajectories are obtained by starting at d and permuting B such that one of the neighbours of d is the unique fixed point. The background colors show the norm of the left extremal vector field, and its discontinuities are clearly visible. Right: A zoomed-in picture of the "D"-shaped region considered in the proof of Lemma 4.B.8 again with parameter $a = 0.3$.

The discontinuities in the norm are important, as they tell us when the control permutation has to be applied. The shapes of these discontinuities are non-trivial, and we show an example in Figure 4.3.

Now the boundary of the reachable set can be computed using solutions following the left and right extremal derivatives. By the previous lemma these solutions exist and are unique. See again Figure 4.3 as well as Figure 4.4. Note that the extremal vector fields never vanish where they are defined, and since they are defined on a contractible domain (if restricted to a Weyl chamber) there are no periodic solutions. This relies on the fact that the state space is two dimensional, see for instance [Str15, Thm. 6.8.2]. Moreover, given a left (right) extremal solution, another solution can only cross it from left to right (right to left); this can be shown as in the proof of [Smi02, Thm. 5.6].

Before we can prove this section's main result, we need the following topological result about reachable sets:

Lemma 4.B.7. *Let any $x \in \Delta^2$ be given. Then $\overline{\mathsf{reach}_{\mathsf{T}}(x)}$ is contractible.*

Proof. Consider the map $F : \overline{\mathsf{reach}_{\mathsf{T}}(x)} \times [0,1] \to \overline{\mathsf{reach}_{\mathsf{T}}(x)}$ defined by

$$F(y,t) = \begin{cases} e^{-\hat{B}f(t)}y & \text{if } t < 1 \\ \frac{1}{3}\mathbf{e} & \text{else,} \end{cases}$$

where $\hat{B}$ is defined as in the proof of Lemma 4.B.2 and $f : [0,1) \to [0,\infty)$ is any homeomorphism. It follows from the same lemma that F is continuous, and hence a (strong) deformation retraction. □

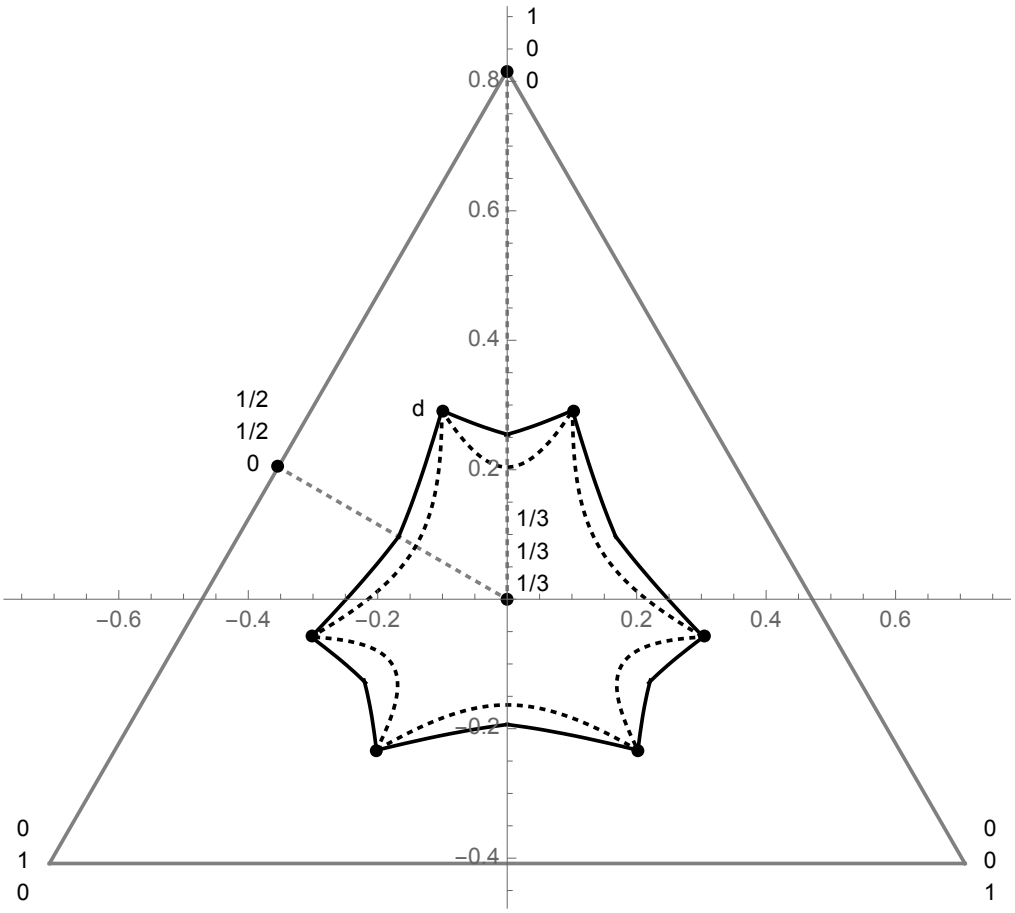

Figure 4.4: The boundary of the reachable set $\overline{\mathsf{reach}_{\mathsf{T}}(d)}$ in the case $a = 0.5$ is shown with solid lines, and the boundary of the stabilizable set $\mathsf{stab}_{\mathsf{T}}$ is shown using dotted lines. Note the curve segments of the boundary of the reachable set are not straight, though the curvatures are hardly visible.

Lemma 4.B.8. *Invoke Assumption (A) and further assume that $d \in \Delta^2_{\downarrow}$ where $\Delta^2_{\downarrow}$ denotes the ordered Weyl chamber of the simplex. The left extremal solution starting from d lies in the complement of the interior of the stabilizable region and terminates in the boundary of the Weyl chamber in finite time, without leaving the (classical) majorization polytope of d. The analogous result holds for the right extremal solution.*

Proof. From [3, Coro. 5] we know that the left extremal solution remains in the majorisation polytope of d. Let us sketch why this solution cannot enter the set of stabilisable points in $\Delta^2_{\downarrow}$. Consider the connected region containing d and $\tau_{23}d$ which is bounded by the majorisation polytope and the set of stabilisable points, and is shaped like a "D" lying on its belly, so let's call it D, see right panel of Figure 4.3.

First note that since there are no fixed points in D, every solution reaches the boundary of D in finite time, and by the above it reaches the curved part of the boundary of D. Now consider the straight part of the boundary, between d and $\tau_{23}d$. The solutions starting from points close to $\tau_{23}d$ will reach the curved boundary of D on the right side. Hence by continuity all points on the straight part of the boundary have solutions ending up on a connected part of the curved boundary. However, as we have seen before, on the boundary of the set of stabilisable points, the cone generated by the achievable

derivatives is a half plane, and hence the left and right extremal vector fields point in opposite directions. Therefore in general only one kind of solution can terminate in each point. By symmetry and the above connectedness, all left extremal solutions must terminate on the right side of D, and analogously for the right extremal solution. As noted above, no solution can leave the region delimited by the extremal solutions. □

Theorem 4.B.9 (Reachable Set Qutrit). *The left and right extremal solutions starting at d separate the Weyl chamber into two parts, and the inner part containing $\frac{1}{3}\mathbf{e}$ equals the intersection of $\overline{\mathsf{reach}_{\mathsf{T}}(d)} = [d]$ with $\Delta^2_{\downarrow}$.*

Proof. Taking the extremal solutions in the ordered Weyl chamber and all of the permuted copies yields a closed curve surrounding and contained in $\overline{\mathsf{reach}_{\mathsf{T}}(d)}$, i.e. they form the "outer boundary" of the reachable set. By Lemma 4.B.7 $\overline{\mathsf{reach}_{\mathsf{T}}(d)}$ is contractible, hence it is equal to the region enclosed by this curve. □

Corollary 4.B.10. *For every $x \in \mathsf{stab}_{\mathsf{T}}$ it holds that $x \sim d$.*

Proof. By Lemma 4.B.2 it holds that $d \rightsquigarrow x$. Theorem 4.B.9 shows that $\overline{\mathsf{reach}_{\mathsf{T}}(d)}$ is the set enclosed by the left and right extremal solutions (and their permuted copies) starting at d and ending in the boundary of the Weyl chamber. Moreover Lemma 4.B.8 shows that this set contains $\mathsf{stab}_{\mathsf{T}}$ and hence $x \rightsquigarrow d$. □

For starting points other than d, a similar result holds.

Corollary 4.B.11. *For any point x outside of $[d]$, we can compute the boundary of $\overline{\mathsf{reach}_{\mathsf{T}}(x)}$ in $\Delta^2_{\downarrow}$ by following the left and right extremal solutions until we hit either the boundary of the Weyl chamber or $[d]$. Moreover, the left extremal solution can only terminate in the right boundary of the Weyl chamber or in the left boundary of $[d]$ and vice-versa.*

Proof. Since $d > 0$, no solution tends to the boundary of the simplex. Hence the left and right extremal solutions must terminate in the boundary of the Weyl chamber or of $[d]$. The fact that the left extremal solution can only terminate in the right boundary of the Weyl chamber or in the left boundary of $[d]$ follows from the fact that integral curves do not intersect and the fact that on the symmetry lines of the simplex, the cone of achievable derivatives opens towards $\frac{1}{3}\mathbf{e}$. □

The more general case of B *not* satisfying Assumption (A) can be treated with similar methods. Note, however, that many of our arguments rely on the fact that the state space is two dimensional, and hence it is not clear how to analytically determine stabilisable and reachable sets in higher dimensions.

4.C Simultaneous Triangulation

Given a finite set of matrices of size $n \times n$ with entries in an algebraically closed field (such as the complex numbers), our goal is to find (i) common eigenvectors and (ii) a basis in which all matrices take on an (upper) triangular shape, or to conclude that no such eigenvectors or basis exist. This then allows us to determine if a system is coolable (Theorem 4.3.7) or stabilizable (Theorem 4.2.7).

Preliminaries

Many mathematical characterizations of the simultaneous triangulability of sets of matrices are collected in [RR00], and of course the famous theorems of Lie and Engel give further conditions [Kna02]. It is beneficial to take a more abstract point of view. The given matrices are simultaneously triangulable if and only if they admit a complete flag[30] of common invariant subspaces. An important result states that if a complete flag of common invariant subspaces exists, then every flag of common invariant subspaces is contained in a complete one, cf. [RR00, Lem. 1.5.2]. This is, in fact, a special case of the Jordan–Hölder Theorem, which holds for arbitrary modules and groups [HGK04, Isa09]. Consequently, in order to find a simultaneous triangulation, it suffices to repeatedly compute (proper, non-trivial) common invariant subspaces. In fact it suffices to be able to compute one-dimensional invariant subspaces, which of course correspond to (the span of) a common eigenvector. This will be the main focus of this appendix.

Surprisingly the literature on finding common eigenvectors is rather slim. Often only a "brute-force" algorithm is provided [Dub09, Al-+23], which however suffers from combinatorial explosion. A seminal result is given by Shemesh [She84] yielding a way to determine the existence of and compute a common eigenvector of two matrices. An extension to an arbitrary number of matrices is presented in [JP15], which is however rather inefficient. Our aim is to provide an efficient algorithm for computing a simultaneous eigenvector of an arbitrary family of matrices, and as a result we obtain an efficient algorithm to compute a simultaneous triangulation, (or to conclude that no such eigenvector or triangulation exists).

A Note on Computational Complexity

We formulate most results for an arbitrary field $\mathbb{K}$, when necessary restricting to algebraically closed fields. When it comes to determining the (worst-case) time-complexity of our algorithms, we assume that all field operations take constant time. When performing exact arithmetic, this is an accurate assumption for finite fields, but when working over the rationals $\mathbb{Q}$ this might break down. For infinite fields like the reals $\mathbb{R}$ or complex numbers $\mathbb{C}$, it is often sufficient to implement the algorithms numerically, i.e. using floating point arithmetic. In particular when determining eigenvalues, numerical algorithms are preferable, and for most linear algebra tasks numerical algorithms perform well and are stable [DDH07]. Again, it is accurate to assume that all field operations take constant time. Note also that for the multiplication of $n \times n$ matrices we use the standard algorithm with complexity $\mathcal{O}(n^3)$, although better complexities can be achieved.

Computing a Common Eigenvector

Consider a list $A_1, \dots, A_k$ of $n \times n$ matrices over an algebraically closed field $\mathbb{K}$. Our goal is to compute a simultaneous eigenvector of these matrices, if it exists. More precisely, we are looking for a vector $v \in \mathbb{K}^n$ such that there are scalars $\lambda_i \in \mathbb{K}$ satisfying $A_i v = \lambda_i v$ for all $i \in \{1, \dots, k\}$. Of course, if such numbers λ_i are known, it is easy to find v as the solution of a linear system of equations. However, even if we know all eigenvalues of all A_i, there are up to n^k combinations to try. This brute-force algorithm, given in [Dub09, Al-+23], becomes very inefficient when k is large. We are looking for an algorithm that has polynomial runtime in n and k.

[30] A *flag* in a finite-dimensional vector space is a nested sequence of strict subspaces, and it is *complete* if it contains a subspace of each dimension.

Algorithm for the Commuting Case

Let us first consider the special case of commuting matrices $A_1, \dots, A_k$, that is, $[A_i, A_j] := A_iA_j - A_jA_i = 0$ for all $i, j \in \{1, \dots, k\}$. To find a common eigenvector, we first compute the eigenspaces of A_1 and denote the smallest of them by S_1. Note that since $\mathbb{K}$ is algebraically closed, there always exists a non-trivial eigenspace. Moreover, as all A_i commute, S_1 is a common invariant subspace.[31] Hence all matrices A_i can be restricted to S_1, and the restrictions still commute. Now consider the restriction $A_2|_{S_1}$ and let S_2 be its smallest eigenspace. Then S_2 is again a common invariant subspace of all A_i. Continuing in this fashion we obtain a nested sequence $\mathbb{K}^n \supseteq S_1 \supseteq S_2 \supseteq \dots \supseteq S_k$ of common invariant subspaces. If some S_j is one dimensional, then any non-zero vector in this S_j is a common eigenvector of all A_i, otherwise any non-zero vector in S_k is a common eigenvector of all A_i.

Assuming $\mathbb{K} = \mathbb{C}$, an approximate numerical algorithm for this task can be implemented with time-complexity $\mathcal{O}(n^3)$, since the Schur decomposition can be performed in $\mathcal{O}(n^3)$ (e.g. using the QR algorithm, cf. [Fra62, Arb16]), and in each iteration the size of the subspace is at least halved (unless some restricted matrix is a multiple of the identity, but this can be checked quickly without affecting the runtime). Over general fields one would first find a root of the characteristic polynomial (possibly over a field extension if the field is not algebraically closed) and then compute the corresponding eigenspace. Such algorithms exist for many different kinds of fields, see for instance [Bos+21, Sec. 24.8.1].

With this special case solved, we can consider the general case of arbitrary matrices A_i. Here the idea is to find a common invariant subspace on which all A_i commute.

Lemma 4.C.1. *Let $A_1, \dots, A_k$ of be a list of $n \times n$ matrices over any field $\mathbb{K}$ and let*

$$T = \bigcap_{i,j=1}^{k} \ker[A_i, A_j].$$

Then it holds that

(i) every common eigenvector of all A_i lies in T, and

(ii) if the A_i span a Lie algebra[32], then T is a common invariant subspace of all A_i.

Proof. (i): Let v be a common eigenvector of all A_i with $A_iv = \lambda_i v$, then $[A_i, A_j]v = (\lambda_i\lambda_j - \lambda_j\lambda_i)v = 0$, and hence $v \in T$. (ii): Now assume that the A_i span a Lie algebra, meaning that there are constants $c_{ij}^l \in \mathbb{K}$ such that $[A_i, A_j] = \sum_{l=1}^k c_{ij}^l A_l$. It is clear that T is a subspace, hence it remains to show that it is invariant under all A_s. But this holds since for $v \in T$, the computation

$$[A_i, A_j]A_sv = \sum_{l=1}^{k} c_{ij}^l A_lA_sv = A_s \sum_{l=1}^{k} c_{ij}^l A_lv = A_s[A_i, A_j]v = 0$$

shows that $A_sv \in T$. This concludes the proof. □

The Lemma is similar to the well-known Shemesh criterion for the existence of a common eigenvector of two (complex) matrices, see [She84, Thm. 3.1], and in general it is more efficient than the extension provided in [JP15, Thm. 1.2], which has time-complexity exponential in k.

[31] Indeed, if $[A, B] = 0$ and $Av = \lambda v$, then $A(Bv) = BAv = \lambda(Bv)$.

[32] Concretely this means that every commutator can be written as a linear combination: $[A_i, A_j] = \sum_{l=1}^k c_{ij}^l A_l$.

Algorithm for the General Case

Using this result we can formulate an algorithm for the common eigenvector problem. We will assume that the matrices $A_1, \ldots, A_k$ are linearly independent, since otherwise we can simply remove the offending elements. Now extend the list $A_1, \ldots, A_k$ by appending matrices $A_{k+1}, \ldots, A_d$ such that the full list forms a basis of the d-dimensional Lie algebra generated by $A_1, \ldots, A_k$. Importantly removing linearly dependent elements does not change the generated Lie algebra, and the common eigenvectors of the matrices $A_1, \ldots, A_k$ are exactly the same as those of the enlarged set $A_1, \ldots, A_d$. Now compute the space $T = \bigcap_{i,j=1}^{d} \ker[A_i, A_j]$. Since, by part (ii) of the Lemma, T is invariant under all A_i with $i \in \{1, \ldots, d\}$, we can restrict them to T, and it holds that $[A_i|_T, A_j|_T] = 0$ for all $i, j \in \{1, \ldots, d\}$. Then using the algorithm for the commutative case, we can find a common eigenvector of all $A_i|_T$ for $i \in \{1, \ldots, k\}$ on T, and this will be the desired eigenvector. Note that this fails if and only if T is trivial, which happens only if there is no common eigenvector by part (i) of the Lemma.

Computing of order d^2 commutators takes time $\mathcal{O}(d^2 n^3)$. Similarly, the subspace T can be found by determining the kernel of the corresponding matrix of size $n\frac{d(d-1)}{2} \times n$, (here we used the anti-symmetry of the commutator) which can be done in the same time-complexity. The time-complexity of generating the Lie algebra is given in App. 4.C, and that of finding a common eigenvector on T was given above (and doesn't influence the overall result). Thus, the total complexity is $\mathcal{O}(dn^3(kn + d))$, which in terms of n can be bounded by $\mathcal{O}(n^8)$ since $k, d \leq n^2$.

Algorithm for Simultaneous Triangulation

As an application we will consider the problem of simultaneously triangulating an arbitrary set of $n \times n$ matrices $A_1, \ldots, A_k$. We consider the following naive algorithm. First compute a common eigenvector of all A_i, and call it v_1. Change to a basis whose first vector is v_1 and restrict to the $(n-1) \times (n-1)$ block in the lower right. Iterate the procedure until the matrices are in triangular form. It turns out that this algorithm actually works whenever the A_i are simultaneously triangulable. This follows from [RR00, Lem. 1.5.2] which states that if a collection of matrices is triangulable, then every flag of invariant subspaces is contained in a complete flag of invariant subspaces. More generally, this is a direct consequence of the Jordan–Hölder Theorem for modules, cf. [HGK04, Thm. 3.2.1]. See also [Dub09]. It is clear from the above that the total time-complexity of the algorithm lies in $\mathcal{O}(n^7(k+n))$.

Generating a Lie Algebra

A key step in our algorithm for computing a common eigenvector of a given set of matrices is that of determining a basis for the Lie algebra generated by the matrices. Algorithms for this task have been presented in [Ell09, BW79, Isi05, SFS01] in the context of control theory, but unfortunately no complexity analysis is provided. Here we give a simple meta-algorithm which includes the standard approach and we analyze its complexity.

In this section $\mathbb{K}$ denotes any field. If $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$, the algorithm can also be implemented numerically, i.e. using floating point arithmetic. However, care has to be taken here, since, for instance, two generic real matrices will generate all of $\mathbb{R}^{n,n}$, see [BW79, Sec. 1].

Meta-algorithm for Generating a Lie Algebra

Given a linearly independent set of matrices $\mathcal{A} = \{A_1, \ldots, A_k\}$ in $\mathbb{K}^{n,n}$, we denote the Lie algebra generated by $\mathcal{A}$ with $\langle \mathcal{A} \rangle_{\mathsf{Lie}}$ and its dimension with d. The meta-algorithm proceeds as follows:

Set $\mathcal{A}_1 = \mathcal{A}$ and $i = 1$. Pick two elements $B_i, B_i' \in \mathcal{A}_i$ and compute the commutator $[B_i, B_i']$. If the commutator does not lie in $\mathrm{span}(\mathcal{A}_i)$, then set $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{[B_i, B_i']\}$, otherwise set $\mathcal{A}_{i+1} = \mathcal{A}_i$.

Increment i and repeat. Terminate the algorithm when all commutators are guaranteed to lie in the span of $\mathcal{A}_i$.

To turn this into a concrete algorithm, a procedure for choosing B_i, B_i' and a termination criterion have to be provided. Below we will show how to check for linear independence such that that each iteration can be implemented with time-complexity $\mathcal{O}(n^4)$. The time-complexity of the entire algorithm is hence determined by the number of iterations. Clearly the algorithm needs at least $d - k$ iterations, but it might take much longer. Note that of course $d \leq n^2$.

The most naive algorithm would try all commutators in $\mathcal{A}_i$, and terminate if none of them yield a new dimension. This leads to a bound of $\mathcal{O}(d^3)$ on the number of iterations and a total time-complexity of $\mathcal{O}(d^3 n^4)$. By simply not computing commutators which have been computed before, and using the anti-symmetry of the commutator, one can bound the number of iterations by $\frac{d(d-1)}{2}$, and hence the total time-complexity is then $\mathcal{O}(d^2 n^4)$, and thus $\mathcal{O}(n^8)$ in term of n. Note that this recovers the complexity given in [ZS11], although the algorithm is slightly different. The algorithm presented in [Ell09], based on the realization that one only needs to commute new elements with the initial matrices in $\mathcal{A}$, yields a bound of $\mathcal{O}(dk)$ on the number of iterations and a total time-complexity of $\mathcal{O}(dkn^4)$. In terms of n alone we again get a bound of $\mathcal{O}(n^8)$.

Update Algorithm for Linear Independence

Consider a matrix $M \in \mathbb{K}^{r,s}$ with $s \leq r$. Determining if the columns of this matrix are linearly independent can be done using Gaussian elimination in time-complexity $\mathcal{O}(rs^2)$. In our Lie algebra generation algorithm above however we find ourselves in the following scenario. We obtain one vector in $\mathbb{K}^r$ after another. We keep the first non-zero vector and discard any zero vectors. Then for every following vector we keep it only if it is independent from all previously kept ones, otherwise we discard it. Using again Gaussian elimination, this can be implemented in $\mathcal{O}(r^2)$ for each vector. The total runtime for s vectors will be slightly worse with $\mathcal{O}(r^2 s)$, but now we can give a result after each vector.

The algorithm proceeds as follows: For the first non-zero vector perform Gaussian elimination to obtain the first standard basis vector and store the row operations performed in a matrix R_1. For every following vector, continue the Gaussian elimination, while updating also the matrix R_i storing the row operations. If we detect linear dependence, discard the vector and revert the matrix R_i to its previous state.

CHAPTER 5

Working the Qubit

5.1 Introduction

Controlling individual qubits is a fundamental task in quantum information technologies, especially in the presence of noise and decoherence. For instance, cooling qubits to their ground state is essential for quantum computation [DiV00]. In this chapter we exploit the reduced control system defined in Chapter 3, and results from Chapter 4, to derive time-optimal controls for a single qubit following a Markovian time evolution and subject to fast unitary control. Chapter 6 will then generalize these ideas to higher dimensional systems.

The special case of the Bloch equations (corresponding exactly to the Lindblad equation with rotational invariance around the z-axis) was studied in [Lap+10, Lap+11, Lap+13], introducing the so-called magic plane and steady state ellipsoid. A special case of the Lindblad equations with bounded controls was studied using geometric methods in [BS09, BCS09]. Another special case of the Bloch equations with incomplete control was addressed in [LPZ24]. The relaxation of certain unital channels and a special case of the Bloch equations was studied in [Muk+13] with bounded and unbounded controls. Some properties of unital systems (also beyond qubits) were found in [Yua10] using the reduced control system. The general Lindblad case was treated in [RBR12, RBR16, Cla+20] using mostly numerical methods.

Our approach uses a different method based on analyzing the generators of the reduced control system and yields a more comprehensive solution to the problem of finding time-optimal controls. The solution is almost completely analytical, except that as a final step one generally needs to use numerical integration, hinting at the fact that in the general case the solution does not admit an analytical expression. Moreover, our approach is very visual and geometric, giving an intuitive understanding of certain features of the obtained solutions. A drawback of our approach is, however, that it does not easily extend to the case of bounded controls.

Outline

Section 5.2 introduces the setting and recalls the reduced control system. Section 5.3 highlights the main tools and methods used in this chapter and contains some preliminary results. The main tasks addressed in this chapter, namely optimal control and stabilization, are introduced and treated in a general fashion in Section 5.4. The remainder of the main part focuses on deriving concrete solutions, starting with the completely general case in Section 5.5. The treatment of the general case leads to the definition of

so-called integral systems for which the solution simplifies considerably. These systems are addressed in Section 5.6, and they include important special cases such as real and coolable systems. The latter are considered in Section 5.6. The study of some special systems (namely unital systems and the Bloch equations) is relegated to Appendix 5.A. Finally, Appendix 5.B contains some technical computations.

5.2 Preliminaries

In this chapter we consider an open two-level quantum system, henceforth called a qubit. Again, the uncontrolled Markovian evolution is described by the (time-independent) Lindblad equation. Introducing fast unitary control to the system we obtain a bilinear control system. The assumptions on the controls lead to the reduced control system as defined in Chapter 3.

Bloch Ball and Lindblad Equation

We start by recalling the basic formalism in more detail. The set of all possible mixed states of a qubit is given by the set of density matrices of size 2×2, which are exactly the positive semi-definite matrices of trace one, denoted $\mathfrak{pos}_1(2) = \{\rho \in \mathbb{C}^{2,2} : \rho \geq 0,\ \mathrm{tr}(\rho) = 1\}$. Using the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which form an orthonormal basis of the set $\mathfrak{isu}(2)$ of traceless Hermitian matrices of size 2×2 with respect to the (rescaled) Hilbert-Schmidt inner product $\langle A, B\rangle := \mathrm{tr}(A^*B)/2$, we obtain the affine-linear isometry $B(1/2) \to \mathfrak{pos}_1(2)$ defined by $(x, y, z) \mapsto \frac{1}{2}\mathbb{1} + x\sigma_x + y\sigma_y + z\sigma_z$, where $B(1/2) \subseteq \mathbb{R}^3$ denotes the ball of radius $1/2$, called the *Bloch ball*,[1] cf. [BŻ17, Section 5.2]. Its surface, representing the pure states, is called the *Bloch sphere*

The special unitary transformations $U \in \mathrm{SU}(2)$ act on the density matrices by conjugation $\rho \mapsto U\rho U^*$. Note that the kernel of this action is $\{\mathbb{1}, -\mathbb{1}\}$. In the Bloch ball picture, these transformations are rotations, belonging to the special orthogonal group $\mathrm{SO}(3)$. The $\mathrm{SU}(2)$ orbits in $\mathfrak{pos}_1(2)$ are exactly the sets of density matrices sharing the same eigenvalues, and they correspond to the concentric spheres of the Bloch ball. This illustrates the fact that the radius of a point in the Bloch ball only depends on the eigenvalues $\lambda, 1-\lambda \in [0, 1]$ of the corresponding density matrix, and in fact determines the eigenvalues up to their order. Indeed, if $\lambda \geq 1/2$ is the larger eigenvalue, the radius is given by $r = \lambda - 1/2 \in [0, 1/2]$. Conversely, given radius r, the eigenvalues are $1/2 \pm r$. We see that a state is pure, meaning $\rho = |\psi\rangle\langle\psi|$, if and only if it corresponds to a point on the surface of the Bloch ball, and the interior of the Bloch ball consists of all mixed states. The center corresponds to the maximally mixed state $\mathbb{1}/2$.

To make this correspondence more precise we consider the isometric embedding $\iota : [0, 1] \to \mathfrak{pos}_1(2)$ given by $\lambda \mapsto \frac{\mathbb{1}}{2} + \left(\lambda - \frac{1}{2}\right)\sigma_z$ which maps $[0, 1]$ to the subset of $\mathfrak{pos}_1(2)$ consisting of diagonal density matrices corresponding exactly to the z-axis of the Bloch ball. Note that if we endow $[0, 1]$ with the metric induced by the absolute value, then ι is isometric since $\sqrt{\mathrm{tr}((\iota(\lambda) - \iota(\lambda'))^2)/2} = |\lambda - \lambda'|$. A nice property of the set $\iota([0, 1])$ is that it intersects all orbits orthogonally and the unitaries which leave $\iota([0, 1])$ invariant act on it either trivially or by reflection about the origin. In $[0, 1]$ this corresponds to the reflection $\lambda \mapsto 1 - \lambda$. To get rid of this final ambiguity one may work on the halved interval $[1/2, 1]$. Note that there is nothing special about the z-axis, except that it corresponds to the diagonal matrices in the standard basis. Any other axis would work for our purposes, since in fact all such axes are related by rotations of the Bloch ball. When defining the reduced control system in the next section we will use $[0, 1]$ as the reduced state space.

[1] Sometimes the Bloch ball is defined such that is has radius 1.

Full and Reduced Control Systems

We recall the definition of the full control system ($\mathcal{D}$), now specialized to a qubit. Let I be an interval of the form $[0, T]$ or $[0, \infty)$. We say that $\rho : I \to \mathfrak{pos}_1(2)$ is a solution to the bilinear control system

$$\dot{\rho}(t) = -\Big(\mathrm{i}\sum_{j=1}^{m} u_j(t)[H_j, \rho(t)] + L(\rho(t))\Big), \quad \rho(0) = \rho_0 \in \mathfrak{pos}_1(2), \tag{$\mathcal{B}$}$$

if it is absolutely continuous and satisfies the equation almost everywhere. Here $-L \in \mathfrak{w}_{\mathsf{KL}}(2)$ denotes the *drift Lindblad generator* (with *Lindblad terms* V_k and including a possible Hamiltonian part H_0) describing the uncontrolled evolution of the system, the $H_j \in \mathfrak{iu}(2)$ for $j = 1, \ldots, m$ denote the *control Hamiltonians*, and the functions $u_j : I \to \mathbb{R}$ are the *control functions*. Throughout we make the following two crucial assumptions: First we only require the control functions to be locally integrable, meaning that we do not assume any bounds. Second we assume that the control Hamiltonians generate the full special unitary Lie algebra $\langle \mathrm{i}H_j : j = 1, \ldots, m\rangle_{\mathsf{Lie}} \supseteq \mathfrak{su}(2)$. Taken together this means that we have fast unitary control over the system.

Under these assumptions, an equivalent reduced control system can be defined, cf. Chapter 3, by focussing on the evolution of the eigenvalues of ρ. More precisely, the reduced state will be $\lambda \in [0, 1]$. First we define the matrices $J_{ij}(U) = \sum_{k=1}^{r} |(U^* V_k U)_{ij}|^2$. For each $U \in \mathrm{SU}(2)$, we obtain an *induced vector field* on $[0, 1]$ defined by

$$\lambda \mapsto -Q_U(\lambda) = J_{12}(U) - \lambda(J_{12}(U) + J_{21}(U)). \tag{5.1}$$

Concretely each $-Q_U$ is an affine linear function on $[0, 1]$. This allows us to define the set-valued function

$$\mathsf{derv} : [0, 1] \to \mathcal{P}(\mathbb{R}), \qquad \mathsf{derv}(\lambda) = \{-Q_U(\lambda) : U \in \mathrm{SU}(2)\},$$

of *achievable derivatives* (where $\mathcal{P}(\cdot)$ denotes the power set). For an example of derv see Figure 5.1. Then the *reduced control system* on $[0, 1]$ is defined by the differential inclusion[2]

$$\dot{\lambda}(t) \in \mathsf{derv}(\lambda(t)), \quad \lambda(0) = \lambda_0 \in [0, 1]. \tag{Δ}$$

The Equivalence Theorem 3.0.4 shows that under the present assumption of fast unitary control, the bilinear control system ($\mathcal{B}$) is equivalent to the reduced control system (Δ) in a precise sense. Importantly, no loss of information is incurred by switching to the reduced control system.

Remark 5.2.1. *In Chapter 3 the reduced control system is defined on the standard simplex Δ^{n-1} in $\mathbb{R}^n$. Using the embedding $[0, 1] \to \Delta^1 \subset \mathbb{R}^2$, given by $\lambda \mapsto (\lambda, 1 - \lambda)^\top$ we pulled back the control system to the interval $[0, 1]$. This turns the (stochastic) linear dynamics on $\mathbb{R}^2$ into affine linear dynamics on $[0, 1]$.*

5.3 Space of Generators and Optimal Derivatives

The induced vector fields $-Q_U$ generate the dynamics of the reduced control system (Δ). Due to the present low-dimensional setting, it turns out that these generators sit in a two-dimensional vector space

[2] There are some slightly different ways to define the reduced control system, see Chapter 3, but the distinction is not relevant for us.

and hence they can easily be visualized. Understanding the exact shape of the set of generators is non-trivial, but it can be done analytically, and this ultimately leads to solutions for the optimal control problem of ($\mathcal{B}$). Furthermore it yields a parametrization of the stabilizable states in the Bloch ball.

Since each $\mathsf{derv}(\lambda)$ is the image of a continuous function on the compact connected set $\mathrm{SU}(2)$, the set $\mathsf{derv}(\lambda)$ must be a closed bounded interval in $\mathbb{R}$. Hence we can define the *optimal derivative function*

$$\mu : [0,1] \to \mathbb{R}, \quad \lambda \mapsto \max \mathsf{derv}(\lambda).$$

To fully understand the graph associated to derv (when seen as a set-valued function) it suffices to study the function μ, since $\lambda \mapsto -\mu(1-\lambda)$ is the corresponding lower boundary due to the reflection symmetry on $[0,1]$ shown in Lemma 5.3.2 below. The optimal derivative function μ enjoys some nice properties:

Lemma 5.3.1. *The function $\mu : [0,1] \to \mathbb{R}$ is continuous, convex and non-increasing. Furthermore, $\mu(1/2)$ is equal to the larger (non-negative) eigenvalue of $\sum_{k=1}^{r}[V_k, V_k^*]/2$. In particular, μ is non-negative on $[0,1/2]$ and non-positive at 1.*

Proof. By definition, the function μ can be seen as the pointwise maximum of the decreasing affine linear functions $-Q_U$ parametrized by $U \in \mathrm{SU}(2)$. From this it follows that μ is continuous, convex, and non-increasing. The last fact follows from $\mu(1/2) = \max_U \frac{1}{2}(J_{12}(U) - J_{21}(U)) = \max_U \frac{1}{2}((J(U) - J(U)^T)\mathbf{e})_1 = \max_U \frac{1}{2}\sum_{k=1}^{r}(U^*[V_k, V_k^*]U)_{11}$, where the maximization is over $\mathrm{SU}(2)$ and where $\mathbf{e} = (1,1)/2$ and we used Lemma 3.0.7. □

Each induced vector field $-Q_U$ is defined by the values taken at the endpoints, namely $J_{12}(U) \geq 0$ at $\lambda = 0$ and $-J_{21}(U) \leq 0$ at $\lambda = 1$. This motivates the definition of the *space of generators* as

$$\mathfrak{Q} = \{(J_{12}(U) - J_{21}(U), J_{12}(U) + J_{21}(U)) : U \in \mathrm{SU}(2)\} \subset \mathbb{R}^2.$$

This set is clearly linearly isomorphic to the set $\{-Q_U : U \in \mathrm{SU}(2)\}$ but has the advantage of being easy to visualize. For an example see Figure 5.1. Understanding the space of generators $\mathfrak{Q}$, and more specifically its boundary, allows us to describe the function μ, which in turn allows us to find solutions to the optimal control problem.

Lemma 5.3.2. *The set $\mathfrak{Q}$ is compact, path-connected, and satisfies $b \geq |a|$ for all $(a,b) \in \mathfrak{Q}$. Moreover, $\mathfrak{Q}$ is symmetric with respect to the reflection $(a,b) \mapsto (-a,b)$.*

Proof. The set $\mathfrak{Q}$ is the image of a continuous function on $\mathrm{SU}(2)$, and thus compact and path-connected. By definition $J_{12}(U), J_{21}(U) \geq 0$, and so $J_{12}(U) + J_{21}(U) \geq |J_{12}(U) - J_{21}(U)|$. For any U it holds that $J_{12}(U\sigma_x) = J_{21}(U)$ and $J_{21}(U\sigma_x) = J_{12}(U)$, proving the symmetry. □

When $(0,0) \in \mathfrak{Q}$ the system is of a special type (called unital stabilizable) which we will explore in Appendix 5.A. In the main part we mostly focus on the case $(0,0) \notin \mathfrak{Q}$. A first relation between the objects $\mathfrak{Q}$ and μ is given by the following result:

Lemma 5.3.3. *If $(0,0) \notin \mathfrak{Q}$ then μ is strictly decreasing.*

Proof. By contraposition, if μ is not strictly decreasing, there are two points, $\lambda_1 < \lambda_2$ in $[0,1]$ such that $\mu(\lambda_1) = \mu(\lambda_2)$ and μ will be constant on $[\lambda_1, \lambda_2]$. Then there must be some horizontal line of the form (5.1) passing through $(\lambda, \mu(\lambda))$ for $\lambda \in (\lambda_1, \lambda_2)$. Let $U \in \mathrm{SU}(2)$ be a corresponding unitary. Then the slope of the line is $-(J_{12}(U) + J_{21}(U))$ and equals 0, and so $J_{12}(U) = J_{21}(U) = 0$, as desired. □

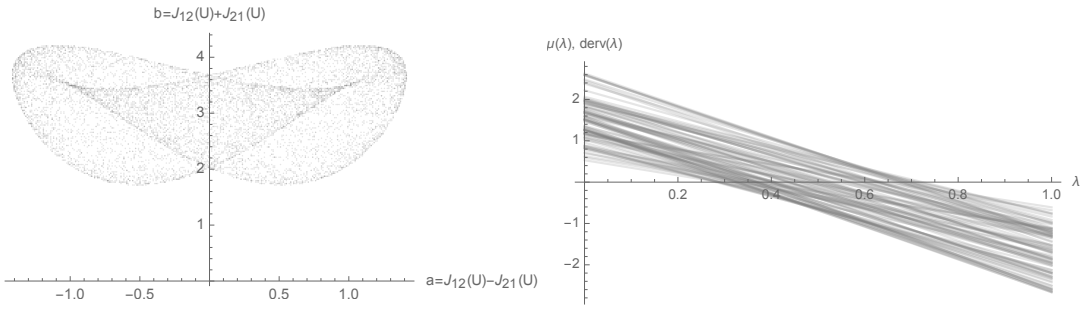

Figure 5.1: We consider a generic Lindblad generator $-L \in \mathfrak{w}_{\mathsf{KL}}$. Left: The space of generators $\mathfrak{Q}$ plotted using randomly sampled points. Clearly $\mathfrak{Q}$ has some intriguing structure, and hence it is the first object we wish to understand. Right: Graph of the set-valued function derv generated by plotting the affine linear functions $-Q_U$ associated to randomly selected unitaries U. The functions $\mu(\lambda)$ and $-\mu(1-\lambda)$ are the upper and lower boundaries of the set of all such lines. The Lindblad terms V_k are $\left(\begin{smallmatrix} 0.0-0.9\mathrm{i} & -0.6+0.6\mathrm{i} \\ 0.+0.8\mathrm{i} & 0.9-\mathrm{i} \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -0.1+0.8\mathrm{i} & 0.3-0.3\mathrm{i} \\ -0.8+0.6\mathrm{i} & 0.3+0.1\mathrm{i} \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} -0.2+0.4\mathrm{i} & 0.6-0.2\mathrm{i} \\ 0.6-0.7\mathrm{i} & 0.2+0.8\mathrm{i} \end{smallmatrix}\right)$.

The space of generators $\mathfrak{Q}$ and the optimal derivative function μ are linked by the following key result:

Proposition 5.3.4. *Let $\lambda \in [0,1]$. Then there is some line of the form (5.1) passing through $(\lambda, \mu(\lambda))$, and a point in $\mathfrak{Q}$ corresponds to such a line if and only if it solves the linear optimization problem*

$$\mu(\lambda) = \max_{(a,b)\in\mathfrak{Q}} \tfrac{1}{2}(a + (1-2\lambda)b).$$

Proof. The optimization problem follows immediately from the definition of μ. The existence of a solution follows from the fact that $\mathfrak{Q}$ is compact. □

Remark 5.3.5. *Note that the relation between the function $\mu : [0,1] \to \mathbb{R}$ and the set $\mathfrak{Q} \subseteq \mathbb{R}^2$ is similar to the Legendre–Fenchel transform of a function and the polar dual of a polytope. In particular, corners of $\mathfrak{Q}$ yield affine linear parts of μ as can be seen for instance in Appendix 5.A. Note also that $\mathfrak{Q}$ need not be convex, as can be seen in Figure 5.1, and μ only depends on the convex hull of $\mathfrak{Q}$.*

The space of generators $\mathfrak{Q}$ fully describes the reduced control system (Δ), and thus allows us to compute for instance reachable and stabilizable sets. In order to compute the fastest path in the Bloch ball, along with the optimal controls, as well as the stabilizable states in the Bloch ball, we need to parametrize $\mathfrak{Q}$ in terms of the corresponding unitaries. More precisely, it suffices to consider a subset of $\mathrm{SU}(2)$ which can map the north pole of the Bloch sphere to any other point on the Bloch sphere. We will give this parametrization in full generality in Section 5.5, and show how it simplifies in special cases of interest. Using the unitaries $U_{x,z} = \exp(\mathrm{i}\pi z\sigma_z)\exp(\mathrm{i}\pi x\sigma_x)$, the parametrization will be of the form

$$(x,z) \mapsto (J_{12}(U_{x,z}) - J_{21}(U_{x,z}), J_{12}(U_{x,z}) + J_{21}(U_{x,z})) \tag{5.2}$$

for $x \in [0,1/2]$ and $z \in [0,1)$. Hence $2\pi x$ corresponds to the polar angle and $2\pi z$ to the azimuthal angle on the Bloch sphere.

5.4 Stabilizable States and Optimal Control

The optimal derivative function $\mu : [0,1] \to \mathbb{R}$ discussed in the previous section gives the fastest increase (or slowest decrease if it is negative) of λ which can be achieved in the reduced control system.

It allows us to directly read off relevant information such as the reachable and stabilizable sets from its graph. Knowledge of the function μ also allows us to determine optimal controls for the full control system, as we will show in this section.

Stabilizable Set

A relevant task in applications is that of stabilizing the system in a certain desired state. We will explicitly determine which states are stabilizable in the full and in the reduced control system. Moreover we will show concretely how such states can be stabilized.

We begin with the reduced control system. For $\lambda \in [0,1]$ we say that λ is *stabilizable*[3] if $0 \in \mathsf{derv}(\lambda)$. Concretely this means that in the reduced control system, the constant path at λ is a solution.

Lemma 5.4.1. *Assume that $(0,0) \notin \mathfrak{Q}$. Then the set of stabilizable states is the non-empty closed interval $[1-\lambda^\star, \lambda^\star]$ where $\lambda^\star \geq 1/2$ is the unique root of μ. If $(0,0) \in \mathfrak{Q}$, then all states $\lambda \in [0,1]$ are stabilizable.*

Proof. The case of $(0,0) \in \mathfrak{Q}$ is clear, so assume that $(0,0) \notin \mathfrak{Q}$. By Lemma 5.3.3 it holds that μ has at most one root. By Lemma 5.3.1 there must be at least one root in the interval $[1/2, 1]$. Denote this root by $\lambda^\star$. The lower bound of derv is given by $\lambda \mapsto -\mu(1-\lambda)$ which has a unique root at $1-\lambda^\star$. Since both μ and the corresponding lower boundary are strictly decreasing, the stabilizable region is $[1-\lambda^\star, \lambda^\star]$. □

We call $\lambda^\star$ the *purest stabilizable state*. The stabilizable set can be obtained graphically from the space of generators as follows.

Lemma 5.4.2. *Assume that $(0,0) \notin \mathfrak{Q}$. It holds that $\lambda \in [0,1]$ is stabilizable if and only if $\lambda = \frac{1}{2}(1+\frac{a}{b})$ for some $(a,b) \in \mathfrak{Q}$, and hence $\lambda^\star = \max_{(a,b)\in\mathfrak{Q}} \frac{1}{2}(1+\frac{a}{b})$.*

Proof. Since $(0,0) \notin \mathfrak{Q}$ the value $\frac{a}{b}$ is well defined for every point $(a,b) \in \mathfrak{Q}$, and every line of the form (5.1) intersects the abscissa in a unique point, namely $\lambda = \frac{1}{2}(1+\frac{a}{b})$. By Lemma 5.3.2 the value $\frac{a}{b}$ is always contained in $[-1,1]$, and the set of all possible $\frac{a}{b}$ is a non-empty closed interval symmetric about 0. In particular, it has a greatest element. □

In practice for general systems, after parametrizing the set $\mathfrak{Q}$ as discussed in the previous section, it is necessary to use a root-finding algorithm (such as the bisection method) to find $\lambda^\star$, cf. [Epp13].

The following result presents a class of systems for which the purest stabilizable state $\lambda^\star$ can be found analytically.

Lemma 5.4.3. *Let V be an arbitrary Lindblad term and consider the system defined by the Lindblad terms V and $\sqrt{\gamma}V^*$ with $\gamma \in [0,1]$. Then the purest stabilizable state is $\lambda^\star = \frac{1}{1+\gamma}$, unless V is normal, in which case it is $\lambda^\star = 1$.*

Proof. By Lemma 5.A.2 (i), if V is normal, then the system is unital stabilizable, and so the stabilizable region is $[0,1]$. Now assume that V is not normal, then, by the same lemma, $(0,0) \notin \mathfrak{Q}$. First we consider the case $\gamma = 0$. Then, as shown in Section 6.3, the system is coolable, and again the stabilizable region is $[0,1]$. Now let $\gamma \in (0,1]$. Note that for any $U \in \mathrm{SU}(2)$, if we denote $t(U) = \frac{J_{21}(U)}{J_{12}(U)}$, the intersection of the corresponding line with the abscissa is $\frac{J_{12}(U)}{J_{12}(U)+J_{21}(U)} = \frac{1}{1+t(U)}$ by Lemma 5.4.2. Hence it suffices to find the minimal value of the ratio $t(U)$. If we denote by $J'(U)$ the matrix corresponding

[3] In Chapters 2 and 4 we distinguish between stabilizable and strongly stabilizable states, but in the present setting the two notions coincide.

to V then $t(U) = \frac{J_{21}(U)}{J_{12}(U)} = \frac{J'_{21}(U)+\gamma J'_{12}(U)}{J'_{12}(U)+\gamma J'_{21}(U)}$. Since all quantities are non-negative, and $\gamma \leq 1$, it follows from the mediant inequality that $\gamma \leq t(U) \leq \frac{1}{\gamma}$. Moreover $t(U) = \gamma$ is achieved when V is in upper triangular form, which is always possible. □

Note that when $\gamma = 1$ the system is unital (cf. Appendix 5.A), and if $\gamma = 0$ we obtain a rank one system, see Section 6.3.

In Lemma 5.4.2 we have described the stabilizable spectra in the reduced control system, that is, in $[0, 1]$. However we can also explicitly describe the set of stabilizable density matrices. In [Lap+13] this was done for the special case of the Bloch equations obtaining an ellipsoid, which we will also recover in Appendix 5.A.

A state $\rho \in \mathfrak{pos}_1(2)$ is called *stabilizable* if there exists a control Hamiltonian H_c turning ρ into a fixed point, that is, if $-L(\rho) - \mathrm{i}[H_c, \rho] = 0$. The following result connects the two notions of stability and is a restatement of Proposition 4.2.2.

Proposition 5.4.4. *Assume that $-L$ is a non-unital Lindblad generator and let $\lambda \neq \frac{1}{2}$. Then the following are equivalent:*

(i) There is some U such that $-Q_U(\lambda) = 0$, that is, λ is stabilizable for the reduced control system (Δ).

(ii) There is some U and H_c such that for $\rho = U\iota(\lambda)U^*$ we have $-L(\rho) - \mathrm{i}[H_c, \rho] = 0$, that is, ρ is stabilizable for the full control system ($\mathcal{B}$).

Moreover, any unitary satisfying one of the above also satisfies the other, and given λ and U one can compute a corresponding compensating Hamiltonian $\mathrm{i}H_c = \mathrm{ad}_\rho^+(L(\rho))$, *where* $\mathrm{ad}_\rho(\cdot) = [\rho, \cdot]$ *and* $(\cdot)^+$ *denotes the Moore–Penrose pseudoinverse.*

Remark 5.4.5. *Whenever $\lambda = 1/2$, it is clear that λ is stabilizable in the reduced system, but in general it is not (exactly) stabilizable in the full system. Indeed, in this case the formula for the compensating Hamiltonian might diverge. The only exception is the case of unital $-L$, where the maximally mixed state is always a fixed point independent of the applied controls.*

The set of stabilizable states in the Bloch ball can be nicely parametrized using a parametrization of $\mathfrak{Q}$. Assume that $(0, 0) \notin \mathfrak{Q}$ and let $(x, z) \mapsto F(x, z)$ be the parametrization of $\mathfrak{Q}$ as in (5.2). By construction, the line represented by $F(x, z)$ corresponds to the affine linear vector field obtained from $-L$ when restricting to and projecting onto the axis passing through the point with polar angle $\theta = 2\pi x$ and azimuthal angle $\phi = 2\pi z$. A stabilizable point on the axis is exactly a zero of this vector field. Together with Lemma 5.4.2 this shows the following result:

Lemma 5.4.6. *The stabilizable set can be parametrized as $r(\theta, \phi) = \frac{1}{2}\frac{J_{12}(U_{x,z})-J_{21}(U_{x,z})}{J_{12}(U_{x,z})+J_{21}(U_{x,z})}$ in spherical coordinates (r, θ, ϕ), where $2\pi x = \theta$ and $2\pi z = \phi$.*

In the Bloch ball, the shape of the set of stabilizable states is some kind of ovoid, and in some special cases it is an ellipsoid. Moreover, the intersection of this set with any plane containing the z-axis (after an appropriate change of basis) is in fact an ellipse, cf. Proposition 5.5.3.

Optimal Controls

The main task of interest is to determine which states are reachable from a given initial state, and to compute the fastest path to reach such a state, together with the corresponding controls. The Equivalence Theorem 3.0.4, allows us to work on the level of the reduced control system (Δ) and then to lift the obtained result to the full control system ($\mathcal{B}$).

Indeed, since the reduced state space $[0,1]$ is one-dimensional, the shortest path is always uniquely defined up to the speed at which the path is traversed. Clearly the maximal (positive) velocity a solution can achieve at λ is the optimal derivative $\mu(\lambda)$, and hence the optimal solution $\lambda(t)$ is obtained by integrating along μ. Since there are only two directions in the one-dimensional reduced state space, there are only two tasks to consider, namely optimal heating (mixing) and optimal cooling (purifying). These problems can be solved together by determining the fastest path from 0 to the purest stabilizable state $\lambda^\star$ in the reduced state space $[0,1]$.

In this section we outline the general approach we will take to solve this problem, and the concrete results will be presented in the subsequent sections. As stated at the end of Section 5.3, we will parametrize the space of generators $\mathfrak{Q}$ using the unitaries $U_{x,z} = \exp(\mathrm{i}\pi z\sigma_z)\exp(\mathrm{i}\pi x\sigma_x)$ where the parameter ranges are $x \in [0, \frac{1}{2}]$ and $z \in [0,1)$. Then, due to Proposition 5.3.4 we are mainly interested in the boundary of $\mathfrak{Q}$. In general we will parametrize the boundary using the curves

$$J^+ \to \mathbb{R}^2,\ \alpha \mapsto (a^+(\alpha), b^+(\alpha)) \qquad J^- \to \mathbb{R}^2,\ \alpha \mapsto (a^-(\alpha), b^-(\alpha))$$

defined on some intervals J^+ and J^- for the upper (corresponding to heating) and lower (corresponding to cooling) part of the boundary respectively. It is important to also determine the unitaries, or more precisely the values of the parameters x, z, which achieve the boundary points of $\mathfrak{Q}$. For this we will determine parametrizations $x^\pm : J^\pm \to \mathbb{R}^2$ and $z^\pm : J^\pm \to \mathbb{R}^2$.

Lemma 5.4.7. *Given parametrizations $a^\pm, b^\pm, x^\pm, z^\pm$ as above, and assuming that $(a^\pm, b^\pm)$ is differentiable with non-zero derivative, one can find parametrizations $\lambda^\pm(\alpha)$, $\mu^\pm(\alpha)$ and $\rho^\pm(\alpha)$ such that $(\lambda^\pm, \mu^\pm)$ parametrizes the graph of $\mu : [0,1] \to \mathbb{R}$ and such that $\rho^\pm(\alpha)$ parametrizes the optimal path through the Bloch ball. Indeed we have the following expressions:*

$$\lambda^\pm(\alpha) = \tfrac{1}{2}\big(1 + \tfrac{a'(\alpha)}{b'(\alpha)}\big), \quad \mu^\pm(\alpha) = \tfrac{1}{2}(a^\pm(\alpha) + b^\pm(\alpha)) - b^\pm(\alpha)\lambda^\pm(\alpha),$$
$$\rho^\pm(\alpha) = U_{x^\pm(\alpha), z^\pm(\alpha)}\, \iota(\lambda^\pm(\alpha)) U^*_{x^\pm(\alpha), z^\pm(\alpha)}.$$

Proof. By Proposition 5.3.4 the value $\mu(\lambda)$ is obtained by maximizing a linear functional over $\mathfrak{Q}$. This maximum must be achieved on the boundary and hence there is some α such that $(a^\pm(\alpha), b^\pm(\alpha))$ achieves the maximum. In this case it holds that the derivative of the parametrization is orthogonal to the direction of maximization, which immediately yields the expression for $\lambda^\pm(\alpha)$ and the expression for $\mu^\pm(\alpha)$ follows at once. Then, by construction $\rho^\pm(\alpha)$ achieves the maximal eigenvalue derivative. □

The parametrizations of the previous lemma are completely analytic and allow to solve a significant part of the general problem. The final goal is to determine the corresponding control functions for the full control system ($\mathcal{B}$). Since these are functions of time all of the above quantities must also be expressed as a function of time. To find the time parametrization $\alpha(t)$ one has to solve the following ordinary differential equation, where we omit the $\pm$ superscript for readability.

$$\alpha' = \frac{\mu(\alpha)}{\lambda'(\alpha)} = \frac{a(\alpha)b'(\alpha)^2 - a'(\alpha)b'(\alpha)b(\alpha)}{b'(\alpha)a''(\alpha) - a'(\alpha)b''(\alpha)}. \tag{5.3}$$

Unfortunately it seems that in general the real time parametrization $\alpha(t)$ cannot be found analytically since the differential equation (5.3) is too complicated. Notable exceptions to this are however the Bloch

equations (cf. Appendix 5.A) and rank one systems (cf. Section 6.3). For this reason numerical methods are indispensible. Indeed, for computational efficiency it can be beneficial to switch to a numerical representation of the functions right away instead of working with the analytical expressions which tend to become extremely convoluted.

