

THIS DISSERTATION IS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

The Control of Gaussian Quantum States

Uther SHACKERLEY-BENNETT



Declaration

I, Uther Shackerley-Bennett confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

- Secs. 5 and 7.3 are entirely my work.
- Sec. 8.3 was collaborative work with Alexander Pitchford, Marco G. Genoni, Daniel Burgarth and Alessio Serafini. Alexander produced the numerics and I did all the analytics. Daniel, Alessio and Marco suggested good routes to take in deriving and producing these results. Alexander's numerics produced the plots. This resulted in the paper, Ref. [1].
- Sec. 8.5 was entirely the work of Theo Sepulchre, a master's student working under my supervision.
- Secs. 9.3 and 10.1 stemmed from collaboration with Alberto Carlini, Alessio Serafini and Vittorio Giovannetti. The three others initially worked on the project and up to a few modes showed that Eq. (9.48) held. I proved this in general, and then extended this to the proof of Theorem 43. I then developed the control results of Sec. 10.1. This resulted in Ref. [2].
- Secs. 9.4 and 10.2 was based on the results of the former paragraph with an extension to entangled states. This has been submitted for publication with Alessio Serafini [3].

Abstract

The rise of quantum technology has put control at the centre of advancements in quantum mechanics. The union of quantum mechanics with mathematical control theory is a meeting that is leading to a much deeper insight into our interaction with the bizarre properties of quantum theory.

Often, the study of discrete variable systems is the focus for making this union. Here, we look at how control theory may be applied to the continuous variable theory of Gaussian states. Special emphasis is given to control of the covariance matrix of these states, as it is here that we find the entanglement and entropic properties of the state.

We begin by exploring some initial results for the geometry of Gaussian states, revealing different manifold structures dependent on symplectic eigenvalue degeneracy. In this geometrical setting a proposal for an extension of Williamson's theorem is put forward and partially completed.

It is often interesting to look at restricted sets of Hamiltonians and ask what transformations can be performed with concatenations of their corresponding unitaries. Controllable systems are those for which the entire group of interest is possible to enact. We explore an uncontrollable system in a single mode and give a physical analysis as to why it behaves this way. This leads to ideas to move forwards for a necessary and sufficient condition for control on the symplectic group that has been conjectured since 1972.

Later, we transfer to the question of open dynamics. We focus on a particular and ubiquitous channel known as 'lossy' or 'the attenuation channel'. An equation is derived describing the evolution for the symplectic invariants of a Gaussian state undergoing such dynamics.

The equation of the former chapter is used to explore the evolution of entropy and entanglement. Optimal protocols are developed for the manipulation of these properties undergoing lossy dynamics.

Acknowledgements

The first people that I would like to thank are my collaborators on all projects, especially Marco Genoni for his inimitable ability to detect my numerous mistakes, Alexander Pitchford, Daniel Burgarth, Alberto Carlini, Vittorio Giovannetti and Alessio Serafini.

My PhD experience was greatly enhanced by numerous discussions with members of my department. Especially Lluís Masanes, who has been incredibly generous with his time; his insights and creativity have been an incredible stimulus during my time at UCL. Alongside these conversations, fellow students like Carlo Sparaciari, Thomas Galley, Alvaro Alhambra, Dan Goldwater and Sofia Qvarfort have livened up the experience.

Of my collaborators, there is one that of course deserves special mention. My supervisor Alessio has pushed forward fun and risky projects that were great fun working on. Spending time indulging in tricky problems together has been enormous fun.

In the background throughout the whole project have been my parents, Martine and Claire, who have put up with a heavy dose of PhD-induced pedantry.

Through the vacillations and vagaries of the three years, my partner Sophie has borne the real brunt, letting me talk for far too long on the mathematical intricacies of control theory.

Contents

1	Introduction	13
2	Gaussian states	17
2.1	Hilbert space	19
2.2	Phase space	25
2.3	Evolution	29
3	The symplectic group	33
3.1	Other bases	34
3.2	Lie algebra	36
3.3	Properties	36
3.4	Subgroups and decompositions	38
3.5	Single-shot symplectics	42
4	The covariance matrix	47
4.1	Decomposition and categorisation	48
4.2	Squeezing measure	51
4.3	Symplectic invariants	52
4.4	Entanglement	55
5	Geometry of Gaussian states	59
5.1	Covariance matrix manifold	59
5.2	Williamson's theorem extension	64
6	Control theory	67
6.1	Preliminaries	69
6.2	Preliminary controllability results	73
6.3	Multiple control field results	76

7	Neutrality	79
7.1	Introduction	79
7.2	In control	82
7.3	Symplectic neutrality	82
8	Single-mode control	85
8.1	Structure of the single-mode symplectic group	86
8.2	Controllability	88
8.3	Visualisation	90
8.3.1	Bounding the reachable set	91
8.3.2	Euler translation	92
8.3.3	Visualising the reachable set	94
8.4	Spacetime analogy	97
8.5	Ever-growing function	100
9	Open Gaussian systems	103
9.1	Input-output formalism	104
9.2	Lossy channels	108
9.3	Symplectic invariant evolution	108
9.4	Entanglement evolution	115
10	Lossy system control	117
10.1	Entropic control	118
10.1.1	Optimal symplectics	118
10.1.2	Decoupling	120
10.2	Entanglement control	123
11	Conclusion	131
A	Groups and topology foray	133
B	Hyperbolic system transformation	135
C	Euler decomposition uniqueness	139
D	Optimisation and bistochastic matrices	141

List of Tables

4.1	A categorisation of the different types of covariance matrix which can describe a Gaussian state.	51
-----	---	----

List of Figures

2.1	Single-mode Wigner functions with different variances, but both complying with the Robertson-Schrödinger uncertainty relation.	29
4.1	An implication diagram linking different conditions for entanglement.	56
5.1	A graphical representation of the manifold of covariance matrices in two modes. The diagonal line depicts the case when the two symplectic eigenvalues, ν_1 and ν_2 , are degenerate. Each point represents a manifold in a different form, either as diffeomorphic to a plane or the product of a 2-sphere and a plane.	63
7.1	A graphical representation of the implications used to prove Theorem 34.	81
8.1	The reachable set of an example single-mode hyperbolic system with $b = 0$ is represented in a cubic plot for different times of evolution.	97
8.2	The reachable set of an example single-mode hyperbolic system is represented in a cubic plot for $b = -0.99$ and $T = 5$	98
8.3	Imagine we begin in a lightcone in one of the angle-bounded regions, depicted by the squares. Enacting $\theta \rightarrow -\phi - \pi/2$ and $\phi \rightarrow -\theta + \pi/2$ will take squares of a single colour to squares of the same colour with the opposite arrow direction, indicating opposite time direction. A transformation of $\theta \rightarrow \theta - \pi$ and $\phi \rightarrow \phi - \pi$ takes us to a square of exactly the same time as this transformation does nothing to the lightcone.	100
9.1	The input-output formalism describes a stream of bath-modes interacting momentarily with the system, labeled S , and then being discarded. Although in many setups these outgoing modes would in fact be measured instead.	105

- 10.1 The solution to Eq. (10.12) is given for the values $\nu_0 = 5$, $\chi = 1$ and two different initial squeezing values. $z = 3$ implies that the entropy monotonically decreases whereas a higher initial squeezing parameter induces an initial ‘bump’ in the entropy. 122
- 10.2 A plot of the trace term in Eq. (10.28) against the variable α_A for the specific values: $a_A = 2$, $b_A = 5$, $c_A = 3$, $z_A = 4$ 126
- 10.3 The graph shows the evolution of $\tilde{\Sigma}$ under Eq. (10.19) for a two-mode state. Different solutions are given as we vary the number of times optimal control was enacted during the evolution. 129

Chapter 1

Introduction

Quantum theory began its life as a remedy to a radiation problem. James Clerk Maxwell's theory of electromagnetism was unable to account for the behaviour of a confined field and so Max Planck introduced the idea of a 'quantum' of energy $h\nu$, where ν is the frequency of the radiation and h was the new quantum constant [4]. This modification of the continuous ontology of Maxwell's fields was applied to a host of other problems which gathered momentum throughout the decade [5]. In 1913 Niels Bohr managed to use the intuition from Planck's idea to provide an understanding of the spectral lines of hydrogen, with a renewed understanding of the Rutherford model, which was itself classical [6]. During the remainder of the decade and into the next, a description of the atom and its spectral properties became the central focus and success story of this early theory of quanta.

Ever-present was a desire to create a more systematic application of quantum theory to broaden its application beyond the atom. The pioneers of the field sought a systematic way to turn *any* classical picture into a quantum description that gave the correct results. The great insight came with Werner Heisenberg's paper of 1925 highlighting the role that non-commutativity should play in the mathematical quantisation of any classical theory [7]. However, the deeper mathematical structure underlying Heisenberg's work was unclear even to him. Following a meeting in 1925 in Göttingen, Heisenberg remarked that "Now the learned Göttingen mathematicians talk so much about Hermitian matrices, but I do not even know what a matrix is" [5]. On one of the great train journeys of history, the young Pascual Jordan took it upon himself to eavesdrop on the conversation of Max Born and Wolfgang Pauli who were discussing Heisenberg's paper. After the train pulled into the station Jordan approached Born and offered him his help which sparked the 1925 *Dreier-Manner-Arbeit* (three-man-work) by Born, Jordan and Heisenberg, where the modern structure of quantum mechanics began to take shape [8].

Throughout the twentieth century the field has blossomed into a cacophany of technological

and philosophical research. Its bizarre nature was settled with Bell's theorem, which remains the linchpin of modern attempts to understand the theory [9]. Despite its metaphysical oddities, as a physical theory it has withstood every test and is the foundation of the Standard Model, which provides an account of the fundamental constituents of matter.

The rise of quantum information theory in the latter half of the last century has provided a new ground on which physicists can flesh out the structure and utility of quantum mechanics. This ground is a place where the application of the theory makes direct contact with questions about its meaning. This is due to the fact that isolating the bizarre properties of quantum mechanics provides intuition for the required content of a non-classical information protocol.

The conceived technologies invented by quantum information theorists, however, are a long way from their real-world implementation. Our power to bring these machines to life remains limited by our ability to manipulate quantum properties. It is thus no surprise that control over the quantum world is the central aim of a vast section of experimental research.

The problem of controlling machinery, however, is of course not new. The industrial revolution spawned a plethora of new technologies that required stabilising, damping and optimising. In the early days, components to do this were added by their inventors based on intuition. The shift came at a conference in 1867 when Maxwell led a call to arms for the community to produce a mathematical structure for these effects. He delivered a paper describing a pre-existing engine component called a 'governor' which acts as a stabilising component in many machines [10]. His description of the governor was through a system of differential equations, which are often cited as the birth equations of modern control theory [11].

This theory developed throughout the nineteenth and twentieth centuries, calling on increasingly sophisticated mathematical techniques. Its collision with quantum theory was inevitable after the latter became viewed as a theory of technological utility. Quantum control theory has a growing literature as people try to increase their ability to manipulate these systems. This has been mirrored by the rise of resource theories that explicitly recognise the utility of certain of the properties of quantum states. A type of state, known as a *Gaussian state*, plays a prominent role in much modern technology, as well as having intriguing mathematical properties. Although only expressible using an infinite-dimensional Hilbert space, the Gaussian subtheory allows one to avoid many of the subtleties contained in the broader exploration of control theory in infinite dimensions.

The thesis is structured so as to introduce the reader to Gaussian states and the symplectic group in Secs. 2–4, which will set the stage for novel results in later sections. The first foray into an extension of the existent theory is given in Sec. 5 where an attempt to develop the geometry of Gaussian states is provided.

The underlying motivation was the success with which geometry has been applied to the understanding of finite-dimensional quantum systems through the Bloch ball representation, as

well as the revolutionising role that geometry has played in abstract control theory, which is introduced in Sec. 6, with some small results in Sec 7 developing the notion of neutrality.

In Sec. 8 the control theory of closed Gaussian systems is explored. Here, we perform a physical analysis of existing mathematical results on the controllability of the symplectic group. We find that the behaviour of nontrivially uncontrollable Gaussian systems is such that unbounded squeezing occurs. This is important in that it provides a characterisation that can be explored for all uncontrollable systems, lending insight into a long-standing open problem in the field.

In Sec. 9 the system is opened to an environment and we consider evolution under lossy dynamics. Fadeev-Leverrier recursion is used to arrive at an evolution equation for the symplectic invariants of the covariance matrix in such a situation. This allows a direct insight into the behaviour of these important quantities.

In Sec. 10 the former equation is used in the context of control. The equation enables the derivation of optimal control strategies for the heating and cooling of Gaussian states, as well as for the preservation of entanglement, which is gradually wiped out by lossy dynamics.

The structure of the thesis is such that the technical background is covered largely before the statement of new results. We begin with an overview of Gaussian states before exploring the structures of the symplectic group and the covariance matrix. These provide the bedrock on which we can discuss new insights in the field.

Chapter 2

Gaussian states

The theory of electromagnetism was present at the birth of quantum theory in Planck's founding paper, where he realised that a thermodynamical and discrete description of light might be a better fit to certain experimental data [4]. Einstein took this further and in 1905 gave a description of the photoelectric effect for which he won the Nobel prize [12]. Following the successes of these early theories, Maxwell's description of radiation was continually questioned. Specifically, the interaction of electrons and electromagnetic radiation provided varying ontologies for how we should conceive of light. One of the most influential experiments in this regard came from Compton's light scattering experiment [13] which was one of the central results used by those arguing that light was made of particles, later called 'photons' [14].

As the theory of quantum fields matured throughout the later twentieth century, the simple ontology of some of the earlier quantum models held firm in the teaching of quantum mechanics, if not in the theory itself. That is, that quantum mechanics was about turning fields into particles. The continued use of this older physical picture led Willis Lamb at a conference in 1960 to declare a ban on the use of the word 'photon'. He suggested a licence be required by anyone wishing to use the word; applications were to be directed to himself [15].

To avoid entering into the mesh of visualisations of quantum theory, many of which rely on pre-1925 intuitions, we restrict ourselves to a mathematical treatment of the field. We consider a Gaussian state to begin as a particular state over the modes of electric and magnetic fields where the amplitudes of these are promoted to operators on a Hilbert space. Any further physical intuition would require an interpretation of the theory.

The electric and magnetic field operators can be mapped and rewritten in a form that looks identical to the position and momentum operators of ordinary quantum mechanics, despite the field theory origins of the equations. These fundamental variables that describe the system have a continuous spectrum which causes the Hilbert space representation to have infinite-

dimensionality. In 1932 Wigner created a new way of representing certain continuous variable systems using a phase space picture. Now referred to as the Wigner function, this provided a novel way to represent a quantum state that did not require a space of infinite dimensions [16].

A Gaussian quantum state is one whose Wigner function is Gaussian. A very early apparition of these states came in 1926 in an exchange between Schrödinger and Hendrik Lorentz. Schrödinger's early understanding of his wave equation was as a description of a continuous expression of the charge density of the electron [5]. In this mental picture, particles were given as 'wave packets' of the field. Lorentz opposed this idea, stating that particles could not be described by waves because of the way in which they inevitably spread. No evidence suggested particles exhibiting this behaviour and so Schrödinger's interpretation should be untenable according to Lorentz [17]. Schrödinger's response was to provide a solution to the wave equation which did not exhibit this behaviour. Due to the way in which it maintained its shape during evolution, it was referred to as a *coherent state* [18]. Coherent states have a Gaussian Wigner function and so form a subset of the Gaussian states. Thus, one of the earliest solutions to the wave equation, before its modern interpretation, was a Gaussian state.

Coherent states are centrally important in the quantum description of the laser, invented in 1960 [19]. They are sometimes referred to as the most classical of the states because the evolution of their average mimics the classical solution to Maxwell's equations [19]. It took some time before the less classical 'squeezed' Gaussian states were observed in 1985 in AT&T Bell labs by Richard E. Slusher [20], although they were considered as early as 1927 by Earle Kennard [18].

The importance of coherent states for quantum optics was recognised in two papers by E. C. George Sudarshan [21] and Roy Glauber [22], published within a few months of each other. In the former paper, Sudarshan shows that coherent states can be used as a basis to provide a new representation of optical states. This new representation was named the *P*-representation in Glauber's paper which had connections with the Wigner representation, as another function on phase space. Glauber's paper won him the Nobel prize, to go alongside his Lamb Licence [23].

Gaussian states appear in a wide range of physical setups, not confined to the quantum optical setting. Optomechanical systems describing micro-scale oscillators interacting with electromagnetic fields use the theory extensively [24]. Furthermore, trapped ion systems depend on the theory [25] as well as new explorations of the gravitational quantum regime using superpositions of massive objects [26].

However, quantum theory is about more than just states. Restricting to Gaussian-preserving measurements and transformations forms the Gaussian subtheory of quantum mechanics. In this subtheory quadratic Hamiltonians, which describe two-point interactions, arise as fundamental. This reveals the cause of the ubiquity of the theory since higher order interactions are difficult to achieve in many situations.

A quick caveat that needs to be borne in mind is that we will always use the term Gaussian

state to refer to Gaussian states of bosonic field modes. The study of these states has been extended to fermionic modes with anti-commuting operators [27], but these will not be discussed in this work. Therefore, it will always be implicit that the bosonic, rather than fermionic, flavour of these states is being referred to.

Over time the Gaussian state formalism, although still barely mentioned in many optics textbooks [20, 28], has been developed into a mathematical subtheory of quantum mechanics, studied independently of any direct interpretation [29, 30]. A development in recent years has been to nestle the framework of Gaussian states into the paradigm of quantum information theory so that its potential use in the development of information technology can be clarified [31]. Before attempting to control these states, we must first learn the formalism to describe them and their properties.

2.1 Hilbert space

The standard definition of Gaussian states is usually with respect to its Wigner function. Although this has pictorial clarity the path through Hilbert spaces can be more instructive and provide a link with finite-dimensional methods – “the longer road sometimes gives more familiarity with the country” [32]. Following Ref. [33], we will begin in the Hilbert space and then work forwards to the phase space representation.

The common starting point in quantum optics, as found in Refs. [20, 28] is to consider an electromagnetic field that is spatially confined. By solving Maxwell’s equations under such boundary conditions, the vibrational modes of the field become discretised into a countably infinite set. Any field in the confined space can be written as a sum of these modes, where each mode has an electric and magnetic component. Taking a finite set of n field modes, we then quantise by promoting these $2n$ degrees of freedom to operators acting on an infinite-dimensional Hilbert space.

Throughout this work the canonical operators of the i th mode will be represented by \hat{x}_i and \hat{p}_i . This abstract treatment could equally well describe a set of n coupled mechanical quantum harmonic oscillators and so they will be referred to as the canonical position and momentum operators for the given mode. We see that restricting to a finite subset of modes allows quantum optics to have the formal form of quantum particle mechanics, despite describing a field theory. Although we deal with a field as the underlying ontology of the theory, the finite degrees of freedom mean that the full complexities of field theory are removed.

The basis of standard quantum mechanics since 1925, and the quantum optical formalism, lies in stating a commutation relation for these canonical operators,

$$[\hat{x}_j, \hat{p}_k] = i\delta_{jk}\hat{\mathbb{1}}, \quad (2.1)$$

where $\hat{\mathbb{1}}$ is the identity operator on the Hilbert space and \hbar has already been set to one, as is commonly done, and δ is the Kronecker delta. Concatenating the $2n$ mode operators into a vector form, $\hat{\mathbf{r}} = (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_n, \hat{p}_n)^\top$, the commutation relation can be rewritten as

$$[\hat{\mathbf{r}}, \hat{\mathbf{r}}^\top] = i\Omega\hat{\mathbb{1}}, \quad (2.2)$$

where the outer product has been used to create a matrix of commutators. We also have

$$\Omega := \bigoplus_{j=1}^n \Omega_1, \quad \Omega_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (2.3)$$

Ω is an object that appears extensively in the Hamiltonian formalism of classical mechanics and appears again now in the quantum formalism.

Definition 1 (Quadratic Hamiltonian). *A Hamiltonian is referred to as quadratic if it is of the form*

$$\hat{H} = \frac{1}{2}\hat{\mathbf{r}}^\top H\hat{\mathbf{r}} + \hat{\mathbf{r}}^\top \mathbf{r}, \quad (2.4)$$

where H is some real and symmetric, $2n \times 2n$ matrix and \mathbf{r} is some real, length $2n$ vector.

The close link between the Gaussian subtheory and classical mechanics immediately provides a clash of nomenclature. Both \hat{H} and ΩH are referred to as *Hamiltonian matrices* [34], due to the fact that \hat{H} is the quantum Hamiltonian operator that acts on the Hilbert space, whereas ΩH acts as a representation of a *Hamiltonian vector field* which forms the Lie algebra of the symplectic group [35], which will be extensively explored later. It should be clear in context when reference is to a Hilbert space operator and when to the matrix ΩH .

Definition 2 (Gaussian state). *Any state of the form*

$$\hat{\rho}_G = \frac{e^{-\beta\hat{H}_+}}{\text{Tr}[e^{-\beta\hat{H}_+}]}, \quad (2.5)$$

with $\hat{H}_+ = \frac{1}{2}\hat{\mathbf{r}}^\top H_+\hat{\mathbf{r}} + \hat{\mathbf{r}}^\top \mathbf{b}$, $H_+ > 0$, $\mathbf{b} \in \mathbb{R}^n$, and $\beta > 0$ is referred to as *Gaussian*. This includes the limiting case of $\beta \rightarrow \infty$, which describes the set of pure Gaussian states.

The restriction to a positive-definite H_+ puts the state in the form of a Gibbs distribution when spectrally decomposed, as well as ensuring that the operator is a bona fide trace-class density matrix. Thus, the set of Gaussian states can be described in words as: the ground and thermal states of positive-definite Hamiltonians [33]. Later we will see that this description is equivalent to them having Gaussian Wigner functions.

Note that the only constraint placed on the vector $\hat{\mathbf{r}}$ was the satisfaction of Eq. (2.2). From this we infer that any transformation of $\hat{\mathbf{r}}$ that conserves this relation will also conserve the form of Eq. (2.5), and hence the state will remain Gaussian. The set of real linear transformations that satisfy the commutation relation must be such that,

$$[\hat{\mathbf{r}}, \hat{\mathbf{r}}^\top] = [S\hat{\mathbf{r}} + \mathbf{r}, \hat{\mathbf{r}}^\top S^\top + \mathbf{r}^\top] = iS\Omega S^\top = i\Omega, \quad (2.6)$$

which beckons the introduction of the symplectic group.

Definition 3 (Symplectic Group). *The set of real, $2n \times 2n$ matrices satisfying*

$$S\Omega S^\top = \Omega, \quad (2.7)$$

where Ω is given in Eq. (2.3), is referred to as the n -mode symplectic group and denoted $\text{Sp}(2n, \mathbb{R})$.

The full set of transformations, $\hat{\mathbf{r}} \rightarrow S\hat{\mathbf{r}} + \mathbf{r}$, defines the *inhomogeneous* symplectic group.

Our aim at this stage is to take the general form of a Gaussian state, as given in Def. 2, and decompose it to reveal the structure of this set of states. In order to enact this decomposition it is necessary to define some operators based on this commutator-preserving action of the inhomogeneous symplectic group. We begin with the inhomogeneous part which displaces $\hat{\mathbf{r}}$ by a constant vector. For this we require Weyl operators.

Definition 4 (Weyl operator). *The set of Hilbert space operators of the form*

$$\hat{D}_{\mathbf{r}} := e^{i\mathbf{r}^\top \Omega \hat{\mathbf{r}}}, \quad (2.8)$$

are known as Weyl operators. Note that $\hat{D}_{-\mathbf{r}} \equiv \hat{D}_{\mathbf{r}}^\dagger$.

Before connecting these with displacement, we require a lemma involving the commutator of powers of matrices.