Finally, the corresponding control Hamiltonian can be found using the following result which is a simplified special case of Proposition 2.3.10, but it can also easily be verified via direct computation. See also [RBR12, Lem. 3.1].

Proposition 5.4.8. *Let $\lambda : [0,T] \to [0,1]$ be a solution to (Δ) not passing through $\frac{1}{2}$ with continuously differentiable control function $U : [0,T] \to \mathrm{SU}(2)$. Define $\rho(t) = U(t)\iota(\lambda(t))U^*(t)$. Then $\mathrm{i}H(t) = -U'(t)U^*(t) + \mathrm{ad}^+_{\rho(t)}(L(\rho(t))$ is a path of skew-Hermitian matrices satisfying $\dot\rho = -(\mathrm{i}\,\mathrm{ad}_H + L)(\rho)$.*

Note that the second term in the definition is analogous to the definition of the compensating Hamiltonian H_c in Proposition 5.4.4.

For our optimal control task, the solution will cross the maximally mixed state $\{\mathbb{1}/2\}$ at one point in time. The problem that occurs in trying to apply Proposition 5.4.8 is that the compensation term might blow up. Luckily, as we will see in the following sections, this does not happen and the control functions tend to be very well behaved. Nonetheless, diverging controls can and do occur in certain special cases, such as the Bloch equations and rank one systems mentioned above. In these cases the optimal solution is not differentiable at one point, where it takes a sharp turn, and the direct term diverges. Still one can cut off this divergence at the price of an arbitrarily small error.

5.5 General Systems

The goal of this section is to implement the program set out in the previous sections. We consider a qubit system described by an arbitrary finite set of Lindblad terms V_k for $k = 1, \dots, r$. The main result is an analytical parametrization of the space of generators $\mathfrak{Q}$, see Figure 5.2 for an example. As a consequence we can determine the stabilizable states in the Bloch ball and the optimal controls in the original control system, cf. Figure 5.3. We will assume that $-L$ is not unital, since this special case is considered in Appendix 5.A.

Parametrization

Let $\{V_k\}_{k=1}^r$ be a finite set of Lindblad terms. Without loss of generality we assume that all V_k are traceless and that $\sum_{k=1}^r [V_k, V_k^*]$ is diagonal. We define the characteristic values

$$\Delta = J_{12}(\mathbb{1}) - J_{21}(\mathbb{1}), \quad \Sigma = J_{12}(\mathbb{1}) + J_{21}(\mathbb{1}), \quad \delta = 2J_{11}(\mathbb{1}) - \Sigma/2.$$

Furthermore we set

$$r_1 e^{\mathrm{i}\phi_1} = \mathrm{i}\sum_{k=1}^r \overline{(V_k)_{12}}(V_k)_{21}, \quad r_2 e^{\mathrm{i}\phi_2} = 4\sum_{k=1}^r \overline{(V_k)_{11}}(V_k)_{21}, \quad \phi = \phi_1 - 2\phi_2 + \pi/2.$$

By choosing the basis appropriately we can always make sure that additionally $\Delta \geq 0$, and in the following we will always assume that this is the case.[4] The first crucial property is that these values are actually well-defined.

[4] Since we assume that the system is non-unital it even holds that $\Delta > 0$.

Lemma 5.5.1. *Assume that $-L \in \mathfrak{w}_{\mathsf{KL}}$ is non-unital. Then the values $|\Delta|$, Σ, δ, r_1, and r_2 are well-defined, and if r_1 and r_2 are non-zero, then ϕ is also well-defined modulo 2π. More precisely this means that these values only depend on the generator $-L$, and not on the choice of V_k or on the basis which diagonalizes $\sum_{k=1}^{r}[V_k, V_k^*]$.*

Proof. We already assumed that the V_k are traceless. This can always be done at the expense of a constant compensating Hamiltonian, hence we only have to show that the constants are invariant under unitary reshuffling, cf. Lemma 4.A.3, or a change of basis keeping $\sum_{k=1}^{r}[V_k, V_k^*]$ diagonal. It is easy to see that all quantities are invariant under unitary reshuffling, and the only allowed unitary basis transformations are those induced by diagonal unitaries, which can change ϕ_1 and ϕ_2 but not ϕ. □

Using these characteristic values we can now give a parametrization of the space of generators $\mathfrak{Q}$:

Proposition 5.5.2. *The unitary $U_{x,z} = \exp(\mathrm{i}\pi z\sigma_z)\exp(\mathrm{i}\pi x\sigma_x)$ for $x \in [0, 1/2]$, and $z \in [0, 1)$ yields the point $(J_{12}(U_{x,z}) - J_{21}(U_{x,z}), J_{12}(U_{x,z}) + J_{21}(U_{x,z})) = (\Delta\cos(2\pi x), F(x, z))$, where*

$$F(x,z) = \Sigma + (\delta + r_1\sin(4\pi z + \phi_1))\sin(2\pi x)^2 - r_2\cos(2\pi x)\sin(2\pi x)\sin(2\pi z + \phi_2)\,,$$

and hence we obtain the following parametrization of the space of generators

$$\mathfrak{Q} = \{(\Delta\cos(2\pi x),\ F(x,z)) : x \in [0,1/2],\ z \in [0,1)\}\,.$$

Proof. The details of this elementary but lengthy computation are given in Appendix 5.B. □

Note that for fixed z, the parametrization can be seen as the graph of a function. See Figure 5.2 for an illustration. The parametrization of $\mathfrak{Q}$ leads directly to a parametrization of the stabilizable states in the Bloch ball as shown in Lemma 5.4.6, see Figure 5.3 for an example.

Proposition 5.5.3. *The parametrization of the set of stabilizable states in the Bloch ball is given in spherical coordinates by*

$$r(\theta,\phi) = \frac{\Delta\cos(\theta)}{2F(\frac{\theta}{2\pi}, \frac{\phi}{2\pi})}.$$

For fixed angle ϕ this is an ellipse in the upper halfplane going through the origin.

Proof. The parametrization follows immediately from Lemma 5.4.6 and it is easy to see that for fixed angle ϕ the formula takes the form given in Lemma 5.B.4 (due to the definition of the polar angle θ all occurences of $\sin$ and $\cos$ in these formulas are swapped). Note also that the shape of the ellipse can then be computed using Lemma 5.B.5. □

Even though the stabilizable set looks somewhat like an ellipsoid, its shape is more complicated. Nevertheless, in some special cases, such as the Bloch equations treated in [Lap+13], it indeed reduces to an ellipsoid, cf. Appendix 5.A.

Optimal Solution

In order to find the optimal derivative function μ we first find the boundary of $\mathfrak{Q}$ by determining the maximal and minimal values of $z \mapsto F(x, z)$ for all $x \in [0, 1/2]$, cf. Proposition 5.3.4. We will use the simplified function

$$G_\phi(\xi,\zeta) = (1-\xi)\sin(4\pi\zeta + \phi - \pi/2) - \xi\sin(2\pi\zeta), \quad \xi,\zeta \in [0,1]. \tag{5.4}$$

See Appendix 5.B for the relevant properties of this function. In particular see Figure 5.10 for some plots for different ϕ.

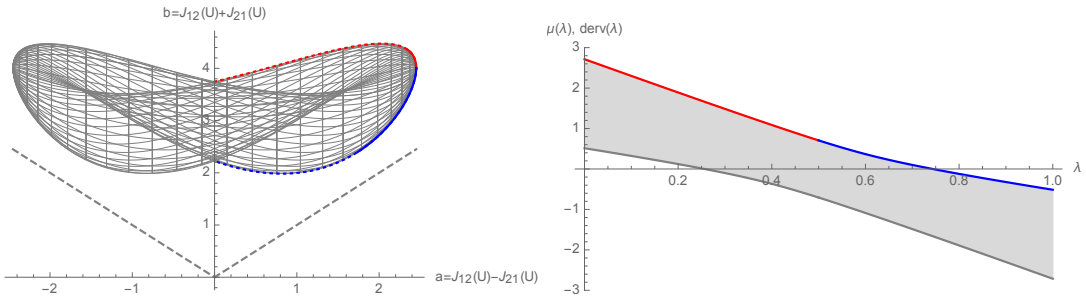

Figure 5.2: Left: Space of generators $\mathfrak{Q}$ parametrized as in Proposition 5.5.2. The boundary, as determined in Corollary 5.5.7, is highlighted in red (upper part) and blue (lower part), with the relevant part solid and the rest dotted. Right: Differential inclusion $\lambda \mapsto \mathsf{derv}(\lambda)$ with the optimal derivatives μ highlighted in red and blue. They are determined using Lemma 5.4.7. The Lindblad terms used are the same as in Figure 5.1.

Lemma 5.5.4. *Assume that $r_1, r_2 \neq 0$ and let $x \in (0, 1/4)$ be given.*[5] *Let $\xi = \frac{1}{1+\frac{r_1}{r_2}\tan(2\pi x)} \in (0,1)$. Then it holds for all $z \in \mathbb{R}$ that*

$$z \in \operatorname*{argmax}_{\tilde{z}} F(x,\tilde{z}) \iff \zeta \in \operatorname*{argmax}_{\tilde{\zeta}} G_\phi(\xi,\tilde{\zeta})$$

where $\zeta = z + \phi_2/(2\pi)$. The same statement holds after replacing argmax *by* argmin.

Proof. We compute

$$\begin{aligned}
\operatorname*{argmax}_{\tilde{z}\in[0,1]} F(x,\tilde{z}) &= \operatorname*{argmax}_{\tilde{z}\in[0,1]} r_1 \sin(2\pi x)\sin(4\pi\tilde{z}+\phi_1) - r_2\cos(2\pi x)\sin(2\pi\tilde{z}+\phi_2) \\
&= \operatorname*{argmax}_{\tilde{z}\in[0,1]} \sin(4\pi\tilde{z}+\phi_1) - \frac{r_2}{r_1}\cot(2\pi x)\sin(2\pi\tilde{z}+\phi_2) \\
&= \operatorname*{argmax}_{\tilde{z}\in[0,1]} \sin(4\pi\tilde{z}+\phi_1) - \frac{\xi}{1-\xi}\sin(2\pi\tilde{z}+\phi_2) \\
&= \operatorname*{argmax}_{\tilde{z}\in[0,1]} (1-\xi)\sin(4\pi\tilde{z}+\phi_1) - \xi\sin(2\pi\tilde{z}+\phi_2) \\
&= \operatorname*{argmax}_{\tilde{z}\in[0,1]} G_\phi(\xi,\tilde{z}+\phi_2/(2\pi)),
\end{aligned}$$

and the computation remains valid after replacing argmax by argmin. □

Remark 5.5.5. *If $r_1 = 0$ or $r_2 = 0$ or both, then the values of z which maximize or minimize $F(x,z)$ can be chosen independently of x. We will call such systems* degenerate. *Note that for the Bloch equations (Appendix 5.A) it holds that $r_1 = r_2 = 0$ and for rank one systems (cf. [6, Sec. V]) it holds that $r_2 = 0$. For the sake of brevity we will not treat degenerate systems in detail, and leave this as an exercise to the reader.*

Corollary 5.5.6. *Assume that $\phi \in (-\pi, 0) \cup (0, \pi)$ and define the function*

$$x^\star(z) = \frac{1}{2\pi}\operatorname{arccot}\left(2\,\frac{r_1}{r_2}\,\frac{\cos(4\pi z+\phi_1)}{\cos(2\pi z+\phi_2)}\right),$$

[5] By symmetry it suffices to consider the right side of $\mathfrak{Q}$.

and the intervals

$$I^+ = \begin{cases} [\frac{3}{4} - \frac{\phi}{4\pi} - \frac{\phi_2}{2\pi}, \frac{3}{4} - \frac{\phi_2}{2\pi}] & \text{if } \phi > 0 \\ [\frac{3}{4} - \frac{\phi_2}{2\pi}, \frac{3}{4} - \frac{\phi}{4\pi} - \frac{\phi_2}{2\pi}] & \text{if } \phi < 0, \end{cases} \quad I^- = \begin{cases} [\frac{1}{4} - \frac{\phi_2}{2\pi}, \frac{1}{4} + \frac{\pi-\phi}{4\pi} - \frac{\phi_2}{2\pi}] & \text{if } \phi > 0 \\ [-\frac{\phi}{4\pi} - \frac{\phi_2}{2\pi}, \frac{1}{4} - \frac{\phi_2}{2\pi}] & \text{if } \phi < 0\,. \end{cases}$$

Let x^+ and x^- denote the restrictions of $x^\star$ to I^+ and I^- respectively. These functions are bijective onto $[0, 1/4]$, and it holds that

$$(x^+)^{-1}(x) = \operatorname*{argmax}_{\tilde{z}} F(x, \tilde{z}), \quad (x^-)^{-1}(x) = \operatorname*{argmin}_{\tilde{z}} F(x, \tilde{z})$$

for all $x \in (0, 1/4)$. Here the argmax *and* argmin *are unique for all $x \in (0, 1/4)$. For $x = 1/4$ there is a spurious second solution and for $x = 0$ we get $F(0, z) = (\Delta, \Sigma)$.*

Proof. This follows directly from Lemma 5.5.4 and Lemma 5.B.1. □

Unfortunately it seems that the inverses of the functions x^+ and x^- defined in the previous lemma cannot be computed analytically. However we can obtain analytical expressions parametrized by z.

Corollary 5.5.7. *Assume that $\phi \in (-\pi, 0) \cup (0, \pi)$ and define the path*

$$\gamma(z) = (\Delta \cos(2\pi x^\star(z)), F(x^\star(z), z)),$$

and let γ^+ and γ^- be the restrictions to I^+ and I^- respectively. Then γ^+ is a parametrization of the upper boundary of the right half of the space of generators $\mathfrak{Q}$, and analogously γ^- parametrizes the lower boundary of the right half. It follows that the boundary point $\gamma(z)$ for $z \in I^+ \cup I^-$ can be obtained using the unitary $U_{x^\star(z),z}$.

The intervals I^+ and I^- are still too large, since on these intervals γ parametrizes part of the boundary of the space of generators which are not relevant for optimal control. Indeed, since $\lambda \in [0, 1]$, we are only interested in the values of z where $\gamma_1'(z)/\gamma_2'(z) \in [-1, 1]$. This parameter region can be computed numerically and we will denote the corresponding closed parameter intervals $J^+ \subseteq I^+$ and $J^- \subseteq I^-$.

With this we can apply the results of Section 5.4 to numerically determine the optimal path through the Bloch ball and the corresponding control functions of the full control system ($\mathcal{B}$), cf. Figure 5.3.

5.6 Integral Systems

In the previous section we had to exclude the cases where ϕ, as defined in Section 5.5, is an integer multiple of π. It turns out that in these cases the general parametrization of $\mathfrak{Q}$ obtained in the previous section simplifies considerably. We call such systems integral and we explore their properties in this section.

Definition 5.6.1. *A system is* integral *if it is non-degenerate (cf. Remark 5.5.5) and $\phi = k\pi$ for some $k \in \mathbb{Z}$. We call $p = (-1)^k$ the* parity *of the system, and we say that the system is* even *or* odd *if k is even or odd respectively.*

Note that due to Lemma 5.5.1 integral systems and their parity are well-defined.

Example 5.6.2. *We say that a Lindblad generator $-L \in \mathfrak{w}_{\mathsf{KL}}(2)$ is* real *if there exists a choice of real Lindblad terms V_k. If all Lindblad terms are real, and the system is non-degenerate, then the system is integral. Indeed, since all V_k are real, the sum $\sum_{k=1}^r [V_k, V_k^*]$ is real and symmetric and hence can be orthogonally diagonalized. Hence $r_1 e^{\mathrm{i}\phi_1}$ is imaginary and $r_2 e^{\mathrm{i}\phi_2}$ is real, so that $\phi_1 + \pi/2$ and $2\phi_2$ are integer multiples of π. Note that both even and odd systems can be obtained in this fashion, and they are separated in $\mathfrak{w}_{\mathsf{KL}}$ by degenerate systems.*

Since they are easy to generate, we will use real systems for the plots shown in this section.

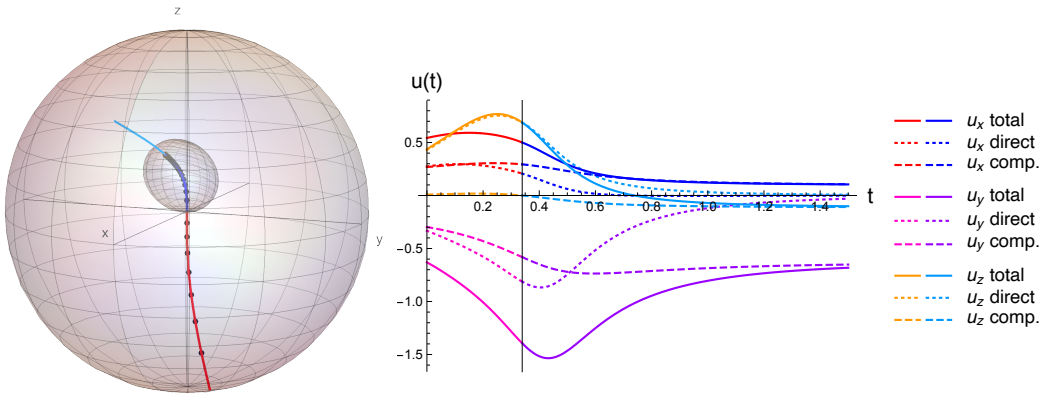

Figure 5.3: Left: Optimal path for heating (red) and cooling (blue) in the Bloch ball together with the set of stabilizable states. The light blue part indicates the part of the optimal path which is not reachable from the interior as it lies beyond the purest stabilizable state. The black dots on the path are equally spaced in time. Right: Optimal control functions u_x, u_y and u_z for heating (left part) and cooling (right part) and their contributing direct and compensating terms. The Lindblad terms used are the same as in Figure 5.1.

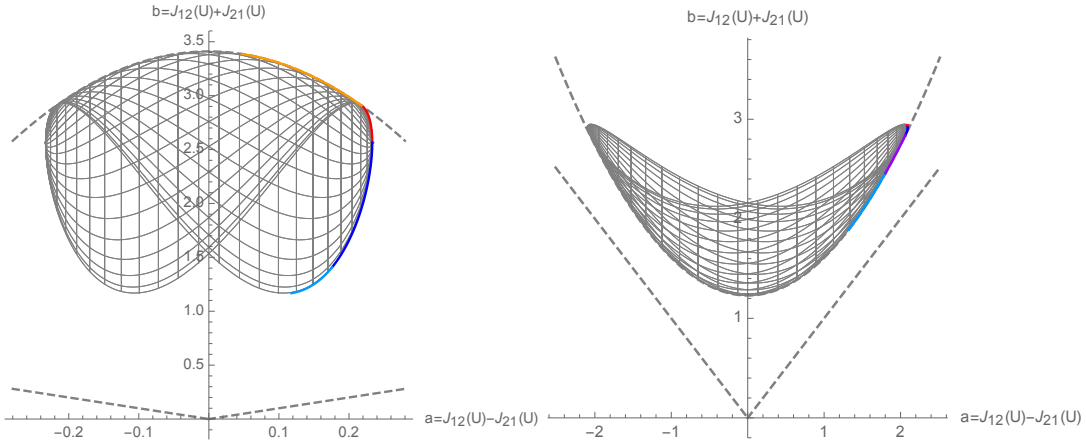

Figure 5.4: Space of generators $\mathfrak{Q}$ for two randomly chosen real systems. The system on the left is odd, and the system on the right is even. The boundary, computed using Lemma 5.6.3, features a parabolic part, located at the top in the odd case and at the bottom in the even case. The Lindblad terms for the odd case are $\left(\begin{smallmatrix} 0.3 & -1.0 \\ -0.6 & 0.9 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0.8 & 0.2 \\ 0.8 & 0.8 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0.1 & -0.4 \\ 0.0 & -0.8 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} -0.3 & 0.6 \\ 0.6 & 0.4 \end{smallmatrix}\right)$. For the even case they are $\left(\begin{smallmatrix} 0.5 & 0.1 \\ -0.4 & -0.1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0.0 & 0.5 \\ -0.9 & 0.0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0.9 & -0.3 \\ 0.7 & 0.6 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 0.4 & -0.2 \\ -1.0 & 0.0 \end{smallmatrix}\right)$.

Optimal Solution

Using Lemma 5.B.2 we can give analytic expressions for the boundary of the space of generators $\mathfrak{Q}$ for integral systems.

Lemma 5.6.3. *Set* $\tilde{x} = \frac{1}{2\pi}\arctan(\frac{r_2}{4r_1})$. *For* $\phi = 0$, *we obtain for every* $x \in [0, 1/4]$ *that*

$$\max_z F(x,z) = \Sigma + \delta\sin(2\pi x)^2 + \sin(2\pi x)(r_1\sin(2\pi x) + r_2\cos(2\pi x))$$
$$\min_z F(x,z) = \begin{cases} \Sigma + \delta\sin(2\pi x)^2 + \sin(2\pi x)(r_1\sin(2\pi x) - r_2\cos(2\pi x)) & \text{if } x \in [0,\tilde{x}] \\ \Sigma + \sin(2\pi x)^2(\delta - r_1) - \cos(2\pi x)^2\frac{r_2^2}{8r_1} & \text{if } x \in [\tilde{x}, 1/4], \end{cases}$$

and analogously for $\phi = \pi$ *we obtain that*

$$\max_z F(x,z) = \begin{cases} \Sigma + \delta\sin(2\pi x)^2 + \sin(2\pi x)(-r_1\sin(2\pi x) + r_2\cos(2\pi x)) & \text{if } x \in [0,\tilde{x}] \\ \Sigma + \sin(2\pi x)^2(\delta + r_1) + \cos(2\pi x)^2\frac{r_2^2}{8r_1} & \text{if } x \in [\tilde{x}, 1/4] \end{cases}$$
$$\min_z F(x,z) = \Sigma + \delta\sin(2\pi x)^2 + \sin(2\pi x)(-r_1\sin(2\pi x) - r_2\cos(2\pi x)).$$

Furthermore, these values can be obtained using the unitary $U_{x,z} = \exp(\mathrm{i}\pi z^{\pm}(x)\sigma_z)\exp(\mathrm{i}\pi x\sigma_x)$ *where*

$$z^+(x) = \tfrac{3}{4} - \tfrac{\phi_2}{2\pi}, \quad z^-(x) = \begin{cases} \frac{1}{4} - \frac{\phi_2}{2\pi} & \text{if } x \in [0,\tilde{x}] \\ \frac{1}{2\pi}\arcsin(\frac{r_2}{4r_1\tan(2\pi x)}) - \frac{\phi_2}{2\pi} & \text{if } x \in [\tilde{x}, 1/4], \end{cases}$$

in the even case and

$$z^-(x) = \tfrac{1}{4} - \tfrac{\phi_2}{2\pi}, \quad z^+(x) = \begin{cases} \frac{3}{4} - \frac{\phi_2}{2\pi} & \text{if } x \in [0,\tilde{x}] \\ \frac{1}{2} + \frac{1}{2\pi}\arcsin(\frac{r_2}{4r_1\tan(2\pi x)}) - \frac{\phi_2}{2\pi} & \text{if } x \in [\tilde{x}, 1/4], \end{cases}$$

in the odd case.

Proof. Using Lemma 5.5.4 and Lemma 5.B.2 we immediately find the optimal value of z as a function of x. Then we find the switching point $\xi = 4/5 \iff x = \frac{1}{2\pi}\arctan(\frac{r_2}{4r_1})$ and we denote this value by $\tilde{x}$. For the even case $\phi = 0$, or equivalently $\phi_1 - 2\phi_2 = -\pi/2$, we compute

$$\begin{aligned} \sin(4\pi z^+(x) + \phi_1) = \sin(3\pi - 2\phi_2 + \phi_1) = 1 \quad & \text{for } x \in [0, 1/4] \\ \sin(2\pi z^+(x) + \phi_2) = \sin(3\pi/2 - \phi_2 + \phi_2) = -1 \quad & \text{for } x \in [0, 1/4] \\ \sin(4\pi z^-(x) + \phi_1) = 1 \quad & \text{for } x \in [0, \tilde{x}] \\ \sin(2\pi z^-(x) + \phi_2) = 1 \quad & \text{for } x \in [0, \tilde{x}], \end{aligned}$$

and using that $1/(4(1/\xi - 1)) = r_2/(4r_1)\cot(2\pi x)$ we get

$$\begin{aligned} \sin(4\pi z^-(x) + \phi_1) &= -\cos(2\arcsin(r_2/(4r_1)\cot(2\pi x))) \\ &= 2(r_2/(4r_1)\cot(2\pi x))^2 - 1 && \text{for } x \in [\tilde{x}, 1/4] \\ \sin(2\pi z^-(x) + \phi_2) &= r_2/(4r_1)\cot(2\pi x) && \text{for } x \in [\tilde{x}, 1/4], \end{aligned}$$

and hence we find

$$F(x, z^-(x)) = \Sigma + \sin(2\pi x)^2(\delta - r_1) - \cos(2\pi x)^2\frac{r_2^2}{8r_1} \quad \text{for } x \in [\tilde{x}, 1/4].$$

The computations for the odd case $\phi = \pi$ are analogous and this yields the result. □

The boundary parametrization is shown in Figure 5.4, and the resulting optimal derivative function μ as well as the optimal path and the set of stabilizable states in the Bloch sphere are presented in Figure 5.5 and Figure 5.6 respectively.

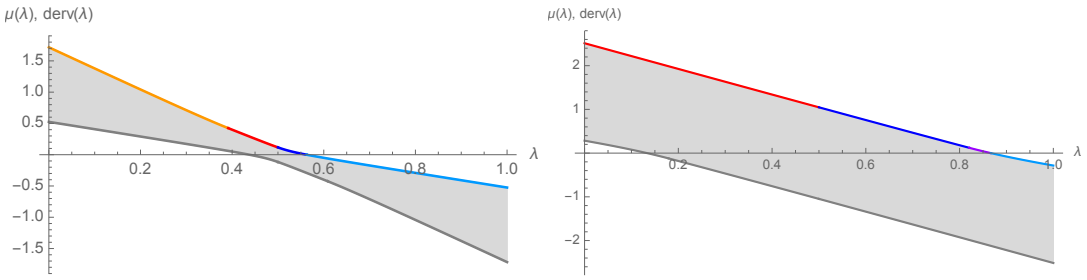

Figure 5.5: For the same two real systems we plot the set-valued function derv of achievable derivatives and the optimal derivatives μ, which can be computed using Lemma 5.4.7. The colors used correspond to those of Figure 5.4.

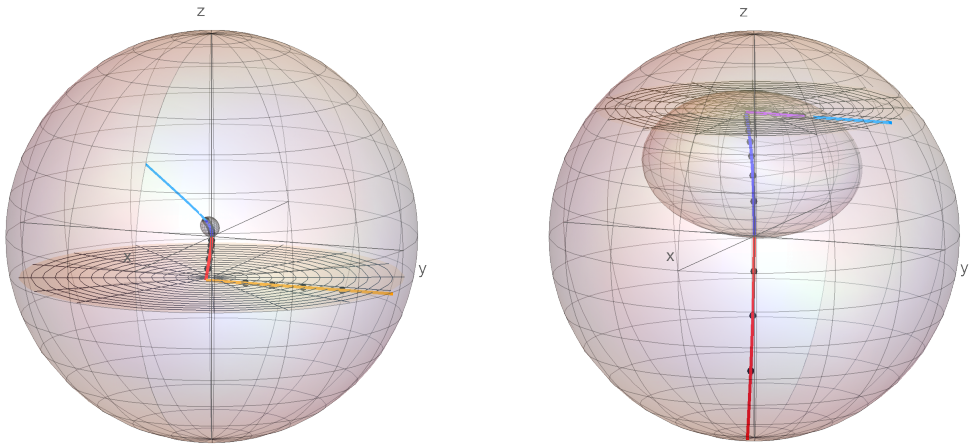

Figure 5.6: For the same two real systems we plot the optimal path through the Bloch ball and the set of stabilizable states. Note that the parabolic parts of the boundary of $\mathfrak{Q}$ correspond to horizontal parts in the Bloch ball, reminiscent of the magic plane in the Bloch equations case, cf. Appendix 5.A. The rest of the path lies in the xz-plane. The colors used correspond to those of Figure 5.4.

Coolable Systems

In Theorem 4.3.7 coolability of Markovian quantum systems with fast unitary control was characterized. As a consequence, in the qubit case the system is asymptotically coolable if and only if the Lindblad terms V_k can be simultaneously unitarily triangulated without being simultaneously diagonal, see also [RBR12, Thm. 5.2]. For non-unital rank one systems this is always satisfied, and such systems will be studied in detail in Section 6.3.

Lemma 5.6.4. *Coolable systems are either degenerate or odd.*

Proof. As usual we choose the Lindblad terms V_k traceless and the basis such that $\sum_{k=1}^r [V_k, V_k^*]$ is diagonal. Then the V_k have the form $V_k = \left(\begin{smallmatrix} u_k & v_k \\ w_k & -u_k \end{smallmatrix}\right)$, where $u_k, v_k, w_k \in \mathbb{C}$ for $k = 1, \dots, r$. Let $u, v, w \in \mathbb{C}^r$ be the vectors with coefficients u_k, v_k and w_k respectively. We will use the shorthand notation $(u, v) = \sum_{k=1}^r \overline{u}_k v_k$ and $|v| = \sqrt{(v, v)}$. Diagonality of $\sum_{k=1}^r [V_k, V_k^*]$ is equivalent to $(v, u) = (u, w)$. It is easy to show that one can always modify the basis (while keeping $\sum_{k=1}^r [V_k, V_k^*]$ diagonal) such that $(v, w) \geq 0$ and we will make this assumption. If $(v, w) = 0$ then $r_1 = 0$ and the

system is degenerate, hence assume that $(v,w) > 0$ and thus $\phi_1 = \frac{\pi}{2}$. Now consider a common eigenvector of all V_k. If it is proportional to $(1,0)^\top$ or $(0,1)^\top$ it is easy to see that $r_2 = 0$. Hence assume that the eigenvector is of the form $(1,\beta)$ with $\beta \neq 0$. We will show that indeed $\beta \in \mathbb{R}$. From the eigenvalue equation we obtain that $2u_k = \frac{w_k}{\beta} - \beta v_k$ for all $k = 1,\dots,r$. By taking the inner product with v and w and using $(v,u) = (u,w)$ we find that $\beta|v|^2 + \frac{1}{\beta}|w|^2 = (\frac{1}{\beta} + \overline{\beta})(v,w)$. By considering the complex argument of each side it is clear that β must be real. But then it follows from the above that $2(v,u) = \frac{1}{\beta}(v,w) - \beta|v|^2$ and so (v,u) is real as well. Thus $\phi_2 = n\pi$ is an integer multiple of π. Thus $\phi = \phi_1 - 2\phi_2 + \frac{\pi}{2} = (1-2n)\pi$ and the system is odd as desired. □

Since non-degenerate coolable systems are integral, the results of the previous section still apply. However, using the simultaneous triangular form of the Lindblad terms another parametrization can be obtained, see Figure 5.7. Choosing all V_k traceless and an appropriate basis, we may assume that they are of the form $V_k = \left(\begin{smallmatrix} u_k & v_k \\ 0 & -u_k \end{smallmatrix}\right)$, where $u_k, v_k \in \mathbb{C}$ are arbitrary for $k = 1,\dots,r$. Let $u, v \in \mathbb{C}^r$ be the vectors with coefficients u_k and v_k, then we define

$$\begin{aligned} c_1(x) &= |v|^2\cos(2\pi x), & c_3(x) &= 2|u|^2\sin^2(2\pi x) + |v|^2(\cos^2(2\pi x)+1)/2, \\ c_2(x) &= -2|(u,v)|\sin(2\pi x), & c_4(x) &= -|(u,v)|\sin(4\pi x). \end{aligned}$$

where again $(u,v) = \sum_{k=1}^r \overline{u}_k v_k$ and $|v| = \sqrt{(v,v)}$.

Lemma 5.6.5. *For a coolable system in the form described above, the space of generators can be parametrized as* $\{(c_1(x) + c_2(x)\sin(2\pi z), c_3(x) + c_4(x)\sin(2\pi z)) : x \in [0,1/2],\ z \in [0,1)\}$, *where each point is obtained using the unitary* $U_{x,z}$. *The non-parabolic part of the boundary is achieved by* $z = \frac{1}{4}$. *Moreover, the parabolic segment lies on a parabola which is tangent to the bisectors and has the form* $a \mapsto \Sigma + \delta + r_2/2 + \frac{1}{4(\Sigma+\delta+r_2/2)}\, a^2$.

Proof. This follows directly from the computations in Appendix 5.B. □

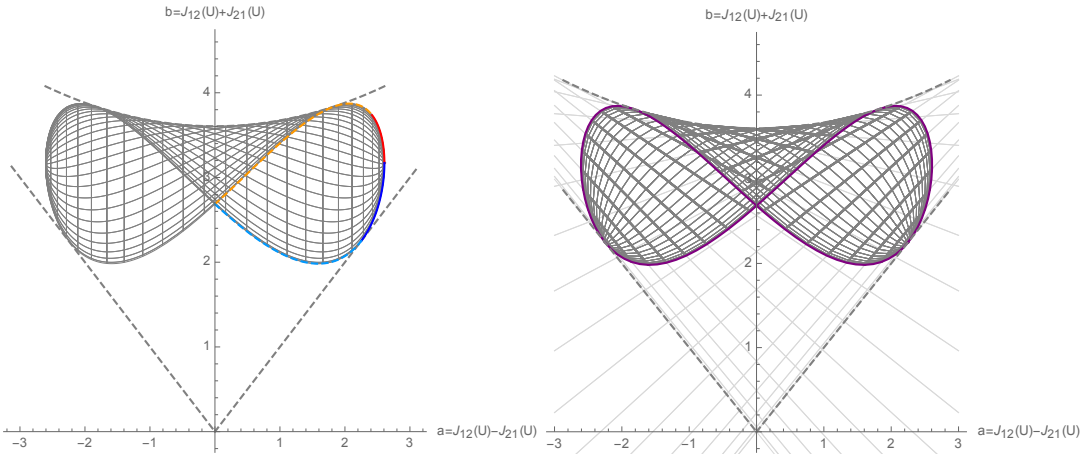

Figure 5.7: Two different parametrizations of the space of generators for a coolable system. The usual parametrization is on the left, and the parametrization of Lemma 5.6.5 is on the right. The Lindblad terms are given by $\left(\begin{smallmatrix} 0.7-0.5\mathrm{i} & 0.6-0.6\mathrm{i} \\ 0.0 & -0.8+0.9\mathrm{i} \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0.8-0.3\mathrm{i} & -1.0 \\ 0.0 & 0.1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 0.3-0.2\mathrm{i} & -0.2+0.7\mathrm{i} \\ 0.0 & 0.2-0.6\mathrm{i} \end{smallmatrix}\right)$.

5.A Special Systems

In this section we study some highly structured systems, namely unital systems and the Bloch equations. The case of rank one systems, i.e. those defined by a single Lindblad term, will be treated in detail in Section 6.3.

Unital Systems

A simple but broad class of examples is given by *unital* systems, which are defined by $-L(\mathbb{1}) = \sum_{k=1}^{r}[V_k, V_k^*] = 0$. This condition is independent of the Hamiltonian part of $-L$, and hence also of the control Hamiltonians. Unital systems in arbitrary (finite) dimension have been addressed in Section 4.4. The qubit case presented here allows for even stronger results. Unital channels were also studied in [Muk+13, Sec. IV, V] and in [Yua10].

Lemma 5.A.1. *The following are equivalent:*

(i) $-L$ *is unital,*

(ii) the optimal derivative function satisfies $\mu(1/2) = 0$,

(iii) the space of generators satisfies $\mathfrak{Q} \subset \{(0, y) : y \geq 0\}$.

Proof. By Lemma 5.3.1, $\mu(1/2) = 0$ if and only if $\sum_{k=1}^{r}[V_k, V_k^*] = 0$. This shows the equivalence of (i) and (ii). Now assume (ii), then every line in $\mathfrak{Q}$ passes through $(1/2, 0)$ due to the central symmetry. This is equivalent to $J_{12}(U) = J_{21}(U)$ for all $U \in \mathrm{SU}(2)$ and hence to (iii). □

Condition (ii) implies that the graph of derv is a cone with origin $(1/2, 0)$, as illustrated in Figure 5.8.

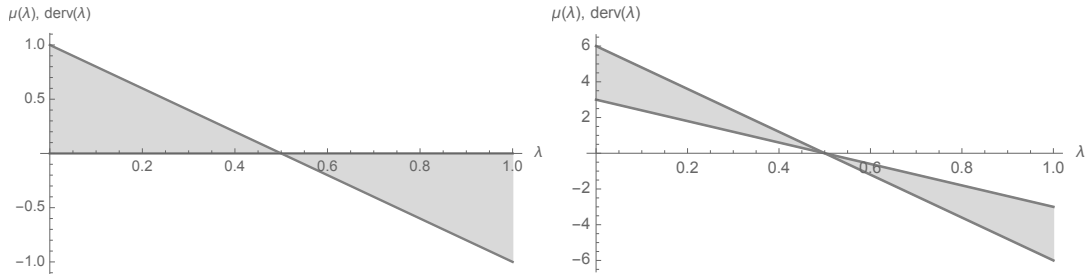

Figure 5.8: The graph of derv for two unital systems. As shown in Lemma 5.A.1, all lines pass through $(1/2, 0)$, and this is indeed characteristic of unital systems. The systems used are defined by $\sqrt{\gamma_k}\sigma_k$ for $k \in x, y, z$. Since the Pauli matrices are normal, such systems are always unital. Left: The system is defined by $\gamma_x = 1$ and $\gamma_y = \gamma_z = 0$, so it is of rank one and by Lemma 5.A.2 it is unital stabilizable. By Lemma 5.A.3 it holds that $\mathsf{derv}(0) = [0, 1]$. Right: The system is defined by $\gamma_x = 4$, $\gamma_y = 2$, and $\gamma_z = 1$. Hence by Proposition 5.A.5 it holds that $\mathsf{derv}(0) = [3, 6]$.

It is clear that for unital systems the stabilizable region is either $\{1/2\}$ or $[0, 1]$. Moreover derv, μ, and $\mathfrak{Q}$ are completely described by the minimal and maximal values of $\mathsf{derv}(0)$.

Lemma 5.A.2. *Let* $-L$ *be an arbitrary Lindblad generator and let* V_k *be a corresponding family of Lindblad terms. Then the following are equivalent:*

(i) all V_k *are normal and commute with each other,*

(ii) the space of generators satisfies $(0,0) \in \mathfrak{Q}$,

(iii) the system is unital and the stabilizable region is all of $[0,1]$.

We call such systems unital stabilizable.

Proof. Since (i) is equivalent to the existence of a unitary $U \in \mathrm{SU}(2)$ simultaneously diagonalizing all V_k, it is also equivalent to (ii). Now assume the conditions above. This implies that all of $[0,1]$ is stabilizable and that the system is unital. This shows (iii). Finally assume (iii). The system is unital and so $[0,1]$ is stabilizable only if $(0,0) \in \mathfrak{Q}$. This concludes the proof. □

Clearly for a unital stabilizable system, μ is completely characterized by its value at 0. The following result yields a simple way to determine this value.

Lemma 5.A.3. *Consider a unital stabilizable system. Then* $\mu(0) = \frac{1}{4}\sum_{k=1}^{r} |\lambda_1(V_k) - \lambda_2(V_k)|^2$ *where the* $\lambda_i(V_k)$ *denote the eigenvalues of* V_k.

Proof. By Lemma 5.A.2 all Lindblad terms V_k are simultaneously unitarily diagonalizable. Using a unitary change of basis, we may assume that they are indeed diagonal. Moreover we may assume that they are traceless. Hence all Lindblad terms are multiples of σ_z. Applying a unitary reshuffling to the V_k one may assume that all but one V_k are zero. Performing this calculation yields the desired result. □

Now we will completely describe μ and $\mathfrak{Q}$ for arbitrary unital systems. For this let $C \in \mathbb{C}^{3,3}$ denote the Kossakowski matrix (cf. [BP02, p. 121]) of the system with respect to the Pauli basis $\{\sigma_x, \sigma_y, \sigma_z\}$.

Lemma 5.A.4. *The system is unital if and only if* C *is real.*

Proof. We will give a sketch of the proof. Using the general "non-diagonal" form of the Lindblad equation (cf. [BP02, p. 121]) with respect to the Pauli basis one finds by a simple computation that the generator is unital, i.e. $-L(\mathbb{1}) = 0$, if and only if the Kossakowski matrix C is symmetric, or equivalently, real. □

The following recovers a result from [Yua10].

Proposition 5.A.5. *Consider a unital system. Let* $\gamma_1 \geq \gamma_2 \geq \gamma_3$ *denote the eigenvalues of* C. *Then* $\mathsf{derv}(0) = [\gamma_2 + \gamma_3, \gamma_1 + \gamma_2]$.

Proof. Applying a unitary basis transformation to the qubit system changes the Kossakowski matrix via a corresponding orthogonal transformation. Since C is real, it can be orthogonally diagonalized and hence, without loss of generality, we can assume that the Lindblad terms are $\sqrt{\gamma_1}\sigma_x$, $\sqrt{\gamma_2}\sigma_y$, and $\sqrt{\gamma_3}\sigma_z$. The result then follows easily from the computation in Appendix 5.B. □

Bloch Equations

Another family of simple Lindblad generators for the qubit is given by those which have a rotation symmetry about some axis, which reduces the dimension of the problem. Without loss of generality we assume that the symmetry is about the z-axis. Such generators correspond to the well-known Bloch equations. It turns out that such systems can be solved analytically, and we will do so in detail in this section. In Lemma 5.A.10 we recover a known result from [Lap+10] about the so-called magic plane for optimal heating and in Lemma 5.A.12 we recover the steady state ellipsoid from [Lap+13]. The optimal controls can also be determined, as was done in Section 2.6. A special case of the Bloch equations was also considered in detail in [Muk+13, Sec. III].

The Bloch equations are equivalent to the Lindblad generator defined by the Lindblad terms $\sqrt{\gamma_+}\sigma_+$, $\sqrt{\gamma_-}\sigma_-$ and $\sqrt{\gamma_z}\sigma_z$ where $\sigma_\pm = (\sigma_x \pm \mathrm{i}\sigma_y)/2$ and $\gamma_+, \gamma_-, \gamma_z \geq 0$. The case considered in [Muk+13, Sec. III] corresponds to the case $\gamma_z = 0$. For convenience we introduce the following parameters:

$$\Delta = |\gamma_+ - \gamma_-|, \quad \Sigma = \gamma_+ + \gamma_-, \quad \delta = 2\gamma_z - \Sigma/2\,.$$

Indeed, the values of Δ, Σ, and δ defined here correspond to those defined in the general case in Section 5.5 with $r_1 = r_2 = 0$. By a change of basis we may and will assume without loss of generality that $\gamma_+ \geq \gamma_-$.

Remark 5.A.6. *The Bloch equations for a single qubit are often written in the following form, see for instance [AM11, Sec. 5.5],*

$$\dot{\mathbf{M}} = \gamma \mathbf{M} \times \mathbf{B} - R(\mathbf{M} - \mathbf{M}_\beta), \quad R = \mathrm{diag}(T_2^{-1}, T_2^{-1}, T_1^{-1})\,,$$

where $\mathbf{M}$ *is the spin magnetization of the system,* γ *is the gyromagnetic ratio,* $\mathbf{B} = (0, 0, B)$ *is the magnetic field, and* R *is the relaxation matrix with* T_1 *the longitudinal and* T_2 *the transversal relaxation time. Furthermore* $\mathbf{M}_\beta = (0, 0, M_\beta)$ *is the steady state magnetization and satisfies* $2M_\beta/(\gamma\hbar) \in [-1, 1]$. *Then we have the relations*

$$T_1^{-1} = \gamma_+ + \gamma_- = \Sigma, \quad T_2^{-1} = (\gamma_+ + \gamma_- + 4\gamma_z)/2 = \Sigma + \delta$$
$$2M_\beta/(\gamma\hbar) = (\gamma_+ - \gamma_-)/(\gamma_+ + \gamma_-) = \Delta/\Sigma\,,$$

Note that the famous relation $2T_1 \geq T_2$ *is equivalent to the non-negativity of the relaxation rates* γ_+, γ_- *and* γ_z *and hence to the complete positivity of the evolution.*

The following results analytically parametrize the space of generators $\mathfrak{Q}$ and deduce the analytical formula for the optimal derivative function μ. See Figure 5.9 for examples. As a consequence we can also determine the purest stabilizable state.

Lemma 5.A.7. *For the Bloch equations, the space of generators* $\mathfrak{Q}$ *is the graph of the parabolic segment*

$$f(a) = \Sigma + \delta\big(1 - \tfrac{a^2}{\Delta^2}\big), \quad a \in [-\Delta, \Delta].$$

The point $(a, f(a))$ *can be obtained using the unitary* $\exp(i\pi x \sigma_x)$ *satisfying* $a = \Delta\cos(2\pi x)$.

Proof. Consider an initial density matrix on the z-axis of the Bloch ball. Then every density matrix in its unitary orbit can be reached by applying an x-rotation followed by a z-rotation. However, since the Lindblad terms are z-symmetric, it suffices to consider only x-rotations. That is

$$\mathfrak{Q} = \{(J_{12}(U) - J_{21}(U), J_{12}(U) + J_{21}(U)) : U = \exp(i\pi x \sigma_x), x \in [0, 1/2]\}\,. \tag{5.5}$$

Then, evaluating the above expression one obtains the points

$$((\gamma_+ - \gamma_-)\cos(2\pi x),\ \tfrac{3(\gamma_+ + \gamma_-)}{4} + \gamma_z + \tfrac{1}{4}(\gamma_+ + \gamma_- - 4\gamma_z)\cos(4\pi x)), \quad x \in [0, 1/2]. \tag{5.6}$$

Setting $a = \Delta\cos(2\pi x)$ and noting that $\cos(4\pi x) = 2(a/\Delta)^2 - 1$ we obtain the desired points. □

Lemma 5.A.8. *For the Bloch equations, the upper bound* $\mu : [0, 1] \to \mathbb{R}$ *takes the form*

$$\mu(\lambda) = \begin{cases} \frac{\Delta^2}{8\delta(1-2\lambda)} + \frac{(\Sigma+\delta)(1-2\lambda)}{2} & \text{if } \lambda \in I \\ \frac{1}{2}(\Delta + (1-2\lambda)\Sigma) & \text{if } \lambda \in [0,1] \setminus I, \end{cases} \quad \text{where } I = \begin{cases} [0, \frac{1}{2} - \frac{\Delta}{4\delta}] & \text{if } \delta \geq \frac{\Delta}{2} \\ [\frac{1}{2} - \frac{\Delta}{4\delta}, 1] & \text{if } \delta \leq -\frac{\Delta}{2} \\ \emptyset & \text{otherwise}\,, \end{cases}$$

in particular $1/2 \notin I$.

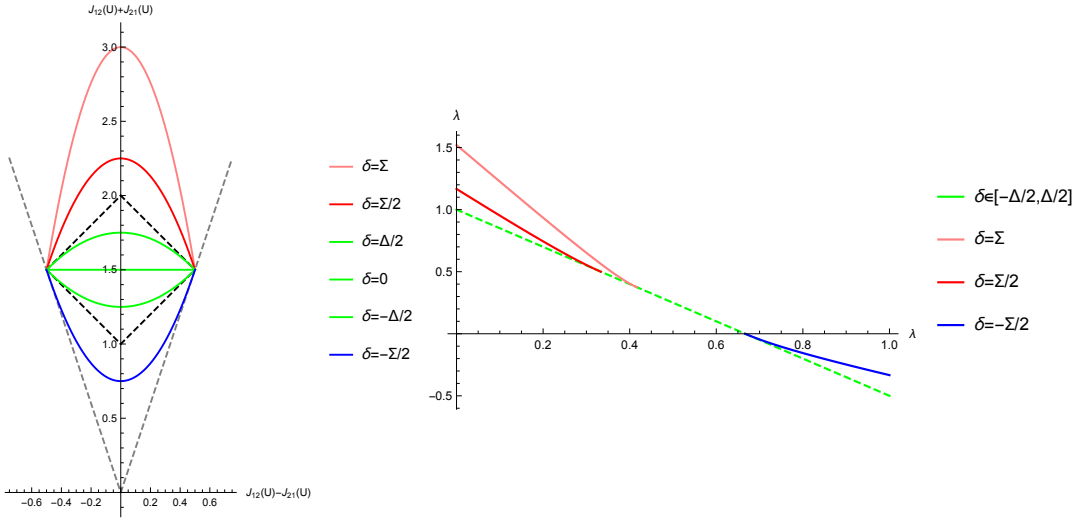

Figure 5.9: Several examples of systems defined by Bloch equations. For all cases we chose $\gamma_+ = 1$ and $\gamma_- = 1/2$. Hence $\Sigma = 3/2$ and $\Delta = 1/2$. We consider different values of γ_z, and hence δ. Left: The space of generators $\mathfrak{Q}$ for fixed Σ and Δ and different δ, as described in Lemma 5.A.7. In each case we get a parabolic segment with endpoints $(\pm\Delta, \Sigma)$ and intersecting the ordinate at $\Sigma + \delta$. By definition $\delta \geq \Sigma/2$ and hence the parabolas are contained between the lines connecting the endpoints with the origin (gray dashed). This allows us to find the purest achievable state in Corollary 5.A.9. Right: The optimal derivative μ for fixed Σ and Δ and different δ, as described in Lemma 5.A.8. For $\delta \in [-\Delta/2, \Delta/2]$ the optimal derivative μ is linear (green dashed) and only depends on γ_+ and γ_- (or Σ and Δ equivalently). For other values of δ the upper bound has to be modified on an interval I defined in the same lemma. Note that this modification does not affect the intersection of μ with the abscissa, again reflecting Corollary 5.A.9.

Proof. Due to Proposition 5.3.4 we have to solve the linear maximization $\max_{(a,b)\in\mathfrak{Q}} \frac{1}{2}(a+(1-2\lambda)b)$. We see that if $\lambda = 1/2$ then the maximum is achieved at $(\Delta, \Sigma) \in \mathfrak{Q}$ with value $\Delta/2$. We have to consider the shape of the parabolic segment depending on δ. It is clear that for $\delta = 0$ this is just a line segment, for $\delta < 0$ it is convex, and for $\delta > 0$ it is concave. Note also that $f'(\Delta) = \frac{-2\delta}{\Delta}$. Hence if $-\Delta/2 < \delta < \Delta/2$ then the maximum will be achieved on $(\Delta, \Sigma) \in \mathfrak{Q}$ for all $\lambda \in [0, 1]$ and so μ will be the affine linear function $\mu(\lambda) = \frac{1}{2}(\Delta + (1 - 2\lambda)\Sigma)$. This proves the case $I = \emptyset$. Now let $\delta \geq \Delta/2$ (the case $\delta \leq -\Delta/2$ is analogous). Then the parabolic segment is concave. For large λ, the maximum will still be at $(\Delta, \Sigma) \in \mathfrak{Q}$, but for λ small enough, the maximum will be achieved in the interior of the parabolic segment. The switching point occurs when the vector $(1, 1-2\lambda)$ is orthogonal to $(1, f'(\Delta))$, that is when $1 + (1 - 2\lambda)(-2\delta/\Delta) = 0$ which is equivalent to $\lambda = \frac{1}{2} - \frac{\Delta}{4\delta}$. Hence it remains to determine μ on $I = [0, \frac{1}{2} - \frac{\Delta}{4\delta}]$. For this we compute

$$\frac{-2\delta a}{\Delta^2}(1 - 2\lambda) = -1 \iff a(\lambda) = \frac{\Delta^2}{2\delta(1-2\lambda)}, \tag{5.7}$$

and plugging in we get on I that $\mu(\lambda) = \frac{1}{2}(a(\lambda) + (1 - 2\lambda)f(a(\lambda)))$, which evaluates to the desired result. □

Corollary 5.A.9. *The purest stabilizable state $\lambda^\star$, that is, the state satisfying $\mu(\lambda^\star) = 0$, is given by $\lambda^\star = \frac{1}{2} + \frac{\Delta}{2\Sigma}$ and $\{\lambda^\star, 1 - \lambda^\star\}$ is the spectrum of the fixed point of the system.*

Proof. Since $\gamma_z \geq 0$ we must have $\delta \geq -\Sigma/2$, and hence $f'(\Delta) = \frac{-2\delta}{\Delta} \leq \frac{\Sigma}{\Delta}$, which shows that the line defined by $(\Delta, \Sigma) \in \mathfrak{Q}$ is the line giving the purest stabilizable state. Hence $\frac{1}{2}(\Delta + (1 - 2\lambda^\star)\Sigma) =$

$0 \iff \lambda^\star = \frac{1}{2} + \frac{\Delta}{2\Sigma}$, as desired. (This also follows from Lemma 5.4.3 and noting that γ_z can be set to 0 without loss of generality.) By z-symmetry, the diagonal states are invariant, and hence by Brouwer's Theorem, there is a fixed point which is diagonal. If this fixed point is pure, then $\gamma_- = 0$ and we are done. Otherwise, the fixed point is unique. If we denote the larger eigenvalue of the fixed point by λ we obtain that $-\gamma_-\lambda + \gamma_+(1-\lambda) = 0$ which shows that $\lambda = \frac{1}{2} + \frac{\Delta}{2\Sigma}$, as desired. □

From these results one can deduce the optimal path through the Bloch ball and the corresponding optimal controls in the original control system ($\mathcal{B}$), see Section 2.6 for details. Here we just recover the so called magic plane and steady state ellipsoid from [Lap+10, Lap+13].