Lemma 5. *Consider three real matrices of the same size, A, B, C . If $[A, B] = C$ and $[B, C] = 0$ then $[A, B^k] = kCB^{k-1}$ for all $k \geq 1$.*

Proof. Suppose $[A, B^k] = kCB^{k-1}$ for some $k \geq 1$. Then

$$\begin{aligned} [A, B^{k+1}] &= AB^{k+1} - B^{k+1}A = AB^{k+1} - B^k AB + B^k AB - B^{k+1}A \\ &= [A, B^k]B + B^k[A, B] = kCB^k + B^kC = (k+1)CB^k. \end{aligned} \quad (2.9)$$

Defining $B^0 = \mathbb{I}$ we know that $[A, B] = CB^0$ and $[B, C] = 0$. The result then follows by induction. \square

Proposition 6. *The action of Weyl operators on $\hat{\mathbf{r}}$ acts as a displacement,*

$$\hat{D}_{\mathbf{r}}\hat{\mathbf{r}}\hat{D}_{\mathbf{r}}^\dagger = \hat{\mathbf{r}} + \mathbf{r}. \quad (2.10)$$

Proof.

$$[\hat{\mathbf{r}}, \hat{D}_{\mathbf{r}}] = \left[\hat{\mathbf{r}}, \sum_{m=0}^{\infty} \frac{(i\mathbf{r}^\top \Omega \hat{\mathbf{r}})^m}{m!} \right] = -\mathbf{r} \sum_{m=1}^{\infty} \frac{(i\mathbf{r}^\top \Omega \hat{\mathbf{r}})^{m-1}}{(m-1)!} = -\mathbf{r}\hat{D}_{\mathbf{r}}, \quad (2.11)$$

by Lemma 5 and using $[\hat{r}_j, is_l \Omega_{lk} \hat{r}_k] = -s_j$, meaning that,

$$\begin{aligned} \hat{D}_{\mathbf{r}}\hat{\mathbf{r}}\hat{D}_{\mathbf{r}}^\dagger &= (\hat{\mathbf{r}}\hat{D}_{\mathbf{r}} + \mathbf{r}\hat{D}_{\mathbf{r}})\hat{D}_{\mathbf{r}}^\dagger \\ &= \hat{\mathbf{r}} + \mathbf{r}, \end{aligned} \quad (2.12)$$

as required. \square

To begin to decompose Eq. (2.5) we note that Proposition 6 implies that

$$\frac{1}{2}(\hat{\mathbf{r}} - \mathbf{r})^\top H(\hat{\mathbf{r}} - \mathbf{r}) = \frac{1}{2}\hat{D}_{\mathbf{r}}^\dagger \hat{\mathbf{r}}^\top H \hat{\mathbf{r}} \hat{D}_{\mathbf{r}}. \quad (2.13)$$

Recalling the meaning of H_+ and \mathbf{b} in Def. 2, and defining

$$\bar{\mathbf{r}} := -H_+^{-1}\mathbf{b}, \quad (2.14)$$

we find that

$$\frac{1}{2}\hat{D}_{\bar{\mathbf{r}}}^\dagger \hat{\mathbf{r}}^\top H_+ \hat{\mathbf{r}} \hat{D}_{\bar{\mathbf{r}}} = \frac{1}{2}(\hat{\mathbf{r}}^\top H_+ \hat{\mathbf{r}} - \bar{\mathbf{r}}^\top H_+ \hat{\mathbf{r}} - \hat{\mathbf{r}}^\top H_+ \bar{\mathbf{r}} + \bar{\mathbf{r}}^\top H_+ \bar{\mathbf{r}}) \quad (2.15)$$

$$= \frac{1}{2}(\hat{\mathbf{r}}^\top H_+ \hat{\mathbf{r}} + \mathbf{b}^\top \hat{\mathbf{r}} + \hat{\mathbf{r}}^\top \mathbf{b} + \mathbf{b}^\top H_+^{-1} \mathbf{b}) \quad (2.16)$$

$$= \frac{1}{2}\hat{\mathbf{r}}^\top H_+ \hat{\mathbf{r}} + \hat{\mathbf{r}}^\top \mathbf{b} + \frac{1}{2}\mathbf{b}^\top H_+^{-1} \mathbf{b}. \quad (2.17)$$

This provides the first dismantling of Eq. (2.5) as

$$\hat{\rho}_G = \hat{D}_{\bar{\mathbf{r}}}^\dagger \frac{e^{-\beta \hat{H}'_+}}{\text{Tr} \left[e^{-\beta \hat{H}'} \right]} \hat{D}_{\bar{\mathbf{r}}}, \quad (2.18)$$

where $\hat{H}'_+ = \frac{1}{2}\hat{\mathbf{r}}^\top H_+ \hat{\mathbf{r}}$. This works because the normalisation of $\hat{\rho}_G$ removes the third term of Eq. (2.17) after exponentiation. The vector $\bar{\mathbf{r}}$ is known as the *vector of first moments* of $\hat{\rho}_G$. To decompose further we use the properties of H_+ and a theorem due to John Williamson.

Theorem 7 (Williamson's theorem [36]). *If M is a $2n$ -dimensional, real, symmetric, positive-*

definite matrix then there exists $S \in \text{Sp}(2n, \mathbb{R})$ such that

$$M = S^\top W S > 0, \quad (2.19)$$

where

$$W = \bigoplus_{i=1}^n \nu_i \mathbb{I}_2, \quad (2.20)$$

where ν_i are referred to as the n symplectic eigenvalues of W . Although S is not unique, the symplectic eigenvalues of M are.

Recalling that $H_+ > 0$, we know from Williamson's theorem that there exists $S \in \text{Sp}(2n, \mathbb{R})$ such that

$$H_+ = S^\top \left(\bigoplus_{i=1}^n \omega_i \mathbb{I}_2 \right) S, \quad (2.21)$$

where ω_i are referred to as the eigenfrequencies of H_+ .

At the level of the Hilbert space, however, we are interested in a slightly larger group structure called the metaplectic group, $\text{Mp}(2n, \mathbb{R})$. This is the double cover of the of the symplectic group and is special in that it has a faithful unitary representation, which the symplectic group does not [37]. Thus we use the metaplectic group representation to discuss the action of the symplectic group on Hilbert space. The relationship between the two groups is not dissimilar to that existing between $\text{SU}(2)$ and $\text{SO}(3)$.

The commutation relation of Eq. (2.1) defines a Lie algebra of operators, known as the Heisenberg algebra [38]. At the Lie group level the Stone-von Neumann theorem states that any two irreducible representations are unitarily equivalent [38]. However, representations at the Lie algebra introduce certain subtleties because the operators are unbounded [39, 40]. Under certain 'integrability' conditions, which will be assumed, the theorem holds at this level. This allows us to posit the existence of unitary operators to connect two vectors satisfying Eq. (2.2). This helps in the following proposition, as outlined in Ref. [37].

Proposition 8. *The transformation of $\hat{\mathbf{r}}$ is such that*

$$\hat{S}^\dagger \hat{\mathbf{r}} \hat{S} = S \hat{\mathbf{r}}, \quad (2.22)$$

where \hat{S} is a unitary representation of the metaplectic group [37] and $S \in \text{Sp}(2n, \mathbb{R})$.

Proof. $\hat{\mathbf{r}}$ provides an Hermitian and irreducible representation of the canonical commutation relation. For $S \in \text{Sp}(2n, \mathbb{R})$, $S \hat{\mathbf{r}}$ also provides an Hermitian and irreducible representation from the property of the symplectic group that it preserves Eq. (2.2), as shown in Eq. (2.6). The Stone-von Neumann theorem then implies that the transition is unitarily implementable. Therefore,

for all $S \in \text{Sp}(2n, \mathbb{R})$ there exists \hat{U}_S acting on the Hilbert space with $S\hat{\mathbf{r}} = \hat{U}_S^\dagger \hat{\mathbf{r}} \hat{U}_S$ such that $U_S^\dagger U_S = \hat{\mathbb{1}}$. \hat{U}_S is arbitrary up to an S -dependent phase factor, i.e.

$$\hat{U}_{S'} \hat{U}_S = e^{i\phi_{S,S'}} \hat{U}_{S'S}. \quad (2.23)$$

$\text{Sp}(2n, \mathbb{R})$ has no faithful unitary representation whereas its universal double cover, the metaplectic group $\text{Mp}(2n, \mathbb{R})$, does [41]. Therefore the phase factor can be reduced to ± 1 . As such, we may choose \hat{U}_S to be given in the unitary representation of the metaplectic group. Operators in this representation will generally be denoted \hat{S} . \square

The set of unitary operators that represent $\text{Mp}(2n, \mathbb{R})$ will be referred to as the group of Gaussian unitaries because under their action, an initial Gaussian state remains Gaussian [29]. Combining Eq. (2.21) and Eq. (2.22),

$$\hat{H}'_+ = \hat{\mathbf{r}}^\top S^\top \left(\bigoplus_{i=1}^n \omega_i \mathbb{I}_2 \right) S \hat{\mathbf{r}} = \hat{S}^\dagger \hat{\mathbf{r}}^\top \left(\bigoplus_{i=1}^n \omega_i \mathbb{I}_2 \right) \hat{\mathbf{r}} \hat{S} = \hat{S}^\dagger \left(\sum_{i=1}^n \hat{H}_i^\mathcal{G} \right) \hat{S}, \quad (2.24)$$

where $\hat{H}_i^\mathcal{G} = \frac{\omega_i}{2} (\hat{x}_i^2 + \hat{p}_i^2)$ is the free Hamiltonian of a single harmonic oscillator. Thus, we further decompose Eq. (2.5) as

$$\hat{\rho}_G = \hat{D}_{\hat{\mathbf{r}}}^\dagger \hat{S}^\dagger \frac{\left(\bigotimes_{i=1}^n e^{-\beta \hat{H}_i^\mathcal{G}} \right)}{\prod_{i=1}^n \text{Tr} \left[e^{-\beta \hat{H}_i^\mathcal{G}} \right]} \hat{S} \hat{D}_{\hat{\mathbf{r}}}, \quad (2.25)$$

using

$$\exp \left(\sum_{i=1}^n \hat{M}_i \right) \equiv \bigotimes_{i=1}^n e^{\hat{M}_i}, \quad (2.26)$$

derivable using a Taylor expansion of the exponential. Eq. (2.25) is sometimes known as a *thermal decomposition* of the Gaussian state because the central element is the thermal state of a free electromagnetic field [29]. One point of interest is that we see that n key pieces of information are contained in the set of values: $\{\beta\omega_i\}$. These provide the core of the state which is then acted on by Gaussian unitaries and Weyl operators, for its construction. Both of these transformations are unable to change these n values and so they have a form of invariance, which will be vitally important later.

Recall that in the definition of Gaussian states we also included the limiting case, $\beta \rightarrow \infty$.

Taking this now in Eq. (2.25) gives

$$\begin{aligned}
\lim_{\beta \rightarrow \infty} \rho_G &= \lim_{\beta \rightarrow \infty} \hat{D}_{\mathbf{r}}^\dagger \hat{S}^\dagger \frac{\left(\bigotimes_{i=1}^n e^{-\beta \hat{H}_i^\mathcal{O}} \right)}{\prod_{i=1}^n \text{Tr} \left[e^{-\beta \hat{H}_i^\mathcal{O}} \right]} \hat{S} \hat{D}_{\mathbf{r}}, \\
&= \lim_{\beta \rightarrow \infty} \hat{D}_{\mathbf{r}}^\dagger \hat{S}^\dagger \frac{\left(\prod_{i=1}^n e^{\frac{1}{2} \beta \omega_i} \right) \left(\bigotimes_{i=1}^n \left(e^{-\frac{1}{2} \beta \omega_j} |0_i\rangle\langle 0_i| + e^{-\frac{3}{2} \beta \omega_i} |1_i\rangle\langle 1_i| + \dots \right) \right)}{\left(\prod_{i=1}^n e^{\frac{1}{2} \beta \omega_i} \right) \prod_{i=1}^n \text{Tr} \left[e^{-\beta \hat{H}_i^\mathcal{O}} \right]} \hat{S} \hat{D}_{\mathbf{r}}, \\
&= \hat{D}_{\mathbf{r}}^\dagger \hat{S}^\dagger |0\rangle\langle 0| \hat{S} \hat{D}_{\mathbf{r}},
\end{aligned} \tag{2.27}$$

where $|0\rangle = \bigotimes_{i=1}^n |0_i\rangle$ is the ground state of the free Hamiltonian $\sum_{i=1}^n \hat{H}_i^\mathcal{O}$, which shows us the form that such pure states will take in this decomposition.

This deconstruction of a Gaussian state into component parts provides us with an idea of the building blocks from which these states are made. To really delve into this it is sensible to move away from the Hilbert space and into the phase space picture originated by Wigner in 1932. In this picture we will see their true Gaussian nature.

2.2 Phase space

The word ‘phase’ has its origin in the Greek *phasis*, meaning ‘appearance’. The link between this word and periodic motion derives from the word being used to describe the ‘phases’, or appearances, of the moon [42]. This clarifies the reason for its ubiquitous use in describing periodic motion. However, it is not at all clear from this, why the word became attached to the phase spaces as found in mechanics, since they have no apparent periodicity.

This use hails back to 1872 when Ludwig Boltzmann was studying the equipartition theory of gas molecules. He noted an analogy between his work and the ‘phase points’ of Lissajous figures, which explicitly deal with periodic motion. Maxwell laid down this phase space structure more formally in 1879 in the modern way, where a state of a system is described by $2n$ position and momentum coordinates. The naming convention for this space varied through the twentieth century, often being referred to as a Γ -space, until the term phase space was eventually settled upon [43].

As shown by Wigner, when the states of a system are constructed out of operators satisfying Eq. (2.1), this enables us to move over to a phase space description. The translation that takes us between the different representations is variously known as the Glauber relation [31], the Fourier-Weyl relation [33] and the Weyl transformation [44]. A proof for the relation can be found in Refs. [33, 45].

Theorem 9 (Fourier-Weyl relation). *A bounded operator \hat{O} acting on a Hilbert space of n bosonic*

modes may be expanded in a basis of Weyl displacement operators as

$$\hat{O} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} d^{2n}\mathbf{r} \operatorname{Tr}[\hat{D}_{\mathbf{r}}\hat{O}]\hat{D}_{\mathbf{r}}^\dagger, \quad (2.28)$$

where $d^{2n}\mathbf{r} \equiv dx_1 dp_1 \dots dx_n dp_n$.

Eq. (2.28) expresses the fact that the Weyl operators form a complete basis for operators on the Hilbert space. Note that this decomposition is possible for any operator that acts on this space, including density matrices, and so

$$\hat{\rho} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^{2n}} d^{2n}\mathbf{r} \operatorname{Tr}[\hat{D}_{\mathbf{r}}^\dagger \hat{\rho}] \hat{D}_{\mathbf{r}}. \quad (2.29)$$

We see that all the information of the state is contained in the coefficients that multiply the Weyl operators, i.e. the continuous function $\operatorname{Tr}[\hat{D}_{\mathbf{r}}^\dagger \hat{\rho}]$ over \mathbf{r} . This provides a new representation of the quantum state.

Definition 10 (Characteristic function). *The characteristic function of a quantum state $\hat{\rho}$, on Fock space, is a complex-valued function,*

$$\chi(\mathbf{r}) = \operatorname{Tr}[\hat{D}_{\mathbf{r}}^\dagger \hat{\rho}], \quad (2.30)$$

where $\hat{D}_{\mathbf{r}}^\dagger$ is a Weyl operator, as per Def. 4.

To extract the Wigner function we take the Fourier transform of the characteristic function [30, 33, 46],

$$\mathcal{W}(\mathbf{r}) := \frac{1}{(2\pi^2)^n} \int_{\mathbb{R}^{2n}} d^{2n}\mathbf{r}' e^{i\mathbf{r}'^\top \Omega \mathbf{r}} \chi(\mathbf{r}'), \quad (2.31)$$

which is a real-valued function on the phase space representing the state that was formerly given in the infinite-dimensional Hilbert space.

If we substitute $\hat{\rho}_G$ of Def. 2 into Eq. (2.30) we will find that the Wigner function representation is itself a Gaussian distribution on phase space, of the form [33],

$$\mathcal{W}_G(\mathbf{r}) = \frac{2^n}{\pi^n \sqrt{\det[\sigma]}} e^{-(\mathbf{r}-\bar{\mathbf{r}})^\top \sigma^{-1} (\mathbf{r}-\bar{\mathbf{r}})}, \quad (2.32)$$

where σ is a $2n \times 2n$ positive-definite matrix. We see that the state is now given as a Gaussian distribution over phase space where σ and $\bar{\mathbf{r}}$ are the two objects necessary for its definition. These two now contain all the information required to describe the state. The infinite dimensions of the Hilbert space has been reduced to a finite vector and matrix to define all the properties of the state, provided it is Gaussian.

Treating the phase space axes as random variables, the Gaussian distribution can be seen as a probability distribution over the points of the phase space. For pure states, this turns out to be peculiar to Gaussian Wigner functions because of Hudson's theorem [47] which states that the only pure state, non-negative Wigner functions are Gaussian. For other pure quantum states the Wigner function is referred to as a *quasi-probability distribution*. It is emphasised in Ref. [48] that this positivity is neither necessary nor sufficient for an underlying classical hidden variable model because the set of measurements and transformations need to be taken into account, as well as the set of states. However, it has recently been shown that in fact the full Gaussian subtheory is entirely equivalent to an epistemically restricted Liouville mechanics [49].

These probability distributions on phase space may be defined by the values of their *moments*. The first moment of a random variable is its mean and the second moment corresponds to its variance. The covariance matrix captures the variance of these variables, as well as the way in which they change with respect to each other, i.e. the covariance.

In general, we consider some random vector $\mathbf{q} = (q_1, \dots, q_p)^\top$ and define the vector of first moments to be

$$\bar{\mathbf{q}} = \langle \mathbf{q} \rangle. \quad (2.33)$$

This definition can be compared with the quantum expression,

$$\bar{\mathbf{r}} = \text{Tr}[\hat{\rho}_G \hat{\mathbf{r}}]. \quad (2.34)$$

$\bar{\mathbf{r}}$ describes the mean position of the Gaussian state in the space - i.e. the location of the peak of the function in the space. As an example we could consider the Gaussian state $|0\rangle\langle 0|$ which, after processing it through the full phase space analysis will turn out to have $\bar{\mathbf{r}} = 0$, i.e. a Wigner function centred at the origin.

If we then act a Weyl operator on the vacuum we describe a new Gaussian state

$$\hat{D}_{\mathbf{r}}^\dagger |0\rangle\langle 0| \hat{D}_{\mathbf{r}} = |\mathbf{r}\rangle\langle \mathbf{r}|, \quad (2.35)$$

which, by substituting into Eq. (2.34) has mean value \mathbf{r} . It is much more common to represent such a state in complex notation. That is, we define $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^\top$ where $\alpha_i = (x_i + ip_i)/\sqrt{2}$, and then denote $|\mathbf{r}\rangle\langle \mathbf{r}|$ as $|\boldsymbol{\alpha}\rangle\langle \boldsymbol{\alpha}|$. This set of *displaced vacua* are exactly the set of coherent states. As mentioned, these states have many interesting properties, describing laser light well and being central to the early understanding of the particle interpretation of wave mechanics.

Classically, the covariance matrix is given as [50]

$$\sigma := \langle (\mathbf{q} - \bar{\mathbf{q}})(\mathbf{q} - \bar{\mathbf{q}})^\top \rangle. \quad (2.36)$$

whereas the quantum version is written as

$$\sigma = \text{Tr}[\{(\hat{\mathbf{r}} - \bar{\mathbf{r}}), (\hat{\mathbf{r}} - \bar{\mathbf{r}})^\top\} \hat{\rho}_G]. \quad (2.37)$$

In both the classical and quantum cases the covariance matrix is positive-definite. For a coherent state, σ is equal to \mathbb{I} but in general it can be any $2n \times 2n$ matrix satisfying

$$\sigma + i\Omega \geq 0. \quad (2.38)$$

This comes out of the formalism and is a direct expression of the uncertainty restriction [51, 52]. This means that the spread of the Gaussian state in phase space, which represents the uncertainty in the variables x_i and p_i , is prevented from being too small.

To gain more intuition behind Eq. (2.38) we consider the single-mode case. Firstly, we may state the form of the covariance matrix as being

$$\sigma = \begin{pmatrix} \text{var}(x) & \text{cov}(x, p) \\ \text{cov}(x, p) & \text{var}(p) \end{pmatrix} = \begin{pmatrix} (\Delta x)^2 & \text{cov}(x, p) \\ \text{cov}(x, p) & (\Delta p)^2 \end{pmatrix}, \quad (2.39)$$

where Δx denotes the standard deviation of the variable x and likewise for p . $\text{cov}(x, p)$ describes the covariance of the two variables which is a measure of how one varies with the other. The uncertainty relation enforces $\sigma + i\Omega$ to be positive-semidefinite which also enforces its determinant to be greater than or equal to zero, providing

$$(\Delta x \Delta p)^2 \geq (\text{cov}(x, p))^2 + 1 \geq 1, \quad (2.40)$$

where we are now able to see the explicit link to the well known, simple Heisenberg relation. Eq. (2.38) is actually more general and corresponds to the Robertson-Schrödinger relation. In Fig. 2.1a we see a distribution in x and p that obeys this relation. In Fig. 2.1b we see that the state can maintain the uncertainty relation by increasing the variance in one of the modes while decreasing it in another. This is a process known as squeezing which we will return to in Sec. 4.

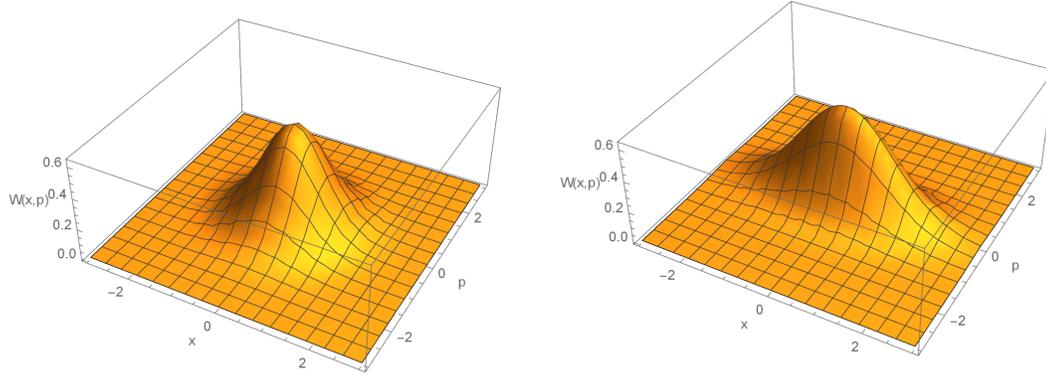
The covariance matrix satisfies the conditions required for its decomposition and so can be written as

$$\sigma = S^\top \left(\bigoplus_{i=1}^n \nu_i \mathbb{I}_2 \right) S, \quad (2.41)$$

where

$$\nu_i := \frac{1 + e^{-\beta\omega_i}}{1 - e^{-\beta\omega_i}}. \quad (2.42)$$

To gain a little more insight into the set of values ν_i we calculate the average value \bar{n}_i of the

(a) Equal variance in both x and p .

(b) Low variance in one direction causing high variance in the other.

Figure 2.1: Single-mode Wigner functions with different variances, but both complying with the Robertson-Schrödinger uncertainty relation.

number operator $\hat{a}_i^\dagger \hat{a}_i$ in a given mode:

$$\bar{n}_i := \text{Tr}[\hat{\rho}_G \hat{a}_i^\dagger \hat{a}_i] = \frac{e^{-\beta\omega_i}}{1 - e^{-\beta\omega_i}}. \quad (2.43)$$

Therefore one finds that ν_i is related to this value via

$$\nu_i = 1 + 2\bar{n}_i, \quad (2.44)$$

providing it with some physical intuition as linearly dependent on the average excitation number in a given mode.

This completes the analysis of the *states* of the Gaussian subtheory of quantum mechanics. We have seen that the phase space description leads itself perfectly to provide a finite description of the degrees of freedom of a Gaussian state. The object that is now studied as ‘the state’ is now the duplet $(\bar{\mathbf{r}}, \sigma)$.

2.3 Evolution

The Gaussian subtheory imposes the restriction that all dynamical processes must conserve the Gaussian nature of the state. Starting with the Heisenberg equation,

$$\dot{\hat{O}} = i[\hat{H}, \hat{O}] + \left(\frac{\partial \hat{O}}{\partial t} \right)_{\hat{H}}, \quad (2.45)$$

we may study the evolution of the vector of operators $\hat{\mathbf{r}}$ under a given Hamiltonian \hat{H} . In anticipation of the final result we restrict this Hamiltonian to being quadratic, as per Def. 1. Note that this is a different Hamiltonian to that which was used to define the state, and so will be written as $\hat{H} = \frac{1}{2}\hat{\mathbf{r}}^\top H \hat{\mathbf{r}} + \hat{\mathbf{r}}^\top \mathbf{c}$. Note that the classical counterpart of these dynamics is integrable.

Given that $\hat{\mathbf{r}}$ has no explicit time dependence and using the Einstein summation convention, the Heisenberg equation for this vector of operators reads,

$$\begin{aligned}\dot{\hat{r}}_i &= i \left[\frac{1}{2} \hat{r}_j H_{jk} \hat{r}_k + \hat{r}_l c_l, \hat{r}_i \right] = \frac{i}{2} H_{jk} (\hat{r}_j [\hat{r}_k, \hat{r}_i] + [\hat{r}_j, \hat{r}_i] \hat{r}_k) + i c_l [\hat{r}_l, \hat{r}_i] \\ &= \Omega_{ik} H_{kj} \hat{r}_j + \Omega_{il} c_l,\end{aligned}\tag{2.46}$$

giving $\dot{\hat{\mathbf{r}}} = \Omega H \hat{\mathbf{r}} + \Omega \mathbf{c}$. Using Eq. (2.46) we may derive the evolution of the first moments, $\bar{\mathbf{r}}$. Calculating $\frac{d}{dt} \text{Tr}[\hat{\mathbf{r}} \hat{\rho}_G]$ immediately provides $\dot{\bar{\mathbf{r}}} = \Omega H \bar{\mathbf{r}} + \Omega \mathbf{c}$, with solution

$$\bar{\mathbf{r}}(t) = e^{\Omega H t} (\bar{\mathbf{r}}(0) - \bar{\mathbf{r}}_{\text{dyn}}) + \bar{\mathbf{r}}_{\text{dyn}}.\tag{2.47}$$

where $\bar{\mathbf{r}}_{\text{dyn}} = -H^{-1} \mathbf{c}$ is set by the dynamics.

The same can be done for σ :

$$\begin{aligned}\dot{\sigma}_{ij} &= \frac{d}{dt} \text{Tr}[\{(\hat{r}_i - \bar{r}_i), (\hat{r}_j - \bar{r}_j)\} \hat{\rho}_G] \\ &= \text{Tr}[(\Omega_{ik} H_{kl} \{(\hat{r}_j - \bar{r}_j), (\hat{r}_l - \bar{r}_l)\} + \Omega_{jk} H_{kl} \{(\hat{r}_i - \bar{r}_i), (\hat{r}_l - \bar{r}_l)\}) \hat{\rho}_G] \\ &\quad + 2c_l \text{Tr}[(\Omega_{il} \hat{r}_j + \Omega_{jl} \hat{r}_i - \Omega_{il} \bar{r}_j - \Omega_{jl} \bar{r}_i) \hat{\rho}_G] \\ &= \Omega_{ik} H_{kl} \sigma_{lj} + \sigma_{il} H_{lk} \Omega_{jk},\end{aligned}\tag{2.48}$$

giving $\dot{\sigma} = \Omega H \sigma + \sigma H \Omega^\top$, which has a solution

$$\sigma(t) = e^{\Omega H t} \sigma(0) (e^{\Omega H t})^\top.\tag{2.49}$$

The matrix $e^{\Omega H t}$ is prominent in the evolution of the first and second moments of a Gaussian state. It is possible to show that it obeys the same symplectic condition as stated before

$$e^{\Omega H t} \Omega (e^{\Omega H t})^\top = \left(\mathbb{I} + \Omega H t + \frac{1}{2} (\Omega H t)^2 + \dots \right) \Omega \left(\mathbb{I} + t H \Omega^\top + \frac{1}{2} (t H \Omega^\top)^2 + \dots \right) = \Omega,\tag{2.50}$$

and so $e^{\Omega H t} \in \text{Sp}(2n, \mathbb{R})$. Given that this evolving element is in $\text{Sp}(2n, \mathbb{R})$ we immediately deduce, from Williamson's theorem, that the symplectic eigenvalues of σ are invariant under closed evolution.

We have seen the real symplectic group crop up on two separate occasions, first to conserve

the form of the canonical commutation relation and now to evolve the vector of first and second moments forwards in time. This shows the deep relationship between preserving the Gaussianity of the state, and preserving the symplectic form, as represented by Ω .

As a preemptive point we can already see some of the key elements of control theory. We have seen that the evolution of the covariance matrix and first moments of Gaussian states involve the action of the symplectic group. Furthermore, we see that these symplectic matrices can be generated by exponentiating ΩH where H links us directly to the quantum Hamiltonian operator. To finally arrive at a notion of control theory, all we require is some human control over H which will be introduced later. A crucial aspect of our investigation will be to explore these links, which will require some mathematical formality. A good place to start is the symplectic group.

Chapter 3

The symplectic group

In the previous section we found that the symplectic group was fundamental to the control of the covariance matrix of Gaussian states. Therefore it is worth spending a little time studying the group itself. The term ‘symplectic’ refers to a wide range of Lie groups over different fields and of different dimensions. The name was coined by Hermann Weyl who originally referred to the ‘complex group’. Eventually he realised that this was likely to cause a not-insignificant amount of confusion and renamed it [53]:

The name “complex group” formerly advocated by me in allusion to line complexes, as these are defined by the vanishing of antisymmetric bilinear forms, has become more and more embarrassing through collision with the word “complex” in the connotation of complex number. I therefore propose to replace it by the corresponding Greek adjective “symplectic.” Dickson calls the group the “Abelian linear group” in homage to Abel who first studied it.

Naming it the Abelian group would have caused an even greater clash and so Weyl’s word stuck. A set of different groups bear this name without a clear consensus on notation. Our particular interest is in the set of real matrices that preserve an anti-symmetric bilinear form, often represented by the matrix Ω , extensively used in the previous section. Even still, a colloquialism will pervade this work in that we will refer to *the* symplectic group, when we really mean the n -mode symplectic group, without specifying n .

The group is a fundamental element in understanding classical Hamiltonian dynamical systems. Given some phase space with coordinate vector $\mathbf{r} = (x_1, p_1, \dots, x_n, p_n)^\top$, the equations of motion are given by

$$\frac{dr_i}{dt} = \Omega \frac{\partial \mathcal{H}}{\partial r_i}, \quad (3.1)$$

where \mathcal{H} is the Hamiltonian of the system. This set of differential equations can be transformed

into a geometrical picture where \mathcal{H} generates a vector field flow on some symplectic manifold. A symplectic manifold is a duplet (\mathcal{M}, ω) where \mathcal{M} is a differentiable manifold and ω is a non-degenerate closed 2-form on \mathcal{M} [36]. A linear Hamiltonian system is one for which the Hamiltonian takes the form [54]

$$\mathcal{H} = \frac{1}{2} \mathbf{r}^\top H \mathbf{r}. \quad (3.2)$$

The lack of linear term reflects the fact that this can be removed via a change of coordinates. Linear Hamiltonian systems evolve on symplectic vector spaces, which arise as tangent spaces to symplectic manifolds [36]. That is, a duplet (\mathbb{V}, ω) where \mathbb{V} is a $2n$ -dimensional vector space and ω is non-degenerate, alternating (or skew-symmetric) bilinear form [54]

$$\omega : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}. \quad (3.3)$$

In our basis, this bilinear form can be represented as a matrix $\omega(\mathbf{u}, \mathbf{v}) \rightarrow \mathbf{u}^\top \Omega \mathbf{v}$, where $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, and is often referred to as the symplectic form. In this context we see the symplectic group as the set of linear transformations that preserve the value of this form. That is to say, transforming the elements of the space such that $\mathbf{u} \rightarrow S\mathbf{u}$ for all $\mathbf{u} \in \mathbb{V}$ preserves the value of $\omega(\cdot, \cdot)$ for any two inputs. The generalised symmetry group for all symplectic manifolds is the set of symplectomorphisms, which form an infinite-dimensional group [36].

In the quantum mechanical commutation relation the symplectic group appears in the same way as it does in systems with classical degrees of freedom. It is via the unitary representation of the metaplectic group that the two notions are brought together to govern dynamics on the Hilbert space. Many results arising in the quantum mechanical case also apply to the classical case, although there is no general equivalence due to quantum phenomena such as the uncertainty principle. This is the element that allows for Gaussian states to describe quantum information theoretic protocols, which are impossible in standard classical mechanics.

3.1 Other bases

As we have seen, Ω is a representation of a more abstract object on some symplectic vector space. Its representation is basis-dependent and so we could consider different forms it could take as a matrix, corresponding to a basis change in the space. There are many different representations that are standardly used, which depend on the problem that is being solved. Here, the main bases will be introduced so that these concepts can be used for solutions later in the text.

The space where the symplectic form is represented by Ω is the space in which the coordinates are in the form $\mathbf{r}_\Omega = (x_1, p_1, \dots, x_n, p_n)^\top$, where we have introduced the subscript on \mathbf{r} for clarity, although this will not be continued in later sections. When \mathbf{r} is in this form we will sometimes refer to the ‘ Ω -basis’ in the text, although Ω is not universally used and so this is not a general

term.

Another very commonly used basis is that for which the coordinates are in the form $\mathbf{r}_J = (x_1, \dots, x_n, p_1, \dots, p_n)^\top$. The transformation between \mathbf{r}_Ω and \mathbf{r}_J is $\mathbf{r}_J = P\mathbf{r}_\Omega$, where

$$P_{kl} = \begin{cases} 1, & k \leq n, l = 2k - 1, \\ 1, & k > n, l = 2(k - n), \\ 0, & \text{otherwise,} \end{cases} \quad (3.4)$$

which is an orthogonal permutation matrix. To find the new representation of the symplectic form in this new basis we return to the place where the form first appeared in the commutation relation, Eq. (2.2). Denoting the vector of operators in the Ω -basis as $\hat{\mathbf{r}}_\Omega$ we derive the new form as:

$$[P\hat{\mathbf{r}}_\Omega, P^\top \hat{\mathbf{r}}_\Omega^\top] = iP\Omega P^\top = i \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix} =: iJ, \quad (3.5)$$

where J is the notation used for the new representation.

A different basis which proves useful later is where we introduce complex numbers. Defining $\alpha_i = \frac{1}{\sqrt{2}}(x_i + ip_i)$, we transform from the basis \mathbf{r}_Ω , as above, to the basis $\mathbf{r}_\Theta = (\alpha_1, \alpha_1^*, \dots, \alpha_n, \alpha_n^*)^\top$ via $\mathbf{r}_\Theta = Q\mathbf{r}_\Omega$, where,

$$Q = \frac{1}{\sqrt{2}} \bigoplus_{j=1}^n \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}, \quad (3.6)$$

and is a unitary matrix. Linking in again with the commutator we find the new representation Θ ,

$$[Q\hat{\mathbf{r}}_\Omega, Q^\dagger \hat{\mathbf{r}}_\Omega^\dagger] = iQ\Omega Q^\dagger = i \bigoplus_{j=1}^n \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} =: i\Theta. \quad (3.7)$$

To see this it is necessary to note that $\hat{\mathbf{r}}_\Omega^\top = \hat{\mathbf{r}}_\Omega^\dagger$.

The final basis we will consider is that written as $\mathbf{r}_{\tilde{\Theta}} = (\alpha_1, \dots, \alpha_n, \alpha_1^*, \dots, \alpha_n^*)$. For this it is easiest to see the transformation from the J -basis via,

$$\tilde{Q} := PQP^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{I} & i\mathbb{I}_n \\ \mathbb{I}_n & -i\mathbb{I}_n \end{pmatrix}, \quad (3.8)$$

and

$$[\tilde{Q}\hat{\mathbf{r}}_J, \tilde{Q}^\dagger \hat{\mathbf{r}}_J^\dagger] = i\tilde{Q}J\tilde{Q}^\dagger = i \begin{pmatrix} -\mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} =: i\tilde{\Theta}. \quad (3.9)$$

We already came across a basis change when we considered $\mathbf{u} \rightarrow S\mathbf{u}$, where S is symplectic and \mathbf{u} is an element of the symplectic vector space. This transformation preserved the bilinear form and so therefore its value for the two input vectors. The basis changes considered here do

not preserve the value of form and so are not related to a symmetry of the system.

3.2 Lie algebra

The relationship between the symplectic Lie algebra and the Lie group is of central importance to the control theory of Gaussian states. The Lie algebra of the symplectic group contains the elements ΩH , where the H represents the Hamiltonian that is directly manipulatable by a controller.

The relationship between the Lie group G and its associated Lie algebra \mathfrak{L} comes via the exponential map [55]

$$\exp : \mathfrak{L} \rightarrow G. \quad (3.10)$$

When dealing with a particular matrix representation of the Lie algebra, this exponential map is the matrix exponential of its elements. Note that this map is neither injective nor surjective in general, although the latter property does hold when the group is compact and connected [56].

In a given representation, the Lie algebra of the symplectic group is a vector space where the elements are matrices X obeying [37]

$$\Omega X + X^\top \Omega = 0, \quad (3.11)$$

where the Lie bracket is the usual matrix commutator. We refer to this vector space as the symplectic Lie algebra and denote it $\mathfrak{sp}(2n, \mathbb{R})$.

Any $X \in \mathfrak{sp}(2n, \mathbb{R})$ exponentiates to an element of the symplectic group. Due to the non-compact nature of the group it is not true that the reverse holds: namely that every Lie group element can be expressed as an exponential of an element of the symplectic Lie algebra. However we will later prove that all symplectic group elements are expressible as $e^X e^Y$ where $X, Y \in \mathfrak{sp}(2n, \mathbb{R})$.

3.3 Properties

To make contact with the pure mathematical literature on the subject it is useful to have clarity when using the wide range of terms employed to describe a system. Appendix A gives a brief overview of some of the categories and properties that are used to describe Lie groups from their group theoretic and geometrical perspectives. Here is provided a general overview of key facts about the symplectic group that aid in the navigation through rougher mathematical waters.

Recall the definition of the symplectic group, in a particular basis, is the set of all real linear $2n \times 2n$ matrices that preserve our representation of a non-degenerate, alternating, bilinear form. Manipulating $S\Omega S^\top = \Omega$ we can show that $S^{-1}\Omega S^{-\top} = \Omega$ and so S^{-1} is also symplectic.

Taking the inverse of both sides of this latter equation provides $S^\top \Omega^{-1} S = \Omega^{-1}$ but since $\Omega^{-1} = \Omega^\top = -\Omega$ this proves that the transpose of a symplectic is also symplectic.

Another key property of the group, especially for the control theorist, is that it is connected [36]. If we consider control theory as directing trajectories on the group manifold then a disconnected group is obviously going to have large inaccessible regions to the flow of the dynamics, depending on where the dynamics begin.

A further crucial property is that it is non-compact, although locally compact. Compactness is a property that provides certain control problems with an easier solution. Specifically it is the unbounded nature of non-compact groups that provide the possibility of a flow on the group being taken out to infinity without visiting other parts of the manifold.

The mathematical literature also makes a clear distinction between Lie groups that are simple and those that are not. The distinction is decided by the nature of the normal subgroups of the group.

Definition 11 (Simple group [57]). *A connected locally compact non-abelian Lie group is simple if it does not have any connected nontrivial normal subgroups.*

The symplectic group does have one nontrivial normal subgroup which is its centre, i.e. the subset of all elements that commute with all other elements, $\mathbb{Z}_2 = \{\mathbb{I}, -\mathbb{I}\}$ [58]. This is not connected and there are no other normal subgroups which makes the symplectic group simple [59]. It is possible to quotient by normal subgroups, and in this case we would obtain the projective symplectic group $\text{PSp}(2n, \mathbb{R})$.

The dimension of the group is the same as the number of degrees of freedom of each symplectic matrix. For the symplectic group this is equal to $n(2n + 1)$, where again n is the number of modes and the matrices are of dimension $2n$ [34].

The eigenvalue structure of the symplectic group and Lie algebra is sufficiently restrictive to make it interesting. Considering $X \in \mathfrak{sp}(2n, \mathbb{R})$ and $S \in \text{Sp}(2n, \mathbb{R})$ we have the following proposition.

Proposition 12. *The characteristic polynomial of X is an even polynomial, therefore if λ is an eigenvalue of a X , then so are $-\lambda, \lambda^*, -\lambda^*$. Furthermore, the characteristic polynomial of a S is a reciprocal polynomial and so if λ is an eigenvalue of a symplectic matrix then so are $\lambda^{-1}, \lambda^*, \lambda^{*-1}$ [54].*

Proof. Recall that $\det[\Omega] = 1$. $p(\lambda) = \det[X - \lambda\mathbb{I}] = \det[\Omega X^\top \Omega - \lambda\mathbb{I}] = \det[\Omega X^\top \Omega + \lambda\Omega\Omega] = \det[\Omega] \det[X + \lambda\mathbb{I}] \det[\Omega] = \det[X + \lambda\Omega] = p(-\lambda)$. Furthermore, using $\det[S] = 1$ we may show that $p(\lambda) = \det[S - \lambda\mathbb{I}] = \det[S^\top - \lambda\mathbb{I}] = \det[-\Omega S^{-1} \Omega - \lambda\mathbb{I}] = \det[-\Omega S^{-1} \Omega + \lambda\Omega\Omega] = \det[-S^{-1} + \lambda\mathbb{I}] = \det[S^{-1}] \det[-\mathbb{I} + \lambda S] = \lambda^{2n} \det[-\lambda^{-1}\mathbb{I} + S] = \lambda^{2n} p(\lambda^{-1})$. \square

The power of the above result is that it allows us to infer and restrict the eigenvalues of a given symplectic matrix, if we have knowledge about some of them. This can greatly diminish the number of values that need to be kept track of in a numerical simulation of control problems.

The Jordan structure of these matrices gives even more information. First, note that if two matrices are similar then they have equal Jordan block structures. We know that for $X \in \mathfrak{sp}(2n, \mathbb{R})$, $X = \Omega^{-1}(-X^\top)\Omega$ and so X and $-X^\top$ are similar. X is always similar to its transpose and to we also know that X and $-X$ are similar [60]. This tells us that the Jordan block structure relating to some eigenvalue λ must be identical in structure to that relating to $-\lambda$. Similar reasoning holds for λ and λ^{-1} as eigenvalues of $S \in \text{Sp}(2n, \mathbb{R})$, using $S^{-1} = \Omega^{-1}S^\top\Omega$.

3.4 Subgroups and decompositions

The subgroups of the symplectic group are crucial in developing a physical understanding of the behaviour of the group, especially in quantum optics. On top of these there are also *subsets* of the group which, although not forming a closed structure under multiplication, provide interest for decompositions.

An important notion in Lie group theory is that of *maximal compact subgroup*. The Cartan-Iwasawa-Malcev theorem states that every compact subgroup of a Lie group G with finitely many connected components is contained in a maximal one and, furthermore, that all maximal compact subgroups of G are conjugate to each other [61]. This statement provides compact subgroups with a notion of maximality. The maximal compact subgroup of $\text{Sp}(2n, \mathbb{R})$ is referred to as the orthogonal symplectic group, which is isomorphic to $U(n)$ [37]. However, as a set of symplectic matrices it is defined as

$$\text{OSp}(2n, \mathbb{R}) := \{S \mid S \in \text{Sp}(2n, \mathbb{R}) \cap O(2n)\}, \quad (3.12)$$

where $O(2n)$ is the orthogonal group of real, $2n \times 2n$ matrices. In the language of quantum optics, which we will explore properly in Sec. 4, these transformations correspond to *passive* or *energy-preserving* operations, or the set of *beam-splitters* and *phase-shifters*. The designation of these transformations as passive, as opposed to active, refers to their preservation of the trace of σ , which corresponds to the average value of the number operator. This is a particular expression of the average energy of the state and so preserving, or not preserving, this value defines these terms.

Another useful subgroup is

$$\mathcal{Z}(2n, \mathbb{R}) := \left\{ \text{diag} \left(\frac{1}{z_1}, z_1, \dots, \frac{1}{z_n}, z_n \right) \mid z \in (0, \infty) \right\}. \quad (3.13)$$

This is clearly a non-compact subgroup due to the infinite range of the elements. In the optics language these matrices are known as *single-mode squeezers* and, in contrast to the former, form a set of active transformations.

An interesting subspace of the algebra comes when we define the subset of symmetric elements. This set exponentiates to a subset of $\mathrm{Sp}(2n, \mathbb{R})$ that consists of the symmetric, positive-definite, symplectic matrices,

$$\mathrm{Sp}_+(2n, \mathbb{R}) := \{S \in \mathrm{Sp}(2n, \mathbb{R}) \mid S^\top = S, S > 0\}. \quad (3.14)$$

Note that this is not a subgroup under the matrix product, because it is not closed and could be used to generate the whole group. As such, we know that the non-compact property of squeezing is contained in this half of the polar decomposition.

Decomposing symplectic matrices into products of matrices is an indispensable tool in working with the group and understanding its behaviour. Some of these decomposition theorems are presented here although a fuller treatment is given in Ref. [37].

Polar decomposition

The polar decomposition of a symplectic matrix splits it into

$$S = RL, \quad \text{s.t. } R \in \mathrm{OSp}(2n, \mathbb{R}), L \in \mathrm{Sp}_+(2n, \mathbb{R}). \quad (3.15)$$

Furthermore, this decomposition is unique up to permutation of the products, i.e. that $S = RL = LR$ [37].

The elements of $\mathrm{OSp}(2n, \mathbb{R})$ are exponentials of the skew-symmetric subset of $\mathfrak{sp}(2n, \mathbb{R})$ and those of $\mathrm{Sp}_+(2n, \mathbb{R})$ are exponentials of symmetric elements. This proves the statement of Sec. 3.2 that any symplectic matrix is expressible as the product of $e^X e^Y$ where $X, Y \in \mathfrak{sp}(2n, \mathbb{R})$.

Furthermore, this decomposition, along with its uniqueness, allows us to make a geometrical statement on top of the algebraic one.

Proposition 13. *The manifold $\mathrm{Sp}(2n, \mathbb{R})$ is diffeomorphic to the Cartesian product of the group $\mathrm{U}(n)$ with a real vector space of dimension $n(n+1)$ [34].*

Proof. $\mathrm{OSp}(2n, \mathbb{R}) \simeq \mathrm{U}(n)$ has manifold structure because it is a Lie group. In Proposition 2.18 of Ref. [36] we see that the exponential map from symmetric elements of $\mathfrak{sp}(2n, \mathbb{R})$ is a diffeomorphism onto $\mathrm{Sp}_+(2n, \mathbb{R})$. Symmetric elements of $\mathfrak{sp}(2n, \mathbb{R})$ form a sub-vector-space of dimension $n(n+1)$. The unique matrix decomposition is therefore equivalent to the manifold decomposition

$$\mathrm{Sp}(2n, \mathbb{R}) \simeq \mathrm{U}(n) \times \mathbb{R}^{n(n+1)}. \quad (3.16)$$

The unique matrix decomposition therefore provides a geometrical decomposition. \square

This decomposition is the first step to gaining some visual intuition for the structure of the symplectic group. Later this result will be used to analyse the set of covariance matrices to develop the geometry of Gaussian states.