Lemma 5.A.10. *If $\delta \geq \Delta/2$, or equivalently, $\gamma_z \geq \gamma_+/2$, then for $\lambda \in I = [0, \frac{1}{2} - \frac{\Delta}{4\delta}]$, the points in the Bloch ball achieving the optimal derivative $\mu(\lambda)$ are given by the plane perpendicular to the z-axis and passing through the density matrix $\operatorname{diag}(\lambda, 1-\lambda)$ with $\lambda = \frac{1}{2} - \frac{\Delta}{4\delta}$. If $\lambda \in [\frac{1}{2} - \frac{\Delta}{4\delta}, \frac{1}{2}]$, then the maximal derivative is reached on the z-axis on the same side of the origin as the magic plane.*

Proof. If $\delta \geq \Delta/2$, then the interval $I = [0, \frac{1}{2} - \frac{\Delta}{4\delta}]$ lies in $[0, \frac{1}{2}]$. For $\lambda \in I$, the upper bound $\mu(\lambda)$ is non-linear. We want to find, for each $\lambda \in I$, the density matrices ρ for which the optimal derivative $\mu(\lambda)$ is achieved. For this we find U such that the optimal derivative is achieved for $U \operatorname{diag}(\lambda, 1-\lambda)U^*$. From (5.7) it follows that for $\lambda \in I$ the optimal point in the space of generators $\mathfrak{Q}$ has $a = \Delta^2/(2\delta(1-2\lambda))$. From (5.5) and (5.6) it follows that a corresponding unitary is $\exp(\mathrm{i}\pi x \sigma_x)$ with $a = \Delta\cos(2\pi x)$. Hence the optimal unitary can be found from $x = \frac{1}{2\pi}\arccos\left(\frac{\Delta}{2\delta(1-2\lambda)}\right)$. Using the z-symmetry of the problem, this shows that for $\lambda \in I$ the optimal derivatives are reached on a plane perpendicular to the z-axis and passing through the point $\operatorname{diag}(\lambda, 1-\lambda)$ with $\lambda = \frac{1}{2} - \frac{\Delta}{4\delta}$. □

Remark 5.A.11. *Lemma 5.A.10 recovers a known result from [Lap+10, Lap+13], which was derived using the PMP. There, the parameters are scaled such that $\Sigma = \Delta$, or equivalently $\gamma_- = 0$, and hence $\gamma_+ = T_1^{-1}$ and $\gamma_z = (2T_2)^{-1} - (4T_1)^{-1}$. The condition from Lemma 5.A.10 then becomes $T_1 \geq \frac{3}{2}T_2$ and the magic plane intersects the z-axis at radius $r = T_2/(2(T_1 - T_2))$ in the lower half. Recall that here the radius of the Bloch sphere is $\frac{1}{2}$.*

Lemma 5.A.12. *The set of stabilizable states in the Bloch disk is an ellipsoid in the upper halfspace which is rotationally symmetric around the z-axis and has a vertical semiaxis of length $\frac{\Delta}{4\Sigma}$ and horizontal semiaxes of length $\frac{\Delta}{4\sqrt{\Sigma(\Sigma+\delta/2)}}$.*

Proof. By Lemma 5.4.6 we obtain the parametrization in polar coordinates $r(\theta, \phi) = \frac{\Delta}{2}\frac{\cos(\theta)}{\Sigma + \delta\sin(\theta)^2}$. The result follows from Lemma 5.B.7. □

Remark 5.A.13. *This recovers another known result from [Lap+13], with parameters rescaled as in Remark 5.A.11. Then the ellipse has vertical semiaxis length $1/4$, i.e. it touches the north pole, and horizontal semiaxis length $\sqrt{T_2/(2T_1)}$. Again recall that here the radius of the Bloch sphere is $\frac{1}{2}$.*

5.B Technical Computations

Detailed Computation for Proposition 5.5.2

The unitary $U_{x,z} = \exp(\mathrm{i}\pi z\sigma_z)\exp(\mathrm{i}\pi x\sigma_x)$ has the form

$$U_{x,z} = \begin{pmatrix} \alpha c & \mathrm{i}\alpha s \\ \mathrm{i}\alpha^* s & \alpha^* c \end{pmatrix}, \quad \text{and hence} \quad U_{x,z}^* = \begin{pmatrix} \alpha^* c & -\mathrm{i}\alpha s \\ -\mathrm{i}\alpha^* s & \alpha c \end{pmatrix},$$

where $\alpha = e^{\mathrm{i}\pi z}$ and $s = \sin(\pi x)$ and $c = \cos(\pi x)$. For a traceless Lindblad term

$$V = \begin{pmatrix} u & v \\ w & -u \end{pmatrix}$$

where $u, v, w \in \mathbb{C}$ are arbitrary we immediately obtain

$$\begin{aligned}(U_{x,z}^* V U_{x,z})_{12} &= 2\mathrm{i}csu + (\alpha^*)^2 c^2 v + \alpha^2 s^2 w \\ (U_{x,z}^* V U_{x,z})_{21} &= -2\mathrm{i}csu + (\alpha^*)^2 s^2 v + \alpha^2 c^2 w\,,\end{aligned}$$

and thus

$$\begin{aligned}J_{12}(U_{x,z}) - J_{21}(U_{x,z}) &= |2\mathrm{i}csu + (\alpha^*)^2 c^2 v + \alpha^2 s^2 w|^2 - |-2\mathrm{i}csu + (\alpha^*)^2 s^2 v + \alpha^2 c^2 w|^2 \\ &= (c^2 - s^2)(|v|^2 - |w|^2) + 2\,\mathrm{Re}(2\mathrm{i}csu(\alpha^2 v^* + (\alpha^*)^2 w^*)) \\ &= (|v|^2 - |w|^2)\cos(2\pi x) + 2\sin(2\pi x)\,\mathrm{Im}(e^{-2\pi\mathrm{i}z}(uv^* - wu^*))\end{aligned}$$

where the first term stems from the norm squared terms and the rest from the cross terms. Next we find

$$\begin{aligned}J_{12}(U_{x,z}) + J_{21}(U_{x,z}) &= |2\mathrm{i}csu + (\alpha^*)^2 c^2 v + \alpha^2 s^2 w|^2 + |-2\mathrm{i}csu + (\alpha^*)^2 s^2 v + \alpha^2 c^2 w|^2 \\ &= 8c^2 s^2 |u|^2 + (c^4 + s^4)(|v|^2 + |w|^2) \\ &\quad + 4c^2 s^2\,\mathrm{Re}(\alpha^4 v w^*) - 4cs(c^2 - s^2)\,\mathrm{Im}(u(\alpha^2 v^* - (\alpha^*)^2 w^*)),\end{aligned}$$

where the first two terms stem from the norm squared terms and the remaining ones from the cross terms. Using the trigonometric identities $\cos(\pi x)^4 + \sin(\pi x)^4 = 1 - \sin(2\pi x)^2/2$ and $\cos(\pi x)^2 - \sin(\pi x)^2 = \cos(2\pi x)$ and $\cos(\pi x)\sin(\pi x) = \sin(2\pi x)/2$ we obtain

$$\begin{aligned}J_{12}(U_{x,z}) + J_{21}(U_{x,z}) &= |v|^2 + |w|^2 + \sin(2\pi x)^2\big(2|u|^2 - \tfrac{|v|^2+|w|^2}{2}\big) \\ &\quad + \sin(2\pi x)^2\,\mathrm{Re}(\alpha^4 v w^*) - 2\sin(2\pi x)\cos(2\pi x)\,\mathrm{Im}(\alpha^2(uv^* + u^* w)).\end{aligned}$$

If we now consider a finite family of traceless Lindblad terms

$$V_k = \begin{pmatrix} u_k & v_k \\ w_k & -u_k \end{pmatrix},$$

and define $\Sigma = \sum_{k=1}^r |v_k|^2 + |w_k|^2$, and $\Delta = \sum_{k=1}^r |v_k|^2 - |w_k|^2$, and $\delta = \sum_{k=1}^r 2|u_k|^2 - (|v_k|^2 + |w_k|^2)/2$, as well as $r_1 e^{\mathrm{i}\phi_1} = \mathrm{i}\sum_{k=1}^r v_k w_k^*$ and $r_2 e^{\mathrm{i}\phi_2} = 2\sum_{k=1}^r (u_k v_k^* + u_k^* w_k)$ then we obtain

$$\begin{aligned}J_{12}(U_{x,z}) - J_{21}(U_{x,z}) &= \Delta\cos(2\pi x) + \sin(2\pi x)\textstyle\sum_{k=1}^r \mathrm{Im}(e^{-2\pi\mathrm{i}z}(u_k v_k^* - w_k u_k^*)) \\ J_{12}(U_{x,z}) + J_{21}(U_{x,z}) &= \Sigma + \delta\sin(2\pi x)^2 + r_1\sin(2\pi x)^2\sin(4\pi z + \phi_1) \\ &\quad - r_2\sin(2\pi x)\cos(2\pi x)\sin(2\pi z + \phi_2),\end{aligned}$$

since $\mathrm{Re}(\alpha^4 v_k w_k^*) = \mathrm{Re}(-\mathrm{i}r_1 e^{\mathrm{i}(4\pi z + \pi_1)}) = r_1\sin(4\pi z + \pi_1)$ and since $2\,\mathrm{Im}(\alpha^2(u_k v_k^* + u_k^* w_k)) = r_2\sin(2\pi z + \phi_2)$. Finally, since $\sum_{k=1}^r [V_k, V_k^*]$ is diagonal if and only if $\sum_{k=1}^r u_k v_k^* = \sum_{k=1}^r w_k u_k^*$, we obtain in this case that

$$J_{12}(U_{x,z}) - J_{21}(U_{x,z}) = \Delta\cos(2\pi x).$$

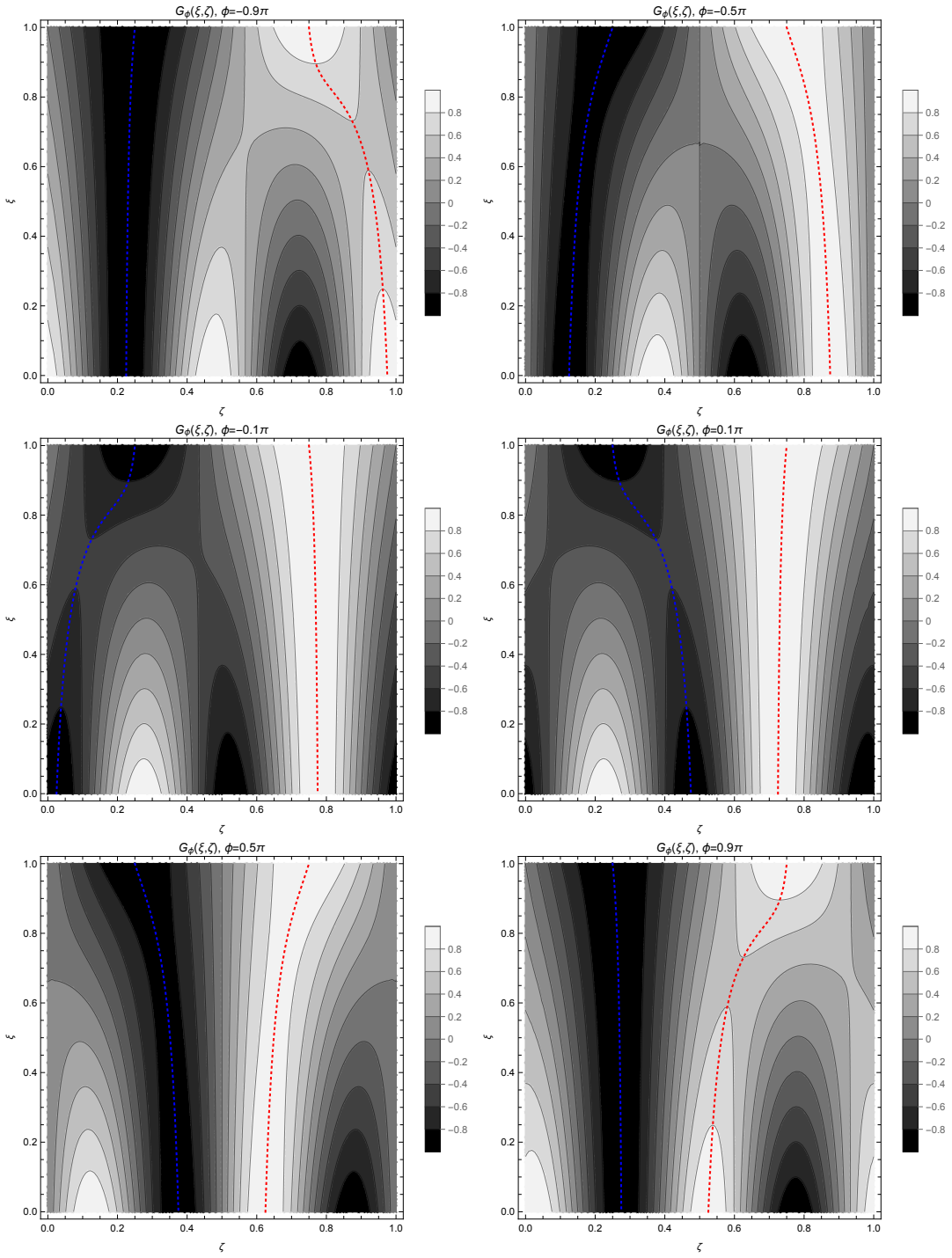

Figure 5.10: Contour plots of the function $G_\phi(\xi, \zeta)$ for different values of ϕ together with the maximizers (red) and minimizers (blue) as a function of ξ.

Study of the Function G_ϕ

In (5.4) we defined the function

$$G_\phi(\xi,\zeta) = (1-\xi)\sin(4\pi\zeta+\phi-\pi/2) - \xi\sin(2\pi\zeta),$$

for $\xi,\zeta \in [0,1]$ and $\phi \in (-\pi,\pi]$. See Figure 5.10 for plots of this function for different values of ϕ.

Lemma 5.B.1. *Let $\phi \in (-\pi,0)\cup(0,\pi)$ and define the function*

$$\xi^\star(\zeta) = \frac{1}{1+\frac{\cos(2\pi\zeta)}{2\cos(4\pi\zeta+\phi-\pi/2)}},$$

and the intervals

$$I^+ = \begin{cases} [3/4-\phi/(4\pi),3/4] \text{ if } \phi>0 \\ [3/4,3/4-\phi/(4\pi)] \text{ if } \phi<0, \end{cases} \qquad I^- = \begin{cases} [1/4,1/4+(\pi-\phi)/(4\pi))] \text{ if } \phi>0 \\ [-\phi/(4\pi),1/4] \text{ if } \phi<0. \end{cases}$$

Let ξ^+ and ξ^- denote the restrictions of $\xi^\star$ to I^+ and I^- respectively. These functions are bijective onto $[0,1]$, and it holds that

$$(\xi^+)^{-1}(\xi) = \operatorname*{argmax}_{\tilde\zeta\in[0,1]} G_\phi(\xi,\tilde\zeta), \qquad (\xi^-)^{-1}(\xi) = \operatorname*{argmin}_{\tilde\zeta\in[0,1]} G_\phi(\xi,\tilde\zeta),$$

for all $\xi\in[0,1]$. Here the argmax *and* argmin *are unique for all $\xi\in(0,1]$ and for $\xi=0$ there is a spurious second solution which we omit.*

The Lemma above runs into problems when $\phi\in\{0,\pi\}$, since in this case the functions ξ^+ and ξ^- are not surjective to $[0,1]$ anymore. However it turns out that in these cases we can compute the desired maximizers and minimizers explicitly.

Lemma 5.B.2. *If $\phi=0$ then we define $\zeta^+(\xi)\equiv 3/4$ and*

$$\zeta^-(\xi) = \begin{cases} \frac14 \pm (\frac{1}{2\pi}\arcsin\frac{1}{4(1/\xi-1)} - \frac14) & \text{if } \xi\in[0,4/5] \\ \frac14 & \text{if } \xi\in[4/5,1], \end{cases}$$

then $\zeta^+(\xi)$ is the unique maximizer of $G_0(\xi,\cdot)$ for all $\xi\in[0,1]$ and the two possibilities for $\zeta^-(\xi)$ are the only minimizers for $G_0(\xi,\cdot)$. Similarly, for $\phi=\pi$ we have $\zeta^-(\xi)\equiv 1/4$ and

$$\zeta^+(\xi) = \begin{cases} \frac34 \pm (\frac{1}{2\pi}\arcsin\frac{1}{4(1/\xi-1)} - \frac34) & \text{if } \xi\in[0,4/5] \\ \frac34 & \text{if } \xi\in[4/5,1]. \end{cases}$$

Ellipses

In this section we derive a useful polar coordinate parametrization of ellipses passing through the origin. For $a,b>0$ and $\phi_0\in(-\pi,\pi]$ we define the following parametrization of an ellipse:

$$E_{a,b,\phi_0}(\phi) = \begin{pmatrix} a(\cos(\phi)-\cos(\phi_0)) \\ b(\sin(\phi)-\sin(\phi_0)) \end{pmatrix}, \qquad \phi\in(-\pi,\pi].$$

This is simply an ellipse with semiaxes of length a and b, translated such that it intersect the origin when $\phi=\phi_0$, and the center of the ellipse is given by $(x_0,y_0) = (-a\cos(\phi_0),-b\sin(\phi_0))$.

Lemma 5.B.3. *In polar coordinates (r, ψ), where r is allowed to be negative, we can parametrize the ellipse E_{a,b,ϕ_0} as*

$$r(\psi) = -2\frac{\frac{\cos(\phi_0)\cos(\psi)}{a} + \frac{\sin(\phi_0)\sin(\psi)}{b}}{\frac{\cos(\psi)^2}{a^2} + \frac{\sin(\psi)^2}{b^2}}, \quad \psi \in (-\pi/2, \pi/2]. \tag{5.8}$$

Proof. Let $\psi \in (-\pi/2, \pi/2)$ be given. The corresponding slope is $m = \tan(\psi)$. We are looking for the intersections of the line $y = mx$ and the ellipse

$$\left(\frac{x - x_0}{a}\right)^2 + \left(\frac{y - y_0}{b}\right)^2 = 1.$$

Note that the origin is always in this intersection. If the line is tangent to the ellipse, this is the only intersection, otherwise there exists exactly one more. We plug in and using $x \neq 0$ we find

$$\left(\frac{1}{a^2} - \frac{m^2}{b^2}\right)x^2 + \left(\frac{-2x_0}{a^2} - \frac{2my_0}{b^2}\right)x + \left(\frac{x_0}{a}\right)^2 + \left(\frac{y_0}{b}\right)^2 = 1$$

and using the definition of (x_0, y_0) this implies that

$$x = 2\frac{\frac{x_0}{a^2} + \frac{my_0}{b^2}}{\frac{1}{a^2} + \frac{m^2}{b^2}}.$$

The distance r of the intersection to the origin can be found by computing

$$r^2 = (1 + m^2)x^2 = 4\left(\frac{\frac{x_0\cos(\psi)}{a^2} + \frac{y_0\sin(\psi)}{b^2}}{\frac{\cos(\psi)^2}{a^2} + \frac{\sin(\psi)^2}{b^2}}\right)^2,$$

and since by the choice of the range of ψ it holds that $\operatorname{sign}(r) = \operatorname{sign}(x)$, we see that

$$r = 2\frac{\frac{x_0\cos(\psi)}{a^2} + \frac{y_0\sin(\psi)}{b^2}}{\frac{\cos(\psi)^2}{a^2} + \frac{\sin(\psi)^2}{b^2}},$$

and by continuity this formula remains true for $\psi = \pm\pi/2$, and this concludes the proof. □

Note that the ellipse can furthermore be rotated around the center by shifting the angular coordinate, i.e., $\psi \mapsto r(\psi - \theta)$. In fact there is a unique angle $\theta^\star$ (modulo 2π), such that the ellipse lies in the upper halfplane, and the parametrization takes on a simplified form.

Lemma 5.B.4. *Let $a, b > 0$ and $\phi \in (-\pi, \pi]$ be given and let $E_{a,b,\phi}$ be the corresponding ellipse, with polar parametrization $r(\psi)$ as in (5.8). Then*

$$\theta^\star = \arctan(b\cos(\phi), a\sin(\phi)) \in (-\pi, \pi]$$

is the unique angle in $(-\pi, \pi]$ such that $\psi \mapsto r(\psi - \theta^\star)$ takes image in the upper halfplane. Moreover it holds that

$$r(\psi - \theta^\star) = \frac{\sigma\sin(\psi)}{\alpha + \beta\cos(\psi)^2 + \gamma\sin(\psi)\cos(\psi)} \tag{5.9}$$

where

$$\alpha = \left(\frac{\cos(\theta^\star)}{b}\right)^2 + \left(\frac{\sin(\theta^\star)}{a}\right)^2 \qquad \beta = \left(\frac{1}{a^2} - \frac{1}{b^2}\right)\cos(2\theta^\star)$$
$$\gamma = \left(\frac{1}{a^2} - \frac{1}{b^2}\right)\sin(2\theta^\star) \qquad \sigma = -2\left(\frac{1}{a}\cos(\phi)\sin(\theta^\star) + \frac{1}{b}\sin(\phi)\cos(\theta^\star)\right).$$

Proof. If we expand the numerator of $r(\psi - \theta)$ we obtain a linear combination of $\cos(\psi)$ and $\sin(\psi)$. For the ellipse to lie in the upper or lower halfplane, the coefficient of $\cos(\psi)$ must be zero. This gives the condition

$$\tan(\theta) = \frac{b\cos(\phi)}{a\sin(\phi)}.$$

This defines θ only modulo π. To make sure that the ellipse is in the upper halfplane, note that the function which maps (a, b, ϕ) to the function $\psi \mapsto r(\psi - \theta^\star)$ is continuous, and the parameter space of (a, b, ϕ) is connected, hence the ellipse will always lie in the same halfspace. Thus it suffices to check one ellipse, e.g., $a = b = 1$ and $\phi = 0$. The remainder of the proof is a straightforward computation using elementary trigonometric identities which we will omit. □

Next we want to find a, b, and ϕ given a parametrization as in (5.9). We start with a special case.

Lemma 5.B.5. *Let $\alpha, \beta, \gamma, \sigma \in \mathbb{R}$ with $\alpha \neq 0$ and consider the parametrization*

$$r(\psi) = \frac{\sigma\sin(\psi)}{\alpha + \beta\cos(\psi)^2 + \gamma\sin(\psi)\cos(\psi)}.$$

Without loss of generality we may assume that[6] *$\beta^2+\gamma^2 = 1$ and $\sigma > 0$. Then, this is the parametrization of an ellipse if and only if $2\alpha + \beta \notin [-1, 1]$, which corresponds to the denominator being non-zero for all ψ. The ellipse lies in the upper halfplane if and only if $2\alpha+\beta > 1$. In this case, the ellipse is exactly $E_{a/s,b/s,\phi}$ where*

$$\theta = \frac{1}{2}\arctan(\gamma, \beta), \quad a = \frac{\sigma}{\sqrt{|\alpha - \sin^2(\theta)|}}, \quad b = \frac{\sigma}{\sqrt{|\alpha + \cos^2(\theta)|}}$$
$$\phi = \arctan(b\cos(\theta), a\sin(\theta)), \quad s = \frac{1}{a}\cos(\phi)\sin(\theta) + \frac{1}{b}\sin(\phi)\cos(\theta).$$

Proof. This can be verified by plugging the values into Lemma 5.B.4. □

Upright Ellipsoids

Next we address the case of ellipsoids in three dimensions. We will only consider ellipsoids in the upper half space intersecting the origin and whose center lies on the z-axis, and we will say that they are *upright*. As a consequence of Lemma 5.B.5 we find the parametrization of ellipses which are upright in the analogous sense.

Corollary 5.B.6. *Let $\alpha, \beta, \sigma \in \mathbb{R}$ with $\sigma > 0$ and $\alpha \neq 0$ as well as $\alpha > \max(0, -\beta)$. Then the curve given in polar coordinates by*

$$r(\psi) = \frac{\sigma\sin(\psi)}{\alpha + \beta\cos(\psi)^2}$$

parametrizes the upright ellipse $E_{a,b,0}$ with

$$a = \frac{\sigma}{2|\alpha|}, \quad b = \frac{\sigma}{2\sqrt{|\alpha(\alpha + \beta)|}}.$$

Lemma 5.B.7. *Let $\alpha, \beta, \eta, \sigma \in \mathbb{R}$ with $\sigma > 0$ and $\alpha \neq 0$. Then the surface parametrized by*

$$r(\theta, \phi) = \frac{\sigma\cos(\theta)}{\alpha + (\beta + \eta\cos(2\phi))\sin(\theta)^2}$$

[6] If $\beta = \gamma = 0$ then the parametrization is that of a circle of diameter σ/α.

is an upright ellipsoid with axes

$$a = \frac{\sigma}{2\sqrt{|\alpha(\alpha+\beta+\eta)|}}, \quad b = \frac{\sigma}{2\sqrt{|\alpha(\alpha+\beta-\eta)|}}, \quad c = \frac{\sigma}{2|\alpha|}.$$

CHAPTER 6

Optimal Cooling

6.1 Introduction

Cooling quantum mechanical systems to a well-defined ground state is an essential task in quantum information technologies such as quantum computing [DiV00]. For trapped ions or cold atoms, the preferred method is laser cooling, such as Doppler cooling, Sisyphus cooling [WDC92], Raman cooling [Mon+95] and many others, as well as using strong coupling [Mac+10] or aided by optimal control [Li+21]. Another popular approach is algorithmic cooling [Boy+02, SMW05, PVC06, ALP19].

Building on the methods and results of the previous chapters, in this chapter we use the reduced control system together with optimal control theory to derive provably time-optimal schemes for cooling Markovian quantum systems with fast unitary control.

Many results have been derived in the qubit case, see Chapter 5 and references therein. Some results for qutrits were obtained in [STK04] using the reduced control system. A system of two coupled spins with control on one of them was considered in [Bas+21].

We recover and generalize the results from [STK04], and we obtain optimal cooling solutions for a certain four-level system. All solutions are obtained using purely analytical means and in each case we prove that the given solution is indeed time-optimal. Notably our time-optimal controls are obtained *without* invoking the Hamilton–Jacobi–Bellman (HJB) Equation or the Pontryagin Maximum Principle (PMP) . This is made possible by using a specialization of the Majorization Theorem 2.5.3, which allows to significantly simplify the search for optimal controls.

Outline

The chapter is organised as follows: Section 6.2 recalls the characterization of coolable systems in simple algebraic terms. The methods are illustrated in Section 6.3 by solving the rank one qubit case is detail. Section 6.4 goes on to study achievable derivatives in higher dimensions and introduces two systems to be solved in the subsequent sections. Section 6.5 presents a method of reducing the achievable derivatives to a subset of optimal derivatives, and finally Section 6.6 determines optimal solutions to the higher dimensional systems introduced in Section 6.4.

6.2 Asymptotically Coolable Systems

Since we assume to have fast unitary control over the system, any pure state can be transformed into any other pure state at no cost. Hence, any pure state can also be transformed into the ground state of the system Hamiltonian. For this reason we will equate cooling the system with purifying it. Moreover, due to the exponential nature of the Lindblad equation, pure states can only be reached asymptotically. Thus one has to clarify what is meant by "cooling the system in the shortest amount of time possible". In particular one has to define a cost (resp. reward) function. The task then becomes to minimize (resp. maximize) this measure in a given time, or to reach a certain value in the shortest possible amount of time.

Some reasonable examples of such measures include the purity of the state, the von Neumann entropy, the largest eigenvalue (which is the maximum fidelity with a pure state) or the minimum energy (with respect to some Hamiltonian with non-degenerate ground state). An important property of these functions is that they are Schur-convex (or concave), i.e. they are monotone with respect to majorization, which will significantly simplify the search for optimal solutions. The details are given in Section 6.5.

The first question to ask is whether the system under consideration is coolable at all, by which we mean that a pure quantum state can be reached from every given initial state, at least in an asymptotic sense. Luckily, we already fully characterized coolability in Theorem 4.3.7. Put simply, the system is asymptotically coolable if and only if there exists a common eigenvector of all V_k which is not a common left eigenvector. Recall also from Appendix 4.C that there is an efficient algorithm for finding such a common eigenvector, if it exists.

In the following we will only deal with asymptotically coolable systems. If the system is not coolable, one first has to determine which states are reachable, and which of them is the coolest by some appropriate measure. The question of reachability was discussed in Chapter 4.

6.3 Optimal Cooling of a Qubit

The simplest special case is that of a single qubit. In this section we focus on rank one systems, which are defined by a single Lindblad term V. These systems are not trivial, but they still allow for a complete description. The general qubit case (including non-coolable systems) was treated in Chapter 5, with coolable systems already addressed in Section 5.6. The solution obtained here shares some similarity with the solution obtained for the special case of the Bloch equations, see Appendix 5.A, which is however simpler to solve as it has a symmetry which allows to reduce the dimensionality of the problem. For later use we define the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let $V \in \mathbb{C}^{2,2}$ be an arbitrary Lindblad term. First we have to check whether such a system is coolable at all. Indeed, it follows immediately from Theorem 4.3.7 that the system is asymptotically coolable if and only if V is not normal (equivalently if and only if $-L$ is not unital).

At first glance, there are eight real parameters defining the problem but the following result shows that all but one parameter can be eliminated.

Lemma 6.3.1. *Let $V \in \mathbb{C}^{2,2}$ be an arbitrary non-normal matrix. Then there exists a (special) unitary matrix $\tilde{U}$, a Hermitian matrix $\tilde{H}$ and numbers $\gamma > 0$ and $\nu \in [0, 1)$ with*

$$\tilde{H} = \tfrac{\mathrm{i}}{4}(\mathrm{tr}(V^*)V - \mathrm{tr}(V)V^*), \quad \tilde{V} = \begin{pmatrix} 0 & 1 \\ \nu & 0 \end{pmatrix},$$

such that $\Gamma_V = \mathrm{i}\,\mathrm{ad}_{\tilde{H}} + \gamma\Gamma_{\tilde{U}^*\tilde{V}\tilde{U}}$.

Proof. Let $\tilde{U}$ be a unitary such that $\tilde{U}[V, V^*]\tilde{U}^*$ is diagonal. Then it is easy to show that $\tilde{U}(V - \mathrm{tr}(V)\mathbb{1}/2)\tilde{U}^*$ is zero on the diagonal. By adjusting $\tilde{U}$ (without renaming) we can additionally make sure that the off-diagonal elements have the same argument and the top right element has the greater modulus. Together this gives $V - \mathrm{tr}(V)\mathbb{1}/2 = \sqrt{\gamma}e^{\mathrm{i}\phi}\tilde{U}^*\tilde{V}\tilde{U}^*$. The freedom of representation of the Lindblad equation (Lemma 4.A.3) then implies $\Gamma_V = \mathrm{i}\,\mathrm{ad}_{\tilde{H}} + \Gamma_{V-\mathrm{tr}(V)\mathbb{1}/2} = \mathrm{i}\,\mathrm{ad}_{\tilde{H}} + \gamma\Gamma_{\tilde{U}^*\tilde{V}\tilde{U}}$ as desired. □

There are two extremal cases. If $\nu = 0$ we obtain a special case of the Bloch equations, and if $\nu = 1$, the matrix V is normal and hence the system is not coolable. For this reason we exclude the case $\nu = 1$.

Space of Generators

From now on we assume that we have a single Lindblad term of the form $\tilde{V}$ as in Lemma 6.3.1 depending only on the parameter $\nu \in [0, 1)$. The general solution will then be recovered at the end, see Remark 6.3.8. All figures in this section use the value $\nu = 1/2$.

Recall from Chapter 5 that in the qubit case the generators $-L_U$ are defined by two non-negative real numbers on the off diagonal and hence they can be easily visualized. Indeed, the main tool in the following will be the *space of generators* $\mathfrak{Q}$ which is linearly isomorphic to the set of all $-L_U$:

$$\mathfrak{Q} = \{(J_{12}(U) - J_{21}(U), J_{12}(U) + J_{21}(U)) : U \in \mathrm{SU}(2)\}.$$

For rank one systems $\mathfrak{Q}$ takes on a rather simple form, see Figure 6.1.

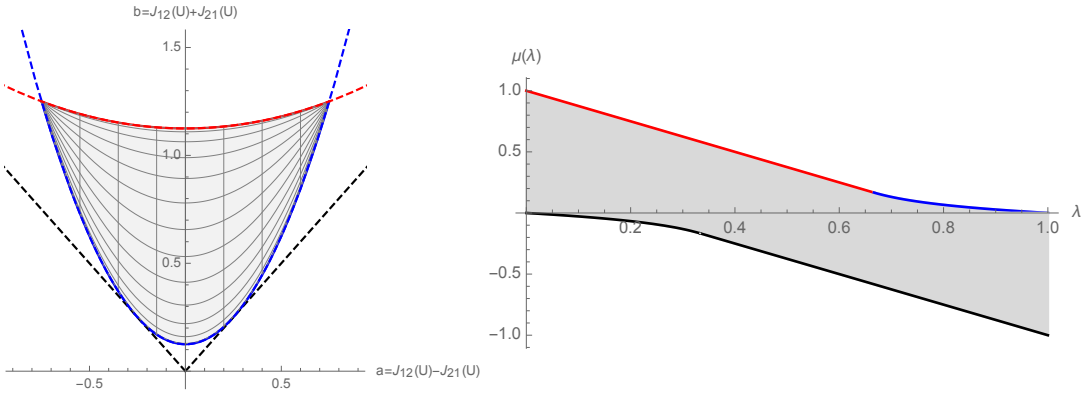

Figure 6.1: Left: The parametrized space of generators $\mathfrak{Q}$ of a rank one system as given in Lemma 6.3.2 and Corollary 6.3.3. The poles of the Bloch sphere are mapped to the corners $(\pm(1-\nu^2), 1+\nu^2)$. The latitude lines are vertical, and the equator lies on the b-axis. The longitude lines are parabolas passing through the poles and intersecting the b-axis between $b = \frac{1}{2}(1 \pm \nu)^2$. Right: Achievable derivatives as a set-valued function of λ with the upper bound μ, given in Lemma 6.3.5, highlighted.

Lemma 6.3.2. *The space of generators can be parametrized as* $\mathfrak{Q} = \{(a, f_z(a)) : a \in [-(1-\nu^2), 1-\nu^2],\ z \in [0, 1)\}$, *where*

$$f_z(a) = 1 + \nu^2 - \frac{1 + \nu^2 - 2\cos(4\pi z)\nu}{2}\left(1 - \left(\frac{a}{1-\nu^2}\right)^2\right),$$

In fact, the point $(a, f_z(a))$ *can be obtained using the unitary* $U_{x,y} = \mathrm{e}^{\mathrm{i}\pi z\sigma_z}\mathrm{e}^{\mathrm{i}\pi x\sigma_x}$ *where* x *satisfies* $1 - a/(1-\nu^2) = 2\sin(\pi x)^2$. *The lower boundary is achieved for* $z = \frac{1}{4}$.

Proof. With $U_{x,y}$ as above we find $(a,b) \in \mathfrak{Q}$ with $a = (1-2r)(1-\nu^2)$, and $b = 1+\nu^2+2r(r-1)|1-e^{4i\pi z}\nu|^2$, where $r = \sin(\pi x)^2$. Hence with some basic trigonometry we get $b = 1+\nu^2-\frac{1}{2}|1-e^{4i\pi z}\nu|^2\big(1-\frac{a^2}{(1-\nu^2)^2}\big)$. □

It is of particular importance to understand the boundary of this set.

Corollary 6.3.3. *For V non-normal, the space of generators is the region enclosed between the two parabolic segments*

$$a \mapsto 1+\nu^2-\tfrac{1}{2}(1\pm\nu)^2\big(1-\big(\tfrac{a}{1-\nu^2}\big)^2\big),$$

on $a \in [-1+\nu^2, 1-\nu^2]$.

Remark 6.3.4. *The two extremal parabolas of Corollary 6.3.3 (considered on $\mathbb{R}$) are the unique parabolas passing through the points $(\pm(1-\nu^2), 1+\nu^2)$ which are tangent to the lines $a \mapsto \pm a$. The points of tangency in $\mathfrak{Q}$ are achieved when $U^*\tilde{V}U$ is upper or lower triangular. Note that for $\nu = 0$ the two parabolas coincide (which is consistent with the Bloch case), and as $\nu \to 1$ all points tend to the y-axis which is consistent with unital systems, cf. Appendix 5.A.*

Optimal Derivatives and Path

In the qubit case it is convenient to represent the reduced state by the first eigenvalue $\lambda \in [0,1]$. The maximal achievable derivative of λ, denoted by $\mu : [0,1] \to \mathbb{R}$, can be obtained from the boundary of $\mathfrak{Q}$ via the relation

$$\mu(\lambda) = \max_{(a,b)\in\mathfrak{Q}} \tfrac{1}{2}(a+(1-2\lambda)b).$$

Details are given in Chapter 5. A computation then yields the following result, see also Figure 6.1 for an illustration.

Lemma 6.3.5. *Let V be non-normal and $\lambda_0 = \frac{1}{2}(1+\frac{1-\nu}{1+\nu})$, then the optimal derivative $\mu : [0,1] \to \mathbb{R}$ takes the form*

$$\mu(\lambda) = \begin{cases} \frac{1}{2}(1-\nu^2-(1+\nu^2)(2\lambda-1)) & \text{if } 0 \le \lambda \le \lambda_0 \\ (\frac{1-\nu}{2})^2(\frac{1}{2\lambda-1}+2\lambda-1) & \text{if } \lambda_0 \le \lambda \le 1. \end{cases}$$

The function μ is continuously differentiable.

From this the optimal path through the Bloch ball can be computed.

Lemma 6.3.6. *Let $t_0 = \log\big(1-\frac{1+\nu^2}{1+\nu}\big)/(1+\nu^2)$. The optimal path through the Bloch ball is given by $\rho^\star(t) = \mathrm{Ad}_{U^\star(t)}(\mathrm{diag}(\lambda^\star(t)))$ with $U^\star(t) = e^{i\pi y^\star(t)\sigma_y}$ and where*

$$\lambda^\star(t) = \begin{cases} \frac{1-e^{-(1+\nu^2)t}}{1+\nu^2} & 0 \le t \le t_0 \\ \frac{1}{2}\big(1+\sqrt{1-ce^{-(1-\nu)^2 t}}\big) & t \ge t_0 \end{cases}$$

with $c = \frac{4\nu}{(1+\nu^2)}\big(\frac{1+\nu}{\nu(1-\nu)}\big)^{\frac{(1-\nu)^2}{(1+\nu)^2}}$, is the unique solution to $\frac{d}{dt}\lambda^\star(t) = \mu(\lambda^\star(t))$ with $\lambda^\star(0) = 0$ and the (continuous) function $y^\star$ is defined by

$$y^\star(t) = \begin{cases} 0 & 0 \le t \le t_0 \\ \frac{1}{\pi}\arcsin\big(\sqrt{\frac{1}{2}(1+\frac{1-\nu}{1+\nu}\frac{1}{1-2\lambda^\star(t)})}\big) & t \ge t_0. \end{cases}$$

Proof. The function $\lambda^\star$ is found by integration and can be checked by differentiating, and $y^\star$ as a function of $\lambda^\star$ is found by computing the values of x and z (cf. Lemma 6.3.2) which give the boundary point of $\mathfrak{Q}$ corresponding to λ. □

The optimal path is illustrated in Figure 6.2. Note that the path is much simpler than one might expect from the formulas.

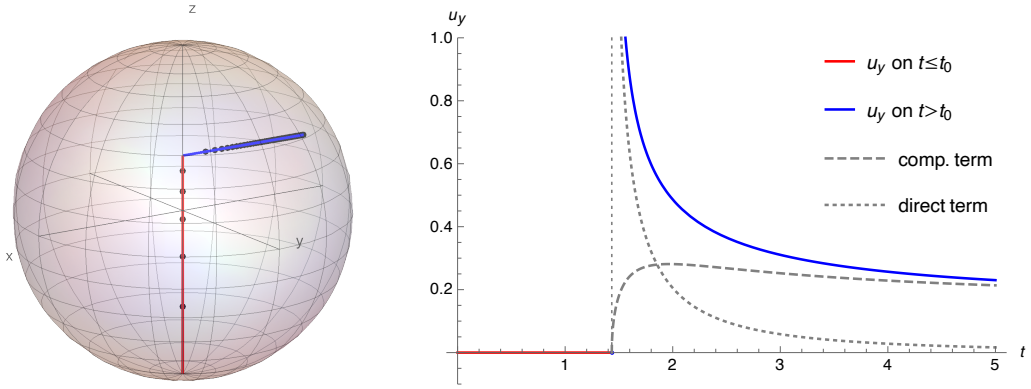

Figure 6.2: Left: Time-optimal path from the boundary of the Bloch ball (pure state) to the center (maximally mixed state) and back. Starting at the south pole, the path follows the z-axis until $\lambda = \lambda_0$. Then the path takes a sharp turn and continues horizontally until it reaches the boundary (which happens only asymptotically). When projected onto the xy-plane, the horizontal part is a straight line lying on the negative x-axis (the mirrored path along the positive x-axis is also optimal). The black dots are equally spaced in time, and accumulate towards the end. The solution shares some similarity with the so-called magic plane result for the Bloch equations obtained in [Lap+10]. Right: Optimal control function u_y (solid) with direct (dotted) and compensation (dashed) contributions using $\nu = 1/2$. The control is identically zero on $[0, t_0]$ and has a singularity at t_0.

Optimal Controls

So far we have found the optimal derivatives of λ and the optimal path of ρ through the Bloch ball. It remains to determine the corresponding optimal controls of the full control system ($\mathcal{D}$). To simplify the problem we assume that the control Hamiltonians are the Pauli matrices and the goal is to determine the corresponding control functions u_x, u_y and u_z.

There are two contributions to the control Hamiltonians. A *direct* term obtained by differentiating the optimal control unitary of the reduced system and a *compensating* term which cancels out the motion tangent to the unitary orbits induced by the drift $-L$, see Proposition 2.3.10. This leads to the following result:

Proposition 6.3.7. *A choice of optimal controls is given by $u_x = u_z \equiv 0$ on $[0, \infty)$. Moreover $u_y \equiv 0$ on $[0, t_0]$ and*

$$u_y(t) = \frac{\frac{1}{2\lambda^\star(t)-1}\frac{1-\nu}{1+\nu}}{\sqrt{(2\lambda^\star(t)-1)^2 - (\frac{1-\nu}{1+\nu})^2}} \cdot \Big(\frac{1-\nu}{2}\Big)^2 \Big(\frac{1}{2\lambda^\star(t)-1} - (2\lambda^\star(t)-1)\Big)$$
$$+ \frac{(1+\nu)\sin(2\pi y^\star(t))}{4}\Big(\frac{2(1-\nu)}{2\lambda^\star(t)-1} + (1+\nu)\cos(2\pi y^\star(t))\Big),$$

on $t \in [t_0, \infty)$ and where $\lambda^\star(t)$ and $y^\star(t)$ are as in Lemma 6.3.6.

Proof. The direct term is $\left(\frac{d}{dt}U_{y^\star(t)}\right)U^*_{y^\star(t)}$ where $U_y = e^{\mathrm{i}\pi y\sigma_y}$ and $y^\star(t)$ is as in Lemma 6.3.6. Using the chain rule this becomes $\mathrm{i}\pi\sigma_y \frac{dy^\star}{d\lambda^\star}(\lambda^\star(t))\mu(\lambda^\star(t))$. The expression follows directly from the evaluation of the derivative. The compensation term takes the form $\mathrm{ad}^+_\rho(L(\rho))$, cf. Proposition 2.3.10, where $\mathrm{ad}_\rho(\cdot) = [\rho, \cdot]$ and $(\cdot)^+$ is the Moore–Penrose pseudo inverse. The result then follows from an elementary computation. □

The obtained optimal controls are illustrated in Figure 6.2.

Remark 6.3.8. *For general non-normal V and with initial state is $|0\rangle\langle 0|$ the optimal controls are the following, using the notation from Lemma 6.3.1. First (almost) instantaneously apply the unitary $\tilde{U}$ to the system. Then apply the control Hamiltonian*

$$H(t) = -\tilde{H} + \gamma u_{\nu,y}(\gamma t)\, \tilde{U}\sigma_y\tilde{U}^*$$

for $t \in [0, \infty)$.

6.4 Achievable Derivatives and Speed Limits

Locally the reduced control system can be understood by studying the set of *achievable derivatives* at λ, denoted $\mathsf{derv}(\lambda) := \{-L_U\lambda : U \in \mathrm{SU}(n)\}$. Due to continuity of the map $U \mapsto -L_U\lambda$ it is clear that the set is compact and path-connected, but the exact shape is difficult to determine in general.

In the qubit case studied above, $\mathsf{derv}(\lambda)$ was just a closed interval, and we were able to give an analytical expression. For more general qubit systems this task becomes more difficult, but it still allows for a partial analytical solution, see Chapter 5. In higher dimensional systems the shape of $\mathsf{derv}(\lambda)$ can be quite arbitrary, but in some special cases it takes the form of a (convex) polytope, which can be seen as a generalization of the qubit case. In general however this is not true, although it is still useful to approximate $\mathsf{derv}(\lambda)$ with polytopes, both from the inside and the outside.

In the remainder of this section we provide some results and examples in this direction. Note that due to the Relaxation Theorem [AC84, Ch. 2.4, Thm. 2] (cf. the relaxed control system ($\bar{\Lambda}$)) we are also interested in the convex hull of $\mathsf{derv}(\lambda)$.

Examples of Polytopes

We present a few cases where the set of achievable derivatives $\mathsf{derv}(\lambda)$ takes the form of a polytope, and some examples where $\mathsf{derv}(\lambda)$ is not convex and $\mathrm{conv}(\mathsf{derv}(\lambda))$ is not a polytope.

First, at the maximally mixed state $\mathbf{e}/n$, the set $\mathsf{derv}(\mathbf{e}/n)$ is always a convex polytope. In fact the vertices are the vectors containing the eigenvalues of $\sum_{k=1}^r [V_k, V_k^*]$ in all possible permutations. This can be shown using Lemma 3.0.7 together with the Schur–Horn Theorem [Sch23, Hor54].

Another special case occurs at the vertices e_i of the simplex Δ^{n-1} under the assumption that there is only one Lindblad term V. Indeed, there exists a value $0 < f^\star \leq \|V\|_\infty$, where $\|\cdot\|_\infty$ is the Schatten ∞-norm (i.e. the largest singular value), such that

$$\mathsf{derv}(e_i) = f^\star \,\mathrm{conv}(0, e_j - e_i : j \neq i).$$

Let $f(U)$ be the sum of the squares of the off-diagonal elements in the first column of $J(U)$. Then $f^\star$ is the maximal value of $f(U)$, and the value 0 is achievable using the Schur decomposition. Using unitaries leaving the first basis vector invariant the entire polytope can be obtained.

It is important to note that generically $\mathsf{derv}(\lambda)$ is not a polytope. For a simple counterexample consider the Lindblad term $V = |1\rangle\langle 2| + \sqrt{2}\,|2\rangle\langle 3|$ on a qutrit. A numerical computation shows that for general λ the set $\mathsf{derv}(\lambda)$ is not convex.

In the remainder of this section we introduce two concrete examples where the polytope structure of $\mathsf{derv}(\lambda)$ occurs, and they will serve as running examples in the following sections.

Qutrit Systems with Spontaneous Emission

Since the qubit case is addressed in detail in Chapter 5, the next logical step is the qutrit. We consider a special form of system with only spontaneous emissions as described in [STK04] and which includes the Λ-system. Such systems are defined by Lindblad terms of the form $\sqrt{\eta_{ij}}\,|i\rangle\langle j|$ for $i, j \in \{1, \dots, n\}$. These systems have the property that $J(U) = \Theta^\top \Gamma \Theta$ where $\Gamma = J(\mathbb{1})$ has entries η_{ij} and Θ is the unistochastic matrix defined by $\Theta_{ij} = |U_{ij}|^2$. We focus on the following two systems:

$$\Gamma^\Lambda = \begin{pmatrix} 0 & 0 & 0 \\ \gamma_1 & 0 & 0 \\ \gamma_2 & 0 & 0 \end{pmatrix}, \quad \Gamma^{\mathrm{V}} = \begin{pmatrix} 0 & \gamma_1 & \gamma_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The first one is the Λ-system studied in [STK04] and the second one is an inverted version, which we call the V-system. We now show that for a V-system $\mathsf{derv}(\lambda)$ is always a convex polytope.

Proposition 6.4.1. *For the* V*-system in three dimensions it holds that* $\mathsf{derv}(\lambda) = \mathrm{conv}(\{-L_P\lambda : P \in S_3\})$*, where* S_3 *is the symmetric group represented by permutation matrices.*

Proof. We find that the generator $-L_U$ equals the convex combination of generators $u_1\Gamma_1 + u_2\Gamma_2 + u_3\Gamma_3$ where the Γ_i are

$$\begin{pmatrix} 0 & \tilde{\gamma}_2 & \tilde{\gamma}_3 \\ 0 & -\tilde{\gamma}_2 & 0 \\ 0 & 0 & -\tilde{\gamma}_3 \end{pmatrix}, \begin{pmatrix} -\tilde{\gamma}_1 & 0 & 0 \\ \tilde{\gamma}_1 & 0 & \tilde{\gamma}_3 \\ 0 & 0 & -\tilde{\gamma}_3 \end{pmatrix}, \begin{pmatrix} -\tilde{\gamma}_1 & 0 & 0 \\ 0 & -\tilde{\gamma}_2 & 0 \\ \tilde{\gamma}_1 & \tilde{\gamma}_2 & 0 \end{pmatrix}$$

and where $(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3) = (\gamma_2, \gamma_1, 0)\Theta$.

It suffices to show that $\Gamma_1\lambda$ lies in the desired polytope. Consider the following linear bijection:

$$(\dot{\lambda}_1, \dot{\lambda}_2, \dot{\lambda}_3) \mapsto (x, y) := (-\dot{\lambda}_2/\lambda_2, -\dot{\lambda}_3/\lambda_3).$$

In these coordinates, the polytope intersected with the first quadrant is defined by

$$x, y \leq \max(\gamma_1, \gamma_2), \quad x + y \leq \gamma_1 + \gamma_2$$

and clearly $\Gamma_1\lambda$ satisfies this since $(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3) \preceq (\gamma_2, \gamma_1, 0)$ as Θ is bistochastic. □

One might be tempted to try and generalize this result to all qutrit systems with spontaneous emission, but unfortunately it fails already for the Λ-system, see Figure 6.3 and the following example.

Example 6.4.2. *Consider the* Λ*-system with* $\gamma_1 = \gamma_2 = 1$ *and let* $\lambda = (1, 0, 0)^\top$*. Using only permutations we obtain the derivatives* $(0, 0, 0)^\top$ *and* $(-2, 1, 1)^\top$*. However, a part of the boundary can be obtained by computing*

$$\begin{pmatrix} x & y & 0 \\ y & x & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x & y & 0 \\ y & x & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} * & y^2 & 0 \\ xy & * & 0 \\ x & y & 0 \end{pmatrix},$$

where $x \in [0, 1]$ *and* $y = 1 - x$*. Clearly the resulting derivatives do not lie in the convex hull of* $(0, 0, 0)^\top$ *and* $(-2, 1, 1)^\top$*.*

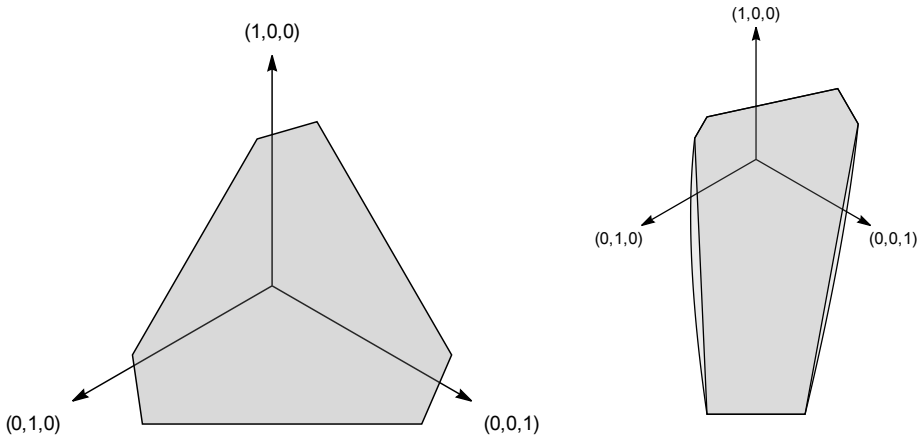

Figure 6.3: Left: Achievable derivatives in the V-system with $\gamma_1 = 1$ and $\gamma_2 = 2$ at the point $\lambda = (0.4, 0.35, 0.25)$. Right: Achievable derivatives in the Λ-system with $\gamma_1 = 1$ and $\gamma_2 = 2$ at the point $\lambda = (0.6, 0.25, 0.15)$. The corners are achieved by permutations, and there are bulges which leave the polytope but belong to $\mathsf{derv}(\lambda)$.

It is somewhat unexpected that the Λ-system does not have this polytope property, since the optimal cooling solution found in [STK04] only requires permutations. This is explained by the fact that the "non-classical" achievable derivatives are suboptimal for cooling, cf. Section 6.5.

Spin-Spin System

In this section we explore a system composed of two qubits with a single Lindblad term $V = \sigma_- \otimes \mathbb{1}$ where $\sigma_- = (\sigma_x - \mathrm{i}\sigma_y)/2$ is the lowering operator. This can be seen as a first approximation for the ubiquitous spin-boson system. We will conjecture an exact description of the (convex hull) of the achievable derivatives and support it with a partial proof and numerical evidence. In Section 6.4 we give a slightly larger upper bound with full proof and in Section 6.6 we derive an optimal cooling procedure for the system.

Conjecture 6.4.3. *For every $\lambda \in \Delta^3$ it holds that*

$$\mathrm{conv}(\mathsf{derv}(\lambda)) = \mathrm{conv}(\{-L_P\lambda : P \in S_4\}) =: \mathcal{P}(\lambda).$$

The $\supseteq$ direction is trivially satisfied, so it remains to show that the right-hand side is an upper bound of $\mathsf{derv}(\lambda)$. First we need to better understand the polytope $\mathcal{P}(\lambda)$. Indeed we can derive a simple inequality description of the polytope:

Lemma 6.4.4. *For regular λ, the polytope $\mathcal{P}(\lambda)$ has 12 vertices and 8 facets, 4 of which are hexagonal and the other 4 are triangular. The hexagonal facets are described by $\dot{\lambda}_i \geq -\lambda_i$. Let a be any of the four eigenvalues and $b \geq c \geq d$ be the remaining ones. Then*

$$\dot{a}(b+d) - \dot{b}(c-b) - b(c+d) \leq 0$$

describes the corresponding triangular facet.

The number of vertices stems from the fact that V is invariant under the permutation $(12)(34)$. The hexagonal facet inequalities are clearly satisfied, and so it remains to show that all achievable derivatives

also satisfy the inequalities corresponding to the triangular facets. So far we only have numerical evidence for this claim. Yet a slightly larger provable bound can be obtained by studying the set of matrices $J(U)$, cf. Section 6.4.

Reduction to Toy Model

We say that $-L$ is *quasi-classical* if for all $\lambda \in \Delta^{n-1}$ it holds that $\mathsf{derv}(\lambda) \subseteq \mathrm{conv}(\{-L_P\lambda : P \in S_n\})$. Hence, by Proposition 6.4.1, the V-system on the qutrit is quasi-classical, but the qubit system studied in Section 6.3 is not. Such quasi-classical systems are particularly nice, because the controls in the reduced control system can be restricted to the permutations (up to convexification).

Recall that in Appendix 4.B we studied the toy model with exactly these controls, see also [3, DES19, End20, SED22]. Note however that these systems are not quasi-classical, meaning that "non-classical" controls might improve the results obtained therein.

On the other hand, these systems have the convenient property that diagonal density matrices remain diagonal, and hence no compensating Hamiltonian is necessary. We will call such systems *diagonally invariant*. Indeed this follows from the formula of the compensating Hamiltonian given in Section 4.2. More generally, a system is diagonally invariant if for all Lindblad terms V_k the associated directed graph (with arcs corresponding to non-zero matrix entries) are disjoint unions of directed paths and directed cycles. In particular, systems with spontaneous emission and the spin-spin system introduced in the previous section have the property that diagonal states remain diagonal.

Upper Bounds and Speed Limits

Working with the toy model, i.e. restricting the controls of the reduced control system to permutations, is a practical way of simplifying the system and allows for the computation of solutions which are not necessarily optimal. In effect, this method approximates $\mathrm{conv}(\mathsf{derv}(\lambda))$ from the inside. Conversely, this section is concerned with approximations from the outside. This is useful as it yields speed limits and upper bounds to optimal solutions.

First we look at some general results before considering specific systems. Useful bounds can be obtained on the level of the $J(U)$ matrices. A somewhat trivial bound can be obtained by setting $\gamma = \sum_{k=1}^{r} \|V_k\|_2^2$. Then all matrices $J = J(U)$ satisfy the inequalities $\sum_{i,j=1}^{n} J_{ij} \leq \gamma$ and $J_{ij} \geq 0$. Using Lemma 3.0.7 we get the following stronger bound.

Lemma 6.4.5. *Let $J = J(U)$, then $J_{ij} \geq 0$ as well as $J\mathbf{e} \preceq \mathrm{spec}(\sum_{k=1}^{r} V_k V_k^*)$, and $J^\top \mathbf{e} \preceq \mathrm{spec}(\sum_{k=1}^{r} V_k^* V_k)$, and*

$$(J + J^\top)\mathbf{e} \preceq \mathrm{spec}(\textstyle\sum_{k=1}^{r} \{V_k, V_k^*\}), \quad (J - J^\top)\mathbf{e} \preceq \mathrm{spec}(\textstyle\sum_{k=1}^{r} [V_k, V_k^*]),$$

and this defines a polytope bound for the set of all $J(U)$.

Here again, $\preceq$ denotes majorization, cf. Section 6.5. Note that by linearity these also define corresponding polytope bound for each $\mathsf{derv}(\lambda)$. An important property of these bounds is that they guarantee that the simplex Δ^{n-1} is preserved, and hence they don't lead to unphysical behavior.

For the spin-spin system we have, so far, only conjectured a polytope bound on $\mathsf{derv}(\lambda)$. By studying the set of matrices $J(U)$, we can prove a slightly larger bound.

Lemma 6.4.6. *Consider the spin-spin system. For every matrix $J = J(U)$ it holds that*

$$J_{ij} \geq 0, \quad (J + J^\top)\mathbf{e} = \mathbf{e}, \quad J_{ii} \leq \tfrac{1}{4}$$

for all $i, j \in \{1, \ldots, n\}$.

Proof. The first two constraints follow immediately from Lemma 6.4.5 and its proof (cf. Lemma 3.0.7), and the last follows from $J(U)_{ii} = |\bar{u}_{1i}u_{3i} + \bar{u}_{1i}u_{3i}|^2 \leq \frac{1}{4}$. □

Note that by strengthening the latter inequalities to $J_{ii} = 0$, we indeed obtain an exact description of $\mathrm{conv}(J(P) : P \in S_4)$. While the polytope bound from Lemma 6.4.6 is always larger than the conjectured bound, numerical results indicate that for different values of λ its volume is only 0–10% larger.

6.5 Optimal Derivatives

In the qubit case the reduced state space $[0, 1] \cong \Delta^1$ is one-dimensional, and hence there is only one optimal derivative for cooling (resp. heating). In higher dimensions this is not true anymore. In this section we show that the set of achievable derivatives $\mathsf{derv}(\lambda)$ can be reduced to a subset of optimal derivatives, which typically still consists of more than one element.

An important way of comparing two mixed quantum states, or rather their eigenvalues, is called majorization. This approach is frequently used in quantum thermodynamics [BŻ17, HO13]. First one defines majorization on vectors [MOA11]. Let $\lambda, \mu \in \Delta^{n-1}$ be given. Then λ is said to *majorize* μ, denoted $\lambda \succeq \mu$ if

$$\sum_{i=1}^{k} \lambda_i^{\downarrow} \geq \sum_{i=1}^{k} \mu_i^{\downarrow}, \quad k = 1, \ldots, n,$$

where $\lambda_i^{\downarrow}$ is the i-th largest element of λ and analogously for μ. The notion carries over to quantum states by defining that a state majorizes another if its eigenvalues majorize those of the other state. A function f is *Schur-convex* if $\lambda \succeq \mu$ implies $f(\lambda) \geq f(\mu)$, and it is *Schur-concave* if its negative is Schur-convex. Due to fast unitary control we consider cost functions which depend only on the eigenvalues of the state. Indeed, all the functions given in Section 6.2 are Schur-convex or Schur-concave functions of the eigenvalues. This follows from the fact that these functions are convex or concave and invariant under permutations. The following result, which specializes Theorem 2.5.3, shows that for the purpose of cooling, a state which majorizes another is always better.

Theorem 6.5.1. *Let $\mu : [0, \infty) \to \Delta^{n-1}$ be a solution to the relaxed control system ($\bar{\Lambda}$) and let $\lambda_0 \in \Delta^{n-1}$ such that $\mu(0) = \mu_0 \preceq \lambda_0$. Then there exists a solution $\lambda : [0, \infty) \to \Delta^{n-1}$ to ($\bar{\Lambda}$) with $\lambda(0) = \lambda_0$ such that $\mu(t) \preceq \lambda(t)$ for all $t \in [0, \infty)$.*

Just like states are compared via majorization, two derivatives, at the same state, can be compared using an infinitesimal version of majorization, also called unordered majorization, cf. [MOA11, Ex. 14.E.6]. Let $v, w \in \mathbb{R}_0$ be two tangent vectors (derivatives) at $\lambda \in \Delta^{n-1}$. Then we say that v *infinitesimally majorizes* w, denoted $v \trianglerighteq w$ if

$$\sum_{i=1}^{k} v_i \geq \sum_{i=1}^{k} w_i, \quad k = 1, \ldots, n.$$

If $\mathsf{derv}(\lambda)$ is a polytope, the subset of optimal derivatives takes a nice form.

As a consequence of Theorem 6.5.1 one can show that derivatives which are not majorized by any other derivatives are *optimal* for cooling. If $\mathsf{derv}(\lambda)$ is a polytope, the subset of optimal derivatives takes a nice form.

Corollary 6.5.2. *Let $\lambda \in \Delta_{\downarrow}^{n-1}$ be regular and assume that $\mathsf{derv}(\lambda)$ is a convex polytope. If at least one point in the relative interior of a face is optimal, then the entire face is optimal. Moreover, the set of optimal faces is connected.*

Proof. Let P be the polytope $\mathsf{derv}(\lambda)$ and let C be the cone of vectors infinitesimally majorizing the origin. Then the optimal elements of P are exactly the bounded faces of $P - C$, and hence they form a subcomplex (cf. [Zie07, Section 5.1]), which, by [Jos+01, Lem. 2.1], is connected. □

Furthermore, as we will see in the following section, for certain quasi-classical systems the optimal derivatives are always given by the same controls in the reduced control system, and do not depend on the state λ.

6.6 Optimal Cooling

Now we have all the tools we need to determine optimal cooling procedures for the V-system and the spin-spin system (assuming Conjecture 6.4.3). Note that in both cases Theorem 4.3.7 shows immediately that the system is coolable.

V-System

Using Corollary 6.5.2 one can show that the optimal derivatives are exactly the convex combinations of

$$\begin{pmatrix} \gamma_2 b + \gamma_1 c \\ -\gamma_2 b \\ -\gamma_1 c \end{pmatrix} \text{ and } \begin{pmatrix} \gamma_1 b + \gamma_2 c \\ -\gamma_1 b \\ -\gamma_2 c \end{pmatrix},$$

where $\lambda = (a, b, c) \in \Delta^{n-1}_{\downarrow}$. These correspond to the two topmost vertices in Figure 6.3. Note that the analogous result holds for the Λ-system, and hence our method also recovers the results of [STK04] using a completely different approach. Note that if $\gamma_1 = \gamma_2$, the problem becomes trivial, so we assume that $\gamma_1 < \gamma_2$. A direct computation shows that the generators for these optimal derivatives commute and hence they can be applied in any order. Hence, following the first derivative for time t_1 and the second for time t_2 the final state is simply

$$\begin{pmatrix} a_0 + (1 - e^{-(\gamma_2 t_1 + \gamma_1 t_2)}) b_0 + (1 - e^{-(\gamma_1 t_1 + \gamma_2 t_2)}) c_0 \\ e^{-(\gamma_2 t_1 + \gamma_1 t_2)} b_0 \\ e^{-(\gamma_1 t_1 + \gamma_2 t_2)} c_0 \end{pmatrix}.$$

Note that following these two derivatives it might happen that the second and third eigenvalue cross, but this does not change anything about the optimality of the derivatives.

In order to go any further we have to clarify the control task, since it necessarily takes infinite time to reach a pure state. One natural choice is to minimize the time necessary to reach a certain largest eigenvalue, although other Schur-convex (or concave) functions such as those mentioned in Section 6.2 are also sensible. Concretely the problem becomes, for any $0 < \varepsilon < b_0 + c_0$, to minimize $T = t_1 + t_2$ under the conditions that $t_1, t_2 \geq 0$ and $e^{-(\gamma_2 t_1 + \gamma_1 t_2)} b_0 + e^{-(\gamma_1 t_1 + \gamma_2 t_2)} c_0 = \varepsilon$. Without the constraint $t_1, t_2 \geq 0$, an elementary computation shows that the optimal solution is

$$t_1 = \frac{\gamma_2 \log(\frac{2b_0}{\varepsilon}) - \gamma_1 \log(\frac{2c_0}{\varepsilon})}{\gamma_2^2 - \gamma_1^2}, \quad t_2 = \frac{\gamma_1 \log(\frac{2b_0}{\varepsilon}) - \gamma_2 \log(\frac{2c_0}{\varepsilon})}{\gamma_1^2 - \gamma_2^2},$$

and the final state satisfies that $b(T) = c(T) = \varepsilon/2$. However, if $\varepsilon > 2b_0 (\frac{c_0}{b_0})^{\gamma_2/(\gamma_2 - \gamma_1)}$, then t_2 becomes negative. In this case the optimal solution has $t_2 = 0$ and t_1 can be computed correspondingly.

This allows us to find the time-optimal controls for the task of reaching a largest eigenvalue of $1 - \varepsilon$. We start by applying a (near) instantaneous unitary transformation to bring the state into diagonal form and with eigenvalues in weakly decreasing order. Then we wait for time t_1 without applying any

controls (recall that the system is diagonally invariant and thus the compensating Hamiltonian vanishes). If $t_2 = 0$ we are done, otherwise we swap the second and third eigenvalue (near) instantaneously and wait for time t_2.

This solution is quite similar to that of [STK04] for the Λ-system, except that we only switch the eigenvalues once. Note also that in contrast to the approach of [STK04], we deduced the optimal solution instead of guessing it and we were able to prove optimality without the application of the Hamilton–Jacobi–Bellman equation.

Spin-Spin System

Finally we determine an optimal cooling procedure for the spin-spin system. Assuming Conjecture 6.4.3 implies that the system is quasi-classical. Corollary 6.5.2 allows us to determine the optimal derivatives. It is easy to show that the only optimal derivatives are the convex combinations of

$$(b, -b, d, -d)^\top, \text{ and } (c, d, -c, -d)^\top.$$

Conveniently, the corresponding generators again commute and hence their order of application is irrelevant. Moreover it is clear that the system is diagonally invariant and hence the compensating Hamiltonian vanishes again. Hence, similarly to the previous case, one can define a Schur-convex (or concave) cost function, such as purity, and find optimal times t_1 and t_2 via direct computation. This is not much more difficult than in the previous case, but the resulting formulas are lengthy and not particularly enlightening, and hence omitted.

Part III

Bipartite Systems with Local Unitary Control

"Long may Louis de Broglie continue to inspire those who suspect that what is proved by impossibility proofs is lack of imagination."
— John Stewart Bell, *On the Impossible Pilot Wave* (1982)

"He had said that the geometry of the dream-place he saw was abnormal, non-Euclidean, and loathsomely redolent of spheres and dimensions apart from ours."
— H. P. Lovecraft, *The Call of Cthulhu* (1926)

Quantum entanglement is one of the primary features of quantum mechanics not present in classical physics. It can be interpreted as a kind of non-classical correlation, which was used to show that quantum mechanics is incompatible with local realism. Consequently, entanglement was found to be a key resource for achieving advantage in quantum computation, quantum cryptography, and quantum sensing. Even so, the precise structure of entanglement, especially in the multipartite and open systems cases, is not well understood. Moreover, entanglement is a fragile property easily destroyed by undesired noise. For the realization of quantum technologies it is therefore essential to be able to precisely control and protect the entanglement present in a quantum system.

This part deals with closed bipartite quantum systems subject to fast local unitary control. In this setting, one can define an equivalent reduced control system on the singular values, which quantify the entanglement present in the system. Analogous results are obtained for bosonic and fermionic systems, where the complex SVD is to be replaced with the Autonne–Takagi factorization and the Hua factorization respectively. With this, we show that such systems are generically controllable and stabilizable, and we derive a speed limit on the evolution of the singular values. Moreover, we derive time-optimal solutions for the generation of maximally entangled states. In particular, the case of two coupled qubits is treated in complete generality, and a more complicated system composed of two qutrits is studied using the Pontryagin Maximum Principle.

Outline Chapter 7 defines the reduced control system on the singular values of the pure bipartite quantum state. Chapter 8 proves some general results for such systems, in particular controllability and stabilizability are shown and a general speed limit is established. Chapter 9 addresses the problem of optimal control of entanglement using the reduced control system.

Acknowledgments This part is based on [7, 8]. The paper [8] is joint work with Léo Van Damme, whose expertise on the Pontryagin Maximum Principle was an essential ingredient.

CHAPTER 7

Reduction to the Singular Values

In this chapter we apply the methods developed in Chapter 2 to closed bipartite quantum systems with fast local unitary control. The connection to symmetric Lie algebras is made via the complex SVD for distinguishable subsystems, and via the Autonne–Takagi factorization and the Hua factorization in the bosonic and fermionic cases. As a consequence we obtain the corresponding equivalent reduced control system defined on the singular values of the bipartite state. The reduced state space is the unit hypersphere, and the reduced dynamics are given by rotations. The remaining chapters of Part III are dedicated to the study of this reduced control system.

While the reduced control system studied in Part II has been formulated previously, the reduced control system defined below seems to be new. Of course the matrix decompositions used here are known [Aut15, Tak25, Hua44] and have been studied in the context of quantum entanglement [HKS13]. The closest idea used in quantum control theory appears to be the KAK-decomposition [KBG01], which can be used to study the operator lift of a two qubit system with fast local unitary control. This method can of course also be used to study the state-level control system, see for instance [Bas+21].

In Section 7.1 we derive the reduced control systems for distinguishable, bosonic and fermionic subsystems respectively, and we prove the corresponding equivalence results. Appendix 7.A gives the detailed relation of the present setting to the symmetric Lie algebra setting of Chapter 2.

7.1 Reduction

First we define what we mean by a closed bipartite quantum system with fast local unitary control, before we derive the equivalent reduced control system obtained by factoring out the local unitary action. The reduced control system describes the dynamics of the singular values of the state, and due to the normalization of the quantum state, the reduced state space turns out to be a hypersphere. The main result of this section is the equivalence, in a precise sense which will be made clear, of the full bilinear control system and the reduced one. In particular, there is no loss of information incurred by passing to the reduced control system. A brief discussion of global phases concludes the section.

Full Control System

Let two finite dimensional Hilbert spaces $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$ of dimensions $d_1, d_2 \geq 2$ representing the subsystems be given. The total Hilbert space of the bipartite system is then $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$.[1] We denote by $\mathfrak{u}(d)$ the unitary Lie algebra consisting of skew-Hermitian matrices in d dimensions. Our goal is to study the full bilinear control system defined by the following controlled Schrödinger equation[2] on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$:

$$|\dot{\psi}(t)\rangle = -\mathrm{i}\Big(H_0 + \sum_{i=1}^{m_1} u_i(t) E_i \otimes \mathbb{1} + \sum_{j=1}^{m_2} v_j(t) \mathbb{1} \otimes F_j\Big) |\psi(t)\rangle\,, \tag{$\mathcal{H}$}$$

where $H_0 \in \mathrm{i}\mathfrak{u}(d_1) \otimes \mathrm{i}\mathfrak{u}(d_2) \cong \mathrm{i}\mathfrak{u}(d_1 d_2)$ is the *drift Hamiltonian* (or *coupling Hamiltonian*), $E_i \in \mathrm{i}\mathfrak{u}(d_1)$ and $F_j \in \mathrm{i}\mathfrak{u}(d_2)$ are the *(local) control Hamiltonians*, and u_i and v_j are the corresponding *control functions*. We make the following key assumptions:

(I) The control functions u_i and v_j are locally integrable, in particular we assume no bounds on the controls.

(II) The control Hamiltonians generate the full local unitary Lie algebra:

$$\langle \mathrm{i}E_i \otimes \mathbb{1},\, \mathbb{1} \otimes \mathrm{i}F_j : i = 1, \ldots, m_1,\, j = 1, \ldots, m_2 \rangle_{\mathsf{Lie}} = \mathfrak{u}_{\mathrm{loc}}(d_1, d_2),$$

where $\mathfrak{u}_{\mathrm{loc}}(d_1, d_2) := (\mathfrak{u}(d_1) \otimes \mathbb{1}) + (\mathbb{1} \otimes \mathfrak{u}(d_2))$ is the Lie algebra of the Lie group $\mathrm{U}_{\mathrm{loc}}(d_1, d_2) := \mathrm{U}(d_1) \otimes \mathrm{U}(d_2)$ of local unitary transformations. Put simply, we have fast control over $\mathrm{U}_{\mathrm{loc}}(d_1, d_2)$.

Remark 7.1.1. *One may also define the group* $\mathrm{SU}_{\mathrm{loc}}(d_1, d_2) := \mathrm{S}(\mathrm{U}(d_1) \otimes \mathrm{U}(d_2))$ *of* local special unitary operations. *Consider the local unitary* $U = e^{\mathrm{i}\phi_1} \mathbb{1} \otimes e^{\mathrm{i}\phi_2} \mathbb{1}$. *Then* $\det(U) = e^{\mathrm{i}d_1 d_2(\phi_1 + \phi_2)}$ *and hence, if* $U \in \mathrm{SU}_{\mathrm{loc}}$, *the value of the applied phase* $e^{\mathrm{i}(\phi_1+\phi_2)}$ *is restricted to a discrete set. Thus, if we do not neglect the global phase of the state* $|\psi\rangle$, *fast control over* $\mathrm{SU}_{\mathrm{loc}}(d_1, d_2)$ *is not sufficient to generate all local unitary state transfers. To simplify the exposition, we assume fast control over* $\mathrm{U}_{\mathrm{loc}}(d_1, d_2)$, *but we will revisit this issue at the end of Section 7.1 to show how this assumption can be weakened.*

This covers the case of two distinguishable subsystems. However, we also wish to treat systems composed of two indistinguishable subsystems. In this case both subsystems have the same dimension $d := d_1 = d_2$. In the *bosonic* case, the state $|\psi\rangle$ is unchanged by swapping the two subsystems, i.e., $\mathrm{U}_{\mathrm{swap}} |\psi\rangle = |\psi\rangle$, where $\mathrm{U}_{\mathrm{swap}} |\psi_1\rangle \otimes |\psi_2\rangle = |\psi_2\rangle \otimes |\psi_1\rangle$. These "symmetric" states lie in the space $\mathrm{Sym}^2(\mathbb{C}^d)$. In the *fermionic* case, swapping yields a phase factor of -1, i.e., $\mathrm{U}_{\mathrm{swap}} |\psi\rangle = -|\psi\rangle$. Such "skew-symmetric" states are contained in the space $\bigwedge^2(\mathbb{C}^d)$. In both cases the set of local unitaries applicable to the system is restricted to symmetric local unitaries $\mathrm{U}^s_{\mathrm{loc}}(d) := \{V \otimes V : V \in \mathrm{U}(d)\}$. The corresponding Lie algebra is $\mathfrak{u}^s_{\mathrm{loc}}(d) := \{\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E : \mathrm{i}E \in \mathfrak{u}(d)\}$ and is isomorphic to $\mathfrak{u}(d)$.[3] The set of all coupling Hamiltonians applicable to such systems is the set $\mathfrak{u}^s(d^2) = \{\mathrm{i}H \in \mathfrak{u}(d^2) : \mathrm{U}_{\mathrm{swap}} H \mathrm{U}^*_{\mathrm{swap}} = H\}$. Hence, in the case of two indistinguishable subsystems, the bilinear control system takes the form:

$$|\dot{\psi}(t)\rangle = -\mathrm{i}\Big(H_0 + \sum_{i=1}^{m} u_i(t)(E_i \otimes \mathbb{1} + \mathbb{1} \otimes E_i)\Big) |\psi(t)\rangle\,, \tag{$\mathcal{H}^s$}$$

where $H_0 \in \mathrm{i}\mathfrak{u}^s(d^2)$ and $E_1, \ldots, E_m \in \mathrm{i}\mathfrak{u}(d)$. Assumption (I) remains unchanged, but Assumption (II) is slightly modified to state:

[1] Abstractly the symbol $\otimes$ denotes the tensor product, but since we always work with concrete vectors and matrices we interpret $\otimes$ as the Kronecker product. This identifies the vector spaces $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and $\mathbb{C}^{d_1 d_2}$ and similarly for matrices.

[2] Recall that we set $\hbar = 1$ and thus write the (uncontrolled) Schrödinger equation as $|\dot{\psi}(t)\rangle = -\mathrm{i}H_0 |\psi(t)\rangle$.

[3] Note however that the map $\mathrm{U}(d) \to \mathrm{U}^s_{\mathrm{loc}}(d)$ given by $V \mapsto V \otimes V$ is a double cover with kernel $\{\mathbb{1}, -\mathbb{1}\}$.

(III) The local control Hamiltonians $\mathrm{i}E_i \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E_i$ for $i = 1, \ldots, m$ generate the full Lie algebra $\mathfrak{u}_{\mathrm{loc}}(d)$.

Note that the only difference between the bosonic and fermionic case is that the initial state of the control system ($\mathcal{H}^s$) lies in $\mathrm{Sym}^2(\mathbb{C}^d)$ and $\bigwedge^2(\mathbb{C}^d)$ respectively.

Remark 7.1.2. *Similarly to Remark 7.1.1, one might consider control Hamiltonians of the form $E \otimes \mathbb{1} + \mathbb{1} \otimes E$ with the additional restriction of $\mathrm{tr}(E) = 0$, defining the Lie algebra $\mathfrak{su}^s_{\mathrm{loc}}(d)$. In this case we again lose control over the global phase of the state. Thus, for simplicity, we do not make this assumption here and refer to the end of Section 7.1 for more details.*

Related Matrix Decompositions

In order to derive and understand the reduced control system (which we will introduce in the next section) obtained by factoring out the local unitary action, we must first understand the mathematical structure of this action. In the case of distinguishable subsystems, the local unitary action corresponds the complex singular value decomposition, an thus the only invariants of a state under this action are its singular values, which therefore are the natural choice for the reduced state. The bosonic and fermionic cases correspond to less well-known matrix decompositions called the Autonne–Takagi and Hua factorization respectively. Importantly, all of these matrix decompositions also correspond to certain symmetric Lie algebras, and this is the key to applying the results on reduced control systems from Chapter 2 to the full bilinear control systems ($\mathcal{H}$) and ($\mathcal{H}^s$).

Let $\{|i\rangle_1\}_{i=1}^{d_1}$ and $\{|j\rangle_2\}_{j=1}^{d_2}$ denote the standard orthonormal bases[4] of $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$ respectively, and let $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ be a state vector. The components $(\psi_{ij})_{i,j=1}^{d_1,d_2}$ of $|\psi\rangle$ are uniquely given by

$$|\psi\rangle = \sum_{i,j=1}^{d_1,d_2} \psi_{ij} \, |i\rangle_1 \otimes |j\rangle_2 =: \sum_{i,j=1}^{d_1,d_2} \psi_{ij} \, |ij\rangle \, .$$

Hence, every bipartite state can be uniquely represented by a matrix.[5] More precisely, we have used the canonical isomorphism

$$\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \to \mathbb{C}^{d_1} \otimes (\mathbb{C}^{d_2})' \cong \mathbb{C}^{d_1,d_2} \qquad |i\rangle_1 \, |j\rangle_2 \mapsto |i\rangle_1 \langle j|_2 \, , \tag{7.1}$$

where $(\cdot)'$ denotes the dual space. We will use $\psi \in \mathbb{C}^{d_1,d_2}$ to denote the matrix corresponding to $|\psi\rangle$ under this isomorphism and vice versa. For distinguishable subsystems, the matrix representing ψ is an arbitrary complex matrix in $\mathbb{C}^{d_1,d_2}$ (the constraint induced by the normalization of the state $|\psi\rangle$ will be discussed in Remark 7.1.3 below). For indistinguishable subsystems, it holds that $d := d_1 = d_2$, and the matrix $\psi \in \mathfrak{sym}(d, \mathbb{C})$ is symmetric in the bosonic case and $\psi \in \mathfrak{asym}(d, \mathbb{C})$ is skew-symmetric in the fermionic case.

Let $V \otimes W \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ be a local unitary, and set $|\phi\rangle = V \otimes W \, |\psi\rangle$. The elements of V are then defined by $V = \sum_{k,i=1}^{d_1} V_{ki} \, |k\rangle\langle i|$, and similarly $W = \sum_{l,j=1}^{d_2} W_{lj} \, |l\rangle\langle j|$. Then $\phi_{kl} = \sum_{i,j=1}^{d_1,d_2} V_{ki} W_{lj} \psi_{ij}$. In matrix form this can be rewritten as $\phi = V \psi W^\top$. Another way to state this is $V \otimes W \, |\psi\rangle = |V \psi W^\top\rangle$. Note that in the case of indistinguishable subsystems we have $V = W$. This suggests a connection to certain matrix diagonalizations, namely:

[4] In the following we will usually omit the index denoting the subsystem, since it will be clear form the order.

[5] Our convention is consistent with [BŻ17, Sec. 9.2] and [HKS13]. In other contexts one often defines the vectorization operation $\mathrm{vec}(\cdot)$ which turns a matrix into a vector by stacking its columns and satisfies $\mathrm{vec}(AXB) = (B^\top \otimes A)\mathrm{vec}(X)$. Our convention is slightly different in that, identifying $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \cong \mathbb{C}^{d_1 d_2}$ via the Kronecker product, we may write $|\psi\rangle = \mathrm{vec}(\psi^\top)$.

- The *complex singular value decomposition* in the distinguishable subsystems case (often referred to as the Schmidt decomposition in the context of quantum mechanics). It states that for any complex matrix $\psi \in \mathbb{C}^{d_1,d_2}$ there exist unitary matrices $V \in \mathrm{U}(d_1)$ and $W \in \mathrm{U}(d_2)$ such that $V\psi W^*$ is real and diagonal. The correspondence is established by

$$|\phi\rangle = V \otimes \overline{W}\,|\psi\rangle \iff \phi = V\psi W^*.$$

- The *Autonne–Takagi factorization* in the bosonic case. It states that for any complex symmetric matrix $\psi \in \mathfrak{sym}(d,\mathbb{C})$, there is a unitary $V \in \mathrm{U}(d)$ such that $V\psi V^\top$ is real and diagonal. The correspondence is then given by

$$|\phi\rangle = V \otimes V\,|\psi\rangle \iff \phi = V\psi V^\top.$$

- The *Hua factorization* in the fermionic case. It states that for any complex skew-symmetric matrix $\psi \in \mathfrak{asym}(d,\mathbb{C})$, there is a unitary $V \in \mathrm{U}(d)$ such that $V\psi V^\top$ is real and *quasi-diagonal* in the following sense: if d is even, then $V\psi V^\top$ is block diagonal with blocks of size 2×2, if d is odd, then there is an additional block of size 1×1 in the lower right corner. Note that the quasi-diagonal matrix is still skew-symmetric, and so the diagonal is zero. The correspondence to local unitary state transformations is as in the bosonic case.

Note that the Autonne–Takagi factorization and the Hua factorization are special cases of singular value decompositions, and hence the resulting (quasi-)diagonal matrix will have the singular values on its (quasi-)diagonal. In the first case we denote by $\Sigma \subset \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ the subspace of "real diagonal states" corresponding to the set of real diagonal matrices $\mathfrak{diag}(d_1, d_2, \mathbb{R})$ under the isomorphism (7.1). Clearly Σ has dimension $d_{\min} := \min(d_1, d_2)$. In the second case we write $\Sigma \subset \mathrm{Sym}^2(\mathbb{C}^d)$ for the d-dimensional subspace corresponding to the real diagonal matrices $\mathfrak{diag}(d, \mathbb{R})$. Similarly, in the third case we write $\Xi \subset \bigwedge^2(\mathbb{C}^d)$ for the $\lfloor d/2 \rfloor$-dimensional subspace of states corresponding to the real (skew-symmetric) quasi-diagonal matrices $\mathfrak{qdiag}(d, \mathbb{R})$. We will use the following maps to send the singular values to their corresponding (quasi-)diagonal state:

$$\begin{aligned} \mathrm{diag} &: \mathbb{R}^{d_{\min}} \to \Sigma, \quad (\sigma_i)_{i=1}^{d_{\min}} \mapsto \textstyle\sum_{i=1}^{d_{\min}} \sigma_i\,|i\rangle \otimes |i\rangle\,, \\ \mathrm{qdiag} &: \mathbb{R}^{\lfloor d/2 \rfloor} \to \Xi, \quad (\xi_i)_{i=1}^{\lfloor d/2 \rfloor} \mapsto \tfrac{1}{\sqrt{2}} \textstyle\sum_{i=1}^{\lfloor d/2 \rfloor} \xi_i(|2i-1\rangle \otimes |2i\rangle - |2i\rangle \otimes |2i-1\rangle). \end{aligned}$$

A convenient shorthand notation is $|\sigma\rangle = \mathrm{diag}(\sigma)$ resp. $|\xi\rangle = \mathrm{qdiag}(\xi)$. We will always use the standard Euclidean inner product on $\mathbb{R}^n$, and on $\mathbb{C}^n$ we will use the real part $\mathrm{Re}(\langle\cdot|\cdot\rangle)$ of the standard inner product. Then, due to the inclusion of the factor $1/\sqrt{2}$ it holds that the maps above are $\mathbb{R}$-linear isometric isomorphisms. Furthermore we denote[6] by $\Sigma_\updownarrow \subset \Sigma$ the cone of states $\mathrm{diag}(\sigma)$ where the diagonal elements $(\sigma_i)_{i=1}^{d_{\min}}$ are non-negative and arranged in non-increasing order, and analogously we write $\Xi_\updownarrow \subset \Xi$ for the quasi-diagonal states $\mathrm{qdiag}(\xi)$ where the $(\xi_i)_{i=1}^{\lfloor d/2 \rfloor}$ are non-negative and arranged in non-increasing order. The set $\Sigma_\updownarrow$ resp. $\Xi_\updownarrow$ is called the *Weyl chamber*.

Conversely to the diagonal embeddings we also define the following orthogonal projections:

$$\begin{aligned} \Pi_\Sigma &: \begin{cases} \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \to \mathbb{R}^{d_{\min}},\ |\psi\rangle \mapsto (\mathrm{Re}(\langle ii|\psi\rangle))_{i=1}^{d_{\min}} & \text{for distinguishable subsystems} \\ \mathrm{Sym}^2(\mathbb{C}^d) \to \mathbb{R}^d,\ |\psi\rangle \mapsto (\mathrm{Re}(\langle ii|\psi\rangle))_{i=1}^{d} & \text{for bosonic subsystems} \end{cases} \\ \Pi_\Xi &: \textstyle\bigwedge^2(\mathbb{C}^d) \to \mathbb{R}^{\lfloor d/2 \rfloor},\ |\psi\rangle \mapsto \sqrt{2}(\mathrm{Re}(\langle 2i|_2\,\langle 2i-1|_1\,|\psi\rangle))_{i=1}^{\lfloor d/2 \rfloor}\,. \end{aligned}$$

More precisely, these are the orthogonal projections on Σ and Ξ followed by diag^{-1} and qdiag^{-1} respectively.

[6] The symbol $\updownarrow$ is a combination of $\downarrow$ and $+$.

Remark 7.1.3. *The normalization of the quantum state entails a normalization of the corresponding singular values. More precisely, the norm of the quantum state $|\psi\rangle$ equals the Frobenius norm of the matrix ψ, and hence $|\psi\rangle$ has unit norm if and only if the singular values $\sigma = (\sigma_i)_{i=1}^{d_{\min}}$ of ψ satisfy $\sum_{i=1}^{d_{\min}} \sigma_i^2 = 1$. Hence the singular values define a point on the unit sphere of dimension $d_{\min} - 1$ in the indistinguishable case, and $d - 1$ in the bosonic case. In the fermionic case there are again d singular values and they lie on the unit sphere of dimension $d - 1$. However there is an additional restriction as the singular values come in pairs of opposite values. Taking only one singular value of each pair and multiplying it by $\sqrt{2}$ (and ignoring the additional 0 singular value in the odd dimensional case) we find that the resulting vector lies on the unit sphere of dimension $\lfloor d/2 \rfloor - 1$. In all cases will call this the* Schmidt sphere*, denoted $S^{d_{\min}-1}$ in the distinguishable case, S^{d-1} in the bosonic case, and $S^{\lfloor d/2 \rfloor - 1}$ in the fermionic case. The maps* diag *and* qdiag *then yield isometric embeddings of the Schmidt sphere into Σ and Ξ respectively. The Schmidt sphere will be the state space of our reduced control system, which we will define in the following section. The Weyl chambers $\Sigma_{\updownarrow}$ and $\Xi_{\updownarrow}$ then yield corresponding Weyl chambers in the Schmidt sphere $S^{d_{\min}-1}_{\updownarrow}$, $S^{d-1}_{\updownarrow}$, and $S^{\lfloor d/2 \rfloor - 1}_{\updownarrow}$.*

Remark 7.1.4. *Often one considers the Schmidt values, which are the squares of the singular values, within the standard simplex, which is then called the Schmidt simplex [BŻ17, Sec. 16.4]. This is easier to visualize, but in our case would lead to unnatural dynamics, which is why we will remain on the sphere. Note also that if one of the systems is a qubit, then the Schmidt sphere is a circle parametrized by the Schmidt angle [BŻ17, p. 440].*

Reduced Control System

Due to Assumptions (I) and (II) (resp. (III)), we can move arbitrarily quickly within the local unitary orbits of the system if we ignore the drift term [Ell09, Prop. 2.7]. With the drift this is still approximately true. In the previous section we have shown that using local unitary transformations, we can always obtain a state of (quasi-)diagonal form which is completely determined by the singular values of the state. In particular, within the bilinear control systems ($\mathcal{H}$) and ($\mathcal{H}^s$) two states are effectively equivalent if and only if they have the same singular values (up to order and signs). This strongly suggests that there should exist a "reduced" control system, defined on the singular values — or rather the Schmidt sphere (cf. Remark 7.1.3). This is indeed the case. The reduced control system is defined in greater generality in Section 2.2 using symmetric Lie algebras, which unify many well-known matrix diagonalizations, such as the ones encountered in the previous section, cf. Chapter 1. No knowledge of symmetric Lie algebras is presupposed here, but the connections are expounded in Appendix 7.A.

Let us briefly motivate the definition of the reduced control system. Let $|\psi\rangle$ be a solution to the full control system ($\mathcal{H}$). Assume that the corresponding matrix ψ can be diagonalized in a differentiable way as $\psi(t) = V(t)\tilde{\sigma}(t)W^\top(t)$, and that it is regular[7]. Here $\tilde{\sigma}(t)$ is the diagonal matrix with diagonal elements $\sigma(t)$. Then by differentiating (cf. Lemma 1.4.1) we obtain that $\dot{\sigma} = -H_{V\otimes W}\sigma$ (and analogously $\dot{\sigma} = -H^s_{V\otimes V}\sigma$ or $\dot{\xi} = -H^a_{V\otimes V}\xi$ in the bosonic and fermionic cases) where

$$\begin{aligned} -H_{V\otimes W} &:= -\Pi_\Sigma \circ (V\otimes W)^* \,\mathrm{i}H_0(V\otimes W) \circ \mathrm{diag}, \\ -H^s_{V\otimes V} &:= -\Pi_\Sigma \circ (V\otimes V)^* \,\mathrm{i}H_0(V\otimes V) \circ \mathrm{diag}, \\ -H^a_{V\otimes V} &:= -\Pi_\Xi \circ (V\otimes V)^* \,\mathrm{i}H_0(V\otimes V) \circ \mathrm{qdiag}\,. \end{aligned}$$

We call $H_{V\otimes W}$, $H^s_{V\otimes V}$ and $H^a_{V\otimes V}$ the *induced vector fields*. The collection of induced vector fields is denoted

$$\mathfrak{H} := \{-H_U : U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)\},\ \mathfrak{H}^s := \{-H^s_U : U \in \mathrm{U}^s_{\mathrm{loc}}(d)\},\ \mathfrak{H}^a := \{-H^a_U : U \in \mathrm{U}^s_{\mathrm{loc}}(d)\}.$$

[7] We say that ψ is regular if its singular values are distinct and non-zero.

Note that these are linear vector fields on $\mathbb{R}^{d_{\min}}$, $\mathbb{R}^d$ and $\mathbb{R}^{\lfloor d/2 \rfloor}$ respectively, and hence they can be represented as matrices in the respective standard basis. We will later see that these are indeed skew-symmetric matrices, and thus the corresponding dynamics preserve the Schmidt sphere. The following proposition gives the explicit expressions.

Proposition 7.1.5. *Let $H_0 \in \mathrm{iu}(d_1 d_2)$ denote an arbitrary coupling Hamiltonian and let $A_k \in \mathrm{iu}(d_1)$ and $B_k \in \mathrm{iu}(d_2)$ for $k = 1, \ldots, r$ be given such that $H_0 = \sum_{k=1}^r A_k \otimes B_k$. Then the induced vector field $H_{V\otimes W}$ on $\mathbb{R}^{d_{\min}}$ takes the form*[8]

$$-H_{V\otimes W} = \sum_{k=1}^{r} \mathrm{Im}(V^* A_k V \circ W^* B_k W).$$

Now assume additionally that $d := d_1 = d_2$ and that $H_0 \in \mathrm{iu}^s(d^2)$ is symmetric. For bosonic systems, on $\mathbb{R}^d$, we obtain

$$-H^s_{V\otimes V} = \sum_{k=1}^{r} \mathrm{Im}(V^* A_k V \circ V^* B_k V).$$

For fermionic systems, on $\mathbb{R}^{\lfloor d/2 \rfloor}$, we obtain

$$-(H^a_{V\otimes V})_{ij} = \sum_{k=1}^{r} \mathrm{Im}((V^* A_k V)_{2i-1,2j-1}(V^* B_k V)_{2i,2j} - (V^* A_k V)_{2i-1,2j}(V^* B_k V)_{2i,2j-1}).$$

Proof. To simplify notation we first consider a product Hamiltonian $H_0 = A \otimes B$. Let $\sigma \in \mathbb{R}^{d_{\min}}$ be given. Let $i, j \in \{1, \ldots, d_{\min}\}$ and compute

$$-(H_{V\otimes W})_{ij} = (-\Pi_\Sigma(\mathrm{i} V^* A V e_j e_j^\top W^\top B^\top (W^*)^\top))_i = \mathrm{Im}((V^* A V)_{ij}(W^* B W)_{ij}).$$

The case of general H_0 for distinguishable subsystems as well as the bosonic case follow by linearity.

Now let us consider the fermionic case. Again for simplicity we consider a product Hamiltonian $H = A \otimes B$ since the general result follows by linearity. Then A and B can be seen as $n \times n$ block matrices with blocks of size 2×2 denoted $A_{(ij)}$ and $B_{(ij)}$ (and, in the odd-dimensional case, an additional row and column). Moreover let $J = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ denote the standard symplectic form. Then we can compute in a similar fashion

$$-(H^a_{V\otimes V})_{ij} = (-\Pi_\Xi(\mathrm{i}(V^* A V)\,\mathrm{qdiag}(e_j)(V^* B^\top V)^\top))_i = \mathrm{Im}(((V^* A V)_{(ij)} J ((V^* B V)_{(ij)})^\top)_{12})\,.$$

This concludes the proof. □

Remark 7.1.6. *The fermionic case can be interpreted as follows. We consider the even dimensional case for simplicity. First define the matrices G_k obtained from A_k and B_k by choosing all the odd indexed rows from A_k and all the even indexed rows from B_k. More precisely, $(G_k)_{ij} = (A_k)_{ij}$ if i is odd and $(G_k)_{ij} = (B_k)_{ij}$ if i is even. Then we divide the G_k into blocks of size 2×2 and we define a new matrix of half the size by replacing each block by the imaginary part of its determinant.*

We see immediately that if the drift Hamiltonian is local, then the induced vector fields vanish:

Corollary 7.1.7. *If the drift Hamiltonian H_0 is local, meaning that $H_0 \in \mathrm{iu}_{\mathrm{loc}}(d_1, d_2)$, then $H_{V\otimes W} = 0$ for every $V \otimes W \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$. The same is true for $H^s_{V\otimes V}$ and $H^a_{V\otimes V}$ whenever $H_0 \in \mathrm{iu}^s_{\mathrm{loc}}(d)$ and $V \otimes V \in \mathrm{U}^s_{\mathrm{loc}}(d)$.*

[8] Here $\circ$ denotes the Hadamard (elementwise) product of two matrices. If the (square) matrices are of different size the resulting matrix will have the size of the smaller one. Similarly Im denotes the elementwise imaginary part.

Finally we can define the reduced control systems:

Definition 7.1.8 (Reduced control systems). *Let $I \subset \mathbb{R}$ be an interval of the form $[0,T]$ with $T \geq 0$ or $[0,\infty)$. We define three reduced control systems:*

For distinguishable subsystems we set:

$$\dot{\sigma}(t) = -H_{U(t)}\sigma(t), \quad \sigma(0) = \sigma_0 \in S^{d_{\min}-1} \tag{Σ}$$

A solution is an absolutely continuous path $\sigma : I \to S^{d_{\min}-1}$ satisfying (Σ) almost everywhere for some measurable control function $U : I \to \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$.

In the bosonic case we set:

$$\dot{\sigma}(t) = -H^s_{U(t)}\sigma(t), \quad \sigma(0) = \sigma_0 \in S^{d-1} \tag{Σ^s}$$

A solution is an absolutely continuous path $\sigma : I \to S^{d-1}$ satisfying (Σ^s) almost everywhere for some measurable control function $U : I \to \mathrm{U}^s_{\mathrm{loc}}(d)$.

In the fermionic case we set:

$$\dot{\xi}(t) = -H^a_{U(t)}\xi(t), \quad \xi(0) = \xi_0 \in S^{\lfloor d/2 \rfloor -1} \tag{Σ^a}$$

A solution is an absolutely continuous path $\xi : I \to S^{\lfloor d/2 \rfloor -1}$ satisfying (Σ^a) almost everywhere for some measurable control function $U : I \to \mathrm{U}^s_{\mathrm{loc}}(d)$.

As mentioned previously, and shown in Lemma 7.1.11 below, the induced vector fields preserve the Schmidt sphere, and hence we may define the reduced control systems directly on the Schmidt sphere.

Remark 7.1.9. *There are several slightly different ways of defining the reduced control system which are given in Section 2.2. The most intuitive definition, given above, is to consider the control system $\dot{\sigma}(t) = -H_{U(t)}(\sigma(t))$ where the control function $U : [0,T] \to \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ is measurable and the solution $\sigma : [0,T] \to S^{d_{\min}-1}$ is absolutely continuous. A more geometric definition uses the differential inclusion $\dot{\sigma}(t) \in \mathsf{derv}(\sigma(t))$, where $\mathsf{derv}(\sigma)$ denotes the set of* achievable derivatives *at σ defined by $\mathsf{derv}(\sigma) = \{-H_U\sigma : U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)\}$. The differential inclusion is exactly equivalent to our definition by Filippov's Theorem, cf. [Smi02, Thm. 2.3]. Often it is convenient to consider a "relaxed" version of the differential inclusion where also convex combinations of achievable derivatives are allowed: $\dot{\sigma}(t) \in \mathrm{conv}(\mathsf{derv}(\sigma(t)))$. This slightly enlarges the set of solutions, but every solution to the relaxed system can still be approximated uniformly on compact time intervals by solutions to our system, cf. [AC84, Ch. 2.4, Thm. 2]. Analogous remarks also hold for the symmetric cases (Σ^s) and (Σ^a).*

The main result of Chapter 2 is the equivalence of the full bilinear control system ($\mathcal{H}$) resp. ($\mathcal{H}^s$) and the reduced control system (Σ), resp. (Σ^s) or (Σ^a), proven in Theorem 2.3.16. In our case this specializes to the following result.

First we need to define the quotient maps $\mathrm{sing}^{\updownarrow} : \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \to \mathbb{R}^{d_{\min}}$ and $\mathrm{qsing}^{\updownarrow} : \bigwedge^2(\mathbb{C}^d) \to \mathbb{R}^{\lfloor d/2 \rfloor}$. Given a (possibly not normalized) vector $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ (or $\mathrm{Sym}^2(\mathbb{C}^d)$), the map $\mathrm{sing}^{\updownarrow}$ yields the singular values of the corresponding matrix $\psi \in \mathbb{C}^{d_1,d_2}$, chosen non-negative and arranged in non-increasing order. Similarly, for $|\xi\rangle \in \bigwedge^2(\mathbb{C}^d)$, the map $\mathrm{qsing}^{\updownarrow}$ yields the singular values of the skew-symmetric matrix ξ, except that we keep only one singular value of each pair and multiply it by $\sqrt{2}$ to keep the normalization. Note that when restricting the domain of $\mathrm{sing}^{\updownarrow}$ and $\mathrm{qsing}^{\updownarrow}$ to (normalized) quantum states, the image will lie in the respective Schmidt sphere, and even in the Weyl chamber $S^{d_{\min}-1}_{\updownarrow}$ resp. $S^{d-1}_{\updownarrow}$ and $S^{\lfloor d/2 \rfloor -1}_{\updownarrow}$.

Recall that here and throughout the chapter we use Assumptions (I) and (II) (resp. (III)), unless stated otherwise.

Theorem 7.1.10 (Equivalence Theorem). *Let $|\psi(t)\rangle$ be a solution on $[0,T]$ to the bilinear control system ($\mathcal{H}$), and let $\sigma^{\updownarrow} : [0,T] \to S^{d_{\min}-1}_{\updownarrow}$ be defined by $\sigma^{\updownarrow} = \mathrm{sing}^{\updownarrow}(|\psi\rangle)$. Then $\sigma^{\updownarrow}$ is a solution to the reduced control system (Σ).*

Conversely, let $\sigma : [0,T] \to S^{d_{\min}-1}$ be a solution to the reduced control system (Σ) with control function $U : [0,T] \to \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ and let $|\sigma\rangle = \mathrm{diag}(\sigma)$ denote the corresponding state. Then $U(t)\,|\sigma(t)\rangle$ can be approximated by solutions to the full control system ($\mathcal{H}$) arbitrarily well. More precisely, for every $\varepsilon > 0$ there exists a solution $|\psi_\varepsilon(t)\rangle$ to ($\mathcal{H}$) such that $\|U\,|\sigma\rangle - |\psi_\varepsilon\rangle\|_\infty \leq \varepsilon$, where $\|\cdot\|_\infty$ denotes the supremum norm.

The analogous results, mutatis mutandis,[a] *also hold in the bosonic and fermionic cases, where the full control system is ($\mathcal{H}^s$) and the reduced control systems are (Σ^s) and (Σ^a) respectively.*

[a] Most results in Part III hold in all three cases with only minimal differences in notation, which are summarized in Table 1.

Proof. The proof is mostly a technicality, as we simply have to show that the full control systems ($\mathcal{H}$) and ($\mathcal{H}^s$) and their respective reduced versions (Σ), (Σ^s) and (Σ^a) can be interpreted as control systems on certain symmetric Lie algebras. We focus on the case of distinguishable subsystems. The corresponding symmetric Lie algebra is that of type AIII. The isomorphisms $\imath^d$ and $\jmath^d$ defined in Appendix 7.A translate the quantum setting into the Lie algebra setting. The results of the appendix then show that all of the conditions of Theorem 2.3.16 are satisfied and that the reduced control system (Σ) indeed corresponds to the reduced control system in the Lie algebraic setting. Taken together, this proves the equivalence in the distinguishable case. The other cases are entirely analogous. □

The Equivalence Theorem 7.1.10 shows that the full bilinear control system ($\mathcal{H}$) resp. ($\mathcal{H}^s$) and the reduced control system (Σ), resp. (Σ^s) or (Σ^a), contain essentially the same information. Hence for every control theoretic notion, such as controllability and stabilizability, there is a specialized equivalence result, see Section 2.4 for an overview. As a first consequence we obtain:

Lemma 7.1.11. *The induced vector fields are skew-symmetric matrices:*

$$\mathfrak{H} \subset \mathfrak{so}(d_{\min}, \mathbb{R}), \quad \mathfrak{H}^s \subset \mathfrak{so}(d, \mathbb{R}), \quad \mathfrak{H}^a \subset \mathfrak{so}(\lfloor d/2 \rfloor, \mathbb{R}).$$

In particular the Schmidt sphere is invariant.

Proof. Due to the Equivalence Theorem 7.1.10, this follows from Proposition 2.4.8. Alternatively this can also be verified by direct computation using the expressions obtained in Proposition 7.1.5. □

Let us also recall the equivalence of reachable sets here, which is arguably the most useful consequence. First we give the definitions of reachable sets in the reduced control system (Σ). The definitions for other control systems are entirely analogous. The *reachable set of σ_0 at time T* is defined as

$$\mathsf{reach}_\Sigma(\sigma_0, T) = \{\sigma(T) : \sigma : [0,T] \to S^{d_{\min}-1} \text{ solves } (\Sigma), \sigma(0) = \sigma_0\}$$

for any $T \geq 0$. By $\mathsf{reach}_\Sigma(\sigma_0) := \bigcup_{T \geq 0} \mathsf{reach}_\Sigma(\sigma_0, T)$ we denote the *all time reachable set of σ_0*, and by $\mathsf{reach}_\Sigma(\sigma_0, [0,T]) := \bigcup_{t \in [0,T]} \mathsf{reach}_\Sigma(\sigma_0, t)$ we denote the *reachable set of σ_0 up to time T*.

The following result is an immediate consequence of the Equivalence Theorem 7.1.10 and Proposition 2.4.3.

Proposition 7.1.12. *Let $T > 0$ be given and assume that $|\psi_0\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and $\sigma_0 \in S^{d_{\min}-1}$ satisfy*[9] *$\sigma_0^{\ddagger} = \mathrm{sing}^{\ddagger}(\psi_0)$. Then it holds that*

$$\mathsf{reach}_{\mathcal{H}}(|\psi_0\rangle, T) \subseteq \{U|\sigma\rangle : \sigma \in \mathsf{reach}_{\Sigma}(\sigma_0, T),\, U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)\} \subseteq \overline{\mathsf{reach}_{\mathcal{H}}(|\psi_0\rangle, T)}\,.$$

In particular, the closures coincide:

$$\overline{\mathsf{reach}_{\mathcal{H}}(|\psi_0\rangle, T)} = \overline{\{U|\sigma\rangle : \sigma \in \mathsf{reach}_{\Sigma}(\sigma_0, T),\, U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)\}}\,.$$

The analogous result, mutatis mutandis*, holds also for the bosonic and the fermionic cases.*

The Equivalence Theorem 7.1.10 guarantees the existence of an approximate lift, but its proof also provides a way to find corresponding control functions. Under some additional assumptions we can give an explicit formula for the controls of an exact lift, see Proposition 2.3.10. In particular this requires the solution to be smooth and regular, and the controls of the bilinear system ($\mathcal{H}$) resp. ($\mathcal{H}^s$) to linearly span the corresponding Lie algebra. Before stating the result we define some notation.

The Lie group of local unitary operations $\mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ and its symmetric counterpart $\mathrm{U}^s_{\mathrm{loc}}(d)$ act on the state spaces $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$, $\mathrm{Sym}^2(\mathbb{C}^d)$ and $\bigwedge^2(\mathbb{C}^d)$ respectively. The corresponding infinitesimal action of the Lie algebras $\mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$ and $\mathfrak{u}^s_{\mathrm{loc}}(d)$ can be determined as in the previous section using the formula $(\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F)|\psi\rangle = |\mathrm{i}(E\psi + \psi\overline{F})\rangle$. In the special case where the state is regular and diagonal this function (more precisely its negative) gets a special name:

$$\begin{aligned}
\mathrm{ad}^d_\sigma &: \mathfrak{u}_{\mathrm{loc}}(d_1, d_2) \to \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}, & \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F &\mapsto -|\mathrm{i}(E\tilde{\sigma} + \tilde{\sigma}\overline{F})\rangle \\
\mathrm{ad}^s_\sigma &: \mathfrak{u}^s_{\mathrm{loc}}(d) \to \mathrm{Sym}^2(\mathbb{C}^d), & \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E &\mapsto -|\mathrm{i}(E\tilde{\sigma} + \tilde{\sigma}\overline{E})\rangle \\
\mathrm{ad}^a_\xi &: \mathfrak{u}^s_{\mathrm{loc}}(d) \to \textstyle\bigwedge^2(\mathbb{C}^d), & \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E &\mapsto -|\mathrm{i}(E\tilde{\xi} + \tilde{\xi}\overline{E})\rangle\,,
\end{aligned}$$

where $\tilde{\sigma} \in \mathfrak{diag}(d_1, d_2, \mathbb{R})$ and $\tilde{\xi} \in \mathfrak{qdiag}(d, \mathbb{R})$ denote the (quasi-)diagonal matrices corresponding to σ and ξ. Although these maps are not bijective, by restricting the domain to the orthocomplement of the kernel and the codomain to the image, inverse maps can be defined. Indeed, this is nothing but the Moore–Penrose pseudoinverse. Explicit expressions are given in Lemmas 7.A.6, 7.A.9 and 7.A.12.

To use these inverse maps, we have to understand the images of the maps ad^d_σ, ad^s_σ, and ad^a_ξ. It turns out that, for regular σ resp. ξ, these images are exactly given by the orthocomplement of the diagonal subspaces Σ, Ξ. We denote the orthogonal projection on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with kernel Σ by $\Pi^{\perp}_{\Sigma}$, and use the same notation on $\mathrm{Sym}^2(\mathbb{C}^d)$. In the matrix picture, this map simply removes the real part of the diagonal elements of ψ. Similarly, on $\bigwedge^2(\mathbb{C}^d)$, the orthogonal projection with kernel Ξ is denoted $\Pi^{\perp}_{\Xi}$ and it removes the real part of the quasi-diagonal of ψ.

With these definitions Proposition 2.3.10 can be specialized as follows:

Proposition 7.1.13. *Let $\sigma : [0, T] \to S^{d_{\min}-1}$ be a solution to the reduced control system (Σ) with control function $U : [0, T] \to \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$. Assume that σ is regular and that U is continuously differentiable. Let $|\psi(t)\rangle = U(t)|\sigma(t)\rangle$ and let $H : [0, T] \to \mathrm{i}\mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$ be given by*

$$-\mathrm{i}H(t) = \dot{U}(t)U^{-1}(t) - \mathrm{Ad}_{U(t)}\big((\mathrm{ad}^d_{\sigma(t)})^{-1} \circ \Pi^{\perp}_{\Sigma}\big)\big(U(t)^*(\mathrm{i}H_0)U(t)\,|\sigma(t)\rangle\big)\,.$$

Then $|\psi\rangle$ satisfies $|\dot{\psi}\rangle = -\mathrm{i}(H_0 + H)|\psi\rangle$. The analogous result also holds mutatis mutandis *for bosonic and fermionic systems.*

[9] Here $\sigma_0^{\ddagger}$ denotes the element of $S^{d_{\min}-1}_{\downarrow}$ whose elements are the absolute values of the elements of σ_0 arranged in non-increasing order.