Euler decomposition

Referred to as the *Euler decomposition* in Ref. [37] and as *Bloch-Messiah* in Ref. [62], this is essentially the singular value decomposition of standard linear algebra. However, it is not identical in that it also imposes a symplectic nature on the three components and can be expressed as

$$S = R_1 Z R_2, \quad \text{s.t. } R_1, R_2 \in \text{OSp}(2n, \mathbb{R}), \quad Z \in \mathcal{Z}(2n, \mathbb{R}). \quad (3.17)$$

This decomposition is one of the most useful ways to deconstruct a symplectic matrix. For physicists it can be interpreted that any symplectic transformation decomposes into a passive transformation, followed by an active single-mode squeezer, followed by another passive.

The decomposition is not unique, which becomes important in Sec. 8. However, by confining the degrees of freedom of the three matrices, it is possible to enforce uniqueness. This is much like the restriction that forces the complex logarithm to be unique, and just as in this case, we will see branch-cuts occur in the process.

Since the Euler decomposition is merely the singular value decomposition in symplectic guise, the diagonal values of Z still make up the set of singular values of the matrix, which are unique. The singular values come in reciprocal pairs and so all that we require to know are the n values greater than or equal to one, which are denoted $z_i(S)$, or z_i when it is clear which matrix is being referred to. These values provide the set of parameters to measure how active a symplectic transformation is. When $z_i = 1$ for all i , the decomposition collapses so that S is a member of $\text{OSp}(2n, \mathbb{R})$. A useful way to extract these values is by calculating

$$\text{eig}[SS^\top] = \text{eig}[R_1 Z R_2 R_2^\top Z R_1^\top] = \text{eig}[Z^2]. \quad (3.18)$$

Passive decomposition

To extend the Euler decomposition further we would be required to decompose the elements of $\text{OSp}(2n, \mathbb{R})$. In Ref. [63] the authors show that any finite-dimensional unitary operator can be factorised into a sequence of two-dimensional beam-splitter and single-mode phase-shifter transformations, where the isomorphism allows these terms to be used in either representation.

For $n = 1$ the set of matrices that form the group are already in the form of single-mode phase-shifters and so our focus is on the two-mode case as it will form the building blocks for n

modes, if we wished to extend the decomposition.

By using the isomorphism between $\text{OSp}(4, \mathbb{R})$ and $\text{U}(2)$ we are able to use the decomposition theorems for $\text{U}(n)$ to deconstruct two-mode orthogonal symplectics. In the J -representation we may write an orthogonal symplectic matrix as [33]

$$R_J = \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}, \quad (3.19)$$

where $XY^\top - YX^\top = 0_n$ and $XX^\top + YY^\top = \mathbb{I}_n$, derivable from the two condition $SJS^\top = J$ and $S^\top S = \mathbb{I}_{2n}$. From here we may move into the $\tilde{\Theta}$ -basis,

$$R_J \rightarrow R_{\tilde{\Theta}} = \tilde{Q}R_J\tilde{Q}^\dagger = \begin{pmatrix} X - iY & 0_n \\ 0_n & X + iY \end{pmatrix} = \begin{pmatrix} U^* & 0_n \\ 0_n & U \end{pmatrix}, \quad (3.20)$$

making the isomorphism to the unitary group explicit.

The decomposition of an element of $\text{U}(2)$ is given in Ref. [64]:

$$U = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\alpha} \end{pmatrix} \begin{pmatrix} e^{-i\psi/2} & 0 \\ 0 & e^{i\psi/2} \end{pmatrix} \begin{pmatrix} \cos \theta/2 & \sin \theta/2 \\ -\sin \theta/2 & \cos \theta/2 \end{pmatrix} \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}, \quad (3.21)$$

where $\alpha, \psi, \theta, \phi \in \mathbb{R}$. Decomposing the elements of $R_{\tilde{\Theta}}$ and converting back to the Ω -basis we derive the decomposition of a general element $R_\Omega \in \text{OSp}(4, \mathbb{R})$ in the Ω -basis as

$$R_\Omega = \begin{pmatrix} R_{-\alpha} & 0 \\ 0 & R_{-\alpha} \end{pmatrix} \begin{pmatrix} R_{\psi/2} & 0 \\ 0 & R_{-\psi/2} \end{pmatrix} \begin{pmatrix} \cos \theta/2 & 0 & \sin \theta/2 & 0 \\ 0 & \cos \theta/2 & 0 & \sin \theta/2 \\ -\sin \theta/2 & 0 & \cos \theta/2 & 0 \\ 0 & -\sin \theta/2 & 0 & \cos \theta/2 \end{pmatrix} \begin{pmatrix} R_{\phi/2} & 0 \\ 0 & R_{-\phi/2} \end{pmatrix}, \quad (3.22)$$

where

$$R_x := \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}, \quad (3.23)$$

for some real x . R_x is a matrix that acts on a single mode and represents a phase-shifter in the quantum optics literature. The matrix involving θ as a parameter represents a beam-splitter [30].

For the n -mode case we would use Ref. [63] to decompose the n -mode orthogonal symplectic into operations acting locally on two modes at a time. These operations could then be decomposed using Eq. (3.22). This decomposition allows us to write down elements of the orthogonal symplectic group and are especially important when writing numerical simulations of two-mode control problems.

The form of a general orthogonal symplectic matrix in Eq. (3.20) allows us to prove another

important property of the symplectic group. From the preservation of the bilinear form we can deduce that

$$\det[S\Omega S^\top] = \det[S] \det[\Omega] \det[S] = \det[S]^2 = \det[\Omega] = 1. \quad (3.24)$$

Thus we know that $\det[S] = \pm 1$. The polar decomposition provides S as a product of $S = RL$. The determinant of the positive-definite matrix L is positive and so the next stage is to consider the determinant of the orthogonal R . For Eq. (3.20) we see that $\det[R] = \det[U]^2 = 1$, and so $\det[S] = 1$ for all $S \in \text{Sp}(2n, \mathbb{R})$.

3.5 Single-shot symplectics

In Sec. 3.2 it was stated that the elements of $\text{Sp}(2n, \mathbb{R})$ are not always expressible as e^X where $X \in \mathfrak{sp}(2n, \mathbb{R})$. The elements that *are* expressible in this form will be referred to as *single-shot* symplectics. Another way to express this is that such elements have a real Hamiltonian logarithm, because $X = \Omega H$ is a Hamiltonian matrix. In terms of control theory, a symplectic that is not single-shot is one for which no constant Hamiltonian exists to enact it. Thus, all the two-shot symplectics require the controller to alter the Hamiltonian in time to be able to enact that transformation.

To develop conditions to decide whether a symplectic matrix is single-shot, we require some knowledge of the different normal forms of a symplectic matrix. A complete necessary and sufficient condition in n modes is still an open question but is partly addressed in Ref. [65].

To state this result we are required to explore some normal forms of symplectic matrices. These allow a categorisation of different symplectic transformations, which allows a systematic path for proofs to take.

Normal forms of symplectic matrices go back to the early work of John Williamson, of the aforementioned Williamson's theorem [66, 67]. Extracted from these early papers, normal forms were written down by D. M. Galin providing them in Ref. [68]. After correcting a few mistakes these normal forms made their way into an appendix of Vladimir I. Arnol'd's *Mathematical Methods of Classical Mechanics* [69, 70]. Other normal forms can be found, for example in Ref. [70], as well as those derived by Alan J. Laub and Kenneth Meyer in Ref. [71]. The normal forms of Laub and Meyer provide the basis for work on conditions for 'shottability' of group elements. That is to say, they provide conditions for whether or not a symplectic matrix is single-shot.

Working in the J -basis, the authors prove that all symplectic matrices S are similar, via a symplectic matrix T , to a matrix V such that

$$V = T^{-1}ST = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (3.25)$$

where

$$\begin{aligned} A &= \text{diag}(A_1, \dots, A_p), & B &= \text{diag}(B_1, \dots, B_p), \\ C &= \text{diag}(C_1, \dots, C_p), & D &= \text{diag}(D_1, \dots, D_p). \end{aligned} \quad (3.26)$$

The $2k \times 2k$ matrices,

$$Y_j = \begin{pmatrix} A_j & B_j \\ C_j & D_j \end{pmatrix}, \quad (3.27)$$

are analogous to Jordan blocks but where we have managed to preserve the symplectic nature of V . A_j, B_j, C_j and D_j are all real, $k \times k$ matrices. As in the case of standard Jordan blocks, each corresponds to a single eigenvalue λ . By categorising the different types of ‘Laub-Meyer’ block, one provides a categorisation of all $2n \times 2n$ symplectic matrices, due to the uniqueness of the decomposition.

The different forms that Y_j can take are written in Ref. [71], to which the reader is referred for more detail. Our interest is restricted to some particular cases and so only these will be reproduced. In the question of shottability we will see that it is the negative real eigenvalues that cause the obstacles to a necessary and sufficient condition. Symplectic matrices that have negative eigenvalues have Laub-Meyer blocks in one of the following forms, corresponding to the eigenvalues λ [71]:

$$(a) \lambda < 0, \lambda \neq -1: Y = \begin{pmatrix} E & 0_k \\ 0_k & E^{-\tau} \end{pmatrix}.$$

$$(b) \lambda = -1: (i) Y = -\mathbb{I}, \quad (ii) Y = \begin{pmatrix} F & 0_k \\ 0_k & F^{-\tau} \end{pmatrix}, \quad (iii) Y = \begin{pmatrix} F & 0_k \\ G & F^{-\tau} \end{pmatrix},$$

where

$$E = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}, \quad E^{-\tau} = \begin{pmatrix} \lambda^{-1} & -\lambda^{-2} & \cdots & (-1)^{k+1} \lambda^{-k} \\ 0 & \lambda^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\lambda^{-2} \\ 0 & \cdots & 0 & \lambda^{-1} \end{pmatrix}, \quad (3.28)$$

and

$$F = - \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 2 & \cdots & 2 & 1 \end{pmatrix}, \quad F^{-\tau} = - \begin{pmatrix} 1 & -2 & \cdots & (-1)^{k+1} 2 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -2 \\ 0 & \cdots & 0 & 1 \end{pmatrix}, \quad (3.29)$$

and

$$G = - \begin{pmatrix} -2 & \cdots & -2 \\ 2 & \cdots & 2 \\ \vdots & \vdots & \vdots \\ (-1)^k 2 & \cdots & (-1)^k 2 \end{pmatrix}. \quad (3.30)$$

Anyone wishing to enact some symplectic transformation may wish to know whether it is possible in a single shot.

Theorem 14 (Conditions for a real Hamiltonian logarithm [65]). *Let a symplectic matrix $S \in \mathbb{R}^{2n \times 2n}$ be given.*

1. *Suppose $-1 \notin \text{eig}(S)$. Then there exists $X \in \mathfrak{sp}(2n, \mathbb{R})$ such that $e^X = S$ if and only if S has an even number of canonical blocks of type (a) of each size relative to every negative eigenvalue.*
2. *Suppose that $-1 \in \text{eig}(S)$, and that relative to the other negative eigenvalues, condition 1 is satisfied. S has a logarithm in $\mathfrak{sp}(2n, \mathbb{R})$ if, relative to -1 , there are only*
 - *blocks of type (b)-(i),*
 - *blocks of type (b)-(ii) with k odd,*
 - *an even number of blocks of each size of type (b)-(iii) or (b)-(ii) with k even.*

with reference to the normal forms provided earlier.

These conditions are only partially necessary and sufficient and so they will only allow us to isolate the types of symplectic matrix that will not have a logarithm in the symplectic Lie algebra, although examples will exist where S does not satisfy these conditions and still have such a logarithm. Furthermore, we see that these conditions isolate the negative eigenvalues as the root of the breakdown of single-shottability since the blocks of type (a) and (b) only refer to these, ignoring the blocks containing complex and imaginary eigenvalues. The significance for control of Gaussian states will be explored in Sec. 5 which attempts an extension of Williamson's theorem to single-shot symplectics, rather than just symplectics.

To cement ideas a little, we can look at an example in a single mode,

$$S_1 = \begin{pmatrix} -z & 0 \\ 0 & -\frac{1}{z} \end{pmatrix} = \exp \begin{pmatrix} 0 & \pi \\ -\pi & 0 \end{pmatrix} \exp \begin{pmatrix} \log[z] & 0 \\ 0 & -\log[z] \end{pmatrix}. \quad (3.31)$$

This has two blocks of type (a) but for two different eigenvalues, and so Theorem 14 states that this is not single-shot. However, we see a valid expression as a product of two single-shot symplectic matrices.

This shottability structure is able to arise because the symplectic group is non-compact. As mentioned before, all connected compact groups have a surjective exponential map. In Sec. 2 we saw that the evolution of a Gaussian state was encoded in H and enacted through $e^{\Omega H t}$.

This however ignores the fact that the Williamson decomposition is not unique and so there is an equivalence class of symplectic matrices which enact the same transformation on a given covariance matrix. We can ask whether an element of the equivalence class is single-shot which is what we return to in Sec. 5. Before we can do this, however, we will be required to understand the properties of the set of covariance matrices that determine the shape of the Gaussian Wigner function on phase space.

Chapter 4

The covariance matrix

Gaussian states are totally specified by the duplet $(\bar{\mathbf{r}}, \sigma)$ which encodes the location and spread of the distribution respectively. The covariance matrix expresses the variance and covariance of the $2n$ random variables of phase space over which some distribution is defined. Recall that coherent states had a trivial covariance matrix and so were totally specified by $\bar{\mathbf{r}}$. However, as far as quantum information protocols are concerned, it is the covariance matrix that encodes the interesting information because it contains all entropic and entanglement properties of the state [31]. To begin our quantum analysis of the covariance matrix, we first need to learn about its properties from the classical theory.

Since σ is positive-definite it is possible to diagonalise using an orthogonal matrix O ,

$$\sigma = ODO^\top. \quad (4.1)$$

The eigenvector associated with the diagonal element D_{ii} can be expressed as $O\mathbf{e}_i$ where \mathbf{e}_i is a vector of zeroes except with a 1 in the i th position. The eigenvalues of σ are real and so we may order them by size, and hence also order the eigenvectors by association, with obvious ambiguity for degeneracies.

The diagonal elements of the covariance matrix correspond to the variance of that particular random variable and the eigenvectors $O\mathbf{e}_i$ correspond to the direction along which this variance is calculated. We find that the maximal element of D is greater than any diagonal element in any basis by noting that the diagonal elements of sigma are of the form,

$$\sigma_{ii} = \sum_j O_{ij}^2 D_{jj}. \quad (4.2)$$

The orthogonality of O implies that

$$\sum_j O_{ij}O_{kj} = \delta_{ik} \implies \sum_j O_{ij}^2 = 1 \quad \forall i. \quad (4.3)$$

This implies that the sum of Eq. (4.2) is convex and hence $D_{jj} \geq \sigma_{ii}$ for all i if D_{jj} is maximal. Similar reasoning show that the minimum eigenvalue in the diagonal basis corresponds to the minimum variance. As such we can see diagonalisation as transforming to a basis in which we capture the direction of greatest variance as a basis vector.

The eigenvector with the greatest eigenvalue is known as the *first principal component* of the distribution and the ray along its direction is known as the *first principal axis*, in principal component analysis. The second eigenvalue of the covariance matrix provides the maximal variance for the directions orthogonal to the direction of the first principal axis. This provides the second principal component and axis. The eigenvalues of the covariance matrix therefore provide the shape of the Gaussian distribution, even when dealing in higher modes where a direct visualisation is not possible.

The quantum covariance matrix has all of these properties, except that it is also bounded by the Robertson-Schrödinger uncertainty relation as given in Eq. (2.38). This, it will be shown, corresponds to a lower bound on the set of symplectic eigenvalues.

Now that we have unpicked the meaning of the covariance matrix and its elements, eigenvalues and eigenvectors, we are able to go deeper into its structure. A good way of seeing this is to use decomposition theorems to rewrite it in different ways. The deconstructions we focus on have physical meaning using the language of quantum optics. The elemental pieces of its structure are the key properties that are used to provide new insights in the latter sections of this work.

4.1 Decomposition and categorisation

The covariance matrix, as we have already seen, is positive-definite, symmetric and real and so the decompositions of previous sections can be applied. Each element of this decomposition will gradually be unpicked to reveal a complete optical description.

The first decomposition that applies is Williamson's, as given in Theorem 7, breaking the covariance matrix into

$$\sigma = SWST, \quad (4.4)$$

where $S \in \text{Sp}(2n, \mathbb{R})$ and $W = \text{diag}(\nu_1, \nu_1, \dots, \nu_n, \nu_n)$, i.e. W denotes the diagonal matrix of symplectic eigenvalues. The set of symplectic eigenvalues contained in this particular normal form are often referred to as *the* symplectic eigenvalues of the Gaussian state, even though we have seen other symplectic eigenvalues arise in Sec. 2, related to H . This decomposition allows

us to imagine every covariance matrix as being an evolved form of W , under some Hamiltonian dynamics, using Sec. 2.3.

This allows us to peel back the meaning behind the uncertainty relation even more, as stated in Ineq. (2.38) as $\sigma + i\Omega \geq 0$. Given that this inequality is invariant under the action of the symplectic group by conjugation, it is equivalent to the inequality

$$W + i\Omega \equiv \bigoplus_{i=1}^n \begin{pmatrix} \nu_i & 1 \\ -1 & \nu_i \end{pmatrix} \geq 0, \quad (4.5)$$

which is in turn equivalent to

$$\nu_i \geq 1, \quad \forall i. \quad (4.6)$$

The uncertainty relation is therefore expressible as a lower bound on the symplectic eigenvalues.

The Euler decomposition of this symplectic matrix provides a notion of what this evolution could have looked like. Expanding S we may write

$$\sigma = R_1 Z R_2 W R_2^T Z^T R_1^T. \quad (4.7)$$

The properties of W and S allow us to categorise the different types of Gaussian state.

The state with covariance matrix $\sigma = W$ is a thermal state of the free Hamiltonian $\sum_{i=1}^n \hat{H}_i^{\mathcal{G}}$, as shown in Sec. 2.2. Such a state is referred to as a ‘chaotic field’ in Ref. [20] but also as the thermal state of a free field. This Hamiltonian describes a set of decoupled harmonic oscillators, each with a different value for $\beta\omega_i$ which is then repackaged in the symplectic eigenvalues ν_i . These states bear a resemblance to the states first looked at by Planck at the beginning of quantum theory and are used to describe blackbody radiation [20].

As we have seen, these symplectic eigenvalues are bounded from below to be greater than or equal to one. If $\sigma = \mathbb{I}$ then this is known as the *vacuum state*, recalling as ever that the vacuum state is shorthand for the ground state of some given Hamiltonian. Here we are implicitly referring to the ground state of the free Hamiltonian.

In Eq. (4.7) we see that the first symplectic element to act is $R_2 \in \text{OSp}(2n, \mathbb{R})$. This subgroup of symplectic transformations is the set that preserves the average energy of the free Hamiltonian. To see this we first expand Eq. (2.37) as

$$\begin{aligned} \sigma_{ij} &= \text{Tr}[(\hat{r}_i \hat{r}_j + \hat{r}_j \hat{r}_i + 2\bar{r}_i \bar{r}_j - 2\hat{r}_i \bar{r}_j - 2\hat{r}_j \bar{r}_i) \hat{\rho}_G] \\ &= \text{Tr}[\{\hat{r}_i, \hat{r}_j\} \hat{\rho}_G] - 2\bar{r}_i \bar{r}_j. \end{aligned} \quad (4.8)$$

Given some quadratic Hamiltonian $\hat{H} = \frac{1}{2} \hat{\mathbf{r}}^T H \hat{\mathbf{r}} + \hat{\mathbf{r}}^T \mathbf{r}$ we can derive the average energy of some

Gaussian state $\hat{\rho}_G$ using Eq. (2.37) to get

$$\text{Tr}[\hat{\rho}_G \hat{H}] = \sum_{ij} H_{ij} \text{Tr}[\{\hat{r}_i, \hat{r}_j\} \hat{\rho}_G] + \text{Tr}[\hat{\rho}_G \hat{r}_i] r_i = H_{ij} \sigma_{ij} + 2H_{ij} \bar{r}_j \bar{r}_i + \bar{r}_i r_i. \quad (4.9)$$

This means that $\text{Tr}[\hat{\rho}_G \hat{H}] = \text{Tr}[\sigma H] + 2\bar{\mathbf{r}}^T H \bar{\mathbf{r}} + \bar{\mathbf{r}}^T \mathbf{r}$. Setting $H = \mathbb{I}$ and $\mathbf{r} = \mathbf{0}$, we obtain the Hamiltonian of the free field. A passive transformation is one that preserves this average energy, and thus preserves $\text{Tr}[\sigma]$. The orthogonal symplectic group is precisely this set of symmetries. In quantum optics these transformations consist of the beam-splitters and phase-shifters in the lab.

In order to not preserve the average energy we must move to the next stage in the decomposition with the Z element. This has already been referred to as a direct sum of single-mode squeezers. We can see the reason for this naming if we consider the affect that Z has on the principal components of the distribution.

Consider a single-mode state with degenerate symplectic eigenvalues, $\sigma = \nu \mathbb{I}$. The maximum eigenvalue of σ provides the maximum variance of the distribution that this describes, i.e. ν . We see a degeneracy here in that the second eigenvalue is also equal to ν , which tells us that the Gaussian state has n -spherical symmetry on a given set of axes, much like the vacuum Wigner function as pictured in Fig. 2.1a, but broader because the variance is larger. If we apply R_2 to this state then it will leave it invariant, in the same way that it does the vacuum. Then applying $Z = \text{diag}(z, 1/z)$ we find that $\max \text{eig}[\sigma] = z^2 \nu$ where $z \geq 1$ without loss of generality. Thus the variance in one principal axis has increased whereas the second, ν/z^2 has decreased. This is an operation known as squeezing and is an active transformation because it does not preserve $\text{Tr}[\sigma]$ [29]. Since the z_i values of S are the singular values of the matrix, then we can define a squeezing transformation as a symplectic that has non-unit singular values. The operation of squeezing is bound by the uncertainty principle, Eq. (2.38). The variance in one direction is allowed to decrease but only when the variance in another direction increases. Squeezing is a technologically useful and desirable feature that is sought in many practical setups [31]. Since the fashion of resource theories in quantum information theory, squeezing has been recast as a limited resource, mirroring the difficulty with which it is produced [62].

Although the definition of a squeezing operation is settled, the definition of a squeezed state is varied in the literature. If we restricted to coherent states, and by ignoring first moments therefore vacuum states, the uncertainty relation is saturated and the variance in all directions is one. Any evolution that involves some S with non-unit singular values will reduce the variance in one direction to be less than unity. Due to the utility of the *sub-vacuum fluctuations* exhibited by this state, a squeezed state is often defined as any state with an eigenvalue less than one [72]. Such states, since their experimental realisation in 1985 [20], have received much attention and it is hoped that they will play a greater role in gravitational wave physics to improve detection [28].

$S \setminus W$	$W = \mathbb{I}$	$W \neq \mathbb{I}$
$z_i(S) = 1 \ \forall i$	Vacuum state	Thermal state of free field
$\exists z_i(S) \neq 1$ for some i	Squeezed vacuum state	Squeezed state

Table 4.1: A categorisation of the different types of covariance matrix which can describe a Gaussian state.

A particular subset of squeezed states called ‘twin-beam’ states are maximally entangled in the Fock basis and are used in continuous variable quantum teleportation protocols [31].

However, we can begin with states that are not coherent and so squeezing the state does not reduce any of its variances to less than unity. In Ref. [73] there is discussion of ‘squeezed Fock states’ and ‘squeezed thermal states’ for example. The proceeding definition follows this example.

Definition 15 (Squeezed Gaussian state). *A squeezed Gaussian state is one for which the Williamson decomposition of its covariance matrix, $\sigma = SWS^\top$, is such that S has non-unit singular values.*

The reason for this definition in the context of this thesis is because we are interested in the behaviour of squeezing in terms of control theory, rather than the threshold at which squeezing becomes a useful quantity.

It has been noted that the Williamson decomposition of a matrix is not unique. Therefore it would seem that categorising based on the properties of S could fail. However, in Sec. 5 it will be shown that the singular values of the diagonalising symplectic are unique. Thus we can unambiguously refer to states with $\sigma = SWS^\top$ as unsqueezed when the singular values of S are unity and squeezed when they are not. We summarise the different types of commonly used Gaussian in Table 4.1.

4.2 Squeezing measure

We have found that the eigenvalues of σ provide the variance along the principal axes. This set of quantities is much easier to find but it leaves the problem of comparison, due to the fact that we are dealing with a set and not a single number. We may provide a crude measure of squeezing by focusing on the minimum eigenvalue of σ as this provides the minimum variance along any given direction. This value is also known as the squeezing variance of the state [74]. To turn this into a squeezing measure that monotonically increases as this variance decreases, and is zero

when this the state is unsqueezed, we define

$$\xi[\sigma] := \frac{1}{\min \text{eig}[\sigma]} - \frac{1}{\min |\text{eig}[\Omega\sigma]|}. \quad (4.10)$$

To understand this measure we notice that the width of a Gaussian state is dictated both by the symplectic eigenvalues of the state, as well as the singular values of the symplectic matrix when the covariance matrix is put in Williamson normal form. In order for this squeezing measure to be zero when all the singular values are zero, it is necessary to subtract off the contribution made by the symplectic eigenvalues. This positive measure allows us to track the maximal extent of the squeeze of a state and to follow its trajectory as it undergoes some varying symplectic control.

In other situations, however, we may like a measure for the squeezing inherent in a particular *symplectic matrix* rather than the covariance matrix on which it has acted. A fine-grained knowledge would come from knowing the singular value decomposition and the set of singular values. For each of the modes, a natural measure that will arise later is

$$\zeta_z := \frac{z^2 + 1/z^2}{2}, \quad (4.11)$$

which monotonically increases with z .

4.3 Symplectic invariants

The set of symplectic eigenvalues $\{\nu_i\}$ arose from the Williamson decomposition of the covariance matrix. Due to their uniqueness it is the case that no symplectic matrix acting by congruence on the covariance matrix will alter its symplectic eigenvalues, and as such they are called *symplectically invariant*.

Although Williamson's theorem can be proved in a constructive manner, there are better ways of finding the set of symplectic eigenvalues. A faster way is to notice that

$$\text{eig}[\Omega\sigma] = \text{eig}[\Omega SW S^\top] = \text{eig}[S^\top \Omega S W] = \text{eig}[\Omega W] = \{\pm i\nu_i\}, \quad (4.12)$$

and so we just seek the modulus of the eigenvalues of the matrix $\Omega\sigma$.

The importance of symplectic invariants is that they will be preserved under closed Gaussian evolution. This immediately implies that this set should uniquely define the entropy of a given Gaussian state, as this is also invariant under such evolution. Two common ways of expressing the entropy of a state are the linear and von Neumann entropies. The von Neumann entropy bears a formal resemblance to the Shannon entropy from classical information theory and is

defined as

$$S_V := -\text{Tr}[\hat{\rho}_G \ln \hat{\rho}_G]. \quad (4.13)$$

In terms of symplectic invariants this takes the form [33]

$$S_V = \sum f(\nu_i), \quad (4.14)$$

where

$$f(x) \equiv \left(x + \frac{1}{2}\right) \ln \left(x + \frac{1}{2}\right) - \left(x - \frac{1}{2}\right) \ln \left(x - \frac{1}{2}\right). \quad (4.15)$$

The linear entropy is a first order approximation to this and is defined as

$$S_L := 1 - \text{Tr}[\hat{\rho}^2], \quad (4.16)$$

which takes the following form in terms of symplectic invariants [33]

$$S_L = 1 - \frac{1}{\sqrt{\det[\sigma]}} = 1 - \frac{1}{\prod_{i=1}^n \nu_i^2}. \quad (4.17)$$

Any measure of entropy for a Gaussian state need only be a function of the n symplectically invariant parameters that define the state.

In Eq. (2.42) we saw that the symplectic eigenvalues were related to the values $\beta\omega_i$ via

$$\nu_i = \frac{1 + e^{-\beta\omega_i}}{1 - e^{-\beta\omega_i}},$$

where the temperature T of the state is related to β by $\beta = 1/(k_B T)$ and where k_B is Boltzmann's constant. This in turn was related to the average mode excitation number \bar{n}_i in Eq. (2.44). We see here that the inverse temperature of the state only ever appears in a product with the mode frequency ω_i and so there is not a unique notion of the temperature of the state without fixing the mode frequencies. Given that ν_i is a monotonically increasing function in $\beta\omega_i$ we can describe a rise in symplectic eigenvalue as an increase in the temperature of a mode, given the implicit assumption that the mode frequency is kept fixed. Decreasing ν_i to its minimal value of one in turn decreases \bar{n}_i to zero. The language of heating and cooling of a state will be extensively used in Sec. 10.

The n symplectic eigenvalues have precedence only by virtue of their being the set of values produced by Williamson decomposition. One could write down any function of these and derive a new set of values which would also be invariant under symplectic transformations. For example, considering the *elementary symmetric functions* of a matrix also provides a set of n values. Given

a set of eigenvalues λ_j of an $m \times m$ matrix M , these are written as [75]

$$\vartheta_k[M] := \sum_{H \in \mathcal{F}_k^m} \prod_{j \in H} \lambda_j, \quad (4.18)$$

where k labels the n functions and

$$H \in \mathcal{F}_k^m \subset P(\mathbb{N}_m) \quad \text{iff} \quad |H| = k, \quad (4.19)$$

where $\mathbb{N}_m = \{1, \dots, m\}$, $P(\cdot)$ denotes the power set and $\vartheta_0[M] := 1$. As an example, for a matrix M with eigenvalues $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$,

$$\vartheta_3[M] = \lambda_1 \lambda_2 \lambda_3 + \lambda_1 \lambda_3 \lambda_4 + \lambda_1 \lambda_2 \lambda_4 + \lambda_2 \lambda_3 \lambda_4. \quad (4.20)$$

Given that the symplectic eigenvalues can be obtained as the modulus of the eigenvalues of $\Omega\sigma$ we see that it is possible to define a new set of symplectic invariants [76]

$$\vartheta_{2i}[\Omega\sigma] = \sum_{H \in \mathcal{F}_i^n} \prod_{j \in H} \nu_j^2. \quad (4.21)$$

This set of symplectic invariants will reappear when we explore multimode entropy in the context of heating and cooling in Sec. 9.

The uncertainty relation expressed before as a lower bound on the symplectic eigenvalues has now become mixed up in these new invariants. In order to rewrite the relation using these new objects we define the function,

$$\Sigma := \sum_{i=0}^n (-1)^{n+i} \vartheta_{2i}, \quad (4.22)$$

which can be shown to be equal to $\Sigma = \prod_{i=1}^n (\nu_i^2 - 1)$, which is easiest to see by working backwards

$$\begin{aligned} \prod_{i=1}^n (\nu_i^2 - 1) &= \prod_{j=1}^n \nu_j^2 - \sum_{H \in \mathcal{F}_{n-1}^n} \prod_{j \in H} \nu_j^2 + \sum_{H \in \mathcal{F}_{n-2}^n} \prod_{j \in H} \nu_j^2 - \dots \\ &= \sum_{i=0}^n (-1)^{n+i} \sum_{H \in \mathcal{F}_i^n} \prod_{j \in H} \nu_j^2 \\ &= \sum_{i=0}^n (-1)^{n+i} \vartheta_{2i}. \end{aligned} \quad (4.23)$$

From this we see that a necessary condition for the satisfaction of Ineq. (4.6) is that

$$\Sigma \geq 0. \quad (4.24)$$

From Eq. (4.23) we see that if an even number of symplectic eigenvalues violate Ineq. (4.6) then Ineq. (4.24) is not going to detect it. A second situation in which this will not be detected is when we have *partial saturation*, i.e. when $\nu_j = 1$ for some j . This is remedied in Ref. [76] by iteratively checking a formula that detects such partial saturation. Such states have some interesting properties, for example in two modes they turn out to be the states with minimal entanglement for a given purity [77, 78].

4.4 Entanglement

Entangled systems are those for which a description of the combined system is not given by an individual description of each component part. That is, the number of variables increases more rapidly than the number of components multiplied by the individual degrees of freedom. The occurrence of such states in quantum mechanics has been recognised as one of the key features of the theory since the famous Einstein, Podolsky, Rosen (EPR) paper of 1935 [79]. Closely related to the crucial property of nonlocality, detecting and quantifying entanglement has been a major focus of quantum theoretical research.

A well known necessary condition for the separability of bipartite states is the ‘positivity of the partial transpose’ (PPT) criterion [80]. This states that a separable bipartite quantum state has non-negative eigenvalues after transposing one of the systems. This was applied to continuous variable systems in Refs. [81, 82]. First we consider a Gaussian state of $(p+q)$ modes, where the first set of p modes is referred to as subsystem A and the latter set of q modes as subsystem B . To enact the partial transposition on subsystem A we transform

$$\sigma \rightarrow T\sigma T =: \tilde{\sigma}, \quad (4.25)$$

where $T = \bigoplus_1^p \text{diag}(1, -1) \oplus \bigoplus_1^q \text{diag}(1, 1)$, which can be shown by returning to the definition of the covariance matrix from the quantum state. It is very quick to see that if σ is separable then this map take us to an equally valid quantum covariance matrix, as intended by partial transposition. The PPT condition is equivalent to

$$\tilde{\sigma} + i\Omega \geq 0. \quad (4.26)$$

T acts by similarity and is symmetric meaning that $\tilde{\sigma}$ remains positive-definite, implying that it also has a Williamson decomposition: $\tilde{\sigma} = \tilde{S}\tilde{W}\tilde{S}^\dagger$. Ineq. (4.26) is in turn equivalent to

$$\tilde{\nu}_j \geq 1, \quad \forall j. \quad (4.27)$$

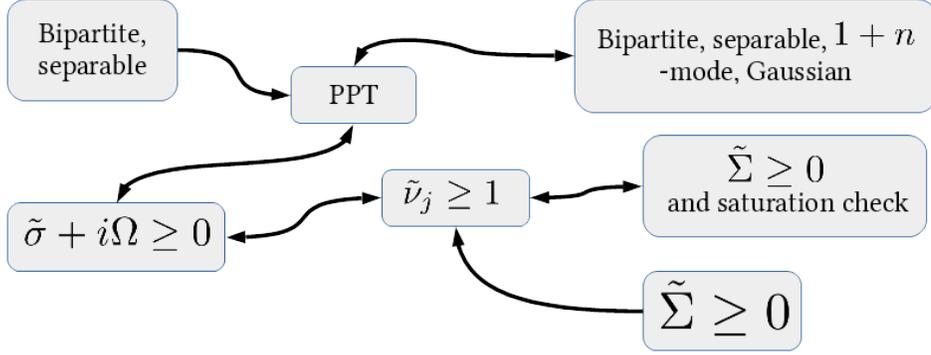


Figure 4.1: An implication diagram linking different conditions for entanglement.

We then define a new object, in analogy with Eq. (4.22):

$$\tilde{\Sigma} := \sum_{k=0}^{p+q} (-1)^{p+q+k} \vartheta_{2k}[\Omega\tilde{\sigma}]. \quad (4.28)$$

The condition

$$\tilde{\Sigma} \geq 0, \quad (4.29)$$

is necessary for Ineq. (4.27) to hold. However it can not be sufficient because of the possibility of an even number of violations or partial saturation, because of the same reasons as before. However, it can be shown that sufficiency holds for *bisymmetric* $(p+q)$ -states [76]. These are states that are invariant under local mode permutations on the subsystems. Some of these results are summarised in Fig. 4.1.

Intuitively we expect that locally-acting symplectics should not have an effect on any entanglement measure of a state. Indeed, it can be shown that local symplectic operations leave the transposed symplectic eigenvalues invariant. Considering a symplectic matrix $\sigma = SW S^\top$ being acted on by some $S_{\text{loc}} = S_A \oplus S_B$, where A and B label the two subsystems, we find that

$$\{\tilde{\nu}_i\} = \text{eig}[\Omega T S_{\text{loc}} \sigma S_{\text{loc}}^\top T] = \text{eig}[S_{\text{loc}}^\top T \Omega T S_{\text{loc}} \sigma] = \text{eig}[\Omega T \sigma T], \quad (4.30)$$

where we used the cyclic invariance of the spectrum and the fact that

$$S_{\text{loc}}^\top T \Omega T S_{\text{loc}} = \begin{pmatrix} S_A^\top & 0 \\ 0 & S_B^\top \end{pmatrix} \begin{pmatrix} \Omega & 0 \\ 0 & -\Omega \end{pmatrix} \begin{pmatrix} S_A & 0 \\ 0 & S_B \end{pmatrix} = T \Omega T. \quad (4.31)$$

If this were not the case then the transposed symplectic eigenvalues would not prove a very good

measure of entanglement.

Chapter 5

Geometry of Gaussian states

The previous sections provided a linear exposition from the Hilbert space definition of Gaussian states through the phase space until we ended with the duplet $(\bar{\mathbf{r}}, \sigma)$. The covariance matrix has received the most attention as it is the main focus of informational theoretical tasks.

As a result, our interests in terms of controlling these states will be geared towards navigating the manifold of covariance matrices. Here, we move towards a new geometrical picture where the set of covariance matrices can be seen as a space of manifolds, labelled by the symplectic eigenvalues and deriving their form from a geometrical interpretation of the matrix decompositions given so far. The set of matrices with fixed symplectic eigenvalues will be termed *isospectral* and is the set on which the symplectic group acts. The shottability question can be recast geometrically to ask whether this action on the manifold can be done with a single control.

5.1 Covariance matrix manifold

The investigation into the manifold structure of Gaussian states has a similar motivation to that for the Bloch ball, widely used in quantum information theory. For qubits the complex Hilbert space can be a less intuitive space in which to work and the same is true for Gaussian states. The higher dimensional generalisation of the Bloch ball was explored in Ref. [83] and here we make an analogous attempt for Gaussian states.

To begin we need to define the concept of a group action because this is where the manifold structure of covariance matrices will emerge. A left action of a Lie group G on a set \mathcal{A} is a map

$$\begin{aligned} \theta : G \times \mathcal{A} &\rightarrow \mathcal{A}, \\ (g, p) &\mapsto g \cdot p, \end{aligned} \tag{5.1}$$

such that $g_1 \cdot (g_2 \cdot p) = (g_1 \cdot g_2) \cdot p$ and $e \cdot p = p$; $g_1, g_2 \in G$, $p \in \mathcal{A}$ and e is the identity in G . A group action is transitive if

$$\forall p, q \in \mathcal{A}, \exists g \in G : p = g \cdot q. \quad (5.2)$$

Given such a group action, the orbit of an element p is the set of elements

$$\mathcal{O}_G(p) = \{g \cdot p \mid g \in G\}. \quad (5.3)$$

Some points of \mathcal{A} may be fixed under the action of a subgroup of G . This is the isotropy group:

$$G_p = \{g \in G \mid g \cdot p = p, p \in \mathcal{A}\}. \quad (5.4)$$

For points that are in the same orbit, this group is independent of p , up to isomorphism. This can be quickly seen by considering some $h \in G_p$ and some other $g \in G$ such that $q = gp$. We see that $ghg^{-1}q = q$ and so all elements in G_p are conjugate to an element of G_q . Also, given $kq = q$, we know that $g^{-1}kqp = p$ and so the converse is true. meaning that the groups are isomorphic.

Lemma 16. *Consider a set \mathcal{A} on which the Lie group G acts transitively such that the isotropy group of a point $p \in \mathcal{A}$ is a closed Lie subgroup of G . Then \mathcal{A} has a unique smooth manifold structure such that the given action is smooth [84].*

Such a set \mathcal{A} as described in Lemma 16 is referred to as a *homogeneous G -space* [84]. Furthermore we define an equivariant diffeomorphism to be a diffeomorphism from a manifold \mathcal{M} to a manifold \mathcal{N} such that it commutes with the group action [84]:

$$\begin{array}{ccc} M & \xrightarrow{F} & N \\ \theta_g \downarrow & & \downarrow \phi_g \\ M & \xrightarrow{F} & N \end{array}.$$

where θ_g denotes the action of g on \mathcal{M} and similarly for ϕ_g and \mathcal{N} . Furthermore, given a subgroup H of G then

$$gH := \{g \cdot h \mid h \in H\}. \quad (5.5)$$

The collection of all distinct sets of this form given H is called the set of left cosets of H in G . These sets may be considered as point elements in a new group expressed as the quotient G/H . This allows us to state the key theorem to reveal the manifold structure of these homogeneous G -spaces.

Lemma 17 (Homogeneous space characterisation theorem [84]). *Let \mathcal{M} be a homogeneous G -space, and let p be any point of \mathcal{M} . Then the map $\varphi : G/G_p \rightarrow \mathcal{M}$ defined by $\varphi(gG_p) = g \cdot p$ is an equivariant diffeomorphism.*

Note that the map φ takes us from an element of G/G_p , the set of left cosets, and takes it in one to one fashion to an element of \mathcal{M} .

Fix some multiset of symplectic eigenvalues $\mathbf{v} = \{\nu_1, \dots, \nu_n\}$ and denote the subset of covariance matrices that have these symplectic eigenvalues $\mathcal{J}(\mathbf{v})$. From Theorem 13 we know that the symplectic group has a manifold structure that is diffeomorphic to $U(n) \times \mathbb{R}^{n(n+1)}$. With the knowledge that $\text{Sp}(2n, \mathbb{R})$ acts transitively on $\mathcal{J}(\mathbf{v})$, we now seek the isotropy group. We know that this is unique since all covariance matrices with the same set of symplectic eigenvalues are in the same orbit under the action of the symplectic group. Therefore it suffices to find the set of matrices K which satisfy

$$W = KWK^\top, \quad (5.6)$$

where K is symplectic and W is, as usual, the diagonal matrix of symplectic eigenvalues.

Proposition 18. *The isotropy group of the $W \in \mathcal{J}(\mathbf{v})$ is $\text{OSp}(2n_1) \times \dots \times \text{OSp}(2n_k)$ where i labels the k distinct symplectic eigenvalues and n_i labels the degeneracy of the k distinct symplectic eigenvalues.*

Proof. We begin with the set of symplectic matrices K as defined in Eq. (5.6), recalling that $K\Omega K^\top = \Omega$. Combining the two conditions we find that we require $[K, W\Omega] = 0$. If K is written as a matrix of 2×2 sub-blocks κ_{ij} :

$$K = \begin{pmatrix} \kappa_{11} & \cdots & \kappa_{1n} \\ \vdots & \ddots & \vdots \\ \kappa_{n1} & \cdots & \kappa_{nn} \end{pmatrix}, \quad (5.7)$$

then the simple form of $W\Omega$ allows us to reduce the condition to a new one on each sub-block:

$$\nu_j \kappa_{ij} \Omega_1 - \nu_i \Omega_1 \kappa_{ij} = 0. \quad (5.8)$$

Writing

$$\kappa_{ij} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (5.9)$$

this provides a set of equations

$$\begin{aligned} (a+d)(\nu_j - \nu_i) &= 0, & (c-d)(\nu_j + \nu_i) &= 0, \\ (a-d)(\nu_j + \nu_i) &= 0, & (c+b)(\nu_i - \nu_j) &= 0. \end{aligned} \quad (5.10)$$

We see that for $\nu_i \neq \nu_j$ these equations provide $\kappa_{ij} = 0$ for all such cases making K block diagonal according to the degeneracy of the symplectic eigenvalues of W . In mathematical form $K = \bigoplus_{i=1}^k K_i$ where $K_i \in \text{OSp}(2n_i, \mathbb{R})$ and the symplectic eigenvalue degeneracy determines the

value of each n_i . A representation that provides a direct sum of groups provides a Cartesian product manifold description. \square

For any given covariance matrix $\sigma = SW S^\top$, Proposition 18 gives the isotropy group as that containing elements, SKS^{-1} , where K is in the direct product group as given before. Applying this to σ we may expand it in the form $\sigma = SKWK^\top S^\top$ where K varies over the isotropy group, as given before, without changing σ . This creates an equivalence class $\{SK \mid K \in \text{OSp}(2n_1) \times \dots \times \text{OSp}(2n_k)\}$ containing all possible symplectic matrices used in the Williamson diagonalisation of σ . Earlier it was stated, but not proven, that all elements of this equivalence class have the same singular values. Proposition 18 shows this because the singular values are invariant under the action of any element of $\text{OSp}(2n, \mathbb{R})$. As a result, the singular values associated with the diagonalising symplectic of a covariance matrix are unique. Furthermore, the proposition leads us to the following result.

Proposition 19. *There exists a diffeomorphism from $\mathcal{J}(\mathbf{v})$ to*

$$\frac{\text{U}(n) \times \mathbb{R}^{n(n+1)}}{\text{U}(n_1) \times \dots \times \text{U}(n_k)}, \quad (5.11)$$

where n_i labels the degeneracy of the i th symplectic eigenvalue.

This comes from a combination of Proposition 13, Lemma 17 and Proposition 18. The dimension of these manifolds is given by $n(2n+1) - \sum_{i=1}^k n_i^2$. When we enact the quotient it is via an equivalence class $(u_1, v) \sim (u_2, v)$, for all $u_1, u_2 \in \text{U}(n)$ and all $v \in \mathbb{R}^{n(n+1)}$, and so is the inverse of the original operation. Thus, we may rewrite the quotient above as

$$\frac{\text{U}(n)}{\text{U}(n_1) \times \dots \times \text{U}(n_k)} \times \mathbb{R}^{n(n+1)}, \quad (5.12)$$

allowing us to explore the geometry of the two parts separately. The left-hand quotient of unitary groups has close similarities with the strata of the Bloch ball in r dimensions. In Ref. [83] the authors show that the orbit of the unitary group on a finite-dimensional density matrix with eigenvalues $\lambda_1, \dots, \lambda_r$ of respective multiplicity n_1, \dots, n_r , has exactly the form given in Eq. (5.12), without the additional $\mathbb{R}^{n(n+1)}$. The difference is that our orbits are labeled by the multiset of symplectic eigenvalues \mathbf{v} whereas theirs are labeled by the eigenvalues of the density matrix.

The crucial difference arises in that the set of eigenvalues is bounded into a convex combination via $\sum_{i=1}^r n_i \lambda_i = 1$, where no such relation holds for the symplectic eigenvalues. Therefore no analogue of the Bloch ball is going to arise in any compact form, even given the compact nature of each of the *strata* given by the quotient.