We call the second term in the definition of H in Proposition 7.1.13 the *(local) compensating Hamiltonian* since it compensated for the local Hamiltonian action induced by the drift term H_0.

If the control directions linearly span the entire Lie algebra $\mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$, then it is easy to find the control functions from Proposition 7.1.13. More generally the problem of finding corresponding controls is studied under the term non-holonomic motion planning, see [Liu97].

Global Phases

Mathematically the quantum state $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ has a global phase which is physically undetectable and hence may be considered irrelevant. We always keep the phase for convenience, noting that the global phase is removed automatically in the reduced control system. This section briefly discusses how Assumptions (II) and (III) can be slightly weakened by neglecting the global phase. We denote by $\mathfrak{su}_{\mathrm{loc}}(d_1, d_2)$ and $\mathfrak{su}^s_{\mathrm{loc}}(d)$ the Lie algebras obtained from $\mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$ and $\mathfrak{u}^s_{\mathrm{loc}}(d)$ by requiring the trace to vanish. Consider the following two weakened controllability assumptions:

(II') The control Hamiltonians generate the local special unitary Lie algebra:

$$\langle \mathrm{i}E_i \otimes \mathbb{1}, \mathbb{1} \otimes \mathrm{i}F_j : i = 1, \ldots, m_1, j = 1, \ldots, m_2 \rangle_{\mathsf{Lie}} = \mathfrak{su}_{\mathrm{loc}}(d_1, d_2).$$

(III') The control Hamiltonians generate the symmetric local special unitary Lie algebra:

$$\langle \mathrm{i}E_i \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E_i : i = 1, \ldots, m \rangle_{\mathsf{Lie}} = \mathfrak{su}^s_{\mathrm{loc}}(d).$$

In both cases adding the (symmetric) local control Hamiltonian $\mathbb{1} \otimes \mathbb{1}$ is sufficient to obtain the stronger Assumptions (II) and (III) respectively. Since this Hamiltonian commutes with everything, the only effect of adding or removing the corresponding term from the control system is a change in the global phase. More concretely, if our control system is ($\mathcal{H}$) (resp. ($\mathcal{H}^s$)) but only satisfies Assumptions (I) and (II') (resp. (III')), then we can add the control Hamiltonian $\mathbb{1} \otimes \mathbb{1}$ so that it satisfies Assumption (II) (resp. (III)). Now we can compute any solution in this extended system and we obtain a corresponding solution in the actual system by setting the control function of $\mathbb{1} \otimes \mathbb{1}$ to zero. The resulting solution will, at all times, be equal to the solution of the extended system up to a global phase.

7.A Relation to Symmetric Lie Algebras

In the main text we have shown that the local unitary actions on bipartite quantum states correspond to certain matrix diagonalizations, and we have stated that they themselves are related to certain symmetric Lie algebras. In this appendix we make these relations explicit and give all the relevant formulas. For a compact overview of the relation of symmetric Lie algebras to matrix diagonalizations see Table 1.3.

Since we want to define a reduced control system on the singular values, a key question is how the singular values change in time. More precisely, given a differentiable path of matrices $\psi(t)$, what can we say about the derivative of the singular values? This question is made more complicated by the fact that the order and signs of the singular values are not unique (and if they are chosen in a unique way, they are not guaranteed to be differentiable). These issues can be resolved, and in fact one can do so in the more general setting of semisimple orthogonal symmetric Lie algebras, see Chapter 1 and in particular Section 1.2. We will recall and adapt the pertinent results as necessary.

The reduction of control systems was treated in detail in Chapter 2 in the setting of semisimple orthogonal symmetric Lie algebras. In order to rigorously prove the Equivalence Theorem 7.1.10, we need to show how the control systems considered here can be interpreted as control systems in such symmetric Lie algebras. See Table 1 for an overview of the notation related to the different control systems.

Complex Singular Value Decomposition (Type AIII)

The complex singular value decomposition is encoded by the symmetric Lie algebra of type AIII, see for instance [Kle06, App. A.7] and [Hel78, Ch. X §2.3]. The standard matrix representation of this Lie algebra is the indefinite special unitary Lie algebra $\mathfrak{g}_{\mathsf{AIII}} = \mathfrak{su}(d_1, d_2)$ with Cartan decomposition $\mathfrak{g}_{\mathsf{AIII}} = \mathfrak{k}_{\mathsf{AIII}} \oplus \mathfrak{p}_{\mathsf{AIII}}$ where[10]

$$\mathfrak{k}_{\mathsf{AIII}} = \left\{ \begin{pmatrix} \mathrm{i}E & 0 \\ 0 & \mathrm{i}F \end{pmatrix} : \mathrm{i}E \in \mathfrak{u}(d_1),\, \mathrm{i}F \in \mathfrak{u}(d_2),\, \operatorname{tr}(E) = -\operatorname{tr}(F) \right\},$$
$$\mathfrak{p}_{\mathsf{AIII}} = \left\{ \begin{pmatrix} 0 & \psi \\ \psi^* & 0 \end{pmatrix} : \psi \in \mathbb{C}^{d_1, d_2} \right\}.$$

A choice of corresponding compact Lie group is $\mathbf{K}_{\mathsf{AIII}} = \mathrm{S}(\mathrm{U}(d_1) \times \mathrm{U}(d_2))$.

Remark 7.A.1. *In general we consider (semi)simple* orthogonal *symmetric Lie algebras $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and so we also have to provide a "compatible" inner product on $\mathfrak{g}$. The inner product is (up to some irrelevant scaling) uniquely defined using the Killing form on $\mathfrak{g}$, cf. [Hel78, Ch. V, Thm. 1.1]. Due to simplicity of $\mathfrak{g}$ the Killing form is (again up to scaling) given by $\operatorname{tr}(AB)$. In the following we will set the inner product on $\mathfrak{k}$ to $-\frac{1}{2}\operatorname{tr}(AB)$ and on $\mathfrak{p}$ to $+\frac{1}{2}\operatorname{tr}(AB)$, and we define that $\mathfrak{k}$ and $\mathfrak{p}$ are orthogonal to each other. Furthermore, we always use the real inner product $\operatorname{Re}(\langle\psi, \phi\rangle) = \operatorname{Re}(\operatorname{tr}(\psi^*\phi))$ on states in $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$.*

The spaces $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and $\mathfrak{p}_{\mathsf{AIII}}$ are identified using the map

$$\imath^d : \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \to \mathfrak{p}_{\mathsf{AIII}}, \quad |\psi\rangle \mapsto \begin{pmatrix} 0 & \psi \\ \psi^* & 0 \end{pmatrix}.$$

Lemma 7.A.2. *The map $\imath^d$ is an $\mathbb{R}$-linear[11] isometric isomorphism. The subspace $\mathfrak{a}_{\mathsf{AIII}} := \imath^d(\Sigma)$ is maximal Abelian and $\imath^d \circ \operatorname{diag} \circ \Pi_\Sigma = \Pi_{\mathfrak{a}_{\mathsf{AIII}}} \circ \imath^d$. The Weyl group $\mathbf{W}_{\mathsf{AIII}}$ is isomorphic to the signed symmetric group $S_{d_{\min}} \wr \mathbb{Z}_2$ and $\mathfrak{w}_{\mathsf{AIII}} := \imath^d(\Sigma^{\ddagger})$ is a Weyl chamber.*

Proof. It is clear that $\imath^d$ is an $\mathbb{R}$-linear isomorphism. With the inner product on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ and $\mathfrak{p}$ defined as in Remark 7.A.1 a simple computation shows that $\imath^d$ is even an isometry:

$$\tfrac{1}{2}\operatorname{tr}(\imath^d(|\psi\rangle), \imath^d(|\phi\rangle)) = \tfrac{1}{2}\operatorname{tr}(\psi\phi^* + \psi^*\phi) = \operatorname{Re}(\operatorname{tr}(\psi^*\phi)).$$

That $\imath^d(\Sigma)$ is maximal Abelian is well-known, cf. Table 1.2. The fact that $\imath^d$ is an isometry also proves that $\imath^d \circ \operatorname{diag} \circ \Pi_\Sigma = \Pi_{\mathfrak{a}_{\mathsf{AIII}}} \circ \imath^d$. That the Weyl group acts by generalized permutations follows from the fact that the singular values are unique up to order and sign and the fact that any generalized permutation can be implemented by choosing V and W appropriately. □

Moreover, we define the following maps:

$$\jmath^d : \mathrm{U}_{\mathrm{loc}}(d_1, d_2) \to \mathrm{Ad}_{\mathbf{K}_{\mathsf{AIII}}}, \qquad V \otimes W \mapsto \mathrm{Ad}_{V \times \overline{W}}$$
$$\jmath^d_\star : \mathfrak{u}_{\mathrm{loc}}(d_1, d_2) \to \mathrm{ad}_{\mathfrak{k}_{\mathsf{AIII}}}, \qquad \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F \mapsto \mathrm{ad}_{\mathrm{i}E \times \overline{\mathrm{i}F}}.$$

Note that $\jmath^d_\star$ is the derivative of $\jmath^d$ at the identity.

[10] Often one denotes $\mathfrak{k}_{\mathsf{AIII}} = \mathfrak{s}(\mathfrak{u}(d_1) \oplus \mathfrak{u}(d_2))$.

[11] Note that in the Lie algebraic context we always work with real vector spaces, even if their standard representation involves complex numbers.

Lemma 7.A.3. *It holds that $\jmath^d$ is a Lie group isomorphism, and so $\jmath^d_\star$ is a Lie algebra isomorphism*[12]*. For $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$, $\mathrm{i}H \in \mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$ and $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$, the isomorphisms $\imath^d$ and $\jmath^d$ satisfy the compatibility conditions*

$$\jmath^d(U)\imath^d(|\psi\rangle) = \imath^d(U\,|\psi\rangle), \quad \jmath^d_\star(\mathrm{i}H)\imath^d(|\psi\rangle) = \imath^d(\mathrm{i}H\,|\psi\rangle). \tag{7.2}$$

Similarly we have the correspondence of the infinitesimal action $\imath^d(\mathrm{ad}^d_\sigma(\mathrm{i}H)) = -\jmath^d_\star(\mathrm{i}H)(\imath^d(|\sigma\rangle))$ and of the induced vector fields where if $-X := \imath^d_\star(\mathrm{i}H_0) = \imath^d \circ (\mathrm{i}H_0) \circ (\imath^d)^{-1}$ then $-H_U = (\imath^d \circ \mathrm{diag})^\star X_{\jmath^d(U)}$. Moreover the map $\mathrm{sing}^\ddagger$ corresponds to the quotient map with image in the Weyl chamber $\mathfrak{p}_{\mathsf{AIII}} \to \mathfrak{a}_{\mathsf{AIII}}/\mathbf{W}_{\mathsf{AIII}} \cong \mathfrak{w}_{\mathsf{AIII}}$, see Section 1.3.

Proof. Even though $V \times \overline{W}$ does not always lie in $\mathbf{K}_{\mathsf{AIII}}$, we can choose $\phi \in \mathbb{R}$ such that $e^{\mathrm{i}\phi}V \times \overline{e^{-\mathrm{i}\phi}W} \in \mathbf{K}_{\mathsf{AIII}}$, and this phase disappears in the tensor product and in the adjoint representation. Indeed, any ϕ satisfying $e^{\mathrm{i}\phi(d_1+d_2)} = \det W/\det V$ will do. Hence $\jmath^d$ is well defined, and one easily verifies that it is an isomorphism. The compatibility condition (7.2) follows from a simple computation. Similarly, the corresponding Lie algebra isomorphism $\jmath^d_\star$ can be written as $\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F \mapsto \mathrm{ad}_{\mathrm{i}(E + \frac{\mathrm{tr}(F)-\mathrm{tr}(E)}{d_1+d_2}\mathbb{1}) \oplus \mathrm{i}(-F + \frac{\mathrm{tr}(F)-\mathrm{tr}(E)}{d_1+d_2}\mathbb{1})}$ to show explicitly that it is well defined. By definition $\mathrm{ad}^d_\sigma(\mathrm{i}H) = -\mathrm{i}H\,|\sigma\rangle$. Using the compatibility condition (7.2) this immediately yields $\imath^d(\mathrm{ad}^d_\sigma(\mathrm{i}H)) = -\jmath^d_\star(\mathrm{i}H)(\imath^d(|\sigma\rangle))$ as desired. Recall that we defined $H_U = \Pi_\Sigma \circ (U^*(\mathrm{i}H_0)U) \circ \mathrm{diag}$ and in Section 2.2 we defined $X_K = \Pi_\mathfrak{a} \mathrm{Ad}^\star_K(X) \circ \iota$. (We are slightly abusing notation here by writing X_{Ad_K}.) Using the compatibility condition (7.2), and $\Pi_\mathfrak{a} \circ \imath^d = \imath^d \circ \mathrm{diag} \circ \Pi_\Sigma$, the claim $H_U = (\imath^d \circ \mathrm{diag})^\star X_{\jmath^d(U)}$ follows from a simple computation. The claim about the quotient map is just a restatement of the uniqueness of the singular values. □

Remark 7.A.4. *Explicitly (7.2) states that $\jmath^d(V \otimes \overline{W})\imath^d(|\psi\rangle) = \imath^d(|V\psi W^*\rangle)$. In a semisimple orthogonal symmetric Lie algebra, every element in $\mathfrak{p}_{\mathsf{AIII}}$ can be mapped into $\mathfrak{a}_{\mathsf{AIII}}$ using the group action of $\mathbf{K}_{\mathsf{AIII}}$, cf. Lemma 1.A.26, which one might call "diagonalization". In our case this means that $\imath^d(|\psi\rangle)$ can be mapped to some element in $\imath^d(|V\psi W^*\rangle) \in \imath^d(\Sigma)$. This exactly corresponds to the complex singular value decomposition. Note however that in the Lie algebra setting we have $V \times W \in \mathrm{S}(\mathrm{U}(d_1) \times \mathrm{U}(d_2))$ and hence there is an additional restriction on the determinants of V and W.*

Remark 7.A.5. *For regular $\sigma \in S^{d_{\min}-1}$ we have defined the map ad^d_σ in the main text and Lemma 7.A.3 shows that it is related to the adjoint representation $\mathrm{ad}_\sigma : \mathfrak{k}_{\mathsf{AIII}} \to \mathfrak{p}_{\mathsf{AIII}}$ (hence the name). Denoting by $\mathfrak{k}^\perp_\sigma$ and $\mathfrak{p}^\perp_\sigma$ the orthogonal complement of the commutant of σ in $\mathfrak{k}_{\mathsf{AIII}}$ and $\mathfrak{p}_{\mathsf{AIII}}$ respectively, it turns out that the restriction $\mathrm{ad}_\sigma : \mathfrak{k}^\perp_\sigma \to \mathfrak{p}^\perp_\sigma$ becomes bijective and hence invertible, see Proposition 1.4.12. In fact, this inverse is simply the Moore–Penrose pseudoinverse. Moreover it holds that $\mathfrak{p}^\perp_\sigma$ is just the orthocomplement of $\mathfrak{a}$. Hence, the pseudoinverse $(\mathrm{ad}^d_\sigma)^{-1}$ is defined on $\Sigma^\perp$ with image in $(\jmath^d_\star)^{-1}(k^\perp_\sigma)$. Note however that $\jmath^d$ is not an isometry and so the orthocomplement has to be calculated in $\mathfrak{k}$. The precise definition is given in the following lemma.*

Lemma 7.A.6. *Let $\sigma \in S^{d_{\min}-1}$ be regular. Then it holds that $\Sigma^\perp = \{|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} : \psi_{ii} \in \mathrm{i}\mathbb{R}, 1 \le i \le d_{\min}\}$, and the map $(\mathrm{ad}^d_\sigma)^{-1}$ can be described explicitly as*

$$(\mathrm{ad}^d_\sigma)^{-1} : \Sigma^\perp \to \mathfrak{u}_{\mathrm{loc}}(d_1, d_2), \quad |A\rangle \mapsto \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F\,,$$

where $E_{ij} = 0$ and $F_{ij} = 0$ whenever $i > d_{\min}$ and $j > d_{\min}$, and for $i \le d_{\min}$ and $j \le d_{\min}$ we get

$$\mathrm{i}E_{ii} = \mathrm{i}F_{ii} = -\frac{A_{ii}}{2\sigma_i}, \quad \mathrm{i}E_{ij} = \frac{\sigma_j A_{ij} + \sigma_i \overline{A}_{ji}}{\sigma_i^2 - \sigma_j^2}, \quad \mathrm{i}F_{ij} = \frac{\sigma_j A_{ji} + \sigma_i \overline{A}_{ij}}{\sigma_i^2 - \sigma_j^2}. \tag{7.3}$$

[12] Contrary to $\imath^d$, the map $\jmath^d_\star$ is not an isometry with respect to the inner products of Remark 7.A.1.

If $d_1 > d_2$ resp. $d_1 < d_2$ we additionally have

$$\mathrm{i}E_{ij} = \begin{cases} -\frac{A_{ij}}{\sigma_j} & \text{if } j \le d_2 < i \\ 0 & \text{if } j > d_2 \end{cases} \quad \textit{resp.} \quad \mathrm{i}F_{ij} = \begin{cases} -\frac{A_{ji}}{\sigma_j} & \text{if } i \le d_1 < j \\ 0 & \text{if } i > d_1. \end{cases}$$

Proof. We only consider the case $d_1 \ge d_2$ for simplicity. Consider the map $\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F \mapsto A := -\mathrm{i}(E\tilde{\sigma} + \tilde{\sigma}\overline{F})$ where $\tilde{\sigma}$ is the corresponding diagonal matrix. Then $A_{ii} = -\mathrm{i}(E_{ii} + F_{ii})\sigma_i \in \mathrm{i}\mathbb{R}$. Hence, when we invert the map above, we will assume that $A_{ii} \in \mathrm{i}\mathbb{R}$. Moreover we need to find the kernel of the map, i.e. solve for $A = 0$ (for all or any regular σ). This happens if $E_{ii} + F_{ii} = 0$, and $E_{ij} = F_{ij} = 0$ for $j \le d_2$ and $E_{ij} \in \mathbb{C}$ for $j > d_2$. The orthocomplement of the kernel is given by $E_{ii} = F_{ii}$, and $E_{ij}, F_{ij} \in \mathbb{C}$ for $j \le d_2$ and $E_{ij} = 0$ for $j > d_2$. Using, for $i, j \le d_{\min}$, that

$$A_{ij} = -\mathrm{i}(E_{ij}\sigma_j + F_{ji}\sigma_i), \quad \overline{A}_{ji} = \mathrm{i}(E_{ij}\sigma_i + F_{ji}\sigma_j),$$

we find

$$\sigma_i A_{ij} + \sigma_j \overline{A}_{ji} = \mathrm{i}F_{ji}(\sigma_j^2 - \sigma_i^2), \text{ and thus } \mathrm{i}F_{ij} = \frac{\sigma_j A_{ji} + \sigma_i \overline{A}_{ij}}{\sigma_i^2 - \sigma_j^2},$$
$$\sigma_j A_{ij} + \sigma_i \overline{A}_{ji} = \mathrm{i}E_{ij}(\sigma_i^2 - \sigma_j^2), \text{ and thus } \mathrm{i}E_{ij} = \frac{\sigma_j A_{ij} + \sigma_i \overline{A}_{ji}}{\sigma_i^2 - \sigma_j^2}.$$

Finally for $j \le d_2 < i$ we find $\mathrm{i}E_{ij} = -A_{ij}/\sigma_j$ and for $j > d_2$ we get $\mathrm{i}E_{ij} = 0$. □

Note that this lemma uniquely defines $\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}F \in \mathfrak{u}_{\mathrm{loc}}(d_1, d_2)$, although there is some freedom in the choice of E and F since we can shift some real multiple of the identity between them.

Autonne–Takagi Factorization (Type CI)

First discovered by Autonne [Aut15] and Takagi [Tak25], the Autonne–Takagi factorization [HJ12, Sec. 4.4] states that for any complex symmetric matrix $A \in \mathfrak{sym}(d, \mathbb{C})$ there exists a unitary matrix $U \in \mathrm{U}(d)$ such that $UAU^\top$ is real and diagonal. The diagonal elements are uniquely defined up to order and signs, and they are in fact the singular values of A.

The corresponding symmetric Lie algebra is that of type CI, usually represented by the real symplectic Lie algebra $\mathfrak{g}_{\mathsf{CI}} = \mathfrak{sp}(d, \mathbb{R})$, see [Kle06, Sec. 4.3] and again [Hel78, Ch. X §2.3]. The Cartan decomposition $\mathfrak{g}_{\mathsf{CI}} = \mathfrak{k}_{\mathsf{CI}} \oplus \mathfrak{p}_{\mathsf{CI}}$ is given explicitly by

$$\mathfrak{k}_{\mathsf{CI}} = \left\{ \begin{bmatrix} A & B \\ -B & A \end{bmatrix} : A = -A^\top, B = B^\top,\ A, B \in \mathbb{R}^{d,d} \right\},$$
$$\mathfrak{p}_{\mathsf{CI}} = \left\{ \begin{bmatrix} C & D \\ D & -C \end{bmatrix} : C = C^\top, D = D^\top,\ C, D \in \mathbb{R}^{d,d} \right\}.$$

The corresponding state space isomorphism $\imath^s$ is given by

$$\imath^s : \mathrm{Sym}^2(\mathbb{C}^d) \to \mathfrak{p}_{\mathsf{CI}},\ |\psi\rangle \mapsto \begin{pmatrix} \mathrm{Re}\,\psi & -\mathrm{Im}\,\psi \\ -\mathrm{Im}\,\psi & -\mathrm{Re}\,\psi \end{pmatrix}.$$

Lemma 7.A.7. *The map $\imath^s$ is an $\mathbb{R}$-linear isometric isomorphism. The subspace $\mathfrak{a}_{\mathsf{CI}} := \imath^s(\Sigma)$ is maximal Abelian and $\imath^s \circ \mathrm{diag} \circ \Pi_\Sigma = \Pi_{\mathfrak{a}_{\mathsf{CI}}} \circ \imath^s$. The Weyl group $\mathbf{W}_{\mathsf{CI}}$ is isomorphic to the generalized permutations $\mathbb{Z}_2 \wr S_d$ and $\mathfrak{w}_{\mathsf{CI}} := \imath^s(\Sigma^\updownarrow)$ is a Weyl chamber.*

Proof. The proof is analogous to that of Lemma 7.A.2 so we just compute the inner product on $\mathfrak{p}_{\mathsf{CI}}$:

$$\frac{1}{2}\operatorname{tr}(\imath^s(|\psi\rangle)\imath^s(|\phi\rangle)) = \operatorname{tr}(\operatorname{Re}(\psi)\operatorname{Re}(\phi) + \operatorname{Im}(\psi)\operatorname{Im}(\phi)) = \operatorname{Re}(\operatorname{tr}(\psi^*\phi)).$$

Hence $\imath^s$ is an isometry and this concludes the proof. □

Now consider the following maps

$$\jmath^s : \mathrm{U}^s_{\mathrm{loc}}(d) \to \mathrm{Ad}_{\mathbf{K}_{\mathsf{CI}}}, \qquad V \otimes V \mapsto \mathrm{Ad}_{\left(\begin{smallmatrix} \operatorname{Re} V & \operatorname{Im} V \\ -\operatorname{Im} V & \operatorname{Re} V \end{smallmatrix}\right)},$$
$$\jmath^s_\star : \mathfrak{u}^s_{\mathrm{loc}}(d) \to \mathrm{ad}_{\mathfrak{k}_{\mathsf{CI}}}, \quad \mathrm{i}H \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}H \mapsto \mathrm{ad}_{\left(\begin{smallmatrix} \operatorname{Re}(\mathrm{i}H) & \operatorname{Im}(\mathrm{i}H) \\ -\operatorname{Im}(\mathrm{i}H) & \operatorname{Re}(\mathrm{i}H) \end{smallmatrix}\right)}.$$

Lemma 7.A.8. *The maps $\jmath^s$ and $\jmath^s_\star$ are Lie isomorphisms satisfying the compatibility conditions*

$$\jmath^s(U)\imath^s(|\psi\rangle) = \imath^s(U|\psi\rangle), \quad \jmath^s_\star(\mathrm{i}H)\imath^s(|\psi\rangle) = \imath^s(\mathrm{i}H|\psi\rangle). \tag{7.4}$$

As in Lemma 7.A.3 we get the correspondence of infinitesimal action, induced vector fields and quotient map.

Proof. Since $V \otimes V = (-V) \otimes (-V)$ we have to check that the map is well defined. But it is clear that $\mathrm{Ad}_{-U} = \mathrm{Ad}_U$ and hence $\jmath^s$ is well defined. That it is an isomorphism follows from [Kle06, Prop. 4.7]. The remainder of the proof is analogous to that of Lemma 7.A.3. □

Just as in Remark 7.A.4, the relation to the Autonne–Takagi factorization can be seen from (7.4), which explicitly states that $\jmath^s(V \otimes V)\imath^s(|\psi\rangle) = \imath^s(|V\psi V^\top\rangle)$, and from the fact that $\imath^s(\Sigma) = \mathfrak{a}_{\mathsf{CI}}$.

As described in Remark 7.A.5 we can explicitly compute the appropriate inverse of the map ad^s_σ. Note that in this case we have $\mathfrak{k}^\perp_\sigma = \mathfrak{k}_{\mathsf{CI}}$.

Lemma 7.A.9. *Let $\sigma \in S^{d-1}$ be regular. Then it holds that $\Sigma^\perp = \{|\psi\rangle \in \mathrm{Sym}^2(\mathbb{C}^d) : \psi_{ii} \in \mathrm{i}\mathbb{R}, 1 \leq i \leq d\}$, and the map $(\mathrm{ad}^s_\sigma)^{-1}$ can be explicitly described as*

$$(\mathrm{ad}^s_\sigma)^{-1} : \Sigma^\perp \to \mathfrak{u}^s_{\mathrm{loc}}(d), \quad |A\rangle \mapsto \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E,$$

where

$$E_{ii} = \frac{\mathrm{i}A_{ii}}{2\sigma_i}, \quad E_{ij} = -\frac{\operatorname{Im}(A_{ij})}{\sigma_i + \sigma_j} - \mathrm{i}\frac{\operatorname{Re}(A_{ij})}{\sigma_i - \sigma_j}.$$

Proof. Let $A = \mathrm{ad}^s_\sigma(\mathrm{i}E) = -\mathrm{i}(E\tilde{\sigma} + \tilde{\sigma}\overline{E})$. Then we have $A_{ii} = -2\mathrm{i}E_{ii}\sigma_i$ and $A_{ij} = -\mathrm{i}(E_{ij}\sigma_j + \sigma_i\overline{E}_{ij})$. Hence $E_{ii} = \frac{\mathrm{i}A_{ii}}{2\sigma_i}$. To invert the second equation we compute

$$A_{ij} = -\mathrm{i}(E_{ij}\sigma_j + \sigma_i\overline{E}_{ij}), \quad \overline{A}_{ij} = +\mathrm{i}(\overline{E}_{ij}\sigma_j + \sigma_i E_{ij})$$

and hence taking sum and difference we get

$$2\operatorname{Re}(A_{ij}) = \mathrm{i}(E_{ij} - \overline{E}_{ij})(\sigma_i - \sigma_j) = -2\operatorname{Im}(E_{ij})(\sigma_i - \sigma_j) \Rightarrow \operatorname{Im}(E_{ij}) = \frac{\operatorname{Re}(A_{ij})}{\sigma_j - \sigma_i},$$
$$2\mathrm{i}\operatorname{Im}(A_{ij}) = -\mathrm{i}(E_{ij} + \overline{E}_{ij})(\sigma_i + \sigma_j) = -2\mathrm{i}\operatorname{Re}(E_{ij})(\sigma_i + \sigma_j) \Rightarrow \operatorname{Re}(E_{ij}) = -\frac{\operatorname{Im}(A_{ij})}{\sigma_i + \sigma_j}.$$

This concludes the proof. □

Hua Factorization (Type DIII)

A skew-symmetric version of the Autonne–Takagi factorization also exists [HJ12, Coro. 4.4.19]. It is called the Hua factorization, and was originally proven in [Hua44, Thm. 7]. It states that for every skew-symmetric complex matrix $A \in \mathfrak{asym}(d, \mathbb{C})$ there exists a unitary $U \in \mathrm{U}(d)$ such that $UAU^\top$ is real and block diagonal with skew-symmetric blocks of size 2×2. If d is odd, then there is an additional 1×1 block containing a zero. We call such matrices quasi-diagonal. Each 2×2 block is determined by a single real number (and its negative) which taken together yield the singular values of A.

This matrix factorization is related to the symmetric Lie algebra of type DIII, usually represented by $\mathfrak{so}^*(2d)$, see [Kle06, App. A.6] and again [Hel78, Ch. X §2.3].[13] In this case we have the following Cartan like decomposition

$$\mathfrak{k}_{\mathsf{DIII}} = \left\{ \begin{pmatrix} \mathrm{i}H & 0 \\ 0 & -\mathrm{i}\overline{H} \end{pmatrix} : \mathrm{i}H \in \mathfrak{u}(d) \right\}, \quad \mathfrak{p}_{\mathsf{DIII}} = \left\{ \begin{pmatrix} 0 & \psi \\ \psi^* & 0 \end{pmatrix} : \psi = -\psi^\top, \psi \in \mathbb{C}^{d,d} \right\}.$$

Clearly the state space isomorphism is

$$\imath^a : \bigwedge\nolimits^2(\mathbb{C}^d) \to \mathfrak{p}_{\mathsf{DIII}}, \quad |\psi\rangle \mapsto \begin{pmatrix} 0 & \psi \\ \psi^* & 0 \end{pmatrix}.$$

Lemma 7.A.10. *The map $\imath^a$ is an $\mathbb{R}$-linear isometric isomorphism. The subspace $\mathfrak{a}_{\mathsf{DIII}} := \imath^a(\Xi)$ is maximal Abelian and $\imath^a \circ \mathrm{qdiag} \circ \Pi_\Xi = \Pi_{\mathfrak{a}} \circ \imath^a$. The Weyl group $\mathbf{W}_{\mathsf{DIII}}$ is isomorphic to the generalized permutations $\mathbb{Z}_2 \wr S_{\lfloor d/2 \rfloor}$ and $\mathfrak{w}_{\mathsf{DIII}} := \imath^a(\Xi^\ddagger)$ is a Weyl chamber.*

Proof. The proof is entirely analogous to that of Lemma 7.A.2. □

The isomorphisms on the Lie group and algebra level are:

$$\begin{aligned} \jmath^a &: \mathrm{U}^s_{\mathrm{loc}}(d) \to \mathrm{Ad}_{\mathbf{K}_{\mathsf{DIII}}}, & V \otimes V &\mapsto \mathrm{Ad}_{\left(\begin{smallmatrix} V & 0 \\ 0 & \overline{V} \end{smallmatrix}\right)} \\ \jmath^a_\star &: \mathfrak{u}^s_{\mathrm{loc}}(d) \to \mathrm{ad}_{\mathfrak{k}_{\mathsf{DIII}}}, & \mathrm{i}H \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}H &\mapsto \mathrm{ad}_{\left(\begin{smallmatrix} \mathrm{i}H & 0 \\ 0 & -\mathrm{i}\overline{H} \end{smallmatrix}\right)} \end{aligned}$$

Lemma 7.A.11. *The maps $\jmath^a$ and $\jmath^a_\star$ are Lie isomorphisms satisfying the compatibility conditions*

$$\jmath^a(U)\imath^a(|\psi\rangle) = \imath^a(U\,|\psi\rangle), \quad \jmath^a_\star(\mathrm{i}H)\imath^a(|\psi\rangle) = \imath^a(\mathrm{i}H\,|\psi\rangle). \tag{7.5}$$

As in Lemma 7.A.3 we get the correspondence of infinitesimal action, induced vector fields and quotient map.

Proof. Since $\mathrm{Ad}_{-U} = \mathrm{Ad}_U$ the map $\jmath^a$ is well-defined, and it is clearly a Lie group isomorphism. The remainder of the proof is analogous to that of Lemma 7.A.3. □

The relation to the Hua factorization can be seen form (7.5), which becomes $\jmath^a(V \otimes V)\imath^a(|\psi\rangle) = \imath^a(|V\psi V^\top\rangle)$, and the fact that $\imath^a(\Xi) = \mathfrak{a}_{\mathsf{DIII}}$.

Lemma 7.A.12. *Let $\xi \in S^{\lfloor d/2 \rfloor - 1}$ be regular. It holds that $\Xi^\perp = \{|\psi\rangle \in \bigwedge^2(\mathbb{C}^d) : \psi_{2i-1,2i} \in \mathrm{i}\mathbb{R},\ i = 1, \ldots, \lfloor d/2 \rfloor\}$. The map $(\mathrm{ad}^a_\xi)^{-1}$ takes the following form:*

$$(\mathrm{ad}^a_\xi)^{-1} : \Xi^\perp \to \mathfrak{u}^s_{\mathrm{loc}}(d), \quad |A\rangle \mapsto \mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E,$$

[13] Note that [Hel78] uses a different but isomorphic matrix representation.

where for $1 \leq i, j \leq \lfloor d/2 \rfloor$ *we get*

$$(\mathrm{i}E)_{(ii)} = -\frac{a_i}{2\xi_i}\mathbb{1}_2, \quad (\mathrm{i}E)_{(ij)} = \frac{\xi_i J\overline{A}_{(ij)} + \xi_j A_{(ij)} J}{\xi_j^2 - \xi_i^2},$$

where $(\mathrm{i}E)_{(ij)}$ *and* $A_{(ij)}$ *indexes the* 2×2 *blocks of the respective matrices,* $\mathbb{1}_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ *and* $J = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. *By* $a_i \in \mathrm{i}\mathbb{R}$ *we denote the value satisfying* $A_{(ii)} = a_i J$. *In the case where* d *is odd we additionally have*

$$\mathrm{i}E_{d,2\mathrm{i}-1} = \mathrm{i}\overline{E}_{2\mathrm{i}-1,d} = \frac{A_{d,2\mathrm{i}}}{\xi_i}, \quad \mathrm{i}E_{d,2\mathrm{i}} = \mathrm{i}\overline{E}_{2\mathrm{i},d} = -\frac{A_{d,2\mathrm{i}-1}}{\xi_i}, \quad \mathrm{i}E_{d,d} = 0.$$

Proof. First consider the even-dimensional case. Let $A := \mathrm{ad}_\xi^a(\mathrm{i}E \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{i}E) = -\mathrm{i}(E\tilde{\xi} + \tilde{\xi}\overline{E})$. For $1 \leq i, j \leq \lfloor d/2 \rfloor$ we compute the bocks $A_{(ii)} = -\mathrm{i}(E_{(ii)}\xi_i J + \xi_i J\overline{E}_{(ii)})$ as well as

$$A_{(ij)} = -\mathrm{i}(E_{(ij)}\xi_j J + \xi_i J\overline{E}_{(ij)}), \quad \overline{A}_{(ij)} = +\mathrm{i}(\overline{E}_{(ij)}\xi_i J + \xi_j J E_{(ij)}).$$

It follows that $\xi_i J\overline{A}_{(ij)} + \xi_j A_{(ij)} J = \mathrm{i}(\xi_j^2 - \xi_i^2)E_{(ij)}$, and hence we find that

$$\mathrm{i}E_{(ii)} = \frac{-a_i}{2\xi_i}\mathbb{1}_2, \quad \mathrm{i}E_{(ij)} = \frac{\xi_i J\overline{A}_{(ij)} + \xi_j A_{(ij)} J}{\xi_j^2 - \xi_i^2}.$$

Here we used that the for $\jmath_\star(\mathrm{i}E\otimes\mathbb{1}+\mathbb{1}\otimes\mathrm{i}E)$ to lie in $\mathfrak{k}_\xi^\perp$ the diagonal blocks $E_{(ii)}$ must be real multiples of the identity. If d is odd there is an additional column at the bottom and row on the right of $\mathrm{i}E$ to be determined. We find that

$$A_{d,2i-1} = -\mathrm{i}E_{d,2i}\xi_i, \quad A_{d,2i} = +\mathrm{i}E_{d,2i-1}\xi_i.$$

The claimed results follow immediately. □

CHAPTER 8

Controllability and Speed Limits

8.1 Introduction

We start by studying general control theoretic concepts, namely controllability and stabilizability. This turns out to be much easier than in Part II, as the systems considered here are generally controllable and stabilizable. Intuitively, this can be attributed to the fact that here we can employ the tools of Lie group theory, instead of having to deal with the more intricate theory of Lie semigroups. Due to controllability, it is interesting to study quantum speed limits (QSL) which give upper bounds on the rate of change of the singular values of the state. The problem of deriving explicit time-optimal controls is addressed in Chapter 9.

Outline and Main Results

In Section 8.2 we use the reduced control system to prove that the full bilinear control system is always controllable and stabilizable. In Section 8.3 we derive a general quantum speed limit for the evolution of the singular values.

8.2 Controllability and Stabilizability

In this section we show that the reduced control system is always controllable and stabilizable. As a consequence, the full control system is also controllable and all states can be stabilized in a certain sense. For notational simplicity, we focus on the case of distinguishable subsystems, noting that the bosonic and fermionic cases are entirely analogous.

First we lift the reduced control system (Σ) to the Lie group $\mathrm{SO}(d_{\min})$. Unless otherwise noted, it is assumed that the initial state is $R(0) = \mathbb{1}$. The *operator lift* can be defined by

$$\dot{R}(t) = -H_{U(t)}R(t), \tag{L}$$

and analogously with H_U^s and H_U^a in the bosonic and fermionic cases. A solution is absolutely continuous and satisfies (L) almost everywhere for some measurable U.

Remark 8.2.1. *Note that the operator lift* (L) *of the reduced control system is a useful but somewhat artificial construction. Indeed, even though* ($\mathcal{H}$) *and* (Σ) *are equivalent, the operator lift* (L) *is* not *equivalent to the operator lift of* ($\mathcal{H}$)*, cf. Remark 2.2.2.*

The reduced control system is *controllable on* $S^{d_{\min}-1}$ if for every two states $\sigma_1, \sigma_2 \in S^{d_{\min}-1}$ it holds that $\sigma_2 \in \mathsf{reach}_{\Sigma}(\sigma_1)$ and it is *controllable on* $S^{d_{\min}-1}$ *in time* T if for every two states $\sigma_1, \sigma_2 \in S^{d_{\min}-1}$ we have that $\sigma_2 \in \mathsf{reach}_{\Sigma}(\sigma_1, [0,T])$. *Approximate controllability* is defined in the same way except that one considers the closures of the respective reachable sets. The analogous definitions also hold for all other control systems.

To understand the properties of the operator lift, we study the set of generators $\mathfrak{H} \subset \mathfrak{so}(d_{\min}, \mathbb{R})$. A key property of $\mathfrak{H}$ is that it is invariant under conjugation by the Weyl group $\mathbf{W} = S_{d_{\min}} \wr \mathbb{Z}_2$, see Lemma 2.A.2 and Appendix 7.A. This fact allows us to prove the following result.

We say that a coupling Hamiltonian H_0 is *effectively local* if all induced vector fields H_U (resp. H_U^s or H_U^a) vanish. For distinguishable and bosonic systems this is the same as a local Hamiltonian, but for fermionic system there are effectively local Hamiltonians which are not local, see Remark 9.2.9.

Proposition 8.2.2. *The Weyl group* $\mathbf{W}$ *acts irreducibly on* $\mathfrak{so}(d_{\min}, \mathbb{R})$. *In particular, if the coupling Hamiltonian* H_0 *is not effectively local, then the operator lift* (L) *is controllable. The analogous result holds,* mutatis mutandis, *in the bosonic and fermionic cases.*

Proof. To show that the Weyl group acts irreducibly, we start with an arbitrary non-zero element $\Omega \in \mathfrak{so}(d_{\min}, \mathbb{R})$ and show that the subrepresentation generated by Ω is all of $\mathfrak{so}(d_{\min}, \mathbb{R})$. If $d_{\min} = 2$ this is trivially true since $\mathfrak{so}(2, \mathbb{R})$ is one-dimensional. So assume that $d_{\min} \geq 3$. Consider the basis $\{e_{ij} = E^{ij} - E^{ji} : 1 \leq i < j \leq d_{\min}\}$, where E^{ij} is the matrix with a 1 in the position (i,j) and 0 elsewhere, and let Ω_{ij} be the coefficients of Ω in this basis. Since Ω is non-zero, at least one of the coefficients is non-zero. Using a permutation in $\mathbf{W}$ we may assume that $\Omega_{12} \neq 0$. Let $W_i \in \mathbf{W}$ be the diagonal matrix whose diagonal equals 1 everywhere except in the i-th position, where it equals -1. Consider the matrix $\Omega' = \frac{\Omega + W_3 \Omega W_3^\top}{2}$. Then $\Omega'_{12} = \Omega_{12}$ and $\Omega'_{i3} = \Omega'_{3j} = 0$. Iterating this procedure with $W_4, \dots, W_{d_{\min}}$ we obtain a multiple of e_{12}, showing that e_{12} lies in the subrepresentation generated by Ω. From this, using the permutations in $\mathbf{W}$, all other basis elements e_{ij} can be obtained. This shows that the representation of $\mathbf{W}$ is irreducible. Controllability of the operator lift then follows from Proposition 2.4.12. □

This result can now be lifted to the full bilinear system using the equivalence of the systems.

Theorem 8.2.3 (Controllability). *If the full control system* ($\mathcal{H}$) *is controllable in time* T*, then the reduced control system* (Σ) *is controllable in time* T *on the Weyl chamber* $S_{\updownarrow}^{d_{\min}-1}$. *Conversely, if the reduced control system* (Σ) *is controllable in time* T *on the Weyl chamber* $S_{\updownarrow}^{d_{\min}-1}$*, then the full control system* ($\mathcal{H}$) *is controllable in time* $T + \varepsilon$ *for all* $\varepsilon > 0$. *Moreover, there exists a finite time* T *such that both systems are controllable in time* T. *The analogous result holds,* mutatis mutandis, *in the bosonic and fermionic cases.*

Proof. Clearly if ($\mathcal{H}$) is controllable in time T, so is (Σ) on the Weyl chamber by Proposition 7.1.12. The same result shows that if (Σ) is controllable in time T on the Weyl chamber, then ($\mathcal{H}$) is approximately controllable in time T. Proposition 8.2.2 shows in particular that (Σ) is directly accessible at every point (see Section 2.4 for the definitions related to accessibility) and hence by Proposition 2.4.10 the bilinear system ($\mathcal{H}$) is accessible (on the set of normalized states) at every regular state. Then [Jur97, Ch. 3 Thm. 2] implies that ($\mathcal{H}$) is controllable in time T for regular initial states. Since regular states are dense and since we can leave the set of non-regular states in an arbitrarily short amount of time ε, the full control system ($\mathcal{H}$) is controllable in time $T + \varepsilon$. That (Σ) is controllable in finite time follows from Proposition 8.2.2. □

Now we turn to stabilizability. In accordance with the definitions given in Section 2.4 we say that a state $\sigma \in S^{d_{\min}-1}$ is *stabilizable* for (Σ) if $0 \in \operatorname{conv}(\mathfrak{H}\sigma)$. A direct consequence of Proposition 8.2.2 is that every state is stabilizable.

Corollary 8.2.4. *Every state is stabilizable for the reduced control systems.*

Proof. If H_0 is effectively local, the statement trivially holds. Otherwise choose some non-zero H_U. Consider the uniform combination $\hat{H}_U = \frac{1}{|\mathbf{W}|}\sum_{w\in\mathbf{W}} wH_Uw^{-1} \in \operatorname{conv}(\mathfrak{H})$ which is clearly $\mathbf{W}$-invariant. If $d_{\min} = 2$, it is clear that $\hat{H}_U = 0$. In higher dimensions $\mathbf{W}$-invariance and irreducibility of the action of $\mathbf{W}$ (Proposition 8.2.2) again show that $\hat{H}_U = 0$. The proof for the bosonic and fermionic cases is the same. □

If a state $\sigma \in S^{d_{\min}-1}$ is stabilizable for (Σ), then in the bilinear control system ($\mathcal{H}$), one can stay close to the local unitary orbit $\mathrm{U}_{\mathrm{loc}}(d_1, d_2)\,|\sigma\rangle$ for an arbitrary amount of time, cf. Proposition 2.4.7.

Recall that a point $\sigma \in \Sigma$ is *strongly stabilizable* if there is $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ such that $H_U\sigma = 0$. Specializing Proposition 2.4.5 we obtain the following result.

Proposition 8.2.5. *Let $\sigma \in S^{d_{\min}-1}$ and $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ and set $|\psi\rangle = U\,|\sigma\rangle$. If there is some $H \in \mathfrak{iu}_{\mathrm{loc}}(d_1, d_2)$ satisfying $(H_0 + H)\,|\psi\rangle = 0$, then $H_U\sigma = 0$ and σ is strongly stabilizable.*

Conversely, let $\sigma \in S^{d_{\min}-1}$ be strongly stabilizable and let $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ be such that $H_U\sigma = 0$. Moreover assume that σ is regular. Then there is some $H \in \mathfrak{iu}_{\mathrm{loc}}(d_1, d_2)$ such that $(H_0 + H)\,|\psi\rangle = 0$ where we again set $|\psi\rangle = U\,|\sigma\rangle$. In fact one can choose

$$-\mathrm{i}H(t) = -U(t)((\mathrm{ad}^d_{\sigma(t)})^{-1} \circ \Pi^{\perp}_{\Sigma})\big(U(t)^*(\mathrm{i}H_0)U(t)\,|\sigma(t)\rangle\big),$$

where $(\mathrm{ad}^d_\sigma)^{-1}$ is given explicitly in Lemma 7.A.6. The analogous result holds, mutatis mutandis, *in the bosonic and fermionic cases.*

Note that the assumption on regularity is necessary in general, cf. Example 2.3.12.

The local Hamiltonian H in the previous result is called a *(local) compensating Hamiltonian*, and indeed this is a special case of Proposition 7.1.13. Note that the expressions in Lemmas 7.A.6, 7.A.9 and 7.A.12, and hence the compensating Hamiltonian, blow up as σ approaches a non-regular state. The following result yields a simple special case in which strong stabilizability is easy to determine.

Lemma 8.2.6. *Let $H_0 = \sum_{i=1}^m A_i \otimes B_i$ and assume that all A_i commute or that all B_i commute. Then there exists $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ such that $H_U \equiv 0$. In particular, in this case every state is strongly stabilizable. The analogous result holds,* mutatis mutandis, *in the bosonic and fermionic cases.*

8.3 Speed Limit and Control Time

By *speed limit* we simply mean an upper bound on the velocity that any solution to the given control system can achieve. Note that the full control system ($\mathcal{H}$) (resp. ($\mathcal{H}^s$)) does not have any such speed limit, since the controls may be unbounded, but, by construction, the reduced control system (Σ) (resp. (Σ^s) and (Σ^a)) always admits a (finite) speed limit, cf. Proposition 2.4.1.

For any matrix $\Omega \in \mathbb{R}^{n,n}$, we write $\|\Omega\|_\infty$ for the largest singular value of Ω. This is exactly the operator norm with respect to the usual Euclidean norm, and hence it is clear that for $\Omega \in \mathfrak{so}(n, \mathbb{R})$, the norm $\|\Omega\|_\infty$ corresponds to the largest velocity that Ω achieves on the unit sphere. This immediately yields the following result:

Lemma 8.3.1. *Let $\sigma : [0,T] \to S^{d_{\min}-1}$ be any solution to (Σ). Then it holds that $\|\dot{\sigma}(t)\| \leq \max_U \|H_U\|_\infty$ almost everywhere.*[1] *The analogous result holds,* mutatis mutandis*, in the bosonic and fermionic cases.*

Hence we need to find a good upper bound for $\|H_U\|_\infty$ over all $U \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$.

Lemma 8.3.2. *Let $H_0 = \sum_{k=1}^r A_k \otimes B_k$. Then*

$$\max_{U \in \mathrm{U}_{\mathrm{loc}}(d_1,d_2)} \|H_U\|_\infty \leq \sqrt{\sum_{k=1}^r \|A_k\|_2^2 \|B_k\|_2^2},$$

where $\|A\|_2 = \sqrt{\mathrm{tr}(A^*A)}$ denotes the Frobenius norm. The same bound holds a fortiori for the bosonic case $\max_{U \in \mathrm{U}^s_{\mathrm{loc}}(d)} \|H^s_U\|_\infty$ and the fermionic case $\max_{U \in \mathrm{U}^s_{\mathrm{loc}}(d)} \|H^a_U\|_\infty$.

Proof. The Frobenius norm $\|\cdot\|_2$ and the spectral norm $\|\cdot\|_\infty$ are related by $\|\cdot\|_\infty \leq \|\cdot\|_2$, see [HJ12, Prob. 5.6.P23]. Using the Cauchy-Schwarz inequality we compute for any $U = V \otimes W \in \mathrm{U}_{\mathrm{loc}}(d_1, d_2)$ that

$$\|H_U\|_\infty^2 \leq \sum_{i,j=1}^{d_{\min}} |(H_U)_{ij}|^2 \leq \sum_{k=1}^r \sum_{i,j=1}^{d_{\min}} |(V^* A_k V)_{ij}|^2 |(W^* B_k W)_{ij}|^2 \leq \sum_{k=1}^r \|A_k\|_2^2 \|B_k\|_2^2.$$

This concludes the proof in the case of distinguishable subsystems. The bound continues to hold in the bosonic and fermionic cases since restricting the drift or the controls cannot lead to faster evolution of the singular values. □

To obtain a lower limit on the time needed to reach any target state from any initial state, we also need to know the largest distance between any pair of points. This is the diameter of the space[2], and due to the Weyl group symmetry every state has an equivalent state in the Weyl chamber. Hence we are particularly interested in the diameter of the Weyl chamber, which is given in the following result.

Lemma 8.3.3. *Consider the unit sphere S^{d-1} embedded in $\mathbb{R}^d$ and let $\mathbf{W} = \mathbb{Z}_2 \wr S_d$ be the Weyl group acting by coordinate reflections and permutations. Then the corresponding Weyl chamber has diameter $\arccos(\frac{1}{\sqrt{d}}) \in [\frac{\pi}{4}, \frac{\pi}{2})$.*

Proof. First recall that the shortest distance on the sphere between two points $x, y \in S^{d-1}$ is given by $\arccos(x^\top y)$. The maximal distance in the Weyl chamber is achieved by two of its corners. Then it is clear that these points are $x = (1, 0, \ldots, 0)$ and $y = (\frac{1}{\sqrt{d}}, \ldots, \frac{1}{\sqrt{d}})$ for the standard Weyl chamber $S^{d-1}_{\downarrow}$. The result follows immediately. □

The *control time* $T^\star$ of a control system is the shortest (infimum) time sufficient to reach any state from any other state. Using the upper bound on the speed of a solution, the lower bound on the diameter of the Weyl chamber and Theorem 8.2.3, we can give a lower bound on the control time:

Theorem 8.3.4 (Singular Value Speed Limit). *The control time $T^\star$ of the full control system ($\mathcal{H}$)*

[1] The maximum exists and is achieved since the map $U \mapsto \|H_U\|_\infty$ is continuous on a compact domain.

[2] Note that distances in the reduced state space are computed as the length of the shortest geodesic joining two points on the sphere. Hence, somewhat unintuitively, the diameter of a unit hypersphere is π, which is the distance between two antipodal points.

(resp. ($\mathcal{H}^s$)) is finite and satisfies

$$T^\star \geq \frac{\pi/4}{\max_U \|H_U\|_\infty} \geq \frac{\pi}{4\sqrt{\sum_k \|A_k\|_2^2 \|B_k\|_2^2}}\,.$$

CHAPTER 9

Optimal Entanglement Generation

9.1 Introduction

In this chapter we apply optimal control theory to the reduced control systems of Chapter 7 in some low dimensional cases. This allows us, for instance, to find the optimal control sequence to prepare a maximally entangled state, or conversely to disentangle a given state. Moreover, we address the problem of stabilizing states with a certain amount of entanglement.

Outline

The following sections use the reduced control system to derive time-optimal controls and stabilizing controls for the full control system in several concrete settings. In particular, Section 9.2 treats those cases where the reduced state space is one-dimensional, for instance the case of two distinguishable qubits. Finally, a higher-dimensional case of two distinguishable qutrits is studied in Section 9.3 using the Pontryagin Maximum Principle to derive time-optimal solutions for preparing maximally entangled states.

9.2 Optimal Control of the Schmidt Angle

Let us now turn to the question of optimal control. Given two quantum states, the challenge is to find a solution connecting them in the least amount of time possible. Additionally we want to be able to stabilize states with a desired amount of entanglement.

In this section we consider the cases in which the reduced state space is one-dimensional. In these cases the state can be described by a single value, called the Schmidt angle χ. We will start with the simplest non-trivial setting of two coupled qubits, before treating the bosonic and the fermionic case. In each case we first derive a speed limit of the reduced control system, and then turn to the full control system to compute the corresponding optimal controls and to show how states with a prescribed set of singular values can be stabilized.

Two Distinguishable Qubits

First we treat the case of two distinguishable qubits with an arbitrary coupling Hamiltonian H_0. Using local unitary control, the state $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ can always be brought into diagonal form $|\psi\rangle = \sigma_1 |11\rangle +$

$\sigma_2 |22\rangle$, where $\sigma_1, \sigma_2 \in \mathbb{R}$ are the singular values of $|\psi\rangle$. Due to the normalization of the state $|\psi\rangle$, it holds that $\sigma_1^2 + \sigma_2^2 = 1$ and hence there exists an angle χ such that $\sigma_1 = \cos(\chi)$ and $\sigma_2 = \sin(\chi)$. Commonly χ is called the *Schmidt angle*. Since the singular values are only defined up to signs and order, it suffices to consider states in the region $\chi \in [0, \frac{\pi}{4}]$ (the Weyl chamber), where 0 represents product states and $\frac{\pi}{4}$ represents maximally entangled states. Since the controls $-H_U$ of the reduced control system (Σ) on this circle are the generators of rotations, i.e. $H_U \in \mathfrak{so}(2, \mathbb{R})$, they are described by their angular velocity $\omega(-H_U) = (H_U)_{12} \in \mathbb{R}$. One can show that the set of achievable angular velocities is a closed interval symmetric around 0. Hence the optimal control task boils down to finding the largest achievable angular velocity of the reduced control system, which we will denote $\omega^\star(H_0)$ or simply $\omega^\star$. These results yield quantum speed limits on the evolution of the singular values (and hence the amount of entanglement) for the corresponding systems.

Using the standard Pauli basis of the 2×2 Hermitian matrices

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad P_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and after removing the local part (cf. Remark 9.A.1) of the coupling Hamiltonian H_0, by Lemma 9.A.2 it can be uniquely written as

$$H_0 = \textstyle\sum_{i,j=1}^{3} C_{ij}\, P_i \otimes P_j$$

where $C = C(H_0) \in \mathbb{R}^{3,3}$ is the coefficient matrix (with the indices $1, 2, 3$ corresponding to x, y, z in that order).

Lemma 9.2.1. *Under a local unitary basis transformation $U = V \otimes W$ the coefficient matrix transforms as*

$$C(U^* H_0 U) = R_V^\top C(H_0) R_W,$$

where R_V, and R_W are the corresponding rotation matrices.

Proof. Recall that applying unitary basis transformations to $\mathfrak{su}(2)$ corresponds to three-dimensional rotations. For a given $U \in \mathrm{SU}(2)$ we write R_U for the resulting rotation in $\mathrm{SO}(3)$. More explicitly, for $\vec{a} \in \mathbb{R}^3$ and $\vec{P} = (P_x, P_y, P_z)$, we have the relation $\mathrm{Ad}_U(\vec{a} \cdot \vec{P}) = (R_U \vec{a})\vec{P}$ where $(R_U)_{i,j=1}^3 = \frac{1}{2}\operatorname{tr}(P_i U P_j U^{-1})$, see [Cor84, Sec. 3.5, Thm. I]. We see that for $U = V \otimes W$ we obtain $U^* H_0 U = \sum_{i,j=1}^3 V^* P_i V \otimes W^* P_j W = \sum_{i,j,k,l=1}^3 C_{ij} (R_V^\top)_{ki} P_k \otimes (R_W^\top)_{lj} P_l = \sum_{k,l=1}^3 (R_V^\top C R_W)_{kl} P_k \otimes P_l$, as desired. □

To avoid confusion, note that here we applied local unitary transformations to the physical state space, whereas in Appendix 9.A we apply basis transformations to the space of Hermitian operators.

With this we can derive the exact quantum speed limit $\omega^\star$ for the evolution of the singular values of the state.

Proposition 9.2.2. *Let s_i for $i = 1, 2, 3$ denote the singular values of C in non-increasing order (and chosen to be non-negative). Then $\omega^\star = \mathit{s}_1 + \mathit{s}_2$.*

Proof. First note that for $U = V \otimes W \in \mathrm{U}_{\mathrm{loc}}(d)$ and coupling Hamiltonian H_0, if we define $\tilde{H}_0 = U^* H_0 U$, then $H_U = \tilde{H}_{\mathbb{1}}$. By Lemma 9.2.1 it holds that the coefficient matrix of $\tilde{H}_0$ in the Pauli basis is $C(\tilde{H}_0) = C(U^* H_0 U) = R_V^\top C R_W$. Using Proposition 7.1.5 and Kostant's convexity theorem (cf. [Kos73]) for the real singular value decomposition we find that $\omega(-H_U) = (\tilde{H}_{\mathbb{1}})_{12} =$

$(R_V^\top C R_W)_{12} + (R_V^\top C R_W)_{21} \leq \vartheta_1 + \vartheta_2$. Moreover, using the real singular value decomposition it is clear that there exist unitaries V, W, and hence rotation matrices R_V, R_W, such that

$$R_V^\top C R_W = \begin{pmatrix} 0 & \vartheta_1 & 0 \\ \vartheta_2 & 0 & 0 \\ 0 & 0 & \vartheta_3 \end{pmatrix},$$

and hence the bound is tight. □

The next step is to compute time-optimal (and later stabilizing) controls in the full control system ($\mathcal{H}$). Let $U = V \otimes W$ be a local unitary achieving the speed limit as in the proof of Proposition 9.2.2. Then the path

$$\sigma : [0, T] \to S^1, \quad t \mapsto (\cos(\omega^\star t), \sin(\omega^\star t))$$

is a solution to the reduced control system $\dot{\sigma} = -H_U \sigma$ with $\sigma(0) = (1, 0)$ which reaches the quantum speed-limit of Proposition 9.2.2.

Using Proposition 7.1.13 one can now determine a lifted solution and the corresponding control functions with possible divergences at the non-regular points 0 and $\frac{\pi}{4}$ corresponding to product states and maximally entangled states respectively. Indeed the (local) control Hamiltonian is[1]

$$H_c = E \otimes \mathbb{1} + \mathbb{1} \otimes F = -\mathrm{i}\,\mathrm{Ad}_U \circ (\mathrm{ad}^d_{\sigma(t)})^{-1} \circ \Pi^\perp_\Sigma(\mathrm{i}\tilde{H}_0\,|\sigma(t)\rangle),$$

where $\mathrm{Ad}_U(\cdot) = U(\cdot)U^*$ and $\tilde{H}_0 = U^* H_0 U$. Moreover $(\mathrm{ad}^d_\sigma)^{-1}$ and $\Pi^\perp_\Sigma$ are given in Lemma 7.A.6. Explicitly, denoting $C' = C(\tilde{H}_0)$, for $\chi \neq k\frac{\pi}{4}$ we obtain that[2]

$$\begin{aligned} V^*EV &= -\begin{pmatrix} \frac{1}{2}(C'_{zz} + (C'_{xx} - C'_{yy})\tan(\chi)) & \sec(2\chi)(C'_{xz} - \mathrm{i}C'_{yz} - (C'_{zx} + \mathrm{i}C'_{zy})\sin(2\chi)) \\ \sec(2\chi)(C'_{xz} + \mathrm{i}C'_{yz} - (C'_{zx} - \mathrm{i}C'_{zy})\sin(2\chi)) & \frac{1}{2}(C'_{zz} + (C'_{xx} - C'_{yy})\cot(\chi)) \end{pmatrix} \\ W^*FW &= -\begin{pmatrix} \frac{1}{2}(C'_{zz} + (C'_{xx} - C'_{yy})\tan(\chi)) & \sec(2\chi)(C'_{zx} - \mathrm{i}C'_{zy} - (C'_{xz} + \mathrm{i}C'_{yz})\sin(2\chi)) \\ \sec(2\chi)(C'_{zx} + \mathrm{i}C'_{zy} - (C'_{xz} - \mathrm{i}C'_{yz})\sin(2\chi)) & \frac{1}{2}(C'_{zz} + (C'_{xx} - C'_{yy})\cot(\chi)) \end{pmatrix}. \end{aligned} \tag{9.1}$$

In our case $C'_{xy} = \vartheta_1$, $C'_{yx} = \vartheta_2$ and $C'_{zz} = \vartheta_3$ and the remaining matrix entries vanish. Hence this becomes $E = F = -\frac{\vartheta_3}{2}\mathbb{1}$, which is independent of the state χ and thus also of time. This proves the following result, where we assume for simplicity that the control Hamiltonians linearly span $\mathfrak{u}_{\mathrm{loc}}(2, 2)$.

Proposition 9.2.3. *Let U be a local unitary achieving the speed limit as in Proposition 9.2.2. For a system composed of two qubits initially in the state $|00\rangle$, the following sequence yields a maximally entangled state and does so in minimal time[3]:*

1. *Apply the local unitary U (almost) instantaneously.*
2. *Apply the constant control Hamiltonian $H_c = -\vartheta_3 \mathbb{1} \otimes \mathbb{1}$ for time $\frac{\pi/4}{\vartheta_1 + \vartheta_2}$.*
3. *(Optionally) apply a local unitary to choose the desired maximally entangled state.*
4. *(Optionally) stabilize the final state with an appropriate local compensating Hamiltonian.*

The reverse direction is analogous.

[1] The direct term vanishes since the optimal control unitary is constant in time.

[2] We use the notation $\sec(\chi) = \frac{1}{\cos(\chi)}$ for the secant and $\csc(\chi) = \frac{1}{\sin(\chi)}$ for the cosecant.

[3] Mathematically, even with unbounded controls, the optimal time can only be reached approximately. Hence one should more accurately speak of infimum time.

Note that the local controls in Step 2. only apply a global phase, and so they may be omitted if the global phase is neglected.

Remark 9.2.4. *The fact that the time-optimal solutions in the reduced control system (Σ) are given by rotations of constant speed $\pm\omega^\star$ significantly simplifies the problem since integrating the solution is essentially trivial, and when lifting the control to the full control system ($\mathcal{H}$) the direct term vanishes. In the setting of Chapter 5 for example this simplification does not occur.*

Another important task is that of stabilizing a state such as the maximally entangled state obtained in the previous section. More precisely, here we want to stabilize a state with a certain set of singular values. One can show abstractly, cf. Proposition 8.2.5, that for regular states this is always possible with a fixed control Hamiltonian. It is easy to see that by choosing a local unitary $U = V \otimes W$ such that $C' = C(U^* H_0 U)$ is diagonal, it holds that $H_U \sigma = 0$ for any $\sigma \in S^1$. Thus, from (9.1) we find that a corresponding local compensating Hamiltonian is given by

$$V^* E V = W^* F W = -\tfrac{1}{2}(C'_{zz} + (C'_{xx} - C'_{yy}) \csc(2\chi))\mathbb{1} + \tfrac{1}{2}(C'_{xx} - C'_{yy}) \cot(2\chi) P_z.$$

Indeed, this also works for a maximally entangled state $\chi = \pi/4$, but diverges when approaching product states $\chi \to 0$, even if we ignore the global phase. In that case a different control Hamiltonian does the job. As $\chi \to 0$ the only terms in (9.1) which might blow up are those containing $\cot(\chi)$. Thus, if U is chosen such that $C'_{xx} = C'_{yy}$, which can always be achieved, this does not happen. The resulting compensating Hamiltonian $H_c = E \otimes \mathbb{1} + \mathbb{1} \otimes F$ is then defined by

$$\begin{aligned} V^* E V &= -\tfrac{C'_{zz}}{2}\mathbb{1} - \sec(2\chi)(C'_{yz} + C'_{zy}\sin(2\chi))P_y, \\ W^* F W &= -\tfrac{C'_{zz}}{2}\mathbb{1} - \sec(2\chi)(C'_{zy} + C'_{yz}\sin(2\chi))P_y. \end{aligned}$$

The examples above show that there might exist many local unitaries $U = V \otimes W$ such that $H_U = 0$, and each choice yields its own compensating Hamiltonian. In particular, by choosing the right U we were able to prevent the controls from blowing up at product states and maximally entangled states.

Two Bosonic Qubits

In addition to the systems composed of two distinguishable subsystems considered above, one may also consider indistinguishable subsystems. Such systems are characterized by the fact that swapping the two subsystems changes the state only up to a global phase. If this phase is $+1$ the system is called bosonic, and if it is -1 the system is called fermionic. The theory goes through with only minor adaptations in this indistinguishable setting, see Chapter 7 for the details. In particular, the coupling and control Hamiltonians have to be symmetric under swapping as well. Recall that we denote the symmetric local unitary Lie group and algebra by $\mathrm{U}^s_{\mathrm{loc}}(d)$ and $\mathfrak{u}^s_{\mathrm{loc}}(d)$ respectively.

Let us now consider the case of two bosonic qubits. Again, the coupling Hamiltonian can be expressed in the Pauli basis as

$$H_0 = \textstyle\sum_{i,j=1}^3 C_{ij}\, P_i \otimes P_j, \quad C \in \mathbb{R}^{3,3}.$$

This time the coefficient matrix C is symmetric. Transforming the coupling Hamiltonian using a local unitary $U = V \otimes V \in \mathrm{U}^s_{\mathrm{loc}}(d)$, it follows from Lemma 9.2.1 that $C(U^* H_0 U) = R_V^\top C(H_0) R_V$.

Here and in the next section we will make use of the following simple lemma about symmetric and Hermitian matrices of size 2×2:

Lemma 9.2.5. *Consider a Hermitian matrix $H \in \mathrm{iu}(2)$ and let $\ell_1 \geq \ell_2$ denote its eigenvalues. Then it holds that $|H_{12}| \leq \frac{\ell_1-\ell_2}{2}$ and there is a unitary $U \in \mathrm{SU}(2)$ such that $|(U^*HU)_{12}| = \frac{\ell_1-\ell_2}{2}$. The analogous statement for a real symmetric 2×2 matrix and orthogonal conjugation also holds.*

Proof. Due to the unitary invariance of the Frobenius norm it holds that $(H_{11})^2 + 2|H_{12}|^2 + (H_{22})^2 = \ell_1^2+\ell_2^2$. Since $H_{11}+H_{22} = \ell_1+\ell_2$ there is some $x \in \mathbb{R}$ such that $H_{11} = \frac{\ell_1+\ell_2}{2}+x$ and $H_{22} = \frac{\ell_1+\ell_2}{2}-x$. Together this gives

$$2|H_{12}|^2 = \ell_1^2 + \ell_2^2 - (\tfrac{\ell_1+\ell_2}{2} + x)^2 - (\tfrac{\ell_1+\ell_2}{2} - x)^2 = \tfrac{(\ell_1-\ell_2)^2}{2} - 2x^2.$$

This proves the desired bound on $|H_{12}|$. It remains to show that there is a unitary U such that the diagonal elements of U^*HU coincide. For this we may assume that H is diagonal. Then we compute

$$\frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}\ell_1 & 0\\ 0 & \ell_2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\ell_1+\ell_2 & \ell_1-\ell_2\\ \ell_1-\ell_2 & \ell_1+\ell_2\end{pmatrix},$$

and this concludes the proof for the Hermitian case. The real symmetric case is almost identical. □

This time the exact speed limit is given in terms of the eigenvalues of the coefficient matrix C:

Proposition 9.2.6. *Let $\ell_i(C)$ for $i = 1, 2, 3$ denote the eigenvalues of C in non-increasing order. Then $\omega^\star = \ell_1(C) - \ell_3(C)$.*

Proof. The idea is similar to the proof of Proposition 9.2.2. If $U = V \otimes V$, then by Lemma 9.2.1 it holds that $C(U^*H_0U) = R_V^\top C R_V$. Again it holds that $\omega(H_U) = (\tilde{H}_{\mathbb{1}})_{12} = C(U^*H_0U)_{12} + C(U^*H_0U)_{21}$. Due to Kostant's Convexity Theorem [Kos73], or more precisely the Schur–Horn Theorem [Sch23, Hor54], it holds that $|C(U^*H_0U)_{11} - C(U^*H_0U)_{22}| \leq \ell_1(C) - \ell_3(C)$, and hence if we denote by $\tilde{C}(U^*H_0U)$ the upper left 2×2 block in $C(U^*H_0U)$, it holds that $|\ell_1(\tilde{C}(U^*H_0U)) - \ell_2(\tilde{C}(U^*H_0U))| \leq \ell_1(C) - \ell_3(C)$. By Lemma 9.2.5, $C(U^*H_0U)_{12} + C(U^*H_0U)_{21} \leq \ell_1(C) - \ell_3(C)$ and by the proof of the same lemma this bound is tight. Explicitly, the bound is achieved by choosing $U = V \otimes V$ such that

$$C(U^*H_0U) = R_V^\top C R_V = \begin{pmatrix}\frac{\ell_1+\ell_3}{2} & \frac{\ell_1-\ell_3}{2} & 0\\ \frac{\ell_1-\ell_3}{2} & \frac{\ell_1+\ell_3}{2} & 0\\ 0 & 0 & \ell_2\end{pmatrix}.$$

This concludes the proof. □

Denoting $C' = C(U^*H_0U) = R_V^\top C R_V$, the compensating Hamiltonian is given by $H_c = E \otimes \mathbb{1} + \mathbb{1} \otimes E$ where

$$V^*EV = -\begin{pmatrix}\frac{1}{2}(C'_{zz}+(C'_{xx}-C'_{yy})\tan(\chi)) & \sec(2\chi)(C'_{xz}-\mathrm{i}C'_{yz}-(C'_{xz}+\mathrm{i}C'_{yz})\sin(2\chi))\\ \sec(2\chi)(C'_{xz}+\mathrm{i}C'_{yz}-(C'_{xz}-\mathrm{i}C'_{yz})\sin(2\chi)) & \frac{1}{2}(C'_{zz}+(C'_{xx}-C'_{yy})\cot(\chi))\end{pmatrix},$$

where we used the results of Appendix 7.A.

With these results, the derivation of time-optimal and stabilizing controls is straightforward and analogous to the previous section.

Two Fermionic Four-Level Systems

In the fermionic case the space of singular values has dimension $\lfloor \frac{d}{2} \rfloor$. Hence, if we consider two coupled four-level systems ($d = 4$), the reduced state space is again a circle. Due to Lemma 9.A.4 we can always write the coupling Hamiltonian in diagonal form. Here we will focus on the rank one case, i.e., $H = A \otimes A$ for some Hamiltonian $A \in \mathrm{iu}(4)$.

Lemma 9.2.7. *In the fermionic $d = 4$ case, the reduced control system can equivalently be formulated for the Schmidt angle χ as $\dot{\chi} = \omega_V^a$, where $\omega_V^a = \mathrm{Im}((V^*AV)_{13}(V^*AV)_{24} - (V^*AV)_{14}(V^*AV)_{23})$ with $V \in \mathrm{U}(4)$.*

Proof. This follows immediately from Proposition 7.1.5. □

We begin by deriving upper and lower bounds for the speed limits of the Schmidt angle χ.

Proposition 9.2.8. *Let ℓ_i denote the eigenvalues of A in non-increasing order. Then it holds that*

$$\tfrac{1}{4}(\ell_1 - \ell_3)(\ell_2 - \ell_4) \leq \omega^* \leq \tfrac{1}{16}(\ell_1 + \ell_2 - \ell_3 - \ell_4)^2,$$

and the two bounds coincide when $\ell_1 + \ell_4 = \ell_2 + \ell_3$.

Proof. First we show that we can assume certain elements of A to vanish. Let $A_{(12)}$ denote the 2×2 block of A in the upper right corner. Then, by Lemma 9.2.7 it holds that $\omega_V^a = \mathrm{Im}(\det((V^*AV)_{(12)}))$. Using a block diagonal unitary change of basis and the complex singular value decomposition we may assume that $A_{(12)}$ is diagonal and additionally that $\omega_{\mathbb{1}}^a = |A_{13}A_{24}|$. Now let $A_{[ij]}$ denote the 2×2 submatrices of A obtained by deleting all but the i-th and j-th row and column. Let $a \geq c$ denote the eigenvalues of $A_{[13]}$ and $b \geq d$ those of $A_{[24]}$. Then by Lemma 9.2.5 it holds that $\omega_V^a = \frac{1}{4}(a-c)(b-d)$, and there exists a unitary transformation V such that the diagonal of V^*AV is (a, b, c, d). Hence we need to solve the optimization problem

$$\max \tfrac{1}{4}(a - c)(b - d) \text{ subject to } (a, b, c, d) \preceq (\ell_1, \ell_2, \ell_3, \ell_4).$$

This can be done using a greedy optimization approach. First we show that the maximum is achieved when $a - c = b - d$. Indeed, if $a - c > b - d$ we can smoothly move in the direction $(-1, 1, 0, 0)$, which preserves majorization and increases the objective value, until equality is achieved. Similarly, if $a - c < b - d$ we move in the direction $(0, 0, -1, 1)$. Finally, by moving in the direction $(-1, 1, -1, 1)$, which does not affect the objective value, we may assume that additionally $a = b$ and $c = d$. Under these additional constraints the maximum is easily seen to be $\frac{1}{16}(\ell_1 + \ell_2 - \ell_3 - \ell_4)^2$, as desired.

By Lemma 9.2.5 the lower bound can be achieved as follows. First diagonalize A to obtain the form $\mathrm{diag}(\ell_1, \ell_2, \ell_3, \ell_4)$. Then, using a unitary mixing the levels 1 and 3, as well as 2 and 4, as in the proof of Lemma 9.2.5, we obtain the form

$$V^*AV = \frac{1}{2}\begin{pmatrix} \ell_1 + \ell_3 & 0 & \ell_1 - \ell_3 & 0 \\ 0 & \ell_2 + \ell_4 & 0 & \ell_2 - \ell_4 \\ \ell_1 - \ell_3 & 0 & \ell_1 + \ell_3 & 0 \\ 0 & \ell_2 - \ell_4 & 0 & \ell_2 + \ell_4 \end{pmatrix},$$

which achieves the claimed lower bound. □

Remark 9.2.9. *To see that the upper bound of Proposition 9.2.8 is not tight in general, consider A of rank one, e.g., $A = \mathrm{diag}(1, 0, 0, 0)$. Then it is easy to verify that $\omega_V^a \equiv 0$ for all $V \in \mathrm{U}(4)$, that is, it is effectively local, cf. Proposition 8.2.2.*

In a basis as described by Proposition 9.2.8, the local compensating Hamiltonian takes the form $E \otimes \mathbb{1} + \mathbb{1} \otimes E$ where $W^*EW = -\tilde{E} - \frac{1}{8}(\ell_1 + \ell_3)(\ell_2 + \ell_4)\mathbb{1}$ and

$$\tilde{E} = \begin{pmatrix} \frac{(\ell_1-\ell_3)(\ell_2-\ell_4)\tan(\chi)}{8} & 0 & \frac{(\ell_1-\ell_3)(\ell_2+\ell_4)}{4} & 0 \\ 0 & \frac{(\ell_1-\ell_3)(\ell_2-\ell_4)\tan(\chi)}{8} & 0 & \frac{(\ell_1+\ell_3)(\ell_2-\ell_4)}{4} \\ \frac{(\ell_1-\ell_3)(\ell_2+\ell_4)}{4} & 0 & \frac{(\ell_1-\ell_3)(\ell_2-\ell_4)\cot(\chi)}{8} & 0 \\ 0 & \frac{(\ell_1+\ell_3)(\ell_2-\ell_4)}{4} & 0 & \frac{(\ell_1-\ell_3)(\ell_2-\ell_4)\cot(\chi)}{8} \end{pmatrix}.$$

This Hamiltonian blows up near product states as $\chi \to k\pi/2$ but not near maximally entangled states.

Let us briefly look at stabilization. By choosing a local basis which diagonalizes A, we obtain the compensating local Hamiltonian $W^*EW = -\operatorname{diag}(\frac{\ell_1\ell_2}{2}, \frac{\ell_1\ell_2}{2}, \frac{\ell_3\ell_4}{2}, \frac{\ell_3\ell_4}{2})$, which is independent of χ.

9.3 Optimal Control of Two Qutrits

In this section we consider a higher dimensional system, namely one composed of two distinguishable three-level systems (qutrits). In this case there are three singular values and hence the state space of the reduced control system is the usual two-dimensional sphere S^2. Compared to the previous section, it is now not at all obvious which path to take between two points on the Schmidt sphere in order to minimize the time. To determine such an optimal path we will make use of the Pontryagin Maximum Principle (PMP) [BSS21].

First we define the reduced control system by characterizing the set of generators $\mathfrak{H}$. Then we use the PMP to find time-optimal solutions to the reduced control system. Finally we (approximately) lift these solutions to the original control system and derive corresponding control functions.

Reducing the Problem

In this section we assume that the two distinguishable subsystems are qutrits, i.e. $d_1 = d_2 = 3$, and that the coupling Hamiltonian $H_0 = A \otimes B$ has rank one. Moreover, we assume that A and B have equidistant eigenvalues.[4] We start with a simple lemma which bounds the size of the off-diagonal elements of A and B in terms of their respective eigenvalues.

Lemma 9.3.1. *Let $A \in \mathrm{iu}(3)$ with equidistant eigenvalues be given and let $a = (A_{32}, A_{13}, A_{21}) \in \mathbb{C}^3$. Then it holds that $\|a\|_2 \leq \frac{1}{2}(\ell_1 - \ell_3)$ where the ℓ_i denote the eigenvalues of A in non-increasing order.*

Proof. Since the expression $\|a\|_2 \leq \frac{1}{2}(\ell_1 - \ell_3)$ is invariant under addition of a multiple of the identity to A, we may assume that A is traceless, and hence the eigenvalues are $-\ell \leq 0 \leq \ell$. Now it holds that $2\ell^2 = \|A\|_2^2 \geq 2\|a\|_2^2$ and hence $\|a\|_2 \leq \ell$ as desired. □

The first step in solving the optimal control problem is to understand the reduced control system. We already know that the reduced control system is defined on the sphere S^2 and that the controls are given by the set $\mathfrak{H} \subset \mathfrak{so}(3)$ consisting of rotation generators. The goal of this section is to understand the precise shape of this set (or at least its convex hull). It is convenient to represent generators in $\mathfrak{so}(3)$ using vectors in $\mathbb{R}^3$. Indeed, for every $-H_U \in \mathfrak{H}$ there is a unique vector $\omega_U \in \mathbb{R}^3$ such that $-H_U\sigma = \omega_U \times \sigma$. Recall that the 1-norm of such a vector is given by $\|\omega\|_1 := |\omega_1| + |\omega_2| + |\omega_3|$.

Proposition 9.3.2. *Assume that A and B have equidistant eigenvalues. Then, for all $U \in \mathrm{U}_{\mathrm{loc}}(3,3)$ it holds that*

$$\|\omega_U\|_1 \leq \omega^\star(H_0) := \frac{(\ell_1(A) - \ell_3(A))(\ell_1(B) - \ell_3(B))}{4}.$$

[4] We say that A has equidistant eigenvalues if $\ell_1(A) - \ell_2(A) = \ell_2(A) - \ell_3(A)$.

Proof. Using Proposition 7.1.5, the Cauchy–Schwarz inequality and Lemma 9.3.1 we compute

$$\|\omega_U\|_1 = \sum_{i=1}^3 |\operatorname{Im}(\tilde{a}_i \tilde{b}_i)| \leq \sum_{i=1}^3 |\tilde{a}_i||\tilde{b}_i| \leq \|\tilde{a}\|_2 \|\tilde{b}\|_2 \leq \omega^\star,$$

where, denoting $U = V \otimes W$, we set $\tilde{a} = ((V^*AV)_{32}, (V^*AV)_{13}, (V^*AV)_{21})$ and similar for $\tilde{B}$. □

Geometrically this has a nice interpretation. Let $O_3 \subset \mathbb{R}^3$ denote the regular octahedron, i.e., the convex hull $O_3 = \operatorname{conv}((\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1))$, see Figure 9.1, and recall from Remark 7.1.9 that one can define a relaxed control system via $\dot{\sigma}(t) \in \operatorname{conv}(\mathfrak{H}\sigma)$, which is still approximately equivalent to (Σ).

Corollary 9.3.3. *The convex hull of the set of induced vector fields considered in* $\mathbb{R}^3 \cong \mathfrak{so}(3)$ *is a regular octahedron:*

$$\operatorname{conv}(\omega_U : U \in \mathrm{U}_{\mathrm{loc}}(3,3)) = \omega^\star(H_0)\, O_3.$$

Thus, in the relaxed control system, two reduced states $\sigma, \tau \in S^2$ *can always be joined in time* $T \in \omega^\star(H_0) \arccos(\sigma \cdot \tau)[1/\sqrt{3}, 1]$.

Proof. The inclusion $\subseteq$ follows immediately from Proposition 9.3.2. To get equality one just has to obtain the vertices of the octahedron, which can be done in a manner similar to the proof of Lemma 9.2.5. The bounds on T follow immediately from the fact that the spheres of radius $1/\sqrt{3}$ and 1 are respectively the inscribed and circumscribed spheres of the regular octahedron. See also Section 8.3. □

For this reason we will work with the relaxed system when we derive time-optimal solutions in the following section. Note also that this bound is stronger than the general bound obtained in Section 8.3.

Solving the Reduced Problem

By rescaling we may assume without loss of generality that $\omega^\star = 1$ and hence the reduced control system becomes

$$\dot{\sigma}(t) = u(t) \times \sigma(t), \quad \sigma(0) = \sigma_0 \in S^2$$

where the control function $u : [0, T] \to \mathbb{R}^3$ is measurable and satisfies

$$\|u(t)\|_1 = |u_x(t)| + |u_y(t)| + |u_z(t)| \leq 1 \tag{9.2}$$

for almost all $t \in [0, T]$.

The time-optimal control problem can be solved using the Pontryagin Maximum Principle (PMP) [AS04, Pon+86, Sac22]. An introduction to the PMP in the context of quantum control theory is given in [BSS21]. The PMP is a first-order necessary condition satisfied by optimal trajectories. We introduce an *adjoint state* $p \in \mathbb{R}^3$ and define the *pseudo-Hamiltonian* $H_{\mathsf{p}}(u, \sigma, p) = p \cdot (u \times \sigma)$. The dynamics of σ and p follow the Hamilton–Jacobi equations, in particular $\dot{p} = -\partial H_{\mathsf{p}}/\partial\sigma = u \times p$. For this kind of control system on the sphere, we can define the variable[5] $l = \sigma \times p$ (note that this implies that l is orthogonal to σ) that allows us to express the pseudo-Hamiltonian in the convenient form:

$$\tilde{H}_{\mathsf{p}}(u, l) = u \cdot l. \tag{9.3}$$

We can show using the Hamilton–Jacobi equations that the dynamics of l follow:

$$\dot{l} = \dot{\sigma} \times p + \sigma \times \dot{p} = (u \times \sigma) \times p + (p \times u) \times \sigma = u \times l. \tag{9.4}$$

[5] Strictly speaking p is a cotangent vector $p \in T^*_\sigma S^2$. When considered as a vector in $\mathbb{R}^3$, it is therefore restricted to be orthogonal to σ at all times. Thus the relation between p and l is bijective.

The PMP states (see [BSS21, Thm. 5]) that the time-optimal control has to maximize this pseudo-Hamiltonian under the constraint (9.2). For this we determine which values of u (under the given constraint) maximize $\tilde{H}_{\mathsf{p}}(u(t), l(t))$ for given l:

Lemma 9.3.4. *Given $l \in \mathbb{R}^3$, the vertex $u = \pm e_i$ of O_3 maximizes $\tilde{H}_{\mathsf{p}}(u, l) = u \cdot l$ if and only if $\pm l_i = \max(|l_x|, |l_y|, |l_z|)$. The set of all maximizers is then simply the convex hull of such vertices, and defines a face of the octahedron.*

Proof. Due to the normalization of u it is clear that $u \cdot l \leq \max(|l_x|, |l_y|, |l_z|)$. Moreover, since the maximization is linear and takes place on the octahedron, which is a convex polytope, the maximum is achieved exactly on a face (which may be a vertex, an edge, or a facet) of the octahedron, and hence defined by a subset of vertices. It is easy to see that the vertices which maximize $u \cdot l$ are exactly the ones given in the statement, and hence the set of maximizers is the convex hull of these vertices. □

Remark 9.3.5. *This result can be visualized intuitively by considering the polar dual of the octahedron O_3, which is nothing but the cube with vertices $(\pm 1, \pm 1, \pm 1)$ denoted C_3. If l lies on a certain face of the cube, then there is a unique corresponding face of the octahedron containing all u which maximize $H_{\mathsf{p}} = u \cdot l$. Additionally it turns out that only the barycenters of the faces of the octahedron yield relevant values of u. This will become clear later when we describe all possible evolutions of l. The duality and the barycenters are shown in Figure 9.1.*

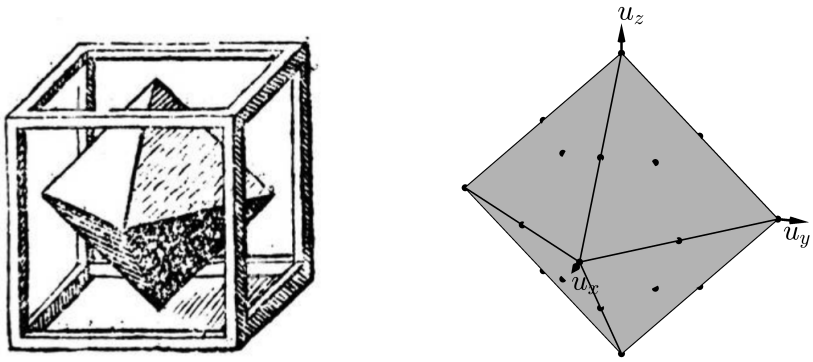

Figure 9.1: Left: The octahedron O_3 and its polar dual, the cube C_3, reproduced from [Kep19]. Note that the d-dimensional faces of O_3 correspond one-to-one with $(2-d)$-dimensional faces of C_3, that is, vertices correspond to facets and edges to edges. Right: The control u can take values at the black points which are the barycenters of the faces of the octahedron O_3, cf. Remark 9.3.5.

The dynamics of l are given by (9.4) and constrained by Lemma 9.3.4. Since the optimal value of u is not always uniquely defined, this yields a differential inclusion instead of a differential equation. Thus the solution is in general not uniquely determined by the initial condition $l(0)$. However, we will see that unique solutions can still be obtained for the optimal control problem.

To understand the evolution of l, note that it has two constants of motion. The first one is due to the fact that l moves on a sphere (recall that $\dot{l} = u \times l$), and the second one, referred to as the *Pontryagin Hamiltonian*, is obtained by substituting any optimal u in the pseudo-Hamiltonian (9.3). They are given by:

$$L^2 = |l_x|^2 + |l_y|^2 + |l_z|^2, \quad \mathcal{H} = \max\{|l_x|, |l_y|, |l_z|\}. \tag{9.5}$$

The first equation corresponds to a sphere of radius L which can be set to 1 without loss of generality, and the second to a cube of side length $2\mathcal{H}$. Any solution of the system must remain on to the intersection of these two surfaces, illustrated in Figure 9.2.

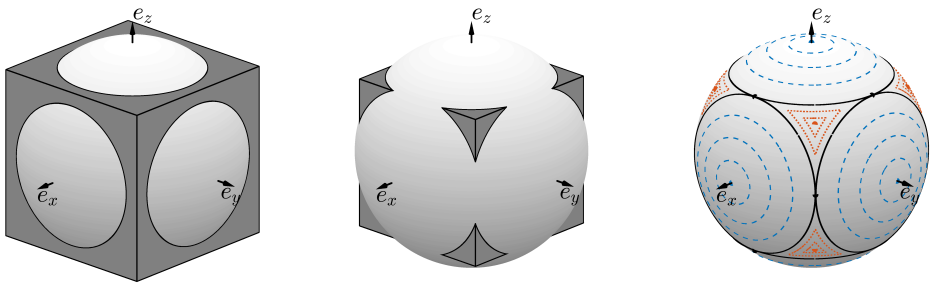

Figure 9.2: Geometric representation of the two first integrals given in Eq. (9.5). The solution $l(t)$ lies on the intersection of these two surfaces. We obtain different families shown on the rightmost panel, namely the constant case (blue dashed lines) for $\mathcal{H} \in (\frac{1}{\sqrt{2}}, 1]$, the switching case (dotted red lines) for $\mathcal{H} \in [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}})$, and the separatrix (black line) for $\mathcal{H} = \frac{1}{\sqrt{2}}$.

The general solution $l(t)$ can be decomposed into two main families, namely the *switching case* and the *constant case*, depending on $\mathcal{H}$ (which depends on $l(0)$). In the constant case, one of the components of $l(t)$ is always dominant. These curves imply constant controls. In the switching case, $l(t)$ follows a concatenation of three circular arcs. In the positive octant, the control jumps between $u_x \to u_z \to u_y \to u_x \to \cdots$, with a duration $\Delta t = \pi/2 - 2\arccos(\mathcal{H}/\sqrt{1-\mathcal{H}^2})$ between two switches. For $\mathcal{H} = \frac{1}{\sqrt{3}}$ the switching solution degenerates into a solution with constant controls of the form $u = (\pm\frac{1}{3}, \pm\frac{1}{3}, \pm\frac{1}{3})$, corresponding to the barycenters of the facets of the octahedron of Figure 9.1. A special case is the black curve separating these two families. If $l(0)$ starts somewhere on this curve, it follows it for a while until it reaches one of the unstable equilibrium points (in black). It can stay on this point for a certain amount of time and then follow any of the trajectories connected to this point. These dynamics correspond to a control that is originally such that, for example, $u = (1, 0, 0)$ during a certain time and then switches to, e.g, $u = (\frac{1}{2}, \frac{1}{2}, 0)$ (if the unstable point is in the xy-plane) and stays for a certain time. It can thus continue with $u = (1, 0, 0)$, or switch to $u = (0, 1, 0)$. The time it stays on the unstable equilibrium depends on the trajectory $\sigma(t)$ in the state space that one wants to achieve, and in particular on the desired final state.

So far we have considered the reduced control system on the entire sphere. However, since the singular values of the quantum state are only defined up to order and sign, there is an additional symmetry, and we may focus on the part of the sphere where $\sigma_z \geq \sigma_x \geq \sigma_y \geq 0$. This is called the *Weyl chamber* and illustrated in Figure 9.3. Indeed, for any solution of the reduced control system one can consider the corresponding path in the Weyl chamber obtained by taking the absolute value of the singular values and ordering them appropriately, and this is also guaranteed to be a solution, see Proposition 2.A.4.

Consider a solution σ starting at the north pole $\sigma(0) = \sigma_0 := (0, 0, 1)$ and let τ denote any desired final state in the Weyl chamber. Then it holds that $l(0)$ lies in the xy-plane. The only possible solution steering σ to τ which satisfies the PMP and remains within the Weyl chamber is to use $u = (-\frac{1}{2}, \frac{1}{2}, 0)$ for time T_1 and then to use $u = (0, 1, 0)$ for time T_2 where

$$T_1 = \sqrt{2}\arccos\left(\sqrt{1-2\tau_y^2}\right), \qquad T_2 = \arccos\left(\frac{\tau_x\tau_y + \tau_z\sqrt{1-2\tau_y^2}}{1-\tau_y^2}\right).$$

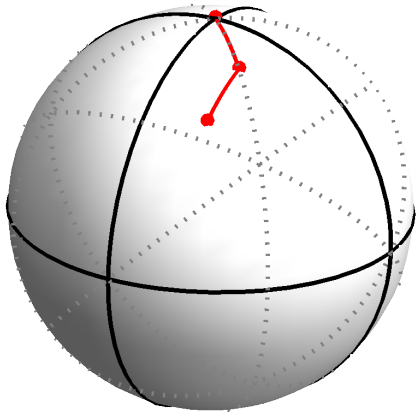

Figure 9.3: The singular values of the quantum state are only defined up to order and signs. On the Schmidt sphere, coordinate permutations and sign flips act via reflections. These reflections divide the sphere into small triangular regions, which can be mapped to each other uniquely and are hence equivalent. We choose one of these triangles (namely the one satisfying $\sigma_z \geq \sigma_x \geq \sigma_y \geq 0$) as state space for the reduced control system and call it the *Weyl chamber*. In red we show the optimal solution starting from the north pole of the Schmidt sphere and ending at an arbitrary point in the Weyl chamber as obtained from the PMP.

In particular, if the final state is $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ (corresponding to a maximally entangled state), then $T_1 = \sqrt{2}\arccos(\frac{1}{\sqrt{3}})$ and $T_2 = 0$. In the following section we will further investigate this solution.

Lifting the Solution

In the final step we derive (approximately) optimal controls for the full control system ($\mathcal{H}$) consisting of two coupled qutrits. For a regular solution $\sigma : [0, T] \to S^2$ with continuously differentiable controls, this can be done exactly (up to some near instantaneous pulses at the beginning and at the end of the solution) using Proposition 7.1.13. For non-regular solutions the existence of an exact lift is not guaranteed. In this section we derive a lifted solution starting at the product state $|\psi_0\rangle = |33\rangle$ and finishing at (or rather arbitrarily close to) the maximally entangled state $|\psi_1\rangle = \frac{1}{\sqrt{3}}(|11\rangle + |22\rangle + |33\rangle)$.

In the previous section we have seen that the optimal solution in the reduced control system starting at $(0, 0, 1)$ and ending at $\frac{1}{\sqrt{3}}(1, 1, 1)$ is simply given by a segment of the corresponding great circle traversed at angular velocity $\omega^\star/\sqrt{2}$, where $\omega^\star$ is given in the beginning of Section 9.3 and depends on the coupling Hamiltonian $H_0 = A \otimes B$. To simplify the notation we assume (without loss of generality) that the eigenvalues of A and B are $1, 0, -1$ and hence $\omega^\star = 1$. Concretely, the optimal solution is given by

$$\sigma(t) = \big(\tfrac{1}{\sqrt{2}}\sin(\tfrac{1}{\sqrt{2}}t), \tfrac{1}{\sqrt{2}}\sin(\tfrac{1}{\sqrt{2}}t), \cos(\tfrac{1}{\sqrt{2}}t)\big),$$

and the maximally entangled state is reached at time $T^\star = \sqrt{2}\arccos(\frac{1}{\sqrt{3}})$. The optimal derivative in the reduced control system is achieved by the generator $-H_U$ using any local unitary $U = V \otimes W$

which satisfies

$$V^*AV = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&1\\0&0&1\\1&1&0\end{pmatrix}, \quad W^*BW = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&-\mathrm{i}\\0&0&-\mathrm{i}\\\mathrm{i}&\mathrm{i}&0\end{pmatrix}. \tag{9.6}$$

Such U exists by assumption on the eigenvalues of A and B. Without applying any controls (other than instantaneously applying U and U^* and the beginning and the end respectively) the evolution of the system is

$$e^{-\mathrm{i}U^*A\otimes BUt}\left|33\right\rangle = -\tfrac{\sin(t)}{2}(\left|11\right\rangle + \left|12\right\rangle + \left|21\right\rangle + \left|22\right\rangle) + \cos(t)\left|33\right\rangle,$$

which, unfortunately, is not a lift of the optimal solution as the singular values are $(\sin(t), 0, \cos(t))$. Indeed, there is no compensating Hamiltonian which yields an exact lift. This is because the local control Hamiltonian applied to the system cannot directly affect the derivative of the singular values at time 0 (see Chapter 1). For more background on this issue see Section 2.3. The origin of the problem is that all states on the solution satisfy $\sigma_x = \sigma_y$, and hence the usual formula for the compensating Hamiltonian does not apply (and might blow-up). This will be made clearer in the following.

One way to fix this problem is to remain on a path which narrowly avoids the degeneracy $\sigma_x = \sigma_y$ and to compute the corresponding compensating Hamiltonian. We will consider a solution which remains on the circular path satisfying $\sigma_x = \sigma_y + \varepsilon\sqrt{2}$ for some small $\varepsilon > 0$. Moreover we use the same generator $-H_U$ with local control unitary $U = V \otimes W$ as above. As in Section 9.2 we can derive the corresponding local compensating Hamiltonian $E \otimes \mathbb{1} + \mathbb{1} \otimes F$. We obtain

$$V^*EV = W^*FW = \frac{\sigma_z}{2\sqrt{2}\varepsilon}P'_y, \quad \text{where } P'_y = \begin{pmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&0\\0&0&0\end{pmatrix},$$

which blows up as ε approaches 0, but is well-behaved for $\varepsilon \neq 0$. Note also the state dependence via σ_z. More precisely, consider the initial state $\frac{\varepsilon}{\sqrt{2}}\left|11\right\rangle - \frac{\varepsilon}{\sqrt{2}}\left|22\right\rangle + \sqrt{1-\varepsilon^2}\left|33\right\rangle$. Since the reduced solution moves on a circle in a plane orthogonal to the $(y-x)$-axis, the reduced solution is given by

$$\big(\tfrac{\sqrt{1-\varepsilon^2}\sin(t/\sqrt{2})+\varepsilon}{\sqrt{2}}, \tfrac{\sqrt{1-\varepsilon^2}\sin(t/\sqrt{2})-\varepsilon}{\sqrt{2}}, \sqrt{1-\varepsilon^2}\cos(t/\sqrt{2})\big).$$

Hence, the control Hamiltonian can be written in the time-dependent (instead of state-dependent) way using $V^*EV = W^*FW = \frac{\sqrt{1-\varepsilon^2}\cos(t/\sqrt{2})}{2\sqrt{2}\varepsilon}P'_y$.

It is also interesting to consider what happens if one applies this control to a solution starting at $U\left|33\right\rangle$. To simplify things we get rid of the time dependence and use the control Hamiltonian given by $V^*EV = W^*FW = \frac{1}{2\sqrt{2}\varepsilon}P'_y$. Some solutions of this form are shown in Figure 9.4. These solutions do not exactly reach a maximally entangled state, but for small ε they get very close. Indeed, consider the cost function

$$C(\varepsilon) = \|\mathrm{sing}(\left|\psi(T^\star)\right\rangle) - \tfrac{1}{\sqrt{3}}(1,1,1)\|_2$$

measuring the Euclidean distance of the reduced state at the final time to the maximally entangled state. This function is plotted in Figure 9.5 and has some interesting properties. In particular, by trial and error one finds that the function becomes almost exactly periodic when transformed as

$$\tilde{C}(x) = \sqrt{3}C\big(\tfrac{1}{2\sqrt{2}x}\big)(2\sqrt{2}x - 1).$$

This allows us to derive a simple approximation for the local minima of $C(\varepsilon)$. Indeed, the points

$$\varepsilon_k = (2\sqrt{2}(x_0 + k\Delta x))^{-1}, \quad \text{where } x_0 \cong 0.0048,\ \Delta x \cong 2.3252, \tag{9.7}$$

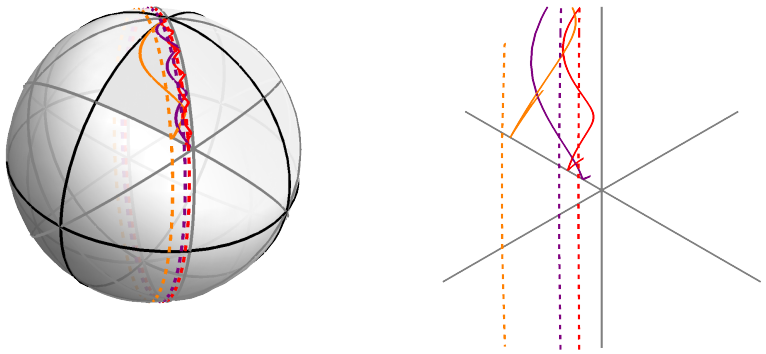

Figure 9.4: Three solutions, with different values of ε, to the optimial control problem projected to the singular values are depicted. In each case the solution (solid) is drawn with the circle satisfying $x - y = \varepsilon\sqrt{2}$ (dashed). In each case the solution starts at $|33\rangle$ and runs for time $T^\star$ to approximate a maximally entangled state. The values for ε chosen are $\varepsilon_1 = 0.12$ (orange), $\varepsilon_2 = 0.0506$ (purple), and $\varepsilon_3 = 0.0276$ (red). The final distances are $C(\varepsilon_1) = 0.140, C(\varepsilon_2) = 0.0215$, and $C(\varepsilon_3) = 0.0436$. Importantly, even though $\varepsilon_2 > \varepsilon_3$, the former solution achieves a better result, since the value of ε_2 is chosen such that it achieves a local minimum of the final distance function.

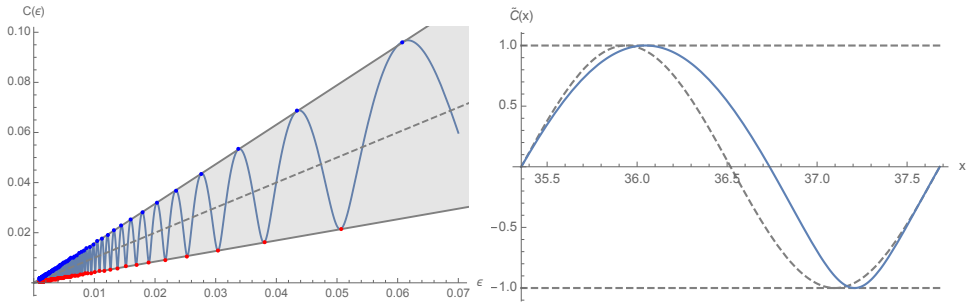

Figure 9.5: Left: Plot of the cost function $C(\varepsilon)$. The gray dashed line is $\varepsilon \mapsto \varepsilon$, and the upper and lower gray lines are $\varepsilon \mapsto (1 \pm 1/\sqrt{3})\varepsilon$. The (approximate) maxima and minima computed using (9.7) are highlighted. Thus the cost decreases linearly with ε, but, as also shown in Figure 9.4 the exact choice of ε can make a significant difference. Right: The transformed cost $\tilde{C}(x)$ is almost perfectly periodic (for x large enough and hence ε small enough). Here we depict a single period of $\tilde{C}(x)$ (blue) alongside a pure sine wave (gray dashed) for comparison.

are close to local minima of C for integer k and close to local maxima for half-integer k. In practice one should always choose ε in such a local minimum as it can significantly decrease the final cost. Indeed, a well chosen ε achieves a cost similar to that of a badly chosen ε which is $\frac{1+1/\sqrt{3}}{1-1/\sqrt{3}} \cong 3.7$ times smaller.

Finally, once a maximally entangled state is reached, we also wish to stabilize the state. This can be achieved by switching into the basis where $A, B = \mathrm{diag}(1, 0, -1)$. Using the local compensating Hamiltonian $E \otimes \mathbb{1} + \mathbb{1} \otimes F$ with $V^*EV = W^*FW = \mathrm{diag}(\frac{1}{2}, 0, \frac{1}{2})$ one finds that $|\psi_1\rangle$ is indeed stabilized.

9.A Decomposition of Drift Hamiltonians

It is most convenient to work with drift Hamiltonians expressed in the form $H_0 = \sum_{i=1}^{m} A_i \otimes B_i$. Given an arbitrary drift Hamiltonian $H_0 \in \mathfrak{iu}(d_1 d_2)$ we want to understand the different ways in which it can be written as a sum of decomposable elements $A_i \otimes B_i$.

It is easy to see that any local term of the form $E \otimes \mathbb{1}$ or $\mathbb{1} \otimes F$ has no effect in the reduced control system. Indeed such terms can be determined uniquely:

Remark 9.A.1. *Every Hermitian matrix $H_0 \in \mathfrak{iu}(d)$ can be written uniquely as $H_0 = \tilde{H}_0 + \frac{\operatorname{tr}(H_0)}{d}\mathbb{1}$ where $\tilde{H}_0 \in \mathfrak{isu}(d)$ is a traceless Hermitian matrix. Indeed, this decomposition is orthogonal with respect to the Hilbert–Schmidt inner product. For product Hamiltonians $A \otimes B$ we analogously get the unique decomposition into four terms $A \otimes B = \tilde{A} \otimes \tilde{B} + \frac{\operatorname{tr}(B)}{d_2} A \otimes \mathbb{1} + \frac{\operatorname{tr}(A)}{d_1} \mathbb{1} \otimes B + \frac{\operatorname{tr}(A)\operatorname{tr}(B)}{d_1 d_2} \mathbb{1} \otimes \mathbb{1}$. Since the three latter terms are local, they can be compensated using local unitary control and hence one may for simplicity assume that all Hamiltonians are traceless. Indeed, while the statements in this section are formulated for general Hermitian matrices, analogous results hold for the traceless case. This is similar to the unique decomposition of Lindblad generators into coherent and dissipative parts, cf. Lemma 4.A.3 and [End23].*

Lemma 9.A.2. *Let $H_0 \in \mathfrak{iu}(d_1 d_2)$ be an arbitrary Hamiltonian and let A_i for $i = 1, \ldots, d_1^2$ be a basis of $\mathfrak{iu}(d_1)$ and similarly for B_j for $j = 1, \ldots, d_2^2$. Then there is a unique coefficient matrix $C \in \mathbb{R}^{d_1, d_2}$ such that*

$$H_0 = \sum_{i,j=1}^{d_1^2, d_2^2} C_{ij} A_i \otimes B_j.$$

If $A_i = \sum_{k=1}^{d_1^2} S_{ik} \tilde{A}_k$, $B_j = \sum_{l=1}^{d_2^2} T_{jl} \tilde{B}_l$ is another choice of bases, with $S \in \mathbb{R}^{d_1, d_1}$ and $T \in \mathbb{R}^{d_2, d_2}$ invertible, then

$$H_0 = \sum_{k,l=1}^{d_1^2, d_2^2} \tilde{C}_{kl} \tilde{A}_k \otimes \tilde{B}_l \quad \text{where } \tilde{C} = S^\top C\, T.$$

As a consequence the rank r of the coefficient matrix is well-defined, i.e. it depends only on H_0. Hence the bases A_i and B_j can always be chosen such that

$$H_0 = \sum_{i=1}^{r} A_i \otimes B_i,$$

and we will say that this representation is "in diagonal form".

Proof. This holds for arbitrary tensor products, cf. [Rom05, Thm. 14.7]. Alternatively this easily follows from the real singular value decomposition and elementary computations. □

Corollary 9.A.3. *Moreover, there exist orthonormal bases A_i and B_j (with respect to any given inner product) such that*

$$H_0 = \sum_{i=1}^{r} \omega_i A_i \otimes B_i,$$

and we will again say that this representation is "in diagonal form". The ω_i are the singular values of the coefficient matrix with respect to any orthonormal basis, and hence they are uniquely defined up to order and sign.

Proof. This follows from Lemma 9.A.2 and the real singular value decomposition. □

Indistinguishable Case

We obtain an analogous result for the case of two indistinguishable subsystems.

Lemma 9.A.4. *Let $H_0 \in \mathfrak{iu}^s(d^2)$ be an arbitrary coupling Hamiltonian in the indistinguishable case and let A_i for $i = 1, \ldots, d^2$ be an orthonormal basis of $\mathfrak{iu}(d)$. Then there is a unique coefficient matrix $C \in \mathbb{R}^{d,d}$ such that*

$$H_0 = \sum_{i,j=1}^{d^2} C_{ij} A_i \otimes A_j.$$

It holds that C is symmetric. If $A_i = \sum_{k=1}^{d_1^2} S_{ik} \tilde{A}_k$ is another choice of basis, with $S \in \mathbb{R}^{d_1,d_1}$ invertible, then

$$H_0 = \sum_{k,l=1}^{d_1^2,d_2^2} \tilde{C}_{kl} \tilde{A}_k \otimes \tilde{A}_l \quad \text{where } \tilde{C} = S^\top C\, S.$$

As a consequence the rank r of the coefficient matrix is well-defined, i.e. it depends only on H_0. Hence the basis A_i can always be chosen such that

$$H_0 = \sum_{i=1}^{r} A_i \otimes A_i,$$

and we will say that this representation is "in diagonal form".

Proof. This follows from the real symmetric eigenvalue decomposition. □

Corollary 9.A.5. *Moreover, there exists an orthonormal basis A_i (with respect to any given inner product) such that*

$$H_0 = \sum_{i=1}^{d^2-1} \omega_i A_i \otimes A_i,$$

and we will say that this representation is "in diagonal form". The ω_i are the eigenvalues of the coefficient matrix with respect to any orthonormal basis, and hence they are uniquely defined up to order.

Remark 9.A.6. *The results of Lemma 9.A.2 and Lemma 9.A.4 are in many ways analogous to the normal form results for the Lindblad equation. In particular, our coefficient matrix C plays the role of the Kossakowski matrix, and by diagonalizing it we obtain a particularly nice form. Whereas the eigenvalues of the Kossakowski matrix represent exponential decay rates, the singular values and eigenvalues of our coefficient matrices represent coupling frequencies.*

Part IV

Epilogue

Conclusion

This thesis developed the method of reduced control systems and applied it to quantum control systems of significant interest. In particular we studied open Markovian quantum systems with fast unitary control and closed bipartite quantum systems with local unitary control.

In Ch. 1 we studied matrix diagonalizations through the unifying lens of symmetric Lie algebras. We showed that many results from the perturbation theory of linear operators can be generalized to this setting. In particular we presented results about continuous (Prop. 1.3.1), differentiable (Thm. 1.4.9), analytic (Thm. 1.5.7), and measurable (Thm. 1.6.6 and Thm. 1.6.12) diagonalizations. Moreover we classified the possible diagonalizations arising from this setting (Thm. 1.7.10).

Ch. 2 built upon these results to define the reduced control system (Sec. 2.2) which formed the main topic of this thesis. Most importantly, we proved the equivalence of the full and reduced control systems. This was done in two parts, a projection (Thm. 2.3.8) and an approximate lift (Thm. 2.3.14). A more applicable version of the equivalence results was given as well (Thm. 7.1.10). We showed how the equivalence manifests for concrete control theoretic properties like reachability (Prop. 2.4.3), stabilizability (Prop. 2.4.5), viability (Prop. 2.4.7), invariance (Prop. 2.4.7), accessibility (Prop. 2.4.10) and controllability (Prop. 2.4.11). Moreover we provided a powerful result showing that majorization can be considered a resource in these systems (Thm. 2.5.3).