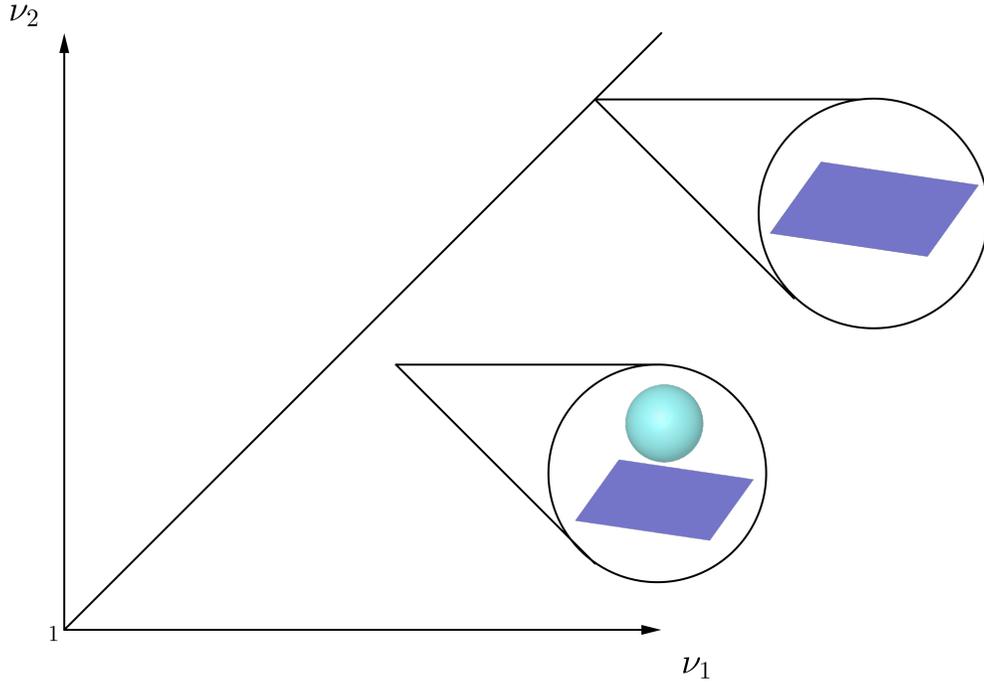


Figure 5.1: A graphical representation of the manifold of covariance matrices in two modes. The diagonal line depicts the case when the two symplectic eigenvalues, ν_1 and ν_2 , are degenerate. Each point represents a manifold in a different form, either as diffeomorphic to a plane or the product of a 2-sphere and a plane.

A naive attempt to create some sort of a geometrical picture in two modes is given in Fig. 5.1. This, however, does not contain the intuitive properties that are so useful in the Bloch ball. Note that in Fig. 5.1 we have included the uncertainty principle in lower bounding the symplectic eigenvalues to 1. We use the fact that $U(2)/(U(1) \times U(1)) \simeq \mathbb{C}\mathbb{P}^1$ is diffeomorphic to the 2-sphere [83]. However, in contrast to qubits, there is no clear to glue these together to form something like the Bloch ball because the symplectic eigenvalues don't obey a convexity condition.

The manifolds that come together in the higher dimensional Bloch ball have a partial ordering that is derived from majorisation conditions. The same will be true of the various manifolds at different points of Fig. 5.1. Each of them will correspond to a set of symplectic eigenvalues which will in turn determine the entropy of that particular state. Majorisation conditions act to compare states at different entropies. However, such conditions are still in their infancy for Gaussian states [85].

5.2 Williamson's theorem extension

The aim of the preceding geometrical analysis was to provide some intuition behind the manifold of covariance matrices, which we will later attempt to navigate. The analysis has not yet given the same clear picture as the Bloch ball although there are still things to be learned from it.

In the previous section we saw that the isotropy group for pure states is $\text{OSp}(2n, \mathbb{R})$. Such covariance matrices are of the form $\sigma = S\mathbb{I}S^\top$. The polar decomposition turns out to completely extract the isotropy group representative as $S = LR$ where $R \in \text{OSp}(2n, \mathbb{R})$ and $L \in \text{Sp}_+(2n, \mathbb{R})$, and so the covariance matrix of such a covariance matrix of a pure state is of the form $\sigma = L^2$. Since $L \in \text{Sp}_+(2n, \mathbb{R})$ are single-shot symplectics, the Williamson decomposition of all pure Gaussian states can be done such that S is single-shot. This reasoning holds for any state with degenerate symplectic eigenvalues.

It is unclear whether mixed states also inherit this single-shot property. The existence of an isotropy group for mixed states as discussed in the previous section at least informs us that there is no one-to-one relation between symplectic matrices and isospectral mixed states. If this were the case then there would exist some state that required a two-shot symplectic for its Williamson decomposition.

Mathematically the question is whether every covariance matrix can be put in the form

$$\sigma = SWS^\top \tag{5.13}$$

where S is of the form e^X , $X \in \mathfrak{sp}(2n, \mathbb{R})$. If this were true then we would have an extension of Williamson's theorem.

By exploring the two-mode case it is hoped that intuition will be provided for a larger n -mode investigation. In Sec. 3.5 we saw that the Dieci conditions provided a sufficient condition for the existence of a real Hamiltonian logarithm in the symplectic Lie algebra. This used a characterisation based on the Laub-Meyer normal form, and so this is where the investigation will start.

For the nondegenerate two-mode case we know that the isotropy group of σ is going to be conjugate to $\text{OSp}(2, \mathbb{R}) \oplus \text{OSp}(2, \mathbb{R})$. This conjugacy is given by the S that Williamson decomposes σ . It is just as valid to put the acting matrix 'through' the symplectic to touch the normal form inner element giving the decomposition

$$\sigma = SKWK^\top S^\top, \tag{5.14}$$

where $K \in \text{OSp}(2, \mathbb{R}) \oplus \text{OSp}(2, \mathbb{R})$. Therefore the problem becomes to find K such that SK is single-shot. If this is possible then we call the matrix S *savable*. We know that $\pm\mathbb{I}$ is always in the range of K and so S is known as *trivially savable* if this suffices. The conjecture for the

extension of Williamson's theorem for two modes will be proven true if this is achievable for every $S \in \text{Sp}(4, \mathbb{R})$.

Of course it is only necessary to focus on those elements of the symplectic group that are two-shot. For the rest, $K = \mathbb{I}$ will obviously suffice. This set of symplectics is at least given by the Dieci conditions, although because they are not necessary this provided set is potentially larger than it needs to be.

Enacting the Laub-Meyer decomposition we may rewrite S as $S = TVT^{-1}$ where $T \in \text{Sp}(4, \mathbb{R})$ and V is one of the normal forms as given in Sec. 3.5. In two modes the only options are $V = Y$, i.e. one block, or V has two blocks Y_1 and Y_2 . When considering two blocks this implies that $k = 1$ for each of the blocks. For one block we only consider k even. The cases that do not satisfy the Dieci condition are when we have:

- One block, Y .
 - a block of type a , $b(ii)$, or $b(iii)$.
- Two blocks, Y_1, Y_2 .
 - a block of type a and another block not of type a
 - some mixture of the b blocks where neither are the same.
 - two $b(iii)$ blocks.
 - a block of type $b(iii)$ and another block not of type $b(iii)$.

Given the list of matrices not satisfying the Dieci condition most of them are trivially savable by setting $K = -\mathbb{I}$ to change the sign of the eigenvalues. In fact the only two that are not are:

- type a and not type a .
- type $b(iii)$ and not type $b(iii)$.

If this second block corresponded to a negative or imaginary eigenvalue then the matrix would be trivially saveable. Therefore the only remaining possibility is for it to be positive. Using Ref. [71] and Sec. 3.5 we write down the two cases that are not trivially savable have Laub-Meyer normal forms,

$$V = \begin{pmatrix} -\lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & -\frac{1}{\lambda_1} & 0 \\ 0 & 0 & 0 & \frac{1}{\lambda_2} \end{pmatrix}, \quad V = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 2 & 0 & 0 & \frac{1}{\lambda_2} \end{pmatrix} \quad (5.15)$$

where $\lambda_1, \lambda_2 > 0$ and not equal to 1 and noting that we are working in the J -basis.

From here it is necessary to show that there exists $K \in \text{OSp}(2, \mathbb{R}) \oplus \text{OSp}(2, \mathbb{R})$ such that $SK = TVT^{-1}K$ is single-shot. Doing this has proven difficult although numerical simulations

have strongly suggested that these types of symplectic are also savable, and so this the extension of Williamson's theorem is conjectured to be true, at least in two modes. Proving this in general for n modes would close the gap a little by showing that both can enact control using a single time-independent Hamiltonian.

Chapter 6

Control theory

Control theory is a broad subject that involves the systematic modelling of the attempt by humans to manipulate physical systems. The formal enquiry is often dated to Maxwell's exploration of governors which "is a part of a machine by means of which the velocity of the machine is kept nearly uniform" [10]. In the paper he gives a call to arms for physicists and mathematicians to study these dynamical systems mathematically.

Generally the picture is this: a space of states is visualised along with dynamics that cause the system to move from some initial state to a final state. A part of the dynamics involves a component that is chosen by some controller. During the evolution this component can be altered, thus changing the dynamics of the system. Within this model, questions are asked about the possibility of reaching specific states, and the time it might take to arrive there. These controls may be dependent on the particular state reached, as in the case of governors, in which case this is referred to as *feedback control* because the state of the system is feeding back to determine the controls.

The modern study of control theory is often taken to begin with Roger Brockett in the nineteen-sixties with his early studies on feedback systems [86]. During the sixties and into the seventies, control theory blossomed into a plethora of investigations concerning the mathematical formalisation of the subject [87]. We will follow a specific body of work that provides the basis for the results in following sections.

The account of a state flowing around a space of states according to dynamics that are affected by a controller has a distinctly geometrical flavour. It is therefore not surprising that geometry was quickly absorbed to aid in formalising control. It became apparent that in many cases the movement of states through the space corresponded to flows induced by right-invariant vector fields, as in Hamiltonian dynamics. This spawned the study of 'right-invariant control systems' which introduce the full theory of Lie groups and algebras to the investigation. Ref. [88] cites

Ref. [89] as the first to apply vector fields to study control problems in 1963.

This field burgeoned with a seminal paper by Velimir Jurdjevic and Hector Sussmann [90] where some crucial results were proven that form the bedrock of much current research, forty-five years on. Control theory absorbed more and more of the theory of Lie groups, requiring ever-greater sophistication of techniques. ‘Strongly regular’ Lie algebra elements received much attention [91] building on old results of Masatake Kuranishi [92, 93]. Specific types of Lie group were explored with specific interest in solvable groups [94] and simple groups [95]. Towards the end of the last century and up to the present, the theory of root spaces has provided deeper insight into the problem [96–98]. A wonderful review of Lie group control can be found in Ref. [99] with Refs. [88, 100] being two major textbooks on the topic.

Control theory was born from questions in classical engineering which has spread into the field of quantum engineering. A series of reviews, summarising contemporary success can be found in Refs. [101–105]. Quantum states, and subsequently their control, can be divided into states living in either finite- or infinite-dimensional Hilbert spaces with the former being dealt with in one of the early textbooks in the field [106]. Given our interest in Gaussian states, however, the work on infinite-dimensional systems requires more time.

In an infinite-dimensional Hilbert the set of pure states live on the unit sphere or *projective* Hilbert space. As an infinite-dimensional manifold it introduces many subtleties into the problem of control. A major insight came from the interest in the subset of analytic vectors, based on Edward Nelson’s 1959 work [107]. Controlling on this subset turns out to be as desirable as any physicist would want, and luckily makes the problem more tractable. The seminal paper providing this analysis is Ref. [108] with a continuing literature including Refs. [109–113].

The set of Gaussian states, however, which includes both pure and mixed states, evolves on a finite-dimensional manifold, which can be immediately inferred by its finite degrees of freedom. However, the related Hilbert space is still forced to be infinite-dimensional. This suggests that the Hilbert space picture is not a natural arena for control. Instead it is the study of Lie groups where many people have gone to gain deeper insight into specific infinite-dimensional systems.

A notable line of exploration in this vein is the work of Jian-Wu Wu et al who have focused on the control of $SU(1, 1)$ which is a symmetry group arising for many systems including Bose-Einstein condensates and in the downconversion process [114, 115]. This will be of crucial interest to us later due to its isomorphism with $Sp(2, \mathbb{R})$.

A different area that arises is the study of optimal control. This concerns the issue of reaching a specific point as fast as possible, given that you know you are able to. These optimisation schemes involve complex mathematical and computational machinery. In recent years the notion of the control landscape has proved useful, with a technical definition on Ref. [116]. It seems that the simple ‘yes-no’ aim of controllability might be far more an interest of pure mathematicians rather than experimentalists. A part of the aim of writing Ref. [117] was to save its reputation

by showing how it links to a more easily tractable control landscape and therefore ties in with the notion of optimal control. However, as yet this result is only proven for finite-dimensional Hilbert spaces.

The group of interest for the control of Gaussian states is of course the symplectic group. Questions of optimality for this group were first studied in Ref. [118]. Control of the first moments with an open system is explored in Ref. [119] and the recent work on conditions to enact any symplectic group element can be found in Ref. [120]. These works base themselves in the context of physics. The most recent purely mathematical attempt to draw out control properties of this group can be found in Ref. [98].

These problems are connected to the relationship between a set of controls at the Lie algebra level and a set of group elements or states. Secs. 7-8 will deal with this aspect of control for Gaussian states. The focus there is on closed evolution and conditions for control when we restrict the Hamiltonians that are accessible to the controller. Optimisation problems will be considered in Secs. 9.2 and 10 in the context of open dynamics with manipulation of at the group level, rather than studying the Lie algebra/group subtleties.

6.1 Preliminaries

The control of quantum systems is often divided into *open-loop* and *closed-loop* control. Open-loop control systems are those for which the control choices of the controller are pre-determined. That is to mean that there is no feedback of measurement outcomes that go into the decisions made during the evolution of the state, which would be termed closed-loop. We will exclusively work on open-loop problems for the remainder of this work.

The mathematical expression that captures these notions generally can be given by [88]

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}(t)), \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0. \quad (6.1)$$

The vector-valued function \mathbf{f} dictating the dynamics depends both on \mathbf{x} and the control functions $\mathbf{u}(t)$. The fact that this \mathbf{u} has no \mathbf{x} dependence indicates that this is an open-loop control equation. \mathcal{U} denotes the set of possible functions that the controller could choose and is referred to as a *control set*. Given an initial value $\mathbf{x}(0)$ one might ask which elements of \mathcal{U} allow one to reach some specific target state, or which elements allow one to reach the target state in a specific time.

Eq. (6.1) is very general, describing a broad class of physical setups. This is as far as we develop for the setup of Sec. 9 but for Sec. 8 we require a narrower subclass of equations. These

arise when we decompose $\mathbf{f}(\mathbf{x}, \mathbf{u}(t))$ so that it takes the form

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) + \sum_{i=1}^m u_i(t) \mathbf{h}_i(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0. \quad (6.2)$$

This situation is very common in a wide range of physical setups. The $\mathbf{g}(\mathbf{x})$ represents a part of the control that is ‘always on’, corresponding to some dynamics that are unaffected by the controller. The second component represents the ability of the controller to manipulate the dynamics by altering a vector of m scalar functions.

To continue to narrow our focus we may linearise by truncating the Taylor expansion such that $\mathbf{g}(\mathbf{x}) \rightarrow \mathbf{g}(\mathbf{0}) + \frac{\partial \mathbf{g}}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{0}} \mathbf{x}$ and $\mathbf{h}_i(\mathbf{x}) \rightarrow \mathbf{h}_i(\mathbf{0}) + \frac{\partial \mathbf{h}_i}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{0}} \mathbf{x}$ giving

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{g}(\mathbf{0}) + \sum_{i=1}^m u_i(t)(B_i\mathbf{x} + \mathbf{h}_i(\mathbf{0})), \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0, \quad (6.3)$$

where $A = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{0}}$ and $B_i = \frac{\partial \mathbf{h}_i}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{0}}$. Eq. (6.3) is referred to as inhomogeneous bilinear [99] or sometimes biaffine [100].

Setting $\mathbf{g}(\mathbf{0}) = \mathbf{h}_i(\mathbf{0}) = \mathbf{0}$ for all i we arrive at

$$\dot{\mathbf{x}} = \left(A + \sum_{i=1}^m u_i(t) B_i \right) \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}_*^d, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0. \quad (6.4)$$

where $\mathbb{R}_*^d = \mathbb{R}^d \setminus \{\mathbf{0}\}$. The new space of states is often set to this new punctured space because a state beginning at the null element will not evolve and this provides issues with the transitivity property described later. The matrix A represents the always-on dynamics which is referred to as the *drift field*. The set of B_i fields are referred to as *control fields* where the *control functions* $u_i(t)$ are set by the controller. The reason they are referred to as fields is because we will later see them as elements of the tangent space of the state manifold which is isomorphic to the Lie algebra of right-invariant vector fields.

There are many different control sets \mathcal{U} that are natural to consider. Four, that are often used are described here:

- \mathcal{U}_u is the class of all functions defined on the interval $[0, \infty)$ with the range taking values in \mathbb{R}^m .
- \mathcal{U}_r is the class of all functions defined on the interval $[0, \infty)$ with the range taking values in the cube: $\{\mathbf{u} \in \mathbb{R}^m : |u_i| \leq 1, i = 1, \dots, m\}$.
- \mathcal{U}_b is the class of all piecewise constant functions defined on the interval $[0, \infty)$ with the range taking values in \mathbb{R}^m and elements with values 1 and -1 . This is also referred to as

‘bang-bang’ control.

- \mathcal{U}_c is the class of all functions defined on the interval $[0, \infty)$ with the range taking values in \mathbb{R}^m and piecewise constant.

All functions are taken to be locally bounded and measurable. Before delving into the Lie group analysis of Eq. (6.4), this provides a good place to pause for an example.

Example. Consider

$$\dot{\mathbf{v}} = (\mathbb{I} - u\Omega)\mathbf{v}, \quad \mathbf{v} \in \mathbb{R}_*^2, \quad u(\cdot) \in \mathcal{U}_u, \quad t \geq 0. \quad (6.5)$$

where $\mathbf{v} = (x, y)^\top$. Changing variables to $x = r \cos \theta$ and $y = r \sin \theta$, we find

$$\dot{r} = \frac{x\dot{x} + y\dot{y}}{\sqrt{x^2 + y^2}} = r, \quad (6.6)$$

and

$$\dot{\theta} = \frac{\dot{x}r^{-1} - xr^{-2}\dot{r}}{\sqrt{1 - \left(\frac{x}{r}\right)^2}} = -u, \quad (6.7)$$

providing

$$r(t) = e^t r(0) \quad \theta(t) = - \int u(t) dt + \theta(0). \quad (6.8)$$

This provides an ever increasing radius and so no chance of reaching every element of \mathbb{R}_*^2 . Swapping \mathbb{I} and Ω we obtain

$$\dot{r} = 2ur, \quad \dot{\theta} = 1, \quad (6.9)$$

where now we may manipulate $u(t)$ to obtain any value of \mathbb{R}_*^2 . For later reference, this property will be referred to as *controllability*.

Before delving further, an important concept to introduce is that of the *transition matrix*. We may relate the evolution of \mathbf{x} to the evolution of a matrix acting on the initial state

$$\mathbf{x}(t) = M(t)\mathbf{x}(0). \quad (6.10)$$

Substituting this into Eq. (6.4) and dropping the initial state we obtain a new equation

$$\dot{M} = \left(A + \sum_{i=1}^m u_i(t) B_i \right) M, \quad M(0) = \mathbb{I}_d, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0. \quad (6.11)$$

This new equation is obviously related to Eq. (6.4) although now the focus is on the set of

transformations rather than the state. Of course, we could forget about the original states entirely and only be interested in the transformations.

It may be the case that we do not want to consider the entirety of \mathbb{R}_*^n but instead some submanifold \mathcal{M} and ask which points we can reach. To lead back to the example of Gaussian states we have already encountered the manifold of covariance matrices which is where these geometrical notions would be important. For quantum mechanics we could consider \mathcal{M} to be the set of density matrices, as explored in the following example.

Example. Consider the Liouville-von Neumann equation

$$\dot{\hat{\rho}} = i[\hat{H}, \hat{\rho}], \quad \hat{\rho} \in \mathcal{D}, \quad \mathbf{u}(\cdot) \in \mathcal{U}, \quad t \geq 0. \quad (6.12)$$

where

$$i\hat{H} = \hat{A} + \sum_{i=1}^m u_i(t)\hat{B}_i, \quad (6.13)$$

and \mathcal{D} is the set of density matrices on some Hilbert space. We have already split the control Hamiltonian into drift and control fields. The ansatz,

$$\hat{\rho}(t) = \hat{U}\hat{\rho}(0)\hat{U}^\dagger \quad (6.14)$$

solves the differential equation,

$$\frac{d\hat{\rho}(t)}{dt} = \frac{d\hat{U}}{dt}\hat{\rho}(0)\hat{U}^\dagger + \hat{U}\hat{\rho}(0)\frac{d\hat{U}^\dagger}{dt} = i\hat{H}\hat{U}\hat{\rho}\hat{U}^\dagger + \hat{U}\hat{\rho}(i\hat{H}\hat{U})^\dagger, \quad (6.15)$$

where \hat{U} must satisfy

$$\frac{d\hat{U}}{dt} = i\hat{H}\hat{U}, \quad (6.16)$$

which allows us to consider the unitary transformations without the states.

As a result, Eqs. (6.4) and (6.11) act as the bedrock on which we can begin to consider the mathematical problems of quantum control theory.

The type of control theory we are considering has connections with another problem that is found in the quantum information literature. The aim of ‘universal quantum computation’ is to find a small set of unitaries that are able, in combination, to enact any unitary of a particular dimension [121]. Here, however, we deal with a set of *Hamiltonians* which correspond to an infinite set of allowed unitaries. Our focus is therefore on the Lie algebra level, rather than the group level.

The key object in the study of control is the Lie algebra generated by the set of possible

Hamiltonians.

$$\Gamma = \left\{ A + \sum_{i=1}^m v_i B_i \mid \mathbf{v} \in \text{range}[\mathbf{u}(t)], \mathbf{u}(\cdot) \in \mathcal{U} \right\}. \quad (6.17)$$

From this we may construct a Lie algebra by introducing the matrix commutator $[X, Y] = XY - YX$ to act as the natural Lie bracket on the set. By taking the repeated commutators until we have a linearly independent set we form a basis for the Lie algebra vector space which will be denoted \mathfrak{L}_Γ with G_Γ the associated Lie group. If we wish to discuss some general group G then its associated Lie algebra will always be denoted as \mathfrak{L} .

The solution to Eq. (6.11) is exponential and so the solutions M will all be elements of G_Γ . Eq. (6.4) can therefore be seen as a set of dynamics where elements of G_Γ act on elements of the manifold \mathcal{M} via Eq. (6.10). Thus we arrive again at the theory of homogeneous G -spaces.

Obviously if the group is not transitive on \mathcal{M} then the control equation will never be able to reach the full set of states which could mean that we should really be considering a submanifold. An example of this would be in considering the Bloch ball as the set of states for unitary evolution. The unitary group is not transitive on the full Bloch ball and so control in this case will be restricted to a consideration on the sub-2-spheres.

Given Eq. (6.11) where we have set the initial matrix \mathbb{I}_d , chosen some control set \mathcal{U} and a set of fields $\{A, B_1, \dots, B_m\}$ defining some Γ , then we may make the following definitions.

Definition 20 (Reachable Set). *Given $\{A, B_1, \dots, B_m\}$, which provides some Γ , as in Eq. (6.17), the values of the set of solutions to Eq. (6.11) ranging over \mathcal{U} forms the reachable set, \mathcal{R} . This will be referred to as \mathcal{R}_Γ if this needs explicitly stating.*

Definition 21 (Controllability). *Γ is said to be controllable on a Lie group G if $\mathcal{R} = G$.*

In that section, the way in which optimisation must occur is made apparent through the way in which the differential equation is written. However, this only dips into the broader theory of optimal control, which comes with a set of different tools and results to the large body of literature on controllability. Optimal control, in general, is closely connected with the calculus of variations but provides a broader and more sophisticated framework for minimising the length of trajectories. The intensity of the space program and the race to the moon went hand in hand with problems of this type and continues to be a central point of interest for experimental and mathematical physics, for its utility and complexity [88].