In Ch. 3 we applied this theory to Markovian quantum systems with fast unitary control, described by ($\mathcal{D}$), to derive the corresponding reduced control system on the eigenvalues of the mixed quantum state (Λ). Of course we immediately obtained the specialized equivalence result (Thm. 3.0.4).

This reduced control system was then exploited in Ch. 4 to derive a number of control theoretic results. We characterized stabilizable points (Prop. 4.2.2) and showed that all points are stabilizable if and only if the Lindblad terms are simultaneously triangularizable (Thm. 4.2.7), a condition which can be checked efficiently (App. 4.C). Moreover we studied the viability of faces of the simplex, and related the notion to lazy subspaces of a Kossakowski–Lindblad generator (Prop. 4.2.4) and we showed that the control system is generically accessible (Prop. 4.2.12). Then we considered notions related to reachability. In particular we characterized asymptotically coolable systems using the existence of a common right eigenvector of the Lindblad terms which is not a common left eigenvector (Thm. 4.3.7). We studied the reachability of faces of the simplex, which is related to the cooling of subsystems (Prop. 4.3.19). Finally we also considered approximate controllability of the system. Specialized results for unital systems were also provided (Sec. 4.4). We gave some results about what we termed relaxation algebras, whose algebraic properties characterize many of the previously mentioned control theoretic properties (Table 4.1). Some concrete reachable and stabilizable sets were computed analytically in App. 4.B within a toy model with restricted controls.

Ch. 5 treated individual qubits using the reduced control system. The main insight was to study the set of generators of the reduced control system. We found a general parametrization of this set and, most importantly, its boundary (Prop. 5.5.2 and Coro. 5.5.7) which allowed us to determine the optimal derivatives and thus also the optimal path for heating and cooling through the Bloch ball. From this we

derived the optimal controls in the full system. As an additional consequence of the parametrization, we immediately obtained a parametrization of the set of stabilizable states 5.5.3. We also specialized these results to (the newly defined) integral systems (Sec. 5.6) which include real and coolable systems, and we also recovered many known results about the Bloch equations and unital systems (Sec. 5.A).

In Ch. 6 we focused on the question of optimal cooling. We explicitly solved the case of a single qubit with a rank one generator using the methods of the previous chapter, and found a solution reminiscent of the solution found for the Bloch equations (Sec 6.3). Then we looked at some higher dimensional systems. In particular we considered certain three-level systems and a four-level system consisting of two qubits. In both cases we were able to determine optimal controls for cooling the system, and we did so by studying the shape of the set of achievable derivatives (Sec. 6.4) and then employing a maximum principle based on the Majorization Theorem to find a much smaller subset of optimal derivatives (Sec 6.5).

In Ch. 7 we switched to closed bipartite quantum systems with fast local unitary control, described by ($\mathcal{H}$). Again we defined a corresponding reduced control system (Σ), this time defined on the singular values of the state. Similarly, for bosonic and fermionic systems ($\mathcal{H}^s$) we obtained analogous reduced control systems (Σ^s) and (Σ^a), this time using the Autonne–Takagi factorization and the Hua factorization respectively. The equivalence for all three cases was shown in Thm. 7.1.10.

The reduced control systems were then applied in Ch. 8 to show that the considered systems are always controllable (Thm. 8.2.3) and each state of the reduced control system is stabilizable (Coro. 8.2.4). Moreover, a quantum speed limit for the evolution of the singular values was provided in Thm. 8.3.4.

Finally we considered time-optimal entanglement generation in Ch. 9. In Sec. 9.2 we gave complete solutions of the problem in the case of two coupled qubits, and also in the bosonic case. We also provided a partial solution for the case of two fermionic four-level systems. In all these cases the reduced state space is one-dimensional. Then we studied a system composed of two distinguishable qutrits, in which case the reduced state space is a two dimensional sphere (Sec. 9.3). Here we used the Pontryagin Maximum Principle to derive time-optimal solutions in the reduced system, which were then lifted approximately to the full system.

Outlook

At this point it is hopefully clear that the reduced control system is a valuable and powerful tool in quantum control theory. It should be equally clear that this thesis can only scratch the surface of this topic. Many questions about the systems studied here remain open, and further applications remain to be discovered.

Generalization of the Reduced Control System

One natural question is how far the idea of the reduced control system can be generalized. One obvious direction, especially for quantum control theory, would be to consider infinite dimensional systems. Indeed, even though the mathematical machinery gets significantly more involved, quantum control theory can also be studied in the infinite dimensional setting. Moreover, the perturbation theory for infinite dimensional linear operators is well-established and would provide the necessary tools for such a generalization.

Another direction is to generalize the ideas to include non-linear spaces and group actions. Indeed, although the setting of symmetric Lie algebras might seem highly restrictive, they essentially encompass all linear group actions with the geometric properties required to define a reduced control system. The most general mathematical setting in which the ideas still work out is likely the setting of infinitesimally polar actions on Riemannian manifolds. The reason is that in this setting the quotient space still has the structure of an orbifold, without which it would be difficult to define a reduced control system.

Numerical Methods for the Reduced Control System

Numerical algorithms play an important role in quantum control theory since analytical solutions are difficult to obtain except for a handful of low dimensional or highly symmetric systems, and adding realistic assumptions to such systems often makes it impossible to find analytical solutions to reachability or optimal control problems. While the reduced control system makes it possible find previously unknown analytical solutions for some special systems, again the general case requires numerical methods. A current work in progress is the development of a numerical algorithm for the discretization of differential inclusions such as the ones obtained via the reduced control system. A complicating factor is that in general describing the differential inclusion itself is complicated and has to be done approximately. A significant advantage of this approach is that globally optimal solutions can be computed whereas most optimal control algorithms currently in use only converge to local optima.

Part V

Appendix

In the appendix we briefly outline the works [10] and [11], which are not part of the core theme of this thesis, but are nevertheless related to quantum control theory. We give a brief overview of the results and sketch the connections to quantum control, illustrating both how fundamental the questions posed by control theory, and how far-reaching the tools it provides, are.

Outline Appendix A studies the thermomajorization polytope, a mathematical object appearing in quantum thermodynamics, and it introduces concepts from transportation theory and connects them to physical properties of the system. In Appendix B we consider a family of randomized gradient descent algorithms on Riemannian manifolds with Morse–Bott functions, and we sketch the connection to adaptive variational quantum algorithms.

Acknowledgments The paper [10] is joint work with Frederik vom Ende who spearheaded the project. The paper [11] is joint work with Christian Arenz, Gunther Dirr and Thomas Schulte-Herbrüggen.

APPENDIX A

The Thermomajorization Polytope

This chapter studies the so-called thermomajorization polytope, also called the d-majorization polytope in the mathematical literature, see [ED22]. In the nascent field of quantum thermodynamics it is common to model state transformations by interactions with an environment system, often called a heat bath, typically requiring that the bath start out in thermal equilibrium and that the total energy be preserved. This leads to the concept of thermal operations, which are considered to be free. An important task is then to determine which states can be obtained from a given initial state using only thermal operations. In the quasi-classical case, i.e. where the initial state is diagonal, this leads to the concept of thermomajorization, a generalization of classical majorization, and it turns out that the states which can be obtained from a given state are exactly those which belong to the corresponding thermomajorization polytope.

The relation to quantum control theory stems from the fact that the thermomajorization polytope can be used to give upper bounds to the set of reachable states in certain systems, as has been done for the toy model introduced in Chapter 4.B, see [3, DES19].

In [10] the structure of this polytope is studied, especially the conditions under which the vertices of the polytope become degenerate. Key concepts are obtained by drawing a hitherto largely unnoticed connection to the theory of transportation polytopes, leading to the definition of "well-structured" and "stable" Gibbs states. Moreover the paper provides many intuitive graphical constructions for the abstract concepts discussed, such as the vertices of the thermomajorization polytope or a certain important class of extremal Gibbs-stochastic matrices. Here we give a brief summary of the main results.

Given an n-level quantum system with system Hamiltonian H_S and background bath temperature $T > 0$, the set $\mathsf{TO}(H_S, T)$ of *thermal operations* contains all quantum channels of the form[1]

$$\rho \mapsto \mathrm{tr}_B(e^{\mathrm{i}H_{\mathrm{tot}}}(\rho \otimes \rho_{B,T})e^{-\mathrm{i}H_{\mathrm{tot}}})$$

where $\rho_{B,T} = e^{-H_B/T}/\mathrm{tr}(e^{-H_B/T})$ is the Gibbs state of the bath with bath Hamiltonian H_B and the total Hamiltonian H_{tot} is required to be energy preserving $[H_{\mathrm{tot}}, H_S + H_B] = 0$. One can show that in the "quasi-classical" case, where H_S is non-degenerate and the initial state is diagonal in the eigenbasis of ρ, the set of states reachable from ρ using thermal operations is fully described using thermomajorization, which we briefly introduce now.

Let $d \in \mathbb{R}^n_{++}$ be given (this means that all $d_i > 0$). We call d the Gibbs vector. A matrix $A \in \mathbb{R}^{n,n}$ is called Gibbs-stochastic (with respect to d), if it is stochastic and preserves the Gibbs vector. Explicitly

[1] Here we set the Boltzmann constant $k_B = 1$.

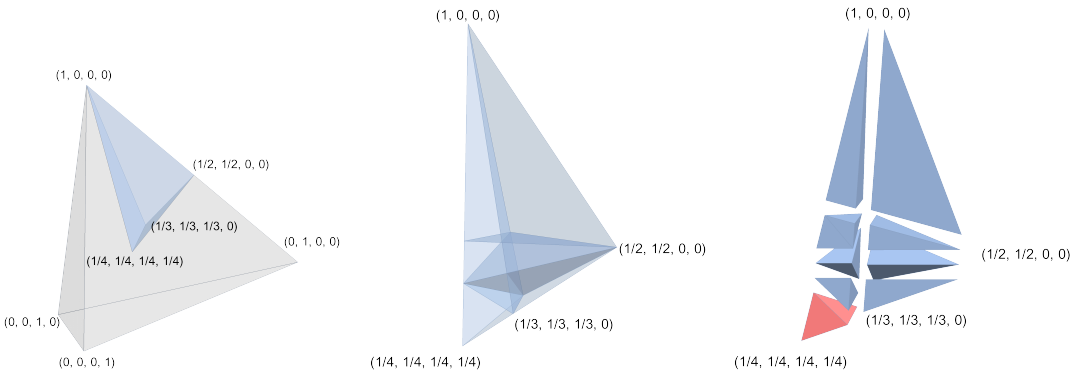

Figure A.1: Illustration of stability and well-structuredness. Consider $d \in \mathbb{R}^4_{++}$ where, w.l.o.g., $\mathbf{e}^\top d = 1$. Then d lies in the relative interior of the standard simplex shown on the left. By reordering its entries in a non-increasing fashion, we can assume that d lies in the ordered Weyl chamber shown in the middle. The unstable points are composed of the walls of the Weyl chamber as well as five planes intersecting the Weyl chamber. These planes cut the Weyl chamber into nine subchambers, and the one which includes the maximally mixed state $\mathbf{e}^\top/4$ (highlighted in red on the right) contains exactly the well-structured Gibbs vectors.

this means that all $A_{ij} \geq 0$, that $\mathbf{e}^\top A = \mathbf{e}^\top$ and that $Ad = d$. Now let $y \in \mathbb{R}^n$ be given. Then the thermomajorization polytope of y with respect to d is defined as

$$M_d(y) = \{Ay : A \text{ is Gibbs-stochastic}\}.$$

If $x \in M_d(y)$ we write $x \preceq_d y$ and say that x is thermomajorized by y. Note that for $d = \mathbf{e}$ we recover ordinary majorization. The connection to thermal operations for diagonal and non-degenerate H_S can be stated as

$$\{\Phi(\mathrm{diag}(y)) : \Phi \in \overline{\mathsf{TO}(H_S, T)}\} = \mathrm{diag}(M_d(y)),$$

where d is now the vector of eigenvalues of the Gibbs state on the system $\rho_{S,T}$. The goal of the paper is to better understand the thermomajorization polytope $M_d(y)$.

A subspace, represented by a subset of indices $P \subseteq \{1, \ldots, n\}$, is said to be in equilibrium if $\frac{y_i}{d_i} = \frac{y_j}{d_j}$ for all $i, j \in P$. If no such subspace exists we say that the system is in total non-equilibrium. If $P = \{1, \ldots, n\}$ then the system is in equilibrium in the usual sense.

Proposition A.0.1. *Let $d \in \mathbb{R}^n_{++}$ and $y \in \mathbb{R}^n$. Then the following statements are equivalent:*

(i) *$M_d(y)$ consists of more than just y.*

(ii) *y is not a multiple of d.*

(iii) *$M_d(y)$ has maximal dimension $n - 1$.*

The result can be rephrased in a more physical language to state the following: If the system is not in equilibrium, it can always be brought into total non-equilibrium using a Gibbs-stochastic matrix.

Consider the map $\mathcal{D} : I \mapsto \sum_{i \in I} d_i$ which takes a set I of indices and returns the sum of the corresponding entries in d. If this map is injective, we say that d is stable.

Theorem A.0.2 (Cyclic State Transfers). *Given $d \in \mathbb{R}^n_{++}$, the following are equivalent:*

(i) *d is stable.*

(ii) *For all $x, y \in \mathbb{R}^n$, if $x \preceq_d y \preceq_d x$, then $x = y$.*

Put differently, cyclic state transfers are possible if and only if the Gibbs state is not stable.

Using the same map $\mathcal{D}$ we can define another notion. If for all I, J with $|I| < |J|$ it holds that $\mathcal{D}(I) < \mathcal{D}(J)$, then we say that d is well-structured. See Figure A.1 for an illustration of stable and well-structured Gibbs vectors. Note that the well structured states contain an open neighborhood of the maximally mixed state. This means that for a given system Hamiltonian there is a critical temperature above which the Gibbs vector becomes well-structured.

An interesting result about the thermomajorization polytope is that there exists a map $E_{d,y} : S_n \to M_d(y)$ which sends permutations to the extreme points of $M_d(y)$. Most importantly, the map reaches all extreme points. In [10] we describe an intuitive visual construction for computing $E_{d,y}$. The following result tells us about injectivity of this map.

Theorem A.0.3 (Degenerate Extreme Points). *Let $d \in \mathbb{R}^n_{++}$ and $y \in \mathbb{R}^n$. If $E_{d,y}$ is not injective, i.e. the number of extreme points of $M_d(y)$ is strictly less than $n!$, then at least one of the following holds:*

(i) *y has a subspace in equilibrium with respect to d.*

(ii) *d is not well structured, i.e. the temperature is below the critical value.*

APPENDIX B

Randomized Gradient Descent

The topic of this chapter is the analysis of a gradient descent algorithm on Riemannian manifolds in which the gradient is only known up to projection onto randomly chosen directions. Such an algorithm has recently been proposed as an adaptive variational quantum algorithm (VQA) for the preparation of the ground state of a given Hamiltonian [MEA23], which has applications for the solution of certain optimization problems on a quantum computer.

To connect this idea to the main topic of the thesis, note that it can be interpreted as a quantum control task, where the goal is to steer the system into the ground state from any given initial state. Due to the use of measured gradients this is a feedback control strategy.

In [11] we describe a gradient descent algorithm where the gradient of a high-dimensional problem is projected, in each step, onto a randomly chosen direction of the tangent space, considering both continuous and discrete probability distributions. The main result is that this randomly projected gradient descent algorithm, applied to a Morse–Bott function with compact sublevel sets on a Riemannian manifold, almost surely escapes saddle points and, as a consequence, it almost surely converges to a local minimum. Moreover, we study the case of a two-dimensional saddle point using analytical methods and numerical simulation. We obtain good approximations for the time necessary to pass the saddle point. Finally, we also show how the previous results can be applied in quantum optimization to the problem of ground state preparation. In particular, the convergence guarantees still hold when the randomization is achieved using an approximate unitary 2-design. Here we just provide a brief exposition of the main result.

The setting is a Riemannian manifold M with a Morse–Bott function $f : M \to \mathbb{R}$. The gradient algorithm is defined by the step rule

$$x_{k+1} = \exp_{x_k}(-\eta_k g(x_k, u_k)), \quad \text{where } g(x,u) = \langle u, \operatorname{grad} f(x)\rangle u,$$

and where $\eta > 0$ is the step size and $\exp$ denotes the Riemannian exponential function. Moreover, the u_k are randomly chosen unit vectors, where the probability distribution may be uniform (the so-called Haar measure) or it may be discrete. If it is discrete, we need to assume that the directions span the entire tangent space, even if all directions collinear with one given direction are removed. Furthermore, we will assume that the function f is ℓ-smooth, that is, M has an atlas of normal charts in which the gradient of f is ℓ-Lipschitz.

The main result of the paper is that the proposed algorithm converges almost surely to a local minimum, and in the specific application to ground state preparation, it even converges to the global minimum.

Theorem B.0.1 (Almost Sure Convergence). *Let M be a Riemannian manifold and let $f : M \to \mathbb{R}$ be an ℓ-smooth Morse–Bott function with compact sublevel sets. Further assume that the initial state x_0 is not a critical point of f. Then, for stepsize $\eta \leq 1/\ell$, the randomly projected gradient descent algorithm converges almost surely to a local minimum.*

The proof proceeds by first showing that the algorithm does not get stuck at (strict) saddle points, which uses on the Morse–Bott property, but might also work more generally. This implies that the algorithm converges to the set of local minima. Again using the Morse–Bott property we show that the algorithm indeed converges to a single point, which must be a local minimum.

The most important aspect of this algorithm is its behavior in the vicinity of saddle points. To get a more quantitative understanding, one may consider the simplest case of the Morse function $f(x, y) = x^2 - y^2$ on $\mathbb{R}^2$ with the Euclidean metric. For small enough stepsize $0 < \eta \ll 1$, and using [KP92, Sec. 6.2] we can approximate this process by the stochastic differential equation (SDE)

$$d\Phi_t = \sin(2\Phi_t)dt + \frac{\sqrt{\eta}}{2}dW_t,$$

where Φ_t represents the angular coordinate and W_t is a Brownian motion. In order to understand how long it takes the algorithm to pass the saddle point, we are interested in computing the hitting time

$$\tau = \inf\{t > 0 : \Phi_t = \pm\tfrac{\pi}{4}\}.$$

Close to the origin, the SDE can be linearized as $d\Phi_t = 2\Phi_t dt + \frac{\sqrt{\eta}}{2}dW_t$, which is a mean repelling Ornstein–Uhlenbeck process. Away from the origin, the deterministic part dominates and we can solve the (deterministic) ODE $\dot{\Phi}_t = \sin(2\Phi_t)$. First we approximate the hitting time distribution of the repelling Ornstein–Uhlenbeck process.

Lemma B.0.2. *Let X_t be the solution of the SDE $dX_t = \kappa X_t dt + \sigma dW_t$ with $\kappa, \sigma > 0$ and $X_0 = 0$. Setting $\tilde{\sigma}(t) = \sigma\sqrt{\frac{e^{2\kappa t}-1}{2\kappa}}$, it holds that $X_t \sim \mathcal{N}(0, \tilde{\sigma}(t)^2)$. If we denote by τ_c the hitting time of $\pm c$ (where $c > 0$), we find the lower bound $\Pr[\tau_c \leq t] \geq \Pr[|X_t| \geq c] = 1 + \operatorname{erf}\left(\frac{-c}{\tilde{\sigma}(t)\sqrt{2}}\right)$ where erf denotes the error function.*

Proof. The SDE is linear and hence has the well-known solution

$$X_t = \sigma \int_0^t e^{\kappa(s-t)} dW_s \cong \mathcal{N}(0, \tilde{\sigma}(t)^2),$$

where $\tilde{\sigma}(t) = \sigma\sqrt{\frac{e^{2\kappa t}-1}{2\kappa}}$, see [KP92, Sec. 4.2 & 4.4]. It is clear that if $|X_t| \geq c$ then $\tau_c \leq t$, which implies the last statement. □

The approximation is very accurate with error smaller than Δt if $\kappa \Delta t \geq 1$ and $c \gg \sigma/\sqrt{\kappa}$, as illustrated in Figure B.1.

As in the proof of Lemma B.0.2, the solution of the SDE with initial condition $X_0 = c$ is given by

$$X_t = e^{\kappa t} c + \sigma \int_0^t e^{\kappa(s-t)} dW_s,$$

and hence

$$e^{\kappa t} c - n\tilde{\sigma}(t) \geq c \implies c \geq n\sigma\sqrt{\frac{e^{2\kappa t}-1}{2\kappa}}\frac{1}{e^{\kappa t}-1} \geq \frac{n\sigma}{\sqrt{2\kappa}} \max\left(1, \sqrt{\tfrac{2}{\kappa t}}\right).$$

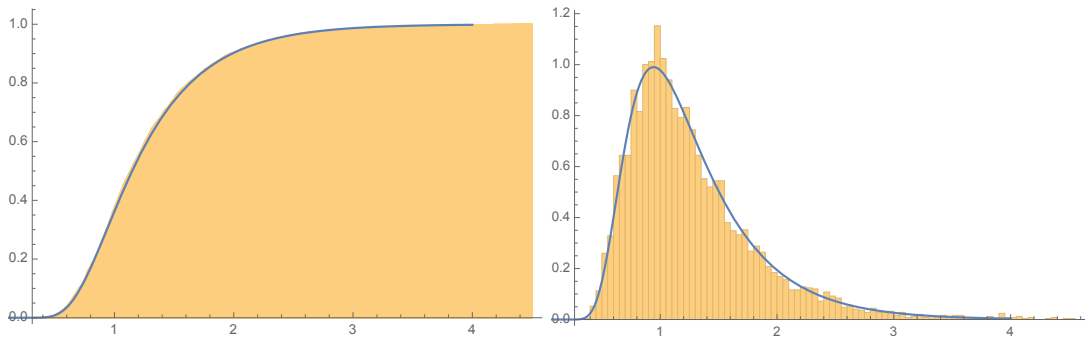

Figure B.1: We used the constants $\kappa = 2$, $\sigma = 3$ and $c = 10$. Shown are the histograms of the c.d.f. and p.d.f of τ_c computed using timesteps of size 0.001 (orange) and the lower bound for the c.d.f. and the resulting p.d.f. obtained in Lemma B.0.2 (blue).

Thus for the error in τ_c to be small compared to Δt, we need $c \gg \frac{\sigma}{\sqrt{\kappa}}$ if $\kappa\Delta t$ is big, and $c \gg \frac{\sigma}{\kappa\sqrt{\Delta t}}$ if $\kappa\Delta t$ is small.

Now if we choose c small enough that the linearization of $\sin$ is accurate but large enough that the approximation of Lemma B.0.2 is good (which is always possible if η is small enough), then we can approximate τ as follows. We fix some value c and note that the ODE $\dot{\phi}_t = \sin(2\phi_t)$ with initial condition $\phi_0 = c$ has the solution $\phi_t = \arctan(e^{2t}\tan(c))$, and hence the hitting time of $\pi/4$ is $\tilde{\tau}_c = -\frac{1}{2}\ln(\tan(c))$. Hence we have approximately $\tau \simeq \tau_c + \tilde{\tau}_c$.

List of Figures

List of Theorems

List of Tables

Index

Abbreviations

AI	Artificial intelligence. xiv
BB84	A protocol for QKD invented by Bennet and Brassard. xv
BNMRZ	Bavarian NMR Center. xi
CPTP	Completely positive and trace-preserving. xxiii
CRAB	Chopped random basis. xxvi
DH	A key-exchange crypto system invented by Diffie and Hellman. xvii
DSA	Digital signature algorithm. xvii
ECC	Elliptic curve cryptography. xvii
EVD	Eigenvalue decomposition. 5, 28, 81
GNSS	Global navigation satellite system. xiv
GRAPE	Gadient ascent pulse engineering. xxvi
HJB	Equation named after Hamilton, Jacobi and Bellman. xxv, xxvi, 157, 168
HSP	Hidden subgroup problem, usually refers to the Abelian version. xvii
MPS	Matrix product state. xv
MRI	Magnetic resonance imaging. xiv, xxi
NISQ	Noisy intermediate-scale quantum. xxvi
NMR	Nuclear magnetic resonance. xiv, xxi, xxiii
NV	Nitrogen-vacancy. xiv, xxi
ODE	Ordinary differential equation. 222
PDE	Partial differential equation. xxv
PEPS	Projected entangled pair state. xv
PET	Positron emission tomography. xiv
PMP	Pontryagin maximum principle. xxv, xxvi, 73, 149, 157, 193, 199, 200
QAOA	The quantum approximate optimization algorithm. xviii
QCCD	Quantum charge-coupled device. xix
QED	Quantum electrodynamics. xxi

QKD Quantum key distribution. xv
QPE Quantum phase estimation. xvii
QSL Quantum speed limit. 187
QUBO Quadratic unconstrained binary optimization. xvii

RSA A public-key crypto system invented by Rivest, Shamir, and Adleman. xv, xvii

SDE Stochastic differential equation. 222
SQUID Superconducting quantum interference devices. xiv, xx
SVD Singular value decomposition. xxiv, 3, 28, 169

VQA Variational quantum algorithm. xviii, 221
VQE Variational quantum eigensolver. xvii

Symbols

$[\cdot,\cdot]$ Commutator of two matrices, more generally Lie bracket. **4**, **5**, 55, 126

$\{\cdot,\cdot\}$ Anti-commutator of two matrices. 85

$\preceq$ Majorization preorder (both in the usual and generalized sense). **71**, 85, 103, 106, **166**, 198

$\trianglelefteq$ Infinitesimal majorization. **166**

$(\cdot)^{\downarrow}$ Vector in non-increasing order (more generally, representative in the standard Weyl chamber). **7**, 61, 72, 75, **82**, 84

$(\cdot)^{\ddagger}$ Vector of absolute values in non-increasing order. **178**

$(\cdot)^{\top}$ Transpose of a matrix. 6, 24

$(\cdot)^{*}$ Hermitian conjugate of a matrix. xxii, 5, 24

$(\cdot)^{+}$ Moore–Penrose pseudo inverse. 135

$(\cdot)^{-}$ Complement of a projection. **16**

$(\cdot)'$ Derivative of a univariate function, also commutant of a set of matrices and dual vector space. 8, 88, 173

$(\cdot)^{\flat}$ Projects a set from $\mathfrak{p}$ to $\mathfrak{a}$. **69**, 70

$(\cdot)^{\sharp}$ Lifts a set from $\mathfrak{a}$ to $\mathfrak{p}$. **69**, 70

$(\cdot)^{\star}$ Pullback, also used for optimal solutions and dual cones. 56, 73, 77

$(\cdot)_{\star}$ Pushforward (sometimes derivative). 65, 66, 181

$(\cdot)_{c}$ Complexification of a linear operator. 15

$|\cdot\rangle$ Ket notation for quantum states. xxii

$[\cdot]$ Equivalence class of approximate reachability relation. **120**

$\sim$ Equivalence in approximate reachability relation. **120**

$\circ$ Hadamard product of matrices, also function composition. 83, 176

$\times$ Cross-product, direct product of groups, or Cartesian product. 181, 200

$\oplus$ Direct sum of vector spaces or Lie algebras. 4

$\otimes$ Tensor product or Kronecker product. xxiv, 172

$\wr$ Wreath product. 5, 188

$\langle\cdot\rangle_{\mathsf{Lie}}$ Lie algebra generated by a set of elements. 55, 70

$\langle\cdot\rangle_{\mathsf{LS}}$ Lie saturate generated by a set of elements. 86

$\langle\cdot\rangle_{\mathsf{alg}}$ Matrix algebra generated by a set of elements. 103

$\hookrightarrow$ Injection. 7

$\twoheadrightarrow$ Surjection. 7

$\rightsquigarrow$ Approximate reachability relation. **102**, 120

$\|\cdot\|_1$ 1-norm of a vector. 199

$\|\cdot\|_2$ 2-norm of a vector or Hilbert–Schmidt norm for matrices. 165, 190, 199

$\|\cdot\|_\infty$ Supremum norm for maps, operator norm for matrices. 64, 96, 162, 178, 189

$\|\cdot\|_{AC}$	Norm of absolutely continuous functions. **75**
$\bigwedge^2(\cdot)$	Alternating tensor algebra. 172
$\mathbb{1}$	Identity matrix. 82
$\mathbf{1}_{(\cdot)}$	Indicator function. 64
$\mathfrak{a}$	Maximal Abelian subspace. 4, 33
$\mathrm{ad}_k(\cdot)$	Adjoint representation of a Lie algebra element k, for matrices this is the commutator. **4**, 55
$\mathrm{Ad}_K(\cdot)$	Adjoint representation of a Lie group element K, for matrices this is conjugation. 4, **5**, 83
$\mathrm{ad}_x^{-1}(\cdot)$	Restricted pseudo-inverse of the adjoint representation. **14**
$\mathfrak{asym}(n,\mathbb{K})$	Skew-symmetric matrices of size $n \times n$ with values in $\mathbb{K}$. **24**, 173, 185
$B(\cdot,\cdot)$	Killing form. 25, **30**
C^k	Class of k times continuously differentiable functions, $k = 1,\dots,\infty$. 4, 6, 14, 62
C^0	Class of continuous functions. 14
C^ω	Class of real analytic functions. 62
$\mathrm{conv}(\cdot)$	Convex hull. 71
D	Derivative in the sense of manifolds or orbifolds. 9
D^-	Contingent derivative of a set-valued map. 72, **76**
$D\pi_{\mathfrak{a}}$	Differential of the quotient map $\pi_{\mathfrak{a}}$. **9**
$\mathsf{derv}(\cdot)$	Set of achievable derivatives. xxiv, **56**, **83**, 116, 131, 177
$\widetilde{\mathsf{derv}}(\cdot)$	Set of all possible projected derivatives. **60**
$\mathrm{diag}(\cdot)$	Maps a vector to its corresponding diagonal matrix. **82**, **174**
$\mathfrak{diag}(m,n,\mathbb{K})$	Diagonal matrices of size $m \times n$ with values in $\mathbb{K}$, short: $\mathfrak{d}(m,n,\mathbb{K})$. **24**
$\mathfrak{diag}(n,\mathbb{K})$	Diagonal matrices of size $n \times n$ with values in $\mathbb{K}$, short: $\mathfrak{d}(n,\mathbb{K})$. **24**, 179
$\mathfrak{diag}_0(n,\mathbb{K})$	Traceless diagonal matrices of size $n \times n$ with values in $\mathbb{K}$, short: $\mathfrak{d}_0(n,\mathbb{K})$. 5, **24**
$\mathfrak{diag}_1(n,\mathbb{K})$	Unit trace diagonal matrices of size $n \times n$ with values in $\mathbb{K}$, short: $\mathfrak{d}_1(n,\mathbb{K})$. 243
δ_{ij}	Kronecker symbol. 83
Δ	Non-zero roots. **36**
Δ_0	Roots including zero. **36**
Δ^{n-1}	The standard simplex of dimension $n-1$ embedded in $\mathbb{R}^n$. xxvii, **82**, 115
$\Delta^{n-1}_{\downarrow}$	Ordered Weyl chamber in the standard simplex of dimension $n-1$ consisting of vectors in non-increasing order. **82**, 91
∂	Topological boundary. 50, 75
$\mathbf{e}$	The vector $(1,\dots,1)^\top$, usually as barycenter of the standard simplex. **82**
$E(\cdot)$	Edge of a Lie wedge. 57
e_i	i-th standard vector, usually as vertex of the standard simplex. **82**
f^d_{in}	Inflow of dimension d. **98**

f^d_{out} Outflow of dimension d. **98**

$\mathfrak{g}$ Lie algebra, usually symmetric. 4, 27
$\mathfrak{g}^\alpha_{\mathbb{C}}$ Root space of α in $\mathfrak{g}_{\mathbb{C}}$. **36**
$\mathfrak{g}_{\mathbb{C}}$ Complexification of Lie algebra $\mathfrak{g}$. 23, 36
Γ_f Graph of the function f. 76
Γ_V Lindblad generator corresponding to Lindblad term V. **81**

$\hbar$ Reduced Planck constant. xxii, 172
H_0 Drift Hamiltonian. xxvi, **81**, 172
$\mathfrak{ham}(n, \mathbb{K})$ Hamiltonian matrices of size $2n \times 2n$ over $\mathbb{K}$. **27**
$\mathfrak{ham}^*(n, \mathbb{K})$ $*$-Hamiltonian matrices of size $2n \times 2n$ over $\mathbb{K}$. **27**
H_c Compensating Hamiltonian. 89, 196
$\mathfrak{herm}(n, \mathbb{K})$ Hermitian matrices of size $n \times n$ with values in $\mathbb{K}$. **24**
$\mathfrak{herm}_0(n, \mathbb{K})$ Traceless Hermitian matrices of size $n \times n$ with values in $\mathbb{K}$. **24**
$\mathfrak{herm}_1(n, \mathbb{K})$ Unit trace Hermitian matrices of size $n \times n$ with values in $\mathbb{K}$. 243
H_j Control Hamiltonian. xxvi, 82

$\mathrm{i}, \mathrm{j}, \mathrm{k}$ Imaginary quaternion basis elements. 24
I_n Identity matrix of size $n \times n$. **24**
$I_{n,m}$ Diagonal matrix with n times 1 and m times -1. **24**
$\mathrm{int}(\cdot)$ Topological interior of a set. 71
$\mathrm{Int}_{\mathfrak{k}}(\mathfrak{g})$ Inner inner automomorphisms of $\mathfrak{g}$ generated by $\mathfrak{k}$. 18, 25, 27, 30, **34**, 55
$\mathrm{Int}(\mathfrak{g})$ Group of inner automomorphisms of $\mathfrak{g}$. 30, **34**, 55
ι Inclusion. 7

J_n Standard symplectic form of size $n \times n$. **24**, 176
$J(U)$ Matrix from which $-L_U$ is defined. **83**

$\mathfrak{k}$ Compact Lie algebra, usually $+1$ eigenspace of a symmetric Lie algebra. 4, **27**
k_B Boltzmann constant. 118
$\mathbb{K}$ Field or skew field, usually $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. 24, 125
$\mathbf{K}_x$ Stabilizer (also called isotropy subgroup) of x in $\mathbf{K}$, similar for $\mathbf{G}_x$ and $\mathbf{W}_x$. **9**, 60
$\mathbf{K}x$ Orbit of x with respect to $\mathbf{K}$, similar for $\mathbf{G}x$ and $\mathbf{W}x$. 58

$-L$ Kossakowski–Lindblad generator. xxiii, **82**
$\mathfrak{L}$ Set of induced vector fields for Kossakowski–Lindblad generator $-L$. **83**
$\mathcal{L}(\cdot)$ Linear maps on a vector space. 108
$\mathrm{Lat}(\cdot)$ Invariant subspace lattice of a set of matrices. **103**
$-L_U$ Induced vector field for Kossakowski–Lindblad generator $-L$. **83**
Λ Diagonal mixed quantum states. 243

$M_{\mathbf{e}}(d)$ Majorization polytope of d. 116
$M_d(y)$ d-Majorization polytope of y. **218**
μ Optimal derivative function. 73, **132**, 160

$N_{\mathbf{K}}(\mathfrak{a})$ Normalizer of $\mathfrak{a}$ in $\mathbf{K}$. **36**, 75

NP The complexity class of decision problems for which solutions have proofs which can be verified deterministically by a Turing machine in polynomial time. xv, xvii

$N_x A$ Normal cone. **77**

$\mathcal{O}(\cdot)$ Big O notation, gives upper bound on the growth rate of a function. xvi, 125

$\tilde{\mathcal{O}}(\cdot)$ Big O notation ignoring poly-logarithmic factors. xvi

P The complexity class of decision problems which can be solved deterministically by a Turing machine in polynomial time. xv

$\mathcal{P}(\cdot)$ Power set. 71, 131

p Parity of an integral qubit system. **140**

$\mathfrak{p}$ -1 eigenspace of a symmetric Lie algebra. 4, **29**

$P(a)$ Weyl polytope of a. **71**

$\mathfrak{p}_x$ Commutant (also centralizer) of x in $\mathfrak{p}$, similar for $\mathfrak{k}_x$ and $\mathfrak{g}_x$. 8, **33**, **58**

$\mathfrak{pos}_1(n)$ Positive semi-definite matrices of size $n \times n$ and of unit trace. xxiii, **81**

pr_i Orthogonal projection onto the i-th component of a Cartesian product. 10, 20

π The quotient map $\pi : \mathfrak{p} \to \mathfrak{p}/\mathbf{K}$. **7**

$\pi_{\mathfrak{a}}$ The quotient map $\pi_{\mathfrak{a}} : \mathfrak{a} \to \mathfrak{a}/\mathbf{W}$. **7**

$\pi_{\mathfrak{a},x}$ The quotient map $\pi_{\mathfrak{a},x} : \mathfrak{a} \to \mathfrak{a}/\mathbf{W}_x$. **9**

π_x The quotient map $\pi_x : \mathfrak{p} \to \mathfrak{p}/\mathbf{K}_x$. **9**

$\Pi_{\mathfrak{a}}$ Orthogonal projection on $\mathfrak{p}$ onto $\mathfrak{a}$. **56**

$\Pi_{\mathrm{diag}}(\cdot)$ Projects matrix to the vector of diagonal elements. **83**

$\Pi_{\Sigma}(\cdot)$ Projects bipartite state to the vector of diagonal elements. 243

$\Pi_{\Xi}(\cdot)$ Projects bipartite state to the vector of quasi-diagonal elements. 243

Π_x Orthogonal projection on $\mathfrak{p}$ onto the commutant of x. 8, **34**, **58**

$\Pi_x^{\perp}$ Orthogonal projection on $\mathfrak{p}$ onto the orthocomplement of the commutant of x. 8, **35**

$\mathrm{qdiag}(\cdot)$ Maps a vector to its corresponding quasi-diagonal matrix. **174**

$\mathfrak{qdiag}(n, \mathbb{K})$ Quasi-diagonal matrices of size $n \times n$ with values in $\mathbb{K}$, short $\mathfrak{qd}(n, \mathbb{K})$. **24**, 179

$\mathrm{qsing}^{\ddagger}$ Singular values chosen non-negative and non-increasing, rescaled and only one of each pair. **177**

R_d Flow ratio of dimension d. **98**

$\mathsf{reach}_{(\cdot)}(x_0, T)$ Reachable set from x_0 in time T. **67**, 94, 178

$\mathsf{reach}_{(\cdot)}(x_0)$ (All-time) reachable set from x_0. **67**, 94, 178

$\mathsf{reach}_{(\cdot)}(x_0, I)$ Reachable set from x_0 on time interval $I = [0, T]$. **67**, 178

$\mathrm{relint}(\cdot)$ Relative interior of a set. 72, 90, 94

$\mathbb{R}_0^n$ Subspace of vectors whose elements add up to 0. **83**, 93, 94

$s(\cdot)$ Involutive automorphism of a symmetric Lie algebra. 4

S_n Symmetric group on n elements. 5, 115

S_x Slice at x. 10, **42**
$S^{d-1}_{\ddagger}$ Ordered non-negative Weyl chamber in the unit sphere in $\mathbb{R}^d$. **175**
$\mathrm{sing}^{\ddagger}$ Singular values chosen non-negative and non-increasing. **177**
$\mathfrak{sl}(n,\mathbb{K})$ Special linear Lie algebra of size $n \times n$ with values in $\mathbb{K}$. 23
$\mathsf{sols}_{(\cdot)}(a_0,[0,T])$ Set of solutions starting at a_0 on time interval $[0,T]$. **75**
$\mathfrak{so}(n)$ Special orthogonal Lie algebra of size $n \times n$ (with values in $\mathbb{R}$). 23
$\mathrm{SO}(n)$ Special orthogonal Lie group of size $n \times n$ (with values in $\mathbb{R}$). 28
$\mathfrak{so}(n,\mathbb{C})$ Special orthogonal Lie algebra of size $n \times n$ with values in $\mathbb{C}$. 23
$\mathfrak{so}^*(2n)$ Symmetric Lie algebra dual to $\mathfrak{so}(2n)$. **24**, 185
$\mathfrak{so}(p,q)$ Indefinite special orthogonal Lie algebra of signature p,q. 24
$\mathfrak{sp}(n)$ Unitary symplectic Lie algebra of size $n \times n$. 23
$\mathrm{Sp}(n)$ Unitary symplectic Lie group of size $n \times n$. 28
$\mathfrak{sp}(n,\mathbb{K})$ Symplectic Lie algebra of size $n \times n$ with values in $\mathbb{K} = \mathbb{R},\mathbb{C}$. 183
$\mathfrak{sp}(p,q)$ Indefinite symplectic Lie algebra of signature p,q. 24
$\mathrm{span}(\cdot)$ Linear span. 127
$\mathrm{spec}^{\downarrow}(\cdot)$ Spectrum of a matrix in non-increasing order. **82**
$\mathsf{stab}_{(\cdot)}$ Set of stabilizable states. **68**, 89, 116
$\mathfrak{stoch}(n)$ Lie wedge of generators of stochastic matrices of size $n \times n$. **83**
$\mathrm{Stoch}(n)$ Semigroup of stochastic matrices of size $n \times n$. **83**
$\mathfrak{su}(n)$ Special unitary Lie algebra of size $n \times n$. 23
$\mathrm{SU}(n)$ Special unitary Lie group of size $n \times n$. 28
$\mathfrak{su}^*(2n)$ Symmetric Lie algebra dual to $\mathfrak{su}(2n)$. **24**
$\mathfrak{su}(p,q)$ Indefinite special unitary Lie algebra of signature p,q. 24
$\mathrm{supp}(\cdot)$ Support, usually of a Hermitian matrix. 112
$\mathrm{Sym}^2(\cdot)$ Symmetric tensor algebra. 172, 183
$\mathfrak{sym}(n,\mathbb{K})$ Symmetric matrices of size $n \times n$ with values in $\mathbb{K}$. **24**, 173, 183
$\mathfrak{sym}_0(n,\mathbb{K})$ Traceless symmetric matrices of size $n \times n$ with values in $\mathbb{K}$. **24**
$\sigma_x, \sigma_y, \sigma_z$ Pauli matrices, denoted P_x, P_y, P_z in Part III. **130**, 158
Σ Diagonal bipartite quantum states. 243

$T_x A$ Tangent space (of a manifold or orbifold) or tangent cone. **76**
$T_x^- A$ Bouligand contingent cone. **76**

u_j, v_j Control functions. 55, 82
$\mathfrak{u}_{\mathrm{loc}}(d_1,d_2)$ Local unitary Lie algebra. 172
$\mathrm{U}_{\mathrm{loc}}(d_1,d_2)$ Local unitary Lie group. xxiv, 172
$\mathfrak{u}^s_{\mathrm{loc}}(d)$ Symmetric local unitary Lie algebra. **172**, 196
$\mathrm{U}^s_{\mathrm{loc}}(d)$ Symmetric local unitary Lie group. **172**, 196

$\mathcal{V}$ Relaxation algebra. 88, 103
$\mathcal{V}^+$ Extended relaxation algebra. 88
$\mathcal{V}_*$ Relaxation $*$-algebra. 114
$\mathcal{V}^+_*$ Extended relaxation $*$-algebra. 114
$V^{\mathbb{C}}$ Complexification of real vector space V. 15
V_k Lindblad term. **81**

$\mathbf{W}$ Weyl group. 4, **36**, 188

$\mathfrak{w}$	Weyl chamber, also Lie wedge. 4, **39**
$\mathfrak{w}_{\mathsf{KL}}$	Kossakowski–Lindblad Lie wedge. xxiii, **82**
Ξ	Quasi-diagonal bipartite quantum states. 243
$Z_{\mathbf{K}}(\mathfrak{a})$	Centralizer of $\mathfrak{a}$ in $\mathbf{K}$. 14, **36**, 75

Table 1: Notation describing the various control systems presented in this thesis.

	General	Distinguishable	Bosonic	Fermionic	Lindbladian
Full System	(F)	$(\mathcal{H})$	$(\mathcal{H}^s)$	$(\mathcal{H}^s)$	$(\mathcal{D})$
Ambient space	$\mathfrak{p}$	$\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$	$\mathrm{Sym}^2(\mathbb{C}^d)$	$\bigwedge^2(\mathbb{C}^d)$	$\mathfrak{herm}_1(n)$
State space	S	$S(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$	$S(\mathrm{Sym}^2(\mathbb{C}^d))$	$S(\bigwedge^2(\mathbb{C}^d))$	$\mathfrak{pos}_1(n)$
State	p	$\lvert\psi\rangle$	$\lvert\psi\rangle$	$\lvert\psi\rangle$	ρ
Fast Controls	$\mathbf{K}$	$\mathrm{U}_{\mathrm{loc}}(d_1, d_2)$	$\mathrm{U}^s_{\mathrm{loc}}(d)$	$\mathrm{U}^s_{\mathrm{loc}}(d)$	$\mathrm{SU}(n)$
Drift	X	$-\mathrm{i}H_0 \in \mathfrak{u}(d_1 d_2)$	$-\mathrm{i}H_0 \in \mathfrak{u}^s(d^2)$	$-\mathrm{i}H_0 \in \mathfrak{u}^s(d^2)$	$-L \in \mathfrak{w}_{\mathsf{KL}}(n)$
ad action	ad_a	ad^d_σ	ad^s_σ	ad^a_ξ	ad_λ
Reduced System	(R)	(Σ)	(Σ^s)	(Σ^a)	(Λ)
Ambient space	$\mathfrak{a}$	$\Sigma \cong \mathfrak{diag}(d_1, d_2)$	$\Sigma \cong \mathfrak{diag}(d)$	$\Xi \cong \mathfrak{qdiag}(d)$	$\Lambda = \mathfrak{diag}_1(n)$
State space	R	$S^{d_{\min}-1}$	S^{d-1}	$S^{\lfloor d/2 \rfloor - 1}$	Δ^{n-1}
State	a	σ	σ	ξ	λ
Weyl group	$\mathbf{W}$	$\mathbb{Z}_2 \wr S_{d_{\min}}$	$\mathbb{Z}_2 \wr S_d$	$\mathbb{Z}_2 \wr S_{\lfloor d/2 \rfloor}$	S_n
Inclusion	ι	diag	diag	qdiag	diag
Projection	$\Pi_{\mathfrak{a}}$	Π_Σ	Π_Σ	Π_Ξ	Π_{diag}
Quotient map	π	$\mathrm{sing}^\ddagger$	$\mathrm{sing}^\ddagger$	$\mathrm{qsing}^\ddagger$	$\mathrm{spec}^\downarrow$
Induced vector fields	$X_K \in \mathfrak{X}$	$-H_{V \otimes W} \in \mathfrak{H}$	$-H^s_{V \otimes V} \in \mathfrak{H}^s$	$-H^a_{V \otimes V} \in \mathfrak{H}^a$	$-L_U \in \mathfrak{L}$
Natural wedge	—	$\mathfrak{so}(d_{\min}, \mathbb{R})$	$\mathfrak{so}(d, \mathbb{R})$	$\mathfrak{so}(\lfloor d/2 \rfloor, \mathbb{R})$	$\mathfrak{stoch}(n)$
Decomposition	—	Complex SVD	Autonne–Takagi fact.	Hua fact.	Hermitian EVD
Lie Algebra	—	AIII	CI	DIII	A

Bibliography

[Aar13] S. Aaronson. *Quantum Computing Since Democritus*. Cambridge: Cambridge University Press, 2013 (cit. on p. xxii).

[AB08] D. Aharonov and M. Ben-Or. "Fault-Tolerant Quantum Computation with Constant Error Rate". In: *SIAM J. Comput.* 38.4 (2008), pp. 1207–1282 (cit. on p. xix).

[AB15] M. Alexandrino and R. Bettiol. *Lie Groups and Geometric Aspects of Isometric Actions*. Cham: Springer, 2015 (cit. on pp. 19, 26, 42, 43).

[AC84] J.-P. Aubin and A. Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*. Berlin Heidelberg: Springer, 1984 (cit. on pp. xxiv, 56, 61, 68, 72, 84, 102, 116, 120, 162, 177).

[Ach+23] R. Acharya et al. "Suppressing Quantum Errors by Scaling a Surface Code Logical Qubit". In: *Nature* 614.7949 (2023), pp. 676–681 (cit. on p. xix).

[Ací+18] A. Acín et al. "The Quantum Technologies Roadmap: A European Community View". In: *New J. Phys.* 20 (2018), p. 080201 (cit. on p. xiv).

[Aha+07] D. Aharonov et al. "Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation". In: *SIAM J. Comput.* 37.1 (2007), pp. 166–194 (cit. on p. xvi).

[AHS96] R. Alur, T. A. Henzinger, and E. Sontag. *Hybrid Systems III: Verification and Control*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1996 (cit. on p. 115).

[AJ14] V. V. Albert and L. Jiang. "Symmetries and Conserved Quantities in Lindblad Master Equations". In: *Phys. Rev. A* 89 (2014), p. 022118 (cit. on pp. 88, 114).

[Al-+23] A. Al-Dweik, R. Ghanam, G. Thompson, and H. Azad. "A Novel Procedure for Constructing Invariant Subspaces of a Set of Matrices". In: *Ann. Mat. Pura Appl.* 202 (2023), pp. 77–93 (cit. on p. 125).

[AL18] T. Albash and D. A. Lidar. "Adiabatic Quantum Computation". In: *Rev. Mod. Phys.* 90.1 (2018), p. 015002 (cit. on p. xvi).

[AL99] D. S. Abrams and S. Lloyd. "Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors". In: *Phys. Rev. Lett.* 83 (1999), pp. 5162–5165 (cit. on p. xvii).

[Alb+15] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar. "Reexamining Classical and Quantum Models for the D-Wave One Processor". In: *Eur. Phys. J. Special Topics* 224.1 (2015), pp. 111–129 (cit. on p. xvi).

[Alb+16] V. V. Albert, B. Bradlyn, M. Fraas, and L. Jiang. "Geometry and Response of Lindbladians". In: *Phys. Rev. X* 6 (2016), p. 041031 (cit. on p. 113).

[Ale+24a] K. Alexander et al. *A Manufacturable Platform for Photonic Quantum Computing*. arXiv: 2404.17570. 2024 (cit. on p. xxi).

[Ale+24b] Y. Alexeev et al. “Quantum-Centric Supercomputing for Materials Science: A Perspective on Challenges and Future Directions”. In: *Future Gener. Comput. Syst.* (2024) (cit. on p. xvii).

[ALJ07] A. Adem, J. Leida, and Y. Juan. *Orbifolds and Stringy Topology*. Cambridge: Cambridge University Press, 2007 (cit. on pp. 44, 45).

[ALP19] Á. Alhambra, M. Lostaglio, and C. Perry. “Heat-Bath Algorithmic Cooling with Optimal Thermalization Strategies”. In: *Quantum* 3 (2019), p. 188 (cit. on p. 157).

[AM11] A. Amann and U. Müller-Herold. *Offene Quantensysteme: Die Primas Lectures*. Berlin Heidelberg: Springer, 2011 (cit. on p. 147).

[Arb16] P. Arbenz. *Lecture Notes on Solving Large Scale Eigenvalue Problems*. 2016 (cit. on p. 126).

[Aru+19] F. Arute et al. “Quantum Supremacy Using a Programmable Superconducting Processor”. In: *Nature* 574.7779 (2019), pp. 505–510 (cit. on pp. xiii, xxi).

[AS04] A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*. Heidelberg: Springer, 2004 (cit. on pp. 53, 200).

[Ass+23] R. Assouly et al. “Quantum Advantage in Microwave Quantum Radar”. In: *Nat. Phys.* 19.10 (2023), pp. 1418–1422 (cit. on p. xiv).

[Aut15] L. Autonne. “Sur les Matrices Hypohermitiennes et sur les Matrices Unitaires”. In: *Ann. Univ. Lyon Nouvelle Ser.* 38.1 (1915) (cit. on pp. 171, 183).

[Bas+21] D. Basilewitsch et al. “Fundamental Bounds on Qubit Reset”. In: *Phys. Rev. Res.* 3 (2021), p. 013110 (cit. on pp. 157, 171).

[Bau85] H. Baumgärtel. *Analytic Perturbation Theory for Matrices and Operators*. Basel: Birkhäuser, 1985 (cit. on p. 3).

[BB14] C. H. Bennett and G. Brassard. “Quantum Cryptography: Public Key Distribution and Coin Tossing”. In: *Theor. Comput. Sci.* 560 (2014). Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84, pp. 7–11 (cit. on pp. xiii, xv, xxiii).

[BCS09] B. Bonnard, M. Chyba, and D. Sugny. “Time-Minimal Control of Dissipative Two-Level Quantum Systems: The Generic Case”. In: *IEEE Trans. Automat. Contr.* 54.11 (2009), pp. 2598–2610 (cit. on p. 129).

[BDZ21] O. Bokanowski, A. Désilles, and H. Zidani. “Relationship Between Maximum Principle and Dynamic Programming in Presence of Intermediate and Final State Constraints”. In: *ESAIM: COCV* 27 (2021), p. 91 (cit. on p. xxv).

[Bel90] S. Bell. “Mapping Problems in Complex Analysis and the $\bar{\partial}$ Problem”. In: *Bull. Amer. Math. Soc. (N.S.)* 22.2 (1990), pp. 233–259 (cit. on p. 120).

[Ben80] P. Benioff. “The Computer as a Physical System: A Microscopic Quantum Mechanical Hamiltonian Model of Computers as Represented by Turing Machines”. In: *J. Stat. Phys.* 22.5 (1980), pp. 563–591 (cit. on pp. xiii, xvi).

[BGK18] S. Bravyi, D. Gosset, and R. König. “Quantum Advantage with Shallow Circuits”. In: *Science* 362.6412 (2018), pp. 308–311 (cit. on p. xviii).

[Bir67] G. Birkhoff. *Lattice Theory*. Providence, Rhode Island: American Mathematical Society, 1967 (cit. on p. 111).

[BL95] D. Boneh and R. J. Lipton. "Quantum Cryptanalysis of Hidden Linear Functions". In: *Advances in Cryptology — CRYPT0' 95*. Ed. by D. Coppersmith. Berlin, Heidelberg: Springer, 1995, pp. 424–437 (cit. on p. xvii).

[Bla+04] A. Blais et al. "Cavity Quantum Electrodynamics for Superconducting Electrical Circuits: An Architecture for Quantum Computation". In: *Phys. Rev. A* 69 (2004), p. 062320 (cit. on p. xxi).

[Bla+21] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff. "Circuit Quantum Electrodynamics". In: *Rev. Mod. Phys.* 93 (2021), p. 025005 (cit. on pp. xix, xxi).

[BN08] B. Baumgartner and H. Narnhofer. "Analysis of Quantum Semigroups with GKS–Lindblad Generators: II. General". In: *J. Phys. A* 41 (2008), p. 395303 (cit. on pp. 88, 90, 97, 99, 106, 108, 112–114).

[Boo+20] K. Boothby, P. Bunyk, J. Raymond, and A. Roy. *Next-Generation Topology of D-Wave Quantum Processors*. arXiv:2003.00133. 2020 (cit. on p. xvi).

[Bor+19] A. Boretti, L. Rosa, J. Blackledge, and S. Castelletto. "Nitrogen-Vacancy Centers in Diamond for Nanoscale Magnetic Resonance Imaging Applications". In: *Beilstein J. Nanotechnol.* 10 (2019), pp. 2128–2151 (cit. on p. xxi).