6.2 Preliminary controllability results

Firstly we state that for Γ to be controllable on G , G must be a connected group. If it is not then there will be no path connecting the initial point to every other point on the group manifold.

From this point onward we will deal exclusively with the piecewise constant control set $\mathcal{U} = \mathcal{U}_c$. This may seem like a wide amount of control for any result hoping to be of interest to experimentalists. However, this set strikes a happy medium in that it allows us to explore what is possible in this limit, providing intuition for the capped case as we will see. This set allows us to provide the solution to Eq. (6.11) as

$$M(t) = \prod_{i=1}^p e^{X_1 t_1} e^{X_2 t_2} \dots e^{X_p t_p}, \quad \sum_{i=1}^p t_i = t, \quad t_i > 0 \forall i, \quad (6.18)$$

where $X_i \in \Gamma$ and p is the number of constant control trajectories that are concatenated.

Observing this solution we see that the reachable set is going to have the properties of a *monoid*, which is a semigroup with identity. The only thing that prevents the reachable set having group structure in general is the positivity of time removing the immediate existence of an inverse for each element.

Theorem 22 (Generation necessity). *For Γ to be controllable on G it is necessary that $\mathfrak{L}_\Gamma = \mathfrak{L}$.*

Proof. If \mathfrak{L}_Γ is not \mathfrak{L} then either it forms a subalgebra or an element of Γ is not in \mathfrak{L} . In the former case there is an element of G that is outside of G_Γ . In the latter case there is an element L of Γ such that e^L is not in G . Either way, $\mathcal{R} \neq G$ and so the system is not controllable. \square

Lemma 23. *Let G be some connected Lie group and let $\{L_i\}$ be a set of elements that generate the Lie algebra of G . Every $g \in G$ can be written as a finite product of elements of the form e^{qL_i} for $q \in \mathbb{R}$, $i = 1, \dots, m$ [90].*

Homogeneous systems, not to be confused with homogeneous G -spaces, are a type of system that frequently arises and leads to some simple controllability results.

Definition 24 (Homogeneous systems [99]). *Control systems of the form of Eq. (6.11) for which $\Gamma = -\Gamma$ are referred to as homogeneous.*

An example of a homogeneous system are those for which the drift field A and the set $\{B_1, \dots, B_m\}$ are linearly dependent. Theorem 23 immediately allows us to prove the following starter control theorem:

Theorem 25 (Homogeneous system controllability). *If the system is homogeneous then a necessary and sufficient condition for controllability is that $\mathfrak{L}_\Gamma = \mathfrak{L}$.*

Proof. Theorem 23 states that any element $g \in G$ can be written as a finite product of exponentials of elements of a generating set of the Lie algebra of G with coefficients $q \in \mathbb{R}$. This is the same form as $M(t)$ takes in Eq. (6.18). Therefore, all we require is that the Lie algebras are the same. \square

Another result that will be key for proving things later is the sufficiency of the reachable set being dense in G , in order to conclude that is $is G$.

Definition 26 (Closure and density [122]). *Consider a topological space \mathcal{T} and a subset $\mathcal{A} \subseteq \mathcal{T}$. The intersection of the family of all closed sets containing \mathcal{A} is denoted $\bar{\mathcal{A}}$ and is the closure of \mathcal{A} . A set \mathcal{A} is dense in \mathcal{T} if $\bar{\mathcal{A}} = \mathcal{T}$.*

Lemma 27. *If $\bar{\mathcal{R}} = G$ then $\mathcal{R} = G$.*

Proof. See Theorem 2.8 in Ref. [99]. □

From here we look at compact groups. Their compact nature provides a result that is akin to the Poincaré recurrence theorem causing a recursive property of the dynamics.

Theorem 28. *For G compact there exists a sequence of times $\{t_k\}_{k=1}^{\infty}$ with $t_k \geq 1$ such that $\lim_{k \rightarrow \infty} \|e^{t_k X} - \mathbb{I}\| = 0$ for all $X \in \mathfrak{L}$ [106].*

Proof. G is compact and therefore it is sequentially compact¹, which means that every sequence in G has a convergent subsequence. First, we consider $e^X, e^{2X}, e^{3X}, \dots$, for some $X \in \Gamma$, which contains a convergent subsequence $e^{m_1 X}, e^{m_2 X}, \dots$, converging on some $g \in G$ where $\{m_1, m_2, \dots\} \subseteq \mathbb{N}$. Furthermore the sequence $e^{-m_1 X}, e^{-m_2 X}, \dots$ will converge to g^{-1} . Therefore

$$\lim_{k \rightarrow \infty} \|e^{(m_{k+1} - m_k)X} - \mathbb{I}\| = \lim_{k \rightarrow \infty} \|e^{t_k X} - \mathbb{I}\| = 0, \quad (6.19)$$

where $t_k := m_{k+1} - m_k$. The minimum value of t_k is 1 because m_{k+1} is always an integer larger than the integer m_k , which proves the statement. □

Theorem 29 (Compact Lie group control). *If G is compact then a necessary and sufficient condition for Γ to be controllable on G is that $\mathfrak{L}_\Gamma = \mathfrak{L}$.*

Proof. Theorem 28 implies that when some M is an element of $\bar{\mathcal{R}}$, for G compact, then so is its inverse, since reaching M along a trajectory and then returning to identity means that the return trajectory can be cut out as M^{-1} . Therefore $\bar{\mathcal{R}}$ has a group structure. This group will be the one that corresponds to \mathfrak{L}_Γ , as Eq. (6.11) just relates trajectories in \mathfrak{L}_Γ with elements in G_Γ . By Lemma 27 this implies that Γ is controllable on G . Combined with Theorem 22 we see that the condition is necessary and sufficient for compact groups. □

We see that the common feature of both driftless systems and drift systems on compact groups is that the elimination of time plays a role in the proof. It is the positivity of the time

¹Compact Lie groups are metrizable by Urysohn's metrisation theorem which states that 'a compact space is metrizable if it is second countable' [122]. All Lie groups are second countable and compactness is the same as sequential compactness for metric spaces. See Appendix A for more detail on terminology.

parameter used in these product solutions that stops the question of controllability being easy – if we could evolve Eq. (6.11) forwards *and* backwards in time then we would be able to reach every element of the Lie group associated to \mathcal{L}_Γ . Homogeneous systems do not have this problem because despite time being positive the generator can always be given a negative coefficient by the property that defines these systems. Compact groups contain the recursive property as given in Theorem 28 which also removes the importance of time in these systems.

The theorems above begin to reveal some of the structure of Lie group control theory. We see that for compact groups and when Γ is homogeneous, the issue of time is eliminated and a necessary and sufficient controllability criterion arises. This suggests the question of whether similar results can be found for non-compact groups with Γ inhomogeneous. Researchers on finite-dimensional quantum systems may cease to be interested because $U(n)$ is compact. However, groups like $Sp(2n, \mathbb{R})$ and $SU(1, 1)$ play an important role in infinite-dimensional quantum systems and are both non-compact.

6.3 Multiple control field results

We introduce a new set

$$\Gamma_B = \left\{ \sum_{i=1}^m v_i B_i \mid \mathbf{v} \in \text{range}[\mathbf{u}(t)], \mathbf{u}(\cdot) \in \mathcal{U} \right\}. \quad (6.20)$$

which is the same as Γ but with the drift field removed. This comes with an associated Lie algebra \mathfrak{L}_{Γ_B} and Lie group G_{Γ_B} , where \bar{G}_{Γ_B} denotes the closure of G_{Γ_B} in G_Γ .

Lemma 30. $\bar{G}_{\Gamma_B} \subseteq \bar{\mathcal{R}}$ [90].

Proof. By Lemma 23 we can express every element of a group as a product of its single shots. Thus to prove that $\bar{G}_{\Gamma_B} \subset \bar{\mathcal{R}}$ it suffices to show that it is true for the single shots. Fix $u_i = (0, \dots, 0, c, 0, \dots, 0)$ where c appears in the i th position. Choose time such that $t = t'/c$. Therefore

$$\lim_{c \rightarrow \infty} e^{t'/c(A+cB_i)} \in \bar{G}_{\Gamma_B} \quad (6.21)$$

and is also in the closure of the reachable set. By taking the limit points of G_{Γ_B} we stay in $\bar{\mathcal{R}}$ which proves the statement. \square

Theorem 31 (Blow away the drift field). *If $\mathfrak{L}_{\Gamma_B} = \mathfrak{L}_\Gamma$ then Γ is controllable on G_Γ .*

Proof. In Lemma 30 we saw that $\bar{G}_{\Gamma_B} \subseteq \bar{\mathcal{R}}$. If $\mathfrak{L}_{\Gamma_B} = \mathfrak{L}_\Gamma$ then $G_{\Gamma_B} = G_\Gamma$. In Ref. [123], we find that the closure of subsemigroups are subsemigroups, i.e. $\bar{\mathcal{R}} \subseteq G_\Gamma$ and so this would prove their equality. From Lemma 27 we see that this is equivalent to controllability of Γ on G . \square

At this stage we reach a divergence between the mathematical and physical communities. The drowning out of the drift field dynamics by powerful control fields seems like a fair physical assumption. This is down to the fact that a transformation being enacted which is very close to another is something that should yield a similar physical effect.

However, closeness at the Lie algebra level is not so physically clear. Two Hamiltonians that are extremely similar may yield very different dynamics when applied to a state. Whether or not such a notion of closeness is allowed will affect reception of the following result that is the cause for most mathematicians ignoring the situation of $m \geq 2$, i.e. more than one control field as written in Eq. (6.4).

Theorem 32 (Generic generation [88]). *Almost all right invariant control systems with $\mathcal{U} = \mathcal{U}_u$, or \mathcal{U}_c , are controllable on G provided that the number of control functions $m \geq 2$.*

This uses the idea that the set of elements $(L_1, L_2) \in \mathfrak{L} \times \mathfrak{L}$ such that L_1 and L_2 generate \mathfrak{L} , is an open and dense subset when \mathfrak{L} is semisimple [88]. Therefore, for $m = 2$, B_1 and B_2 are, in a sense, close to two elements that generate \mathfrak{L} . Therefore, using Theorem 31 the system is controllable.

A recent advance for the $m = 1$ case can be found in Ref. [98] where root spaces play a central role in developing a sufficient condition on Γ for controllability. As yet, however, there does not exist a general necessary and sufficient condition for controllability of non-compact groups.

As we have seen, however, for compact groups and driftless systems this question has been answered. It is known that the Lie algebra rank criterion is not a sufficient for the controllability non-compact groups and so the current aim is in finding another property that might be added to create a necessary and sufficient condition.

The property of ‘neutrality’ was proven to be sufficient in 1972 but its necessity for control is still only a conjecture² [90]. This property has an intuitive link with the idea of recursivity of the Hamiltonian dynamics and, as such, is a novel way in which time may be removed as an issue, as in compact groups and driftless systems. It is hoped that some physical insight can be drawn from this property to shed light on this open conjecture.

²In the paper it is just stated that they do not know if the result is also necessary. The conjecture was stated in private correspondence between V. Jurdjevic and the present author in April 2015.

Chapter 7

Neutrality

The concept of neutrality in the context of control comes in Ref. [90] where the authors show that the existence of constant recursive trajectories on the manifold would provide a new condition for controllability. The clarification of this theorem, and link to linear algebra was provided in Ref. [124], although the name was not coined there. Indeed, it is referred to as *compactness* in Ref. [125]. In this section we will explore the condition of neutrality and its relation to control. Towards the end we will begin to focus on the symplectic group as a key object of study in the thesis.

7.1 Introduction

There are many different ways to define a neutral matrix which is shown by the list of equivalences below. These are useful at different times, depending on how we want to use them.

Definition 33 (Neutral matrix). *A matrix $N \in \mathbb{R}^{m \times m}$ is neutral if*

$$\exists P \in \mathbb{R}^{n \times n} \quad \text{st} \quad PNP^{-1} = M \quad \text{and} \quad M^T = -M. \quad (7.1)$$

Lemma 34. *Neutrality of N is equivalent to each of the following properties [100]*

1. *A nonzero matrix $N \in \mathbb{R}^{n \times n}$ if $N^T Q + QN = 0$ for some positive-definite Q .*
2. *$\text{spec}[N]$ lies on the imaginary axis and N is diagonalisable over \mathbb{C} .*
3. *The closure of $e^{\mathbb{R}N}$ is compact. $N \neq 0$.*
4. *There exists $\sigma > 0$ and a sequence of times $\{t_k\}_1^\infty$ with $t_k \geq \sigma$ such that $\lim_{k \rightarrow \infty} \|e^{t_k N} - \mathbb{I}\| = 0$.*

Proof. We find a set of inferences which will provide the equivalence of the above statements.

- Def \iff 1.

$$\begin{aligned} N^\top Q + QN &= N^\top P^\top P + P^\top PN = 0 \iff \\ (P^{-1})^\top N^\top P^\top &= (PNP^{-1})^\top = -(PNP^{-1}) \end{aligned} \quad (7.2)$$

where we set $Q = P^\top P$ which is the Cholesky decomposition of Q due to its positive-definiteness [126].

- Def \iff 2. $PNP^{-1} = M$ as in the definition. M is real and skew-symmetric matrix if and only if iM is Hermitian which is true if and only if $iM = UDU^\dagger$, where D diagonal and real. Therefore the spectrum of M is imaginary and hence so is the spectrum of N .
- 1 \implies 3. Given $N^\top Q + QN = 0$ we have already seen that $\text{Tr } N = 0$ because a skew-symmetric matrix has trace zero. Therefore $\det e^{rN} = 1$ for all $r \in \mathbb{R}$. The determinant is a continuous function mapping the set to a compact subset of \mathbb{R} . Continuous functions map compact spaces to compact spaces and so the closure of $e^{\mathbb{R}N}$ is compact.
- 3 \implies 2. First we show that compactness implies diagonalisability. First define $J_k \in \mathbb{R}^{d \times d}$

$$J_k := \begin{pmatrix} 0 & \cdots & \overbrace{1}^k & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (7.3)$$

where $J_0 := \mathbb{I}$ and $J_k := 0_{d \times d}$ for $k > d - 1$. Note that $J_k J_l = J_{k+l} \forall k, l \in \mathbb{N}$. By the binomial theorem

$$(\lambda \mathbb{I} + J_1)^n = \sum_{m=0}^n \binom{n}{m} \lambda^{n-m} J_m, \quad (7.4)$$

and so

$$\begin{aligned} e^{(\lambda \mathbb{I} + J_1)t} &= \sum_{n=0}^{\infty} \sum_{m=0}^n \binom{n}{m} \lambda^{n-m} J_m \frac{t^n}{n!} = e^{\lambda t} + \sum_{m=1}^{d-1} J_m \sum_{n=m}^{\infty} \binom{n}{m} \frac{\lambda^{n-m} t^n}{n!} \\ &= e^{\lambda t} + \sum_{m=1}^{d-1} J_m \frac{t^m}{m!} \sum_{n=0}^{\infty} \frac{\lambda^n t^n}{n!} = e^{\lambda t} \left(\mathbb{I} + \sum_{m=1}^{d-1} J_m \frac{t^m}{m!} \right). \end{aligned} \quad (7.5)$$

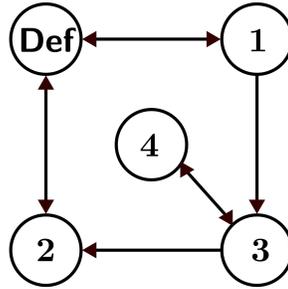


Figure 7.1: A graphical representation of the implications used to prove Theorem 34.

In the first line we took out an $m = 0$ term, used the fact that $m > d - 1$ causes $J_m = 0_{d \times d}$ and noted that $\binom{n}{m} = 0$ for $n < m$. The final expression is not equal to \mathbb{I} for all $t \neq 0$. A Jordan block is of the form $\lambda \mathbb{I} + J_1$ and so the above result shows that the identity is not reachable for $t > 0$ if N is not diagonalisable. Now, given that N must be diagonalisable, we focus on its spectrum. Any real eigenvalues will exponentiate in time to infinity and so can not recur. Therefore, all eigenvalues must be imaginary. This provides the implication required.

- **3** \iff **4**. Since $e^{\mathbb{R}N}$ is a subset set of the $n \times n$ real matrices, which is a metric space, compactness is equivalent to sequential compactness. Therefore the proof of Theorem 28 applies here. That provides **3** \implies **4**. The reverse is clear in that if $e^{rN} = \mathbb{I}$ for some r then the set is compact because $e^{(r+r')N} = e^{r'N}$ and so the exponential map is a continuous map from a closed segment of the real number line to a set of matrices. The domain is compact and therefore so is the range.

□

To illustrate the implications of the proof, Fig. 7.1 is provided as a graphical representation. We see that the concept of neutrality is completely related to the notion of recursion via the exponential map. In physical systems we expect recursion to relate to recursive dynamics, for example simple harmonic oscillation. This is exposed mathematically in Sec. 7.3. However, before delving too deeply into the physics of this problem we shall first look at its relation to control theory.

7.2 In control

Near the genesis of geometrical control theory, a seminal paper by Jurdjevic and Sussmann [90] provided a sufficient result related to neutrality that will be reproven here.

Theorem 35. *Given a control system on Eq. (6.11) with associated Γ and \mathcal{U}_c , if there exists a control \mathbf{u} such that $A + \sum_{i=1}^m u_i B_i$ is neutral then Γ is controllable on G_Γ .*

Proof. By Lemma 34 our neutral control allows us to get arbitrarily close to \mathbb{I} and so we can make the broader statement that there exists some $X \in \Gamma$ such that $e^{tX} \in \tilde{G}_{\Gamma_B}$, given that $\mathbb{I} \in \tilde{G}_{\Gamma_B}$. Thus we know that $e^{-tX} \in \tilde{G}_{\Gamma_B}$ and also that $e^{-ntX} \in \tilde{G}_{\Gamma_B}$ because of its group properties. We know that for all $q \in \mathbb{R}$ there exists some $n \in \mathbb{N}$ such that $e^{(nt+q)X} \in \mathcal{R}$, for n large enough – the large enough n is to counteract the q . We also know by Lemma 30 that $\tilde{G}_{\Gamma_B} \subseteq \bar{\mathcal{R}}$. Therefore

$$e^{(nt+q)X} e^{-ntX} = e^{qX} \in \bar{\mathcal{R}} \quad (7.6)$$

Therefore $\bar{\mathcal{R}}$ is the group generated by $\{X, B_i\}$, by Lemma 23, which is just G_Γ . Having proven that a neutral element of Γ implies $\bar{\mathcal{R}} = G_\Gamma$ then we just require Lemma 27 to state that $\mathcal{R} = G_\Gamma$. \square

Theorem 35 provides a sufficient result for controllability for Eq. (6.11). To this date it is not known whether it forms a necessary and sufficient condition for all Lie groups. Next we are going to link this concept with our symplectic conditions.

7.3 Symplectic neutrality

So far the discussion about control has been for any Lie group. We have seen that necessary and sufficient conditions for controllability on compact Lie groups and homogeneous systems exist. The existence of a neutral element in Γ has been shown to be a sufficient condition for controllability on G_Γ with no final outcome as to its necessity.

To solve this problem in general would be a significant advance in understanding the nature of controllability. From this point in the thesis we will narrow our focus to our particular Lie group of interest: the symplectic group $\text{Sp}(2n, \mathbb{R})$. It is hoped that exploring neutrality in this context will give some clues as to the route forwards.

In terms of the dynamics, the place where we take interest in neutral matrices is at the Lie algebra level $\mathfrak{sp}(2n, \mathbb{R})$, as discussed in Sec. 6. In that section we saw that elements of the algebra were of the form $X = \Omega H$ where H is symmetric.

First we seek to combine the algebra condition, $\Omega X + X^\top \Omega = 0$ with the neutrality condition, Def. 33. This leads to the following result:

Theorem 36. *A necessary and sufficient condition for a $X \in \mathfrak{sp}(2n, \mathbb{R})$ to be neutral is that there exists some real, symmetric, positive-definite Q such that*

$$[\Omega Q, X] = 0. \quad (7.7)$$

Proof. If X is neutral then $X^\top = -QXQ^{-1}$ for $Q > 0$ by Lemma 34. From the symplectic Lie algebra condition, Eq. (3.11), $X^\top = -\Omega X \Omega^\top$. Putting these together we prove the necessity of the statement. The argument works in reverse for sufficiency. \square

Note that elements of the Lie algebra $\mathfrak{osp}(2n, \mathbb{R})$ associated to the Lie group $\text{OSp}(2n, \mathbb{R})$ already satisfy $[\Omega, X] = 0$ and this full algebra is made up of neutral matrices. As it stands, intuition behind neutrality is lacking as so we will explore further.

To develop our insight we may explore the impact of the neutrality of X on the matrix H , as all the degrees of freedom for the X matrix lie in this. Given that X is neutral, we know from Lemma 34 that X must be diagonalisable and have pure imaginary eigenvalues. These conditions will dictate the normal form that it takes, some of which are discussed in Sec. 4. The pure imaginary property will obviously limit the eigenvalues and the diagonalisability will require that all Jordan blocks are of order 1, since Jordan blocks are not diagonalisable. The proofs of Ref. [71] relate the Jordan and Laub-Meyer blocks meaning that this also implies that we are restricted to $k = 1$, in the notation of Sec. 3.5.

Hence we find that for a neutral matrix $X \in \mathfrak{sp}(2n, \mathbb{R})$ there will exist some $T \in \text{Sp}(2n, \mathbb{R})$ such that $X = \Omega H = TYT^{-1}$ such that

$$Y = \bigoplus_{i=1}^n \begin{pmatrix} 0 & -\lambda_i \\ \lambda_i & 0 \end{pmatrix}, \quad (7.8)$$

where Y is taken from Ref. [71] and presented in the Ω -basis. Given $T\Omega T^\top = \Omega$, we find that

$$T^\top H T = T^\top \Omega^\top T Y = -\Omega Y = \bigoplus_{i=1}^n \begin{pmatrix} -\lambda_i & 0 \\ 0 & -\lambda_i \end{pmatrix}. \quad (7.9)$$

This normal form for H is inherited from that for X and corresponds to a set of n uncoupled simple harmonic oscillators. The dynamics of such a system is, of course, recursive and it is shown here that, up to the symplectic transformation T , it is the unique recursive system for symplectic dynamics. This shows that neutral elements of the Lie algebra are characterised by their relation to simple harmonic motion.

This leads to some interpretation of Eq. (7.7) and its $\mathfrak{osp}(2n, \mathbb{R})$ counterpart. Given X we seek a natural Q that will fulfill the role in Eq. (7.7), ensuring that it is positive-definite, symmetric and real. We find that taking the absolute value of the matrix ΩY and then using T to take it

out of normal form does the trick, i.e. letting

$$Q = (T^{-1})^\top |\Omega Y| T, \quad (7.10)$$

we find that Eq. (7.7) is satisfied. Q is thus constructed from X by extracting the matrix H from $X = \Omega H$, finding its normal form which will appear as decoupled simple harmonic oscillators, taking the absolute value of this matrix and then undoing the normal form diagonalisation. We note that the $\mathfrak{osp}(2n, \mathbb{R})$ condition is just the statement that this is always possible for trivial simple harmonic motion where $\lambda_i = 1$ for all i , in Eq. (7.8).

This set of neutral elements in $\mathfrak{sp}(2n, \mathbb{R})$, \mathcal{N} , forms a proper subset of the algebra. The image of \mathcal{N} under the exponential map is a also proper subset of the group, $e^{\mathcal{N}} \subset \text{Sp}(2n, \mathbb{R})$. Note that although the exponential of a simple harmonic motion Hamiltonian will be an element of $\text{OSp}(2n, \mathbb{R})$, a general exponential of a neutral element will have non-trivial singular values, i.e. squeezing, since we only require similarity to such dynamics. As such there is no immediate characterisation of this subset of the symplectic group. The question of the exact link between neutrality and control is not solved as of yet. To continue our discussion we will focus on the single-mode case to see if we can draw out any physical intuition for the dynamics of these systems.