[Bos+21] W. Bosma, J. J. Cannon, C. Fieker, and A. Steel, eds. *Handbook of Magma Functions*. 2.26-4. 2021 (cit. on p. 126).

[Boy+02] P. Boykin et al. "Algorithmic Cooling and Scalable NMR Quantum Computers". In: *Proc. Natl. Acad. Sci.* 99 (2002), pp. 3388–3393 (cit. on p. 157).

[BP02] H.-P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford: Oxford University Press, 2002 (cit. on pp. xxiii, 110, 146).

[BP04] U. Boscain and B. Piccoli. *Optimal Syntheses for Control Systems on 2-D Manifolds*. Mathématiques et Applications. Berlin, Heidelberg: Springer, 2004 (cit. on p. 68).

[BR01] H. J. Briegel and R. Raussendorf. "Persistent Entanglement in Arrays of Interacting Particles". In: *Phys. Rev. Lett.* 86 (2001), pp. 910–913 (cit. on p. xvi).

[Bra+99] S. L. Braunstein et al. "Separability of Very Noisy Mixed States and Implications for NMR Quantum Computing". In: *Phys. Rev. Lett.* 83 (1999), pp. 1054–1057 (cit. on pp. xiii, xxi).

[Bro11] L. Brouwer. "Über Abbildung von Mannigfaltigkeiten". In: *Math. Ann.* 71 (1911), pp. 97–115 (cit. on p. 93).

[Bro23] M. Brooks. "The Race to Find Quantum Computing's Sweet Spot". In: *Nature* 617.7962 (2023), S1–S3 (cit. on p. xiv).

[Bru+19] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage. "Trapped-Ion Quantum Computing: Progress and Challenges". In: *Appl. Phys. Rev.* 6 (2019), p. 021314 (cit. on p. xix).

[BS09] B. Bonnard and D. Sugny. "Time-Minimal Control of Dissipative Two-Level Quantum Systems: The Integrable Case". In: *SIAM J. Control Optim.* 48.3 (2009), pp. 1289–1308 (cit. on p. 129).

[BSS21] U. Boscain, M. Sigalotti, and D. Sugny. "Introduction to the Pontryagin Maximum Principle for Quantum Optimal Control". In: *PRX Quantum* 2 (2021), p. 030203 (cit. on pp. 199–201).

[Bun+91] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. Nichols. "Numerical Computation of an Analytic Singular Value Decomposition of a Matrix Valued Function". In: *Numer. Math.* 60 (1991), pp. 1–39 (cit. on pp. 3, 14, 18).

[BV97] E. Bernstein and U. Vazirani. "Quantum Complexity Theory". In: *SIAM J. Comput.* 26.5 (1997), pp. 1411–1473 (cit. on p. xiii).

[BW79] W. M. Boothby and E. N. Wilson. "Determination of the Transitivity of Bilinear Systems". In: *SIAM J. Control Optim.* 17 (1979), pp. 212–221 (cit. on p. 127).

[BWS16] V. Bergholm, F. Wilhelm, and T. Schulte-Herbrüggen. *Arbitrary n-Qubit State Transfer Implemented by Coherent Control and Simplest Switchable Local Noise*. arXiv:1605.06473. 2016 (cit. on p. 102).

[BŻ17] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. 2nd edition. Cambridge: Cambridge University Press, 2017 (cit. on pp. xxiv, 130, 166, 173, 175).

[Cam+20] P. Campagne-Ibarcq et al. "Quantum Error Correction of a Qubit Encoded in Grid States of an Oscillator". In: *Nature* 584.7821 (2020), pp. 368–372 (cit. on p. xxi).

[Cao+19] Y. Cao et al. "Quantum Chemistry in the Age of Quantum Computing". In: *Chem. Rev.* 119.19 (2019), pp. 10856–10915 (cit. on p. xvii).

[Car92] M. do Carmo. *Riemannian Geometry*. Basel: Birkhäuser, 1992 (cit. on p. 43).

[CCM11] T. Caneva, T. Calarco, and S. Montangero. "Chopped Random-Basis Quantum Optimization". In: *Phys. Rev. A* (2011) (cit. on p. xxvi).

[Cer+21] M. Cerezo et al. "Variational Quantum Algorithms". In: *Nat. Rev. Phys.* 3.9 (2021), pp. 625–644 (cit. on p. xviii).

[CGK98] I. L. Chuang, N. A. Gershenfeld, and M. Kubinec. "Experimental Implementation of Fast Quantum Searching". In: *Phys. Rev. Lett.* 80 (1998), pp. 3408–3411 (cit. on pp. xiii, xxi).

[CH13] L. Childress and R. Hanson. "Diamond NV Centers for Quantum Computing and Quantum Networks". In: *MRS Bull.* 38.2 (2013), pp. 134–138 (cit. on p. xxi).

[Che+14] Y. Chen et al. "Qubit Architecture with High Coherence and Fast Tunable Coupling". In: *Phys. Rev. Lett.* 113 (2014), p. 220502 (cit. on p. 115).

[Chi+21] A. M. Childs et al. "Theory of Trotter Error with Commutator Scaling". In: *Phys. Rev. X* 11.1 (2021) (cit. on p. xvii).

[Chi93] A. Chi-Chih Yao. "Quantum Circuit Complexity". In: *Proc. of 1993 IEEE 34th Annual Foundations of Computer Science*. 1993, pp. 352–361 (cit. on p. xvi).

[Cho+10] C. W. Chou, D. B. Hume, T. Rosenband, and D. J. Wineland. "Optical Clocks and Relativity". In: *Science* 329.5999 (2010), pp. 1630–1633 (cit. on p. xiv).

[Cho+16] J.-y. Choi et al. "Exploring the Many-Body Localization Transition in Two Dimensions". In: *Science* 352.6293 (2016), pp. 1547–1552 (cit. on p. xvi).

[CK00] F. Colonius and W. Kliemann. *The Dynamics of Control*. Boston: Birkhäuser, 2000 (cit. on pp. xxiv, 120).

[CL66] E. Collingwood and A. Lohwater. *The Theory of Cluster Sets*. Cambridge: Cambridge University Press, 1966 (cit. on p. 47).

[Cla+20] W. Clark, A. Bloch, L. Colombo, and P. Rooney. “Time-Minimum Control of Quantum Purity for 2-Level Lindblad Equations”. In: *Discrete Contin. Dyn. Syst. - S.* 13.4 (2020), pp. 1061–1073 (cit. on p. 129).

[Cla13] F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Graduate Texts in Mathematics. London: Springer, 2013 (cit. on p. 77).

[CNV07] O. Cârjă, M. Necula, and I. Vrabie. *Viability, Invariance and Applications*. Amsterdam: Elsevier, 2007 (cit. on pp. 69, 76).

[Cor84] J. F. Cornwell. *Group Theory in Physics*. Vol. 1. London: Academic Press, 1984 (cit. on p. 194).

[CPW15] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf. “Undecidability of the Spectral Gap”. In: *Nature* 528.7581 (2015), pp. 207–211 (cit. on p. xvi).

[CZ95] I. Cirac and P. Zoller. “Quantum Computations with Cold Trapped Ions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 4091–4094 (cit. on p. xix).

[Dad85] J. Dadok. “Polar Coordinates Induced by Actions of Compact Lie Group”. In: *Trans. Amer. Math. Soc.* 288 (1985), pp. 125–137 (cit. on p. 26).

[Dah10] G. Dahl. “Majorization Permutahedra and $(0, 1)$-Matrices”. In: *Lin. Alg. Appl.* 432 (2010), pp. 3265–3271 (cit. on p. 105).

[Dal+22] A. J. Daley et al. “Practical Quantum Advantage in Quantum Simulation”. In: *Nature* 607.7920 (2022), pp. 667–676 (cit. on p. xv).

[DAl21] D. D’Alessandro. *Introduction to Quantum Control and Dynamics*. 2nd edition. New York: Chapman and Hall/CRC, 2021 (cit. on p. xxv).

[Dav70] E. Davies. “Quantum Stochastic Processes II”. In: *Commun. Math. Phys.* 19 (1970), pp. 83–105 (cit. on p. 88).

[Dav80] E. Davies. “Uniqueness of the Standard Form of the Generator of a Quantum Dynamical Semigroup”. In: *Rep. Math. Phys.* 17 (1980), pp. 249–255 (cit. on p. 110).

[DCM11] P. Doria, T. Calarco, and S. Montangero. “Optimal Control Technique for Many-Body Quantum Dynamics”. In: *Phys. Rev. Lett.* 106 (2011), p. 190501 (cit. on p. xxvi).

[DDH07] J. Demmel, I. Dumitriu, and O. Holtz. “Fast Linear Algebra is Stable”. In: *Numer. Math.* 108 (2007), pp. 59–91 (cit. on p. 125).

[Des+16] J. Deschamps, F. Fagnola, E. Sasso, and V. Umanità. “Structure of Uniformly Continuous Quantum Markov Semigroups”. In: *Rev. Math. Phys.* 28 (2016), p. 1650003 (cit. on p. 108).

[DES19] G. Dirr, F. vom Ende, and T. Schulte-Herbrüggen. “Reachable Sets from Toy Models to Controlled Markovian Quantum Systems”. In: *Proc. IEEE Conf. Decision Control (IEEE-CDC)* 58 (2019), p. 2322 (cit. on pp. 81, 87, 102, 115, 116, 118, 165, 217).

[Deu85] D. Deutsch. “Quantum Theory, the Church-Turing Principle, and the Universal Quantum Computer”. In: *Proc. Royal Soc. London A* 400 (1985), pp. 97–117 (cit. on pp. xiii, xvi).

[DH08] G. Dirr and U. Helmke. “Lie Theory for Quantum Control”. In: *GAMM-Mitteilungen* 31 (2008), pp. 59–93 (cit. on pp. xxv, 102).

[Dir+09] G. Dirr, U. Helmke, I. Kurniawan, and T. Schulte-Herbrüggen. “Lie-Semigroup Structures for Reachability and Control of Open Quantum Systems: Kossakowski-Lindblad Generators Form Lie Wedge to Markovian Channels”. In: *Rep. Math. Phys.* 64 (2009), pp. 93–121 (cit. on pp. xxiii, 82, 86).

[DiV00] D. DiVincenzo. “The Physical Implementation of Quantum Computation”. In: *Fortschr. Phys.* 48 (2000), pp. 771–783 (cit. on pp. 95, 129, 157).

[DJ92] D. Deutsch and R. Jozsa. “Rapid Solution of Problems by Quantum Computation”. In: *Proc. Royal Soc. London A* 439.1907 (1992), pp. 553–558 (cit. on p. xiii).

[DP89] D. Deutsch and R. Penrose. “Quantum Computational Networks”. In: *Proc. R. Soc. Lond. A.* 425.1868 (1989), pp. 73–90 (cit. on p. xvi).

[DRC17] C. L. Degen, F. Reinhard, and P. Cappellaro. “Quantum Sensing”. In: *Rev. Mod. Phys.* 89 (2017), p. 035002 (cit. on pp. xiv, xxiii).

[Dub09] C. Dubi. “An Algorithmic Approach to Simultaneous Triangularization”. In: *Lin. Alg. Appl.* 430 (2009), pp. 2975–2981 (cit. on pp. 125, 127).

[DV07] A. Datta and G. Vidal. “Role of Entanglement and Correlations in Mixed-State Quantum Computation”. In: *Phys. Rev. A* 75.4 (2007), p. 042310 (cit. on p. xxiii).

[Dya19] M. I. Dyakonov. “The Case Against Quantum Computing”. In: *IEEE Spectrum* (2019) (cit. on p. xiv).

[Dya20] M. I. Dyakonov. *Will We Ever Have a Quantum Computer?* SpringerBriefs in Physics. Cham: Springer, 2020 (cit. on p. xiv).

[ED22] F. vom Ende and G. Dirr. “The d-Majorization Polytope”. In: *Lin. Alg. Appl.* 649 (2022), pp. 152–185 (cit. on pp. 116, 217).

[EH12] A. Elgart and G. A. Hagedorn. “A Note on the Switching Adiabatic Theorem”. In: *J. Math. Phys.* 53.10 (2012), p. 102202 (cit. on p. xvi).

[Ein14] A. Einstein. “Beiträge zur Quantentheorie”. In: *Verh. d. D. phys. Ges.* 16 (1914), pp. 820–828 (cit. on p. xvi).

[EL77] D. Evans and J. Lewis. *Dilations of Irreversible Evolutions in Algebraic Quantum Theory*. Dublin, 1977 (cit. on p. 110).

[Ell09] D. Elliott. *Bilinear Control Systems: Matrices in Action*. London: Springer, 2009 (cit. on pp. xxv, 55, 127, 128, 175).

[End+19] F. vom Ende, G. Dirr, M. Keyl, and T. Schulte-Herbrüggen. “Reachability in Infinite-Dimensional Unital Open Quantum Systems with Switchable GKS–Lindblad Generators”. In: *Open Syst. Inf. Dyn.* 26 (2019), p. 1950014 (cit. on pp. 103, 107).

[End20] F. vom Ende. “Reachability in Controlled Markovian Quantum Systems: An Operator-Theoretic Approach”. PhD thesis. TU Munich, 2020. DOI: 10.14459/2020md1559809 (cit. on pp. 81, 87, 165).

[End23] F. vom Ende. *Understanding & Generalizing Unique Decompositions of Generators of Dynamical Semigroups*. arXiv:2310.04037. 2023 (cit. on p. 206).

[Epp13] J. F. Epperson. *An Introduction to Numerical Methods and Analysis*. 2nd edition. New Jersey: Wiley, 2013 (cit. on p. 134).

[ETS20] ETSI. *Migration Strategies and Recommendations to Quantum Safe Schemes*. Technical Report. 2020 (cit. on p. xv).

[Far+00] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. *Quantum Computation by Adiabatic Evolution*. arXiv:quant-ph/0001106. 2000 (cit. on p. xvi).

[Far01] D. Farenick. *Algebras of Linear Transformations*. New York: Springer, 2001 (cit. on pp. 104, 111).

[Fey82] R. P. Feynman. “Simulating Physics with Computers”. In: *Int. J. Theo. Phys.* 21 (1982), pp. 467–488 (cit. on pp. xiii, xv).

[FGG14] E. Farhi, J. Goldstone, and S. Gutmann. *A Quantum Approximate Optimization Algorithm*. arXiv:1411.4028. 2014 (cit. on p. xviii).

[Fra62] J. G. F. Francis. “The QR Transformation – Parts 1 and 2”. In: *Comput. J.* 4 (1961–1962), 265–271 and 332–345 (cit. on p. 126).

[GAN14] I. Georgescu, S. Ashhab, and F. Nori. “Quantum Simulation”. In: *Rev. Mod. Phys.* 86 (2014), pp. 153–185 (cit. on p. xv).

[Gar+01] A. Garroway et al. “Remote Sensing by Nuclear Quadrupole Resonance”. In: *IEEE Trans. Geosci. Remote Sens.* 39 (2001), pp. 1108–1118 (cit. on p. xiv).

[Gar19] C. L. Garrido Alzar. “Compact Chip-Scale Guided Cold Atom Gyrometers for Inertial Navigation: Enabling Technologies and Design Study”. In: *AVS Quantum Sci.* 1.1 (2019), p. 014702 (cit. on p. xiv).

[GE21] C. Gidney and M. Ekerå. “How to Factor 2048 Bit RSA Integers in 8 Hours Using 20 Million Noisy Qubits”. In: *Quantum* 5 (2021), p. 433 (cit. on pp. xv, xvii).

[Gib19] E. Gibney. “Quantum Gold Rush: The Private Funding Pouring into Quantum Start-Ups”. In: *Nature* 574.7776 (2019), pp. 22–24 (cit. on p. xiv).

[Gis+02] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. “Quantum Cryptography”. In: *Rev. Mod. Phys.* 74 (2002), pp. 145–195 (cit. on p. xxiii).

[GKS76] V. Gorini, A. Kossakowski, and E. Sudarshan. “Completely Positive Dynamical Semigroups of N-Level Systems”. In: *J. Math. Phys.* 17 (1976), pp. 821–825 (cit. on pp. xxiii, 81, 108, 110).

[Gla+15] S. Glaser et al. “Training Schrödinger’s Cat: Quantum Optimal Control”. In: *Eur. Phys. J. D* 69 (2015), p. 279 (cit. on pp. xxi, xxv).

[Gla03] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Stochastic Modelling and Applied Probability. New York: Springer, 2003 (cit. on p. xvii).

[GLR06] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications*. Classics in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics, 2006 (cit. on pp. 104, 111, 114).

[GR21] J. Gauthier and F. Rossi. “A Universal Gap for Non-Spin Quantum Control Systems”. In: *Proc. Amer. Math. Soc.* 149 (2021), pp. 1203–1214 (cit. on p. 53).

[Gro96] L. K. Grover. “A Fast Quantum Mechanical Algorithm for Database Search”. In: *Proc. of the 28th Annual Symposium on the Theory of Computing*. New York: ACM Press, 1996, pp. 212–219 (cit. on pp. xiii, xiv).

[Gro97] L. K. Grover. “Quantum Mechanics Helps in Searching for a Needle in a Haystack”. In: *Phys. Rev. Lett.* 79 (1997), pp. 325–328 (cit. on pp. xiii, xiv).

[Hah50] E. Hahn. “Spin Echoes”. In: *Phys. Rev.* 80 (1950), pp. 580–601 (cit. on p. xxi).

[Hal13] B. C. Hall. *Quantum Theory for Mathematicians*. New York: Springer, 2013 (cit. on p. xxii).

[Ham17] D. Hambling. “China’s Quantum Submarine Detector Could Seal South China Sea”. In: *New Scientist* (2017). URL: https://www.newscientist.com/article/2144721-chinas-quantum-submarine-detector-could-seal-south-china-sea/ (cit. on p. xiv).

[Hat02] A. Hatcher. *Algebraic Topology*. Cambridge: Cambridge University Press, 2002 (cit. on p. 102).

[Hel78] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. New York: Academic Press, 1978 (cit. on pp. 23–25, 27, 29–31, 33, 34, 39, 40, 44, 181, 183, 185).

[Hen+20] L. Henriet et al. "Quantum Computing with Neutral Atoms". In: *Quantum* 4 (2020), p. 327 (cit. on p. xix).

[Her+22] D. Herman et al. *A Survey of Quantum Computing for Finance*. arXiv:2201.02773. 2022 (cit. on p. xvii).

[HG22] C. Hoofnagle and S. Garfinkel. *Law and Policy for the Quantum Age*. Cambridge: Cambridge University Press, 2022 (cit. on p. xiv).

[HGK04] M. Hazewinkel, N. Gubareni, and V. Kirichenko. *Algebras, Rings and Modules*. Vol. 1. New York: Springer, 2004 (cit. on pp. 104, 125, 127).

[HHL89] J. Hilgert, K.-H. Hofmann, and J. Lawson. *Lie Groups, Convex Cones, and Semigroups*. Clarendon Press, Oxford, 1989 (cit. on pp. xxv, 57, 86).

[HHT23] T. Hoefler, T. Häner, and M. Troyer. "Disentangling Hype from Practicality: On Realistically Achieving Quantum Advantage". In: *Commun. ACM* 66 (2023), pp. 82–87 (cit. on pp. xiv, xv, xviii).

[HJ12] R. Horn and C. Johnson. *Matrix Analysis*. 2nd edition. Cambridge: Cambridge University Press, 2012 (cit. on pp. 183, 185, 190).

[HJ87] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge: Cambridge University Press, 1987 (cit. on p. 106).

[HKS13] A. Huckleberry, M. Kuś, and A. Sawicki. "Bipartite Entanglement, Spherical Actions, and Geometry of Local Unitary Orbits". In: *J. Math. Phys.* 54.2 (2013), p. 022202 (cit. on pp. 171, 173).

[HMR21] F. Hensel, M. Moor, and B. Rieck. "A Survey of Topological Machine Learning Methods". In: *Front. Artif. Intell.* 4 (2021) (cit. on p. xviii).

[HN12] J. Hilgert and K. Neeb. *Structure and Geometry of Lie Groups*. Springer Monographs in Mathematics. Berlin: Springer, 2012 (cit. on p. 86).

[HN19] H. Heidari and V. Nabaei. "SQUID Sensors". In: *Magnetic Sensors for Biomedical Applications*. Wiley & Sons, 2019. Chap. 5, pp. 163–212 (cit. on p. xiv).

[HN93] J. Hilgert and K.-H. Neeb. *Lie Semigroups and Their Applications*. Berlin: Springer, 1993 (cit. on pp. xxv, 57, 58).

[HO13] M. Horodecki and J. Oppenheim. "Fundamental Limitations for Quantum and Nanoscale Thermodynamics". In: *Nat. Commun.* 4 (2013), p. 2059 (cit. on p. 166).

[Hof+09] M. Hofheinz et al. "Synthesizing Arbitrary Quantum States in a Superconducting Resonator". In: *Nature* 459 (2009), pp. 546–549 (cit. on p. 115).

[Hor54] A. Horn. "Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix". In: *Am. J. Math.* 76 (1954), pp. 620–630 (cit. on pp. 85, 103, 162, 197).

[HR06] S. Haroche and J.-M. Raimond. *Exploring the Quantum: Atoms, Cavities, and Photons*. New York: Oxford University Press, 2006 (cit. on p. xix).

[Hua44] L. K. Hua. "On the Theory of Automorphic Functions of a Matrix Variable. I: Geometrical Basis". In: *Amer. J. Math.* 66 (1944), pp. 470–488 (cit. on pp. 171, 185).

[Hum72] J. Humphreys. *Introduction to Lie Algebras and Representation Theory*. New York: Springer, 1972 (cit. on p. 44).

[Isa09] I. Isaacs. *Algebra: a Graduate Course*. Providence, Rhode Island: Amer. Math. Soc., 2009 (cit. on p. 125).

[Isi05] A. Isidori. *Nonlinear Control Systems*. 3rd edition. London: Springer, 2005 (cit. on p. 127).

[JL03] R. Jozsa and N. Linden. "On the Role of Entanglement in Quantum-Computational Speed-up". In: *Proc. R. Soc. Lond. A.* 459 (2003), pp. 2011–2032 (cit. on p. xxiii).

[JLB23] M. Johnsson, L. van Luijk, and D. Burgarth. *Exact and Lower Bounds for the Quantum Speed Limit in Finite Dimensional Systems*. arXiv:2304.06617. 2023 (cit. on p. 53).

[JM98] J. A. Jones and M. Mosca. "Implementation of a Quantum Algorithm on a Nuclear Magnetic Resonance Quantum Computer". In: *J. Chem. Phys.* 109.5 (1998), pp. 1648–1653 (cit. on pp. xiii, xxi).

[Joh+11] M. W. Johnson et al. "Quantum Annealing with Manufactured Spins". In: *Nature* 473.7346 (2011), pp. 194–198 (cit. on p. xvi).

[Jos+01] M. Joswig, V. Kaibel, M. E. Pfetsch, and G. M. Ziegler. "Vertex-Facet Incidences of Unbounded Polyhedra". In: *Adv. Geom.* 1 (2001), pp. 23–36 (cit. on p. 167).

[Jos62] B. D. Josephson. "Possible New Effects in Superconductive Tunnelling". In: *Phys. Lett.* 1 (1962), pp. 251–253 (cit. on p. xx).

[Jos74] B. D. Josephson. "The Discovery of Tunnelling Supercurrents". In: *Rev. Mod. Phys.* 46 (1974), pp. 251–254 (cit. on p. xx).

[JP15] A. Jamiołkowski and G. Pastuszak. "Generalized Shemesh Criterion, Common Invariant Subspaces and Irreducible Completely Positive Superoperators". In: *Lin. Multilin. Alg.* 63 (2015), pp. 314–325 (cit. on pp. 125, 126).

[JRS07] S. Jansen, M.-B. Ruskai, and R. Seiler. "Bounds for the Adiabatic Approximation with Applications to Quantum Computation". In: *J. Math. Phys.* 48.10 (2007), p. 102111 (cit. on p. xvi).

[Jur97] V. Jurdjevic. *Geometric Control Theory*. Cambridge: Cambridge University Press, 1997 (cit. on pp. xxv, 55, 188).

[Kat50] T. Kato. "On the Adiabatic Theorem of Quantum Mechanics". In: *J. Phys. Soc. Jpn.* 5.6 (1950), pp. 435–439 (cit. on p. xvi).

[Kat80] T. Kato. *Perturbation Theory for Linear Operators*. Berlin: Springer, 1980 (cit. on pp. 3, 6, 7, 12, 13, 16–18, 49, 50, 115).

[KBG01] N. Khaneja, R. Brockett, and S. Glaser. "Time Optimal Control in Spin Systems". In: *Phys. Rev. A* 63 (2001), p. 032308 (cit. on pp. 53, 171).

[KC23] A. Khandelwal and M. G. Chandra. "Quantum Machine Learning with Quantum Topological Data Analysis". In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. Vol. 2. 2023, pp. 10–14 (cit. on p. xviii).

[KDH12] I. Kurniawan, G. Dirr, and U. Helmke. "Controllability Aspects of Quantum Dynamics: A Unified Approach for Closed and Open Systems". In: *IEEE Trans. Automat. Contr.* 57 (2012), pp. 1984–1996 (cit. on p. 93).

[Kep19] J. Kepler. *Harmonices Mundi Libri V.* Lincii Austriae [Linz]: sumptibus Godofredi Tampachii Bibl. Francof., excudebat Ioannes Plancus, 1619 (cit. on p. 201).

[KGB02] N. Khaneja, S. Glaser, and R. Brockett. "Sub-Riemannian Geometry and Time-Optimal Control of Three-Spin Systems: Quantum Gates and Coherence Transfer". In: *Phys. Rev. A* 65 (2002), p. 032301 (cit. on p. 53).

[Kha+05] N. Khaneja et al. "Optimal Control of Coupled Spin Dynamics: Design of NMR Pulse Sequences by Gradient Ascent Algorithms". In: *J. Magn. Reson.* 172 (2005), pp. 296–305 (cit. on p. xxvi).

[Kha02] H. Khalil. *Nonlinear Systems*. 3rd edition. Upper Saddle River, New Jersey: Prentice-Hall, 2002 (cit. on p. 65).

[Kit03] A. Y. Kitaev. "Fault-Tolerant Quantum Computation by Anyons". In: *Ann. Phys.* 303.1 (2003), pp. 2–30 (cit. on p. xviii).

[Kit96] A. Y. Kitaev. "Quantum Measurements and the Abelian Stabilizer Problem". In: *Electron. Colloquium Comput. Complex.* TR96 (1996) (cit. on p. xvii).

[KL10] P. Kok and B. W. Lovett. *Introduction to Optical Quantum Information Processing*. New York: Cambridge University Press, 2010 (cit. on p. xix).

[Kle06] M. Kleinsteuber. "Jacobi-Type Methods on Semisimple Lie Algebras: A Lie Algebraic Approach to the Symmetric Eigenvalue Problem". PhD thesis. TU Munich, 2006 (cit. on pp. 4, 23, 28, 181, 183–185).

[KLM01] E. Knill, R. Laflamme, and G. J. Milburn. "A Scheme for Efficient Quantum Computation with Linear Optics". In: *Nature* 409.6816 (2001), pp. 46–52 (cit. on p. xxi).

[KLZ98] E. Knill, R. Laflamme, and W. H. Zurek. "Resilient Quantum Computation". In: *Science* 279.5349 (1998), pp. 342–345 (cit. on p. xix).

[KM03] A. Kriegl and P. Michor. "Differentiable Perturbation of Unbounded Operators". In: *Math. Ann.* 327 (2003), pp. 191–201 (cit. on p. 6).

[KMW02] D. Kielpinski, C. Monroe, and D. J. Wineland. "Architecture for a Large-Scale Ion-Trap Quantum Computer". In: *Nature* 417 (2002), pp. 709–711 (cit. on p. xix).

[KN96] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Vol. I–II. New York: Wiley Interscience, 1996 (cit. on pp. 14, 29).

[KN98] T. Kadowaki and H. Nishimori. "Quantum Annealing in the Transverse Ising Model". In: *Phys. Rev. E* 58 (1998), pp. 5355–5363 (cit. on p. xvi).

[Kna02] A. W. Knapp. *Lie Groups Beyond an Introduction*. 2nd edition. Boston: Birkhäuser, 2002 (cit. on pp. xxv, 4, 40, 93, 125).

[Koc+07] J. Koch et al. "Charge-Insensitive Qubit Design Derived from the Cooper-Pair Box". In: *Phys. Rev. A* 76 (2007), p. 042319 (cit. on p. xx).

[Koc+14] G. Kochenberger et al. "The Unconstrained Binary Quadratic Programming Problem: A Survey". In: *J. Comb. Optim.* 28 (2014) (cit. on p. xvii).

[Koc+22] C. Koch et al. "Quantum Optimal Control in Quantum Technologies. Strategic Report on Current Status, Visions and Goals for Research in Europe". In: *EPJ Quantum Technol.* 9 (2022), p. 19 (cit. on p. xxv).

[Kok+07] P. Kok et al. "Linear Optical Quantum Computing with Photonic Qubits". In: *Rev. Mod. Phys.* 79 (2007), pp. 135–174 (cit. on p. xxi).

[Kos73] B. Kostant. "On Convexity, the Weyl Group and the Iwasawa Decomposition". In: *Ann. Sci. Éc. Norm. Supér.* 6 (1973), pp. 413–455 (cit. on pp. 60, 194, 197).

[KP02] S. Krantz and H. Parks. *A Primer of Real Analytic Functions*. 2nd edition. Boston: Birkhäuser, 2002 (cit. on p. 114).

[KP92] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Heidelberg: Springer, 1992 (cit. on p. 222).

[Kra+08] B. Kraus et al. "Preparation of Entangled States by Quantum Markov Processes". In: *Phys. Rev. A* 78 (2008), p. 042307 (cit. on pp. 87, 88, 97, 108, 113).

[Kup23] I. Kuprov. *Spin: From Basic Symmetries to Quantum Optimal Control*. Cham: Springer, 2023 (cit. on p. xxi).

[Laf92] T. Laffey. "A Structure Theorem for Some Matrix Algebras". In: *Linear Algebra Appl.* 162–164 (1992), pp. 205–215 (cit. on pp. 114, 115).

[Lap+10] M. Lapert et al. "Singular Extremals for the Time-Optimal Control of Dissipative Spin $\frac{1}{2}$ Particles". In: *Phys. Rev. Lett.* 104.8 (2010), p. 083001 (cit. on pp. 73, 129, 146, 149, 161).

[Lap+11] M. Lapert, Y. Zhang, S. J. Glaser, and D. Sugny. "Towards the Time-Optimal Control of Dissipative Spin-1/2 Particles in Nuclear Magnetic Resonance". In: *J. Phys. B: At. Mol. Opt. Phys.* 44.15 (2011), p. 154014 (cit. on p. 129).

[Lap+13] M. Lapert, E. Assémat, S. J. Glaser, and D. Sugny. "Understanding the Global Structure of Two-Level Quantum Systems with Relaxation: Vector Fields Organized Through the Magic Plane and the Steady-State Ellipsoid". In: *Phys. Rev. A* 88 (2013), p. 033407 (cit. on pp. 129, 135, 138, 146, 149).

[Law99] J. Lawson. "Geometric Control and Lie Semigroup Theory". In: *Differential Geometry and Control*. Ed. by G. Ferreyra. Providence, Rhode Island: Amer. Math. Soc., 1999, pp. 207–221 (cit. on pp. xxv, 57, 58, 68, 86).

[LBS89] V. Lakshmikantham, D. Bainov, and P. Simeonov. *Theory of Impulsive Differential Equations*. Series in Modern Applied Mathematics, Vol. 6. Singapore: World Scientific, 1989 (cit. on p. 115).

[LC19] G. H. Low and I. L. Chuang. "Hamiltonian Simulation by Qubitization". In: *Quantum* 3 (2019), p. 163 (cit. on p. xvii).

[LCW98] D. Lidar, I. Chuang, and K. Whaley. "Decoherence-Free Subspaces for Quantum Computation". In: *Phys. Rev. Lett.* 81 (1998), pp. 2594–2597 (cit. on pp. 90, 98).

[Lee13] J. M. Lee. *Introduction to Smooth Manifolds*. 2nd edition. Graduate Texts in Mathematics. New York: Springer, 2013 (cit. on pp. 20, 21, 30, 34).

[Let19] J. Letourneur. *Traité de Perspective*. Eyrolles, 2019 (cit. on p. 1).

[Lev08] M. H. Levitt. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. 2nd edition. Wiley & Sons, 2008 (cit. on pp. xix, xxi).

[LGZ16] S. Lloyd, S. Garnerone, and P. Zanardi. "Quantum Algorithms for Topological and Geometric Analysis of Data". In: *Nat. Commun.* 7.1 (2016), p. 10138 (cit. on p. xviii).

[Li+19] M. Li, T. Chen, J. J. Gooding, and J. Liu. "Review of Carbon and Graphene Quantum Dots for Sensing". In: *ACS Sens.* 4.7 (2019), pp. 1732–1748 (cit. on p. xiv).

[Li+21] X.-Q. Li et al. "Fast Laser Cooling Using Optimal Quantum Control". In: *Phys. Rev. A* 104 (2021), p. 043106 (cit. on p. 157).

[Li+24] L. Li et al. "Heterogeneous Integration of Spin-Photon Interfaces with a CMOS Platform". In: *Nature* 630.8015 (2024), pp. 70–76 (cit. on p. xxi).

[Lid13] D. A. Lidar. *Quantum Error Correction*. Cambridge: Cambridge University Press, 2013 (cit. on p. xix).

[Lin76] G. Lindblad. "On the Generators of Quantum Dynamical Semigroups". In: *Commun. Math. Phys.* 48 (1976), pp. 119–130 (cit. on pp. xxiii, 81, 108).

[Liu97] W. Liu. "An Approximation Algorithm for Nonholonomic Systems". In: *SIAM J. Control Optim.* 35 (1997), pp. 1328–1365 (cit. on pp. 65, 180).

[Llo96] S. Lloyd. "Universal Quantum Simulators". In: *Science* 273 (1996), pp. 1073–1078 (cit. on p. xiii).

[Llo99] S. Lloyd. "Rolf Landauer (1927-99)". In: *Nature* 400.720 (1999) (cit. on p. xiii).

[LMS93] S. Leela, F. McRae, and S. Sivasundaram. "Controllability of Impulsive Differential Equations". In: *J. Math. Anal. Appl.* 177 (1993), pp. 24–30 (cit. on p. 115).

[LPZ24] L. V. Lokutsievskiy, A. N. Pechen, and M. I. Zelikin. *Time-Optimal State Transfer for an Open Qubit*. arXiv:2401.10099. 2024 (cit. on p. 129).

[Lud+15] A. D. Ludlow et al. "Optical Atomic Clocks". In: *Rev. Mod. Phys.* 87 (2015), pp. 637–701 (cit. on p. xix).

[Mac+10] S. Machnes et al. "Superfast Laser Cooling". In: *Phys. Rev. Lett.* 104 (2010), p. 183001 (cit. on p. 157).

[Mac+11] S. Machnes et al. "Comparing, Optimizing, and Benchmarking Quantum-Control Algorithms in a Unifying Programming Framework". In: *Phys. Rev. A* 84 (2011), p. 022305 (cit. on p. xxvi).

[Mad+22] L. S. Madsen et al. "Quantum Computational Advantage with a Programmable Photonic Processor". In: *Nature* 606.7912 (2022), pp. 75–81 (cit. on p. xiii).

[Mar+12] E. Martín-López et al. "Experimental Realization of Shor's Quantum Factoring Algorithm Using Qubit Recycling". In: *Nature Photonics* 6.11 (2012), pp. 773–776 (cit. on p. xv).

[Mar09] J. M. Martinis. "Superconducting Phase Qubits". In: *Quantum Inf. Process.* 8 (2009), pp. 81–103 (cit. on p. xx).

[McC+03] N. McCarthy, D. Ogilvie, N. Zobin, and V. Zobin. "Convex Geometry of Coxeter-Invariant Polytopes". In: *Trends in Banach Spaces and Operator Theory*. Providence, Rhode Island: Amer. Math. Soc., 2003, pp. 153–180 (cit. on p. 78).

[MDC85] J. M. Martinis, M. H. Devoret, and J. Clarke. "Energy-Level Quantization in the Zero-Voltage State of a Current-Biased Josephson Junction". In: *Phys. Rev. Lett.* 55 (1985), pp. 1543–1546 (cit. on p. xx).

[MEA23] A. B. Magann, S. E. Economou, and C. Arenz. "Randomized Adaptive Quantum State Preparation". In: *Phys. Rev. Res.* 5 (2023), p. 033227 (cit. on p. 221).

[MKS05] A. Mailybaev, O. Kirillov, and A. Seyranian. "Coupling of Eigenvalues of Complex Matrices at Diabolic and Exceptional Points". In: *J. Phys. A Math. Theor.* 38 (2005), pp. 1723–1740 (cit. on p. 3).

[MMD05] D. S. Mackey, N. Mackey, and D. M. Dunlavy. "Structure Preserving Algorithms for Perplectic Eigenproblems". In: *Electronic J. Linear Algebra* 13 (2005), pp. 10–39 (cit. on p. 28).

[MN07] J. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. 3rd edition. Wiley & Sons, Chichester, 2007 (cit. on p. 110).

[MOA11] A. Marshall, I. Olkin, and B. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2nd edition. New York: Springer, 2011 (cit. on pp. 103, 105, 107, 166).

[Moi77] E. E. Moise. *Geometric Topology in Dimensions 2 and 3*. Vol. 47. Graduate Texts in Mathematics. New York: Springer, 1977 (cit. on p. 120).

[Mon+95] C. Monroe et al. "Resolved-Sideband Raman Cooling of a Bound Atom to the 3D Zero-Point Energy". In: *Phys. Rev. Lett.* 75 (1995), pp. 4011–4014 (cit. on p. 157).

[Mon15] A. Montanaro. "Quantum Speedup of Monte Carlo Methods". In: *Proc. R. Soc. A: Math. Phys. Eng. Sci.* 471.2181 (2015), p. 20150301 (cit. on p. xvii).

[Mon16] A. Montanaro. "Quantum Algorithms: An Overview". In: *npj Quantum Inf.* 2.1 (2016), p. 15023 (cit. on p. xvii).

[Muk+13] V. Mukherjee et al. "Speeding up and Slowing down the Relaxation of a Qubit by Optimal Control". In: *Phys. Rev. A* 88 (2013), p. 062326 (cit. on pp. 129, 145–147).

[NC10] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. 10th anniversary edition. Cambridge: Cambridge University Press, 2010 (cit. on pp. xvii, xxii).

[Neu56] J. von Neumann. "Probabilistic Logics and the Synthesis of Reliable Organisms From Unreliable Components". In: *Automata Studies. (AM-34), Volume 34*. Ed. by C. E. Shannon and J. McCarthy. Princeton: Princeton University Press, 1956, pp. 43–98 (cit. on p. xix).

[Nie04] M. A. Nielsen. "Optical Quantum Computation Using Cluster States". In: *Phys. Rev. Lett.* 93 (2004), p. 040503 (cit. on p. xxi).

[NIS23] NIST. *Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process*. Status Report. 2023 (cit. on p. xv).

[Nor97] J. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 1997 (cit. on pp. 115, 120).

[NPT99] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai. "Coherent Control of Macroscopic Quantum States in a Single-Cooper-Pair Box". In: *Nature* 398 (1999), pp. 786–788 (cit. on p. xx).

[NW17] R. Niederhagen and M. Waidner. *Practical Post-Quantum Cryptography*. White Paper. Fraunhofer SIT, 2017 (cit. on p. xv).

[OBr07] J. L. O'Brien. "Optical Quantum Computing". In: *Science* 318.5856 (2007), pp. 1567–1570 (cit. on p. xxi).

[Orl+99] T. P. Orlando et al. "Superconducting Persistent-Current Qubit". In: *Phys. Rev. B* 60 (1999), p. 15398 (cit. on p. xx).

[PCZ22] F. Pan, K. Chen, and P. Zhang. "Solving the Sampling Problem of the Sycamore Quantum Circuits". In: *Phys. Rev. Lett.* 129.9 (2022), p. 090502 (cit. on p. xiii).

[Per+14] A. Peruzzo et al. "A Variational Eigenvalue Solver on a Photonic Quantum Processor". In: *Nat. Commun.* 5 (2014), p. 4213 (cit. on p. xvii).

[Pon+86] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Vol. 4. L.S. Pontryagin Selected Works. New York: Gordon and Breach Science Publishers, 1986 (cit. on p. 200).

[PVC06] M. Popp, K. Vollbrecht, and J. Cirac. "Ensemble Quantum Computation and Algorithmic Cooling in Optical Lattices". In: *Fortschr. Phys.* 54 (2006), pp. 686–701 (cit. on p. 157).

[QR14] Y. Quintana and J. Rodríguez. "Measurable Diagonalization of Positive Definite Matrices". In: *J. Approx. Theory* 185 (2014), pp. 91–97 (cit. on pp. 3, 19, 20).

[Rai11] A. Rainer. “Quasianalytic Multiparameter Perturbation of Polynomials and Normal Matrices”. In: *Trans. Am. Math. Soc.* 363 (2011), pp. 4945–4977 (cit. on pp. 3, 13).

[RB01] R. Raussendorf and H. J. Briegel. “A One-Way Quantum Computer”. In: *Phys. Rev. Lett.* 86 (2001), pp. 5188–5191 (cit. on pp. xvi, xxi).

[RB02] R. Raussendorf and H. J. Briegel. “Computational Model Underlying the One-Way Quantum Computer”. In: *Quant. Inf. Comput.* 2 (2002), pp. 443–486 (cit. on p. xvi).

[RBR12] P. Rooney, A. M. Bloch, and C. Rangan. *Decoherence Control and Purification of Two-Dimensional Quantum Density Matrices under Lindblad Dissipation*. arXiv:1201.0399. 2012 (cit. on pp. 129, 137, 143).

[RBR16] P. Rooney, A. M. Bloch, and C. Rangan. “Flag-Based Control of Quantum Purity for $n = 2$ Systems”. In: *Phys. Rev. A* 93 (2016), p. 063424 (cit. on p. 129).

[RBR18] P. Rooney, A. Bloch, and C. Rangan. “Steering the Eigenvalues of the Density Operator in Hamiltonian-Controlled Quantum Lindblad Systems”. In: *IEEE Trans. Automat. Contr.* 63 (2018), pp. 672–681 (cit. on pp. 53, 81, 87, 89, 107).

[RCC89] P. Ray, B. K. Chakrabarti, and A. Chakrabarti. “Sherrington–Kirkpatrick Model in a Transverse Field: Absence of Replica Symmetry Breaking Due to Quantum Fluctuations”. In: *Phys. Rev. B* 39 (1989), pp. 11828–11832 (cit. on p. xvi).

[Rel69] F. Rellich. *Pertubation Theory of Eigenvalue Problems*. New York: Gordon and Breach, 1969 (cit. on pp. 3, 12, 13).

[RNK12] D. M. Reich, M. Ndong, and C. P. Koch. “Monotonically Convergent Optimization in Quantum Control Using Krotov’s Method”. In: *J. Chem. Phys.* 136.10 (2012), p. 104103 (cit. on p. xxvi).

[Rom05] S. Roman. *Advanced Linear Algebra*. New York: Springer, 2005 (cit. on p. 206).

[Ros+16] G. Rosenberg et al. “Solving the Optimal Trading Trajectory Problem Using a Quantum Annealer”. In: *IEEE J. Sel. Top. Signal Process.* 10.6 (2016), pp. 1053–1060 (cit. on p. xvii).

[Ros16] G. Rosenberg. *Finding Optimal Arbitrage Opportunities Using a Quantum Annealer*. White Paper. 1Qbit, 2016 (cit. on p. xvii).

[RR00] H. Radjavi and P. Rosenthal. *Simultaneous Triangularization*. Universitext. New York: Springer, 2000 (cit. on pp. 93, 103, 111, 125, 127).

[Sac22] Y. Sachkov. *Introduction to Geometric Control*. Optimization and Its Applications. Springer, 2022 (cit. on p. 200).

[Sat56] I. Satake. “On a Generalization of the Notion of Manifold”. In: *Proc. Natl. Acad. Sci.* 42.6 (1956), pp. 359–363 (cit. on p. 45).

[SAZ19] G. Styliaris, Á. Alhambra, and P. Zanardi. “Mixing of Quantum States under Markovian Dissipation and Coherent Control”. In: *Phys. Rev. A* 99 (2019), p. 042333 (cit. on pp. 81, 87, 107).

[SC99] R. Schack and C. M. Caves. “Classical Model for Bulk-Ensemble NMR Quantum Computation”. In: *Phys. Rev. A* 60 (1999), pp. 4354–4362 (cit. on pp. xiii, xxi).

[Sch+21] P. Scholl et al. “Quantum Simulation of 2D Antiferromagnets with Hundreds of Rydberg Atoms”. In: *Nature* 595.7866 (2021), pp. 233–238 (cit. on p. xvi).

[Sch23] I. Schur. "Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie". In: *Sitzungsber. Berl. Math. Ges.* 22 (1923), pp. 9–20 (cit. on pp. 85, 103, 162, 197).

[SED22] T. Schulte-Herbrüggen, F. vom Ende, and G. Dirr. "Exploring the Limits of Open Quantum Dynamics I: Motivation, First Results from Toy Models to Applications". In: *Proc. MTNS*. arXiv:2003.06018. 2022, p. 1069 (cit. on pp. 81, 87, 165).

[SFS01] S. G. Schirmer, H. Fu, and A. I. Solomon. "Complete Controllability of Quantum Systems". In: *Phys. Rev. A* 63 (2001), p. 063410 (cit. on p. 127).

[SG04] R. A. Servedio and S. J. Gortler. "Equivalences and Separations Between Quantum and Classical Learnability". In: *SIAM J. Comput.* 33.5 (2004), pp. 1067–1092 (cit. on p. xviii).

[Sha20] J. H. Shapiro. "The Quantum Illumination Story". In: *IEEE Aerosp. Electron. Syst. Mag.* 35.4 (2020), pp. 8–20 (cit. on p. xiv).

[She84] D. Shemesh. "Common Eigenvectors of Two Matrices". In: *Lin. Alg. Appl.* 62 (1984), pp. 11–18 (cit. on pp. 125, 126).

[Sho94] P. Shor. "Algorithms for Quantum Computation: Discrete Logarithms and Factoring". In: *Proceedings of the Symposium on the Foundations of Computer Science, 1994, Los Alamitos, California*. IEEE Computer Society Press, New York, 1994, pp. 124–134 (cit. on pp. xiii, xiv, xvii).

[Sho96] P. W. Shor. "Fault-tolerant quantum computation". In: *Proceedings of 37th Conference on Foundations of Computer Science*. 1996, pp. 56–65 (cit. on p. xix).

[Sho97] P. Shor. "Polynomial-Time Algorithms for Prime Factorisation and Discrete Logarithm on a Quantum Computer". In: *SIAM J. Comput.* 26 (1997), pp. 1484–1509 (cit. on pp. xiii, xiv, xvii).

[Sim94] D. Simon. "On the Power of Quantum Computation". In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. 1994, pp. 116–123 (cit. on pp. xiii, xvii).

[SK14] V. Scarani and C. Kurtsiefer. "The Black Paper of Quantum Cryptography: Real Implementation Problems". In: *Theor. Comput. Sci.* 560 (2014), pp. 27–32 (cit. on p. xv).

[SL05] A. Shabani and D. Lidar. "Theory of Initialization-Free Decoherence-Free Subspaces and Subsystems". In: *Phys. Rev. A* 72 (2005), p. 042303 (cit. on p. 112).

[SM99] A. Sørensen and K. Mølmer. "Quantum Computation with Ions in Thermal Motion". In: *Phys. Rev. Lett.* 82 (1999), pp. 1971–1974 (cit. on p. xix).

[Smi02] G. Smirnov. *Introduction to the Theory of Differential Inclusions*. Providence, Rhode Island: Amer. Math. Soc., 2002 (cit. on pp. xxiv, xxv, 56, 61, 69, 71, 72, 75, 76, 83, 116, 122, 177).

[SMW05] L. J. Schulman, T. Mor, and Y. Weinstein. "Physical Limits of Heat-Bath Algorithmic Cooling". In: *Phys. Rev. Lett.* 94 (2005), p. 120501 (cit. on p. 157).

[SN21] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. 3rd edition. Cambridge: Cambridge University Press, 2021 (cit. on p. xxii).

[Son98] E. Sontag. *Mathematical Control Theory*. New York: Springer, 1998 (cit. on pp. xxiv, 55, 65, 68, 70, 72).

[SP18] M. Schuld and F. Petruccione. *Supervised Learning with Quantum Computers*. Cham: Springer, 2018 (cit. on p. xviii).

[Spo77] H. Spohn. "An Algebraic Condition for the Approach to Equilibrium of an Open N-Level System". In: *Lett. Math. Phys.* 2 (1977), pp. 33–38 (cit. on p. 88).

[SSV13] J. A. Smolin, G. Smith, and A. Vargo. "Oversimplifying Quantum Factoring". In: *Nature* 499.7457 (2013), pp. 163–165 (cit. on p. xv).

[STK04] S. Sklarz, D. Tannor, and N. Khaneja. "Optimal Control of Quantum Dissipative Dynamics: Analytic Solution for Cooling the Three-Level Λ System". In: *Phys. Rev. A* 69 (2004), p. 053408 (cit. on pp. 53, 81, 157, 163, 164, 167, 168).

[Str15] S. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chamistry and Engineering*. 2nd edition. Boca Raton: CRC Press, 2015 (cit. on p. 122).

[SW10] S. Schirmer and X. Wang. "Stabilizing Open Quantum Systems by Markovian Reservoir Engineering". In: *Phys. Rev. A* 81 (2010), p. 062306 (cit. on pp. 87, 88, 108, 113).

[Tak25] T. Takagi. "On an Algebraic Problem Related to an Analytic Theorem of Carathéodory and Fejér and on an Allied Theorem of Landau". In: *Jap. J. Math.* 83.1 (1925) (cit. on pp. 171, 183).

[Tan19] E. Tang. "A Quantum-Inspired Classical Algorithm for Recommendation Systems". In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2019. Association for Computing Machinery, 2019, pp. 217–228 (cit. on p. xviii).

[TB16] M. A. Taylor and W. P. Bowen. "Quantum Metrology and its Application in Biology". In: *Physics Reports* 615 (2016), pp. 1–59 (cit. on p. xiv).

[TC21] A. Tirelli and N. C. Costa. "Learning Quantum Phase Transitions Through Topological Data Analysis". In: *Phys. Rev. B* 104 (2021), p. 235146 (cit. on p. xviii).

[Tes12] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Vol. 140. Graduate Studies in Mathematics. Providence, Rhode Island: Amer. Math. Soc., 2012 (cit. on p. 60).

[TM92] W. G. Teich and G. Mahler. "Stochastic Dynamics of Individual Quantum Systems: Stationary Rate Equations". In: *Phys. Rev. A* 45.5 (1992), pp. 3300–3318 (cit. on p. 81).

[Tur37] A. M. Turing. "On Computable Numbers, with an Application to the Entscheidungsproblem". In: *Proc. London Math. Soc.* s2-42.1 (1937), pp. 230–265 (cit. on p. xvi).

[TV09] F. Ticozzi and L. Viola. "Analysis and Synthesis of Attractive Quantum Markovian Dynamics". In: *Automatica* 45 (2009), pp. 2002–2009 (cit. on pp. 87, 88, 97, 99, 108, 115).

[US 22] U.S. Department of Energy. "DOE National Laboratory Makes History by Achieving Fusion Ignition". In: (2022). URL: https://www.energy.gov/articles/doe-national-laboratory-makes-history-achieving-fusion-ignition (cit. on p. xiii).

[Vid03] G. Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". In: *Phys. Rev. Lett.* 91.14 (2003), p. 147902 (cit. on p. xxiii).

[VMC08] F. Verstraete, V. Murg, and J. Cirac. "Matrix Product States, Projected Entangled Pair States, and Variational Renormalization Group Methods for Quantum Spin Systems". In: *Adv. Phys.* 57.2 (2008), pp. 143–224 (cit. on p. xv).

[Wat18] J. Watrous. *The Theory of Quantum Information*. Cambridge: Cambridge University Press, 2018 (cit. on p. 110).

[WDC92] D. J. Wineland, J. Dalibard, and C. Cohen-Tannoudji. "Sisyphus Cooling of a Bound Atom". In: *J. Opt. Soc. Am. B* 9 (1992), pp. 32–42 (cit. on p. 157).

[Web+10] J. R. Weber et al. “Quantum Computing with Defects”. In: *Proc. Natl. Acad. Sci.* 107.19 (2010), pp. 8513–8518 (cit. on p. xxi).

[Wen17] G. Wendin. “Quantum Information Processing with Superconducting Circuits: a Review”. In: *Rep. Prog. Phys.* 80 (2017), p. 106001 (cit. on p. xx).

[WHT16] D. Wecker, M. B. Hastings, and M. Troyer. “Training a Quantum Optimizer”. In: *Phys. Rev. A* 94.2 (2016), p. 022309 (cit. on p. xviii).

[Wie83] S. Wiesner. “Conjugate Coding”. In: *SIGACT News* 15.1 (1983), pp. 78–88 (cit. on p. xv).

[Win+23] K. Wintersperger et al. “Neutral Atom Quantum Computing Hardware: Performance and End-User Perspective”. In: *EPJ Quantum Technol.* 10 (2023), p. 32 (cit. on p. xix).

[Won+19] C. Wong, C. Wilen, R. McDermott, and M. Vavilov. “A Tunable Quantum Dissipator for Active Resonator Reset in Circuit QED”. In: *Quant. Sci. Technol.* 4 (2019), p. 025001 (cit. on p. 115).

[Yin+13] Y. Yin et al. “Catch and Release of Microwave Photon States”. In: *Phys. Rev. Lett.* 110 (2013), p. 107001 (cit. on p. 115).

[Yua10] H. Yuan. “Characterization of Majorization Monotone Quantum Dynamics”. In: *IEEE Trans. Automat. Contr.* 55 (2010), pp. 955–959 (cit. on pp. 53, 81, 87, 103, 129, 145, 146).

[Zab20] J. Zabczyk. *Mathematical Control Theory*. 2nd edition. Boston: Birkhäuser, 2020 (cit. on pp. xxiv, xxv).

[Zăl02] C. Zălinescu. *Convex Analysis in General Vector Spaces*. Singapore: World Scientific, 2002 (cit. on p. 19).

[Zar88] R. N. Zare. *Angular Momentum*. New York: Wiley Interscience, 1988 (cit. on p. 116).

[Zho+20] H.-S. Zhong et al. “Quantum Computational Advantage Using Photons”. In: *Science* 370.6523 (2020), pp. 1460–1463 (cit. on p. xiii).

[Zie07] G. M. Ziegler. *Lectures on Polytopes*. New York: Springer, 2007 (cit. on p. 167).

[ZS11] R. Zeier and T. Schulte-Herbrüggen. “Symmetry Principles in Quantum System Theory”. In: *J. Math. Phys.* 52 (2011), p. 113510 (cit. on p. 128).

[ZYK08] R. Zeier, H. Yuan, and N. Khaneja. “Time-Optimal Synthesis of Unitary Transformations in a Coupled Fast and Slow Qubit System”. In: *Phys. Rev. A* 77 (2008), p. 032332 (cit. on p. 53).