Chapter 8

Single-mode control

Studying control problems in low dimensions can lead to physical intuition behind how such problems may be solved in general. Here, we restrict to $\mathrm{Sp}(2, \mathbb{R})$ to explore different dynamics on this group. A visualisation of the reachable set will be constructed for a single mode to see what form it takes under different types of set, Γ .

For this low dimension a set of group isomorphisms come into play which allow the importation of other control results. For our particular case we find that

$$\mathrm{Sp}(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R}) \simeq \mathrm{SU}(1, 1). \quad (8.1)$$

The bijective property of isomorphisms means that the reachable set for systems evolving on one of these groups will be in one-to-one correspondence with the reachable sets on the other groups.

Control on $\mathrm{SU}(1, 1)$ was explored in Ref. [115] where the authors prove that the existence of a neutral element in Γ , along with $\mathfrak{L}_\Gamma = \mathfrak{su}(1, 1)$, is necessary and sufficient for Γ to be controllable on $\mathrm{SU}(1, 1)$. The study of $\mathrm{SU}(1, 1)$ systems do not call for an immediate generalisation whereas, with the isomorphism to $\mathrm{Sp}(2, \mathbb{R})$ there is a much greater physical push to increase the number of modes to see if the result still holds. Clarifying paths forward for this generalisation led to the project of visualising the behaviour of uncontrollable systems as presented here. The main aim of this section is to enact a physical analysis of this result with the aim of clarifying routes forwards for the multimode case. For the remainder of this chapter we will restrict to the single-mode case, $n = 1$.

8.1 Structure of the single-mode symplectic group

The focus in this branch of control theory is always the relation between trajectories that occur in the Lie algebra and how they carve out trajectories at the group level. Examining this link can be recast physically as the examination of how changes of Hamiltonian affect changes of transformation and how to explore this in a rigorous way.

We begin by exploring a classification scheme that exists at the Lie algebra level of *ellipticity*, *hyperbolicity* and *parabolicity*. As well as having interpretations as physical transformations, these three types also relate to control theoretical properties as will be seen.

Given some $X \in \mathfrak{sp}(2n, \mathbb{R})$ we call it

- Elliptic if $\text{Tr}[X^2] < 0$,
- Parabolic if $\text{Tr}[X^2] = 0$,
- Hyperbolic if $\text{Tr}[X^2] > 0$.

This is a partition of the algebra into three classes.

The characteristic polynomial of a 2×2 matrix X is given by

$$\det[X - \lambda \mathbb{I}] = \lambda^2 - \text{Tr}[X]\lambda + \det[X]. \quad (8.2)$$

Lemma 37 (Cayley-Hamilton Theorem [127]). *The matrix X obeys its own characteristic equation.*

We have already seen that elements of $\mathfrak{sp}(2, \mathbb{R})$ have vanishing trace and so Eq. (8.2) shows that

$$X^2 = -\det[X]\mathbb{I}. \quad (8.3)$$

From this, the conditions for ellipticity, hyperbolicity and parabolicity as given above take a new form in the single-mode case, as follows: An element of $\mathfrak{sp}(2, \mathbb{R})$ is

- Elliptic if $\det[X] > 0$,
- Parabolic if $\det[X] = 0$,
- Hyperbolic if $\det[X] < 0$.

Using Eq. (8.2) and the eigenvalue condition for neutral matrices, we see that all elliptic and some parabolic, matrices are neutral. The set of parabolics that are not neutral are those that are not diagonalisable.

To complete the analysis we may now see whether a similar structure also exists at the group level. This would be self-evident if the exponential map were bijective but that is not the case

for the symplectic group for any number of modes. Indeed, as we saw in Sec. 3.2 the exponential map is not surjective for the symplectic group. We will analyse the links between the algebra and the group a little more closely.

For this we introduce a basis for the algebra:

$$K_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad K_2 = \frac{1}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K_3 = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (8.4)$$

which satisfies

$$[K_1, K_2] = -K_3, \quad [K_2, K_3] = K_1, \quad [K_3, K_1] = K_2. \quad (8.5)$$

We see that K_1 and K_2 are hyperbolic whereas K_3 is elliptic. From this we see that a general element of $\mathfrak{sp}(2, \mathbb{R})$ can be written as

$$X = \begin{pmatrix} -b & a \\ c & b \end{pmatrix}, \quad (8.6)$$

where $a, b, c \in \mathbb{R}$. Using this we find that [128]

$$e^X = \sum_{k=0}^{\infty} \frac{X^k}{(2k)!} = \sum_{k=0}^{\infty} \frac{X^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{X^{2k+1}}{(2k+1)!} = \cos[\sqrt{\det X}] \mathbb{I} + \frac{\sin[\sqrt{\det X}]}{\sqrt{\det X}} X, \quad (8.7)$$

given that $X^2 = -\det[X] \mathbb{I}$, $X^3 = -\det[X] X$, etc. Note that we allow the argument of the cosine to be imaginary leading to the hyperbolic cosine and sine. We know that $\text{Tr}[X] = 0$ and so taking the trace on the left and right hand sides provides $\text{Tr}[e^X] \geq -2$. There are elements of the group that have trace less than -2 and so we see a proof of our statement that not all elements of the symplectic group have a real logarithm in $\mathfrak{sp}(2n, \mathbb{R})$, even for a single mode. Specifically we have a condition showing that symplectic matrices whose trace is less than -2 do not possess such a logarithm. For the subset of single-shot symplectic matrices we may provide a group level partition that reflects the partition in the Lie algebra:

- Elliptic. $|\text{Tr}[S]| < 2$,
- Parabolic. $|\text{Tr}[S]| = 2$,
- Hyperbolic. $|\text{Tr}[S]| > 2$.

This provides a partition at the group level where hyperbolics are generally taken to span the two-shots as well, although given that the hyperbolic cosine is unbounded from above, this also covers many single-shot symplectics. This categorisation has a geometrical interpretation where elliptics are related to rotations of figures in a plane, parabolics to shear distortion (turning a square into a rhombus), and hyperbolics to squeezing, in the literal rather than optical sense [129].

From here we delve into a visualisation of the controllability problem in a single mode.

8.2 Controllability

We adapt Eq. (6.11) to a form that explicitly deals with symplectic matrices. Therefore our equation is of the form

$$\dot{S} = \left(A + \sum_{i=1}^m u_i(t) B_i \right) S, \quad S(0) = \mathbb{I}_{2n}, \quad \mathbf{u}(\cdot) \in \mathcal{U}_c, \quad t \geq 0. \quad (8.8)$$

Γ will be defined as before where it is made up of drift and control fields that are elements of the algebra and \mathcal{R} will denote the reachable set of the system.

We repeat many of the results from Ref. [115] with the aim of extending the result with a visual, physical analysis. There, it is proven that a neutral element in Γ , along with Lie algebra generation, is necessary and sufficient for controllability on $\text{Sp}(2, \mathbb{R})$, using the isomorphism between this group and $\text{SU}(1, 1)$. It is also shown that parabolic elements deny the possibility for Γ to generate the full algebra and so lead to trivial uncontrollability. Since all elliptic, and some parabolic, matrices are neutral, this proves that ellipticity and Lie algebra generation is necessary and sufficient for controllability on this group.

It was proven in Ref. [115] that the only nontrivial system of study for a single mode involves a single control field, i.e. $m = 1$. The low dimensionality of the problem means that in the other cases we are either trivially controllable or trivially uncontrollable in that generation is not satisfied. Thus the system we study is given by

$$\dot{S} = (A + u(t)B) S, \quad S(0) = \mathbb{I}_2, \quad \mathbf{u}(\cdot) \in \mathcal{U}_c, \quad t \geq 0. \quad (8.9)$$

The main result from Ref. [115] that helps is the following

Theorem 38. *If Γ only contains hyperbolic elements such that $\mathfrak{L}_\Gamma = \mathfrak{sp}(2n, \mathbb{R})$ then Eq. (8.9) is similar, via a symplectic transformation, to*

$$\dot{S}(t) = (-K_1 + bK_3 + u(t)K_2)S(t), \quad S(0) = \mathbb{I}_2, \quad (8.10)$$

where b is some real constant with modulus strictly less than one.

Recalling that parabolic elements cause the Lie algebra generation to fail, hyperbolic systems are the only ones which are nontrivially uncontrollable. This theorem then enables us to see how they act in general by relating them all to a specific system, Eq. (8.10). Its proof is given in

Appendix B. Eq. (8.10) becomes the new control equation with dynamical algebra

$$\tilde{\Gamma} = \{-K_1 + bK_3 + vK_2 \mid v \in \mathbb{R}\}, \quad (8.11)$$

and reachable set $\tilde{\mathcal{R}}$. This new control equation has much less freedom than Eq. (8.9) and so we can focus in on its more explicit properties.

We do this by writing down a general 2×2 element of $\text{Sp}(2, \mathbb{R})$ in the form

$$S = \begin{pmatrix} x_1 + x_3 & x_2 + x_4 \\ x_4 - x_2 & x_1 - x_3 \end{pmatrix}, \quad x_i \in \mathbb{R}. \quad (8.12)$$

Note that this general form in fact holds for any 2×2 real matrix. Using this form allows us to state the following result of Ref. [115]. We include the proof here as it leads on to the translation using the Euler decomposition which we will enact.

Theorem 39. *Consider some solution $S(t)$ to Eq. (8.10), parametrised as according to Eq. (8.12). The following function of S ,*

$$f(S) := (x_1 - x_4)^2 - (x_2 - x_3)^2 \quad (8.13)$$

has the properties $f(S) \geq 1$ and $\dot{f}(S) \geq 0$, for all choices of $u(t)$. Equality in the former case is only reached for $S = \mathbb{I}$.

Proof. Substituting the parametrisation of Eq. (8.12) into Eq. (8.10) we are provided with the set of coupled differential equations on the four parameters of S ,

$$\dot{x}_1 = \frac{1}{2}(bx_2 - x_4 - ux_3), \quad \dot{x}_2 = \frac{1}{2}(-bx_1 + x_3 - ux_4), \quad (8.14)$$

$$\dot{x}_3 = \frac{1}{2}(-bx_4 + x_2 - ux_1), \quad \dot{x}_4 = \frac{1}{2}(bx_3 - x_1 - ux_2). \quad (8.15)$$

Subtracting \dot{x}_4 from \dot{x}_1 , with a multiplication by $2(x_1 - x_4)$ provides

$$\frac{d}{dt}(x_1 - x_4)^2 = b(x_1 - x_4)(x_2 - x_3) + (x_1 - x_4)^2 + u(x_1 - x_4)(x_2 - x_3). \quad (8.16)$$

By a similar method, we have

$$\frac{d}{dt}(x_2 - x_3)^2 = -b(x_1 - x_4)(x_2 - x_3) - (x_2 - x_3)^2 + u(x_1 - x_4)(x_2 - x_3). \quad (8.17)$$

By subtracting Eq. (8.17) from Eq. (8.16) we obtain

$$\begin{aligned} \frac{d}{dt} ((x_1 - x_4)^2 - (x_2 - x_3)^2) &= 2b(x_1 - x_4)(x_2 - x_3) + ((x_1 - x_4)^2 + (x_2 - x_3)^2) \\ &= (1 - |b|) ((x_1 - x_4)^2 + (x_2 - x_3)^2) + |b| ((x_1 - x_4) - \text{sign}[b](x_2 - x_3))^2. \end{aligned} \quad (8.18)$$

Upon inspection we see that this derivative is always positive and so $\dot{f}(S) \geq 0$ for every trajectory solving Eq. (8.10). The initial value of the system is \mathbb{I} and $f(\mathbb{I}) = 1$ with $\dot{f}(S = \mathbb{I}) = 1$, from Eq. (8.18), providing $f(S) > 1$ for all other parts of the trajectory. \square

Of course it is possible that both A and B could be non-elliptic and the system still be controllable. For example letting

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (8.19)$$

we see that neither are elliptic themselves but $A + uB$ for $u < -1$ is. Given that Γ in this case has an elliptic element and that the pair generate the algebra, we immediately know that this system is controllable.

The derivation of the function f and its behaviour under hyperbolic dynamics proves that hyperbolic systems are uncontrollable. This is due to the monotonic behaviour of f which precludes it taking certain values which are the values for certain symplectic matrices, thus rendering them unreachable. This leads to a necessary and sufficient condition for controllability on $\text{Sp}(2, \mathbb{R})$ which is the main result of Ref. [115]. The continuation here is to characterise the reachable set of hyperbolic systems.

8.3 Visualisation

We have seen that a neutral element of Γ is necessary for the controllability of systems on a single mode. As has been mentioned, it is unknown whether this property extends to multimode systems. Rather than attempting this directly it is interesting to ask what it is about neutral elements that makes them necessary.

At this point it may seem obvious: the set of neutral elements of $\mathfrak{sp}(2, \mathbb{R})$ is unavailable and so it may be thought that the set of elements that these exponentiate to should be obviously inaccessible. However, there is no apparent reason why we should not be able to reach a passive element along two or more trajectories generated by hyperbolic elements. For example the matrices

$$S_1 = \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 4 & 10 \end{pmatrix}, \quad \text{and} \quad S_2 = \begin{pmatrix} -2 & -11 \\ 1 & 5 \end{pmatrix} \quad (8.20)$$

are both hyperbolic, in that $\text{Tr}[S_1] = \text{Tr}[S_2] = 3$, whereas the trace of their product is $3/4$ and is therefore an elliptic element. It is this concatenation of elements that causes some of the deeper subtleties of control research.

8.3.1 Bounding the reachable set

The controllability proof of the previous section uses the fact that any hyperbolic system associated with Eq. (8.9) may be transformed into Eq. (8.10). This transformation was then used to derive an ever-growing function that was used to prove the uncontrollability of such a system. It would be interesting to use this function to find a more intuitive bound for the reachable set of Eq. (8.10).

Theorem 40. *Given Eq. (8.9) with its associated Γ only containing hyperbolic elements, the reachable set \mathcal{R} does not contain any element that is symplectically similar to an element of $\text{OSp}(2, \mathbb{R})$, excepting \mathbb{I} .*

Proof. First we show that no element of $\tilde{\mathcal{R}}$ exists such that

$$S = TR_\theta T^{-1}, \quad (8.21)$$

where $T \in \text{Sp}(2, \mathbb{R})$, $R_\theta \in \text{SO}(2) = \text{OSp}(2, \mathbb{R})$. We will proceed by contradiction. Consider the existence of some $S \in \tilde{\mathcal{R}}$ that satisfies Eq. (8.21). This implies that

$$S^m \in \tilde{\mathcal{R}} \quad \forall m \in \mathbb{N}, \quad (8.22)$$

because $\tilde{\mathcal{R}}$ has monoidal structure. Furthermore, note that

$$\begin{aligned} \|S^m - \mathbb{I}\| &= \|T(R_\theta^m - \mathbb{I})T^{-1}\| \\ &\leq \|T\| \|T^{-1}\| \|R_\theta^m - \mathbb{I}\|, \end{aligned} \quad (8.23)$$

using the Euclidean norm,

$$\|M\| := \sqrt{\text{Tr}[M^T M]}. \quad (8.24)$$

The time-independence of T means that $\|T\| \|T^{-1}\|$ is constant. Since $R_\theta \in \text{SO}(2)$, there must exist some m such that

$$\|R_\theta^m - \mathbb{I}\| < \varepsilon, \quad \forall \varepsilon > 0 \quad (8.25)$$

and so there exists m such that

$$\|S^m - \mathbb{I}\| < \varepsilon, \quad \forall \varepsilon > 0. \quad (8.26)$$

Theorem 39 provides a lower bound $f(S) > 1$ which is increasing at the start of the dynamics. After some finite time of evolution the value of $f(S)$ will be some value greater than 1. We can follow S^m along and get beneath this value, violating $\dot{f}(S) \geq 0$. Therefore, no S satisfying Eq. (8.21) can be an element of $\tilde{\mathcal{R}}$. The reachable set \mathcal{R} corresponding to Eq. (8.9) is similar under a symplectic transformation to $\tilde{\mathcal{R}}$. We know that no element of $\tilde{\mathcal{R}}$ is similar under a symplectic transformation to any member of $\text{SO}(2) \setminus \mathbb{I}$. This property is therefore inherited by \mathcal{R} . \square

8.3.2 Euler translation

The visualisation of the reachable set is greatly improved if we change coordinates and introduce the Euler decomposition. The difficulty that comes with using this decomposition is that it is not unique. However, its advantage comes by virtue of its physical interpretation being clear, as discussed in Sec. 3. Uniqueness can, however, be given by an appropriate restriction of the ranges of the parameters in the three matrices. The Euler decomposition in a single mode takes the form,

$$S = R_\theta Z R_\phi, \quad (8.27)$$

where $Z = \text{diag}(z, 1/z)$ and $z > 1$, as well as,

$$R_\theta = \begin{pmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{pmatrix}, \quad (8.28)$$

and similarly for R_ϕ . We see that the group $\text{OSp}(2, \mathbb{R})$ is identical to $\text{SO}(2)$. We may make this decomposition unique by restricting the angle parameters to

$$-\pi + \theta_0 \leq \theta < \pi + \theta_0, \quad -\frac{\pi}{2} + \phi_0 \leq \phi < \frac{\pi}{2} + \phi_0, \quad (8.29)$$

where θ_0 and ϕ_0 are arbitrary numbers that fix the centres of the ranges. The proof of this is given in Appendix C. The case of $Z = \mathbb{I}$ is special in that the equation reduces to the product of two passive symplectics which are elements of $\text{SO}(2)$ which only requires a single parameter to describe. Hence we simply fix ϕ and let θ vary over the 2π range. Note that this restriction for the Euler decomposition introduces branch-cut style effects into the physical analysis of the reachable set.

We begin with two expressions for $S \in \text{Sp}(2, \mathbb{R})$, the first from Eq. (8.12) and the latter from

the Euler decomposition:

$$S = \begin{pmatrix} x_1 + x_3 & x_2 + x_4 \\ x_4 - x_2 & x_1 - x_3 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{\cos[\theta] \cos[\phi]}{z} - z \sin[\theta] \sin[\phi] & -\frac{\cos[\theta] \sin[\phi]}{z} - z \sin[\theta] \cos[\phi] \\ \frac{\sin[\theta] \cos[\phi]}{z} + z \cos[\theta] \sin[\phi] & -\frac{\sin[\theta] \sin[\phi]}{z} + z \cos[\theta] \cos[\phi] \end{pmatrix}. \quad (8.30)$$

Equating the two expressions we find that f in the new coordinates is

$$f(S) = \cos[2\theta] \cos[2\phi] - \sin[2\theta] \left(\frac{1}{2} \left(z^2 + \frac{1}{z^2} \right) \sin[2\phi] - \frac{1}{2} \left(z^2 - \frac{1}{z^2} \right) \right). \quad (8.31)$$

For the proofs in the following two lemmas, we define

$$\delta := \frac{1}{2} \left(z^2 + \frac{1}{z^2} \right) \sin[2\phi] - \frac{1}{2} \left(z^2 - \frac{1}{z^2} \right) - \sin[2\phi]. \quad (8.32)$$

so that

$$f(S) \equiv \cos[2(\theta - \phi)] - \delta \sin[2\theta]. \quad (8.33)$$

Lemma 41. *If $f(S) > 1$ then $\sin[2\theta] > 0$.*

Proof. Let $\sin[2\theta] \leq 0$. Eq. (8.33) tells us that $f(S) > 1$ only has solutions if $\delta > 0$, i.e. that

$$\left(z^2 + \frac{1}{z^2} - 2 \right) \sin[2\phi] > z^2 - \frac{1}{z^2}. \quad (8.34)$$

For the full range of z for which this inequality holds, we note that $z^2 + 1/z^2 - 2 \geq 0$ and that $\sin[2\phi] \leq 1$. Therefore any z satisfying the previous inequality will also satisfy

$$z^2 + \frac{1}{z^2} - 2 > z^2 - \frac{1}{z^2}, \quad (8.35)$$

for some θ . This implies that $z < 1$ which is ruled out by the restriction, set at Eq. (8.27). By contradiction the statement is proven. \square

Lemma 42. *If $f(S)$ is lower bounded such that $f(S) > d$, where $d \geq 1$, then*

$$z > \sqrt{\frac{d+1}{2}}. \quad (8.36)$$

Proof. Given the bound on $f(S)$, by Lemma 41 we may state that $\sin[2\theta] > 0$. Now, Eq. (8.33) tells us that $f(S) > d$ only has solutions if $\delta < 1 - d$, i.e. that

$$\left(z^2 + \frac{1}{z^2} - 2 \right) \sin[2\phi] < z^2 - \frac{1}{z^2} - (d - 1). \quad (8.37)$$

Again $z^2 + \frac{1}{z^2} - 2 \geq 0$ and $\sin[2\phi] \leq 1$ so any z satisfying this inequality will also satisfy

$$z^2 + \frac{1}{z^2} - 2 < z^2 - \frac{1}{z^2} + 1 - d, \quad (8.38)$$

for some θ . This implies that

$$z^2 > \frac{d+1}{2}, \quad (8.39)$$

completing the proof. \square

We see that as $f(S)$ does not decrease which provides a nondecreasing lower bound on z , the squeezing parameter.

8.3.3 Visualising the reachable set

The analysis so far has provided us with some general bounds that *any* single-mode, uncontrollable system with $\mathfrak{L}_\Gamma = \mathfrak{sp}(2, \mathbb{R})$, must obey. The main analytical result in terms of visualisation is that no element symplectically similar to a passive matrix may be reached – apart from the identity which is the initial position on the group. Furthermore we saw that z will always be non-decreasing.

Eq. (8.10) provides the form to which all hyperbolic systems may be transformed. In terms of drift and control fields it can be written as

$$A = \begin{pmatrix} 0 & -(1+b) \\ -(1-b) & 0 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (8.40)$$

where $b \in \mathbb{R}$ and $|b| < 1$ to make it hyperbolic. Note that we have dropped the factor of $1/2$ since this can be absorbed by the time variable. Through Eq. (2.4) which relates elements of the Lie algebra to Hamiltonian operators on the phase space, these matrices correspond to the Hamiltonians

$$\hat{H}_A = \frac{(1-b)\hat{x}^2 - (1+b)\hat{p}^2}{2}, \quad \hat{H}_B = -\frac{\hat{x}\hat{p} + \hat{p}\hat{x}}{2}. \quad (8.41)$$

Using the conditions of Sec. 8.1, A can be shown to be either parabolic, hyperbolic or elliptic if $|b|$ is, respectively, equal to, less than, or greater than 1. We provide an analysis of the hyperbolic case for $b = 0$.

The non-uniqueness of the Euler decomposition is again going to cause some subtleties with a graphical illustration of the reachable set and so some time will be required to unpick this. If we consider a three-dimensional plot where the axes label the z , θ and ϕ parameters then every point on the plot will correspond, uniquely to a symplectic matrix. However, the reverse is not true. To compensate for this we return to the angle bound of Appendix C to restrict the latter two axes. This creates a one-to-one mapping for all matrices with $z(S) \neq 1$. However, we then find

degeneracies at $z = 1$ because we will have two angles deciding a point on the one-dimensional passive line.

One solution to this problem would be to alter the cubic geometry of the plot and pinch it into an egg-timer shape. However, since we already know from Theorem 40 that only the identity element of the passive set is going to be reachable, this seems a little drastic. As a result it remains for us to uncover the natural point on the passive plane which will correspond to the identity, and then we can look at the results of the numerics.

Recalling that the angles are bounded such that we must choose appropriate bounds. Following Eq. (8.29), we set these to be

$$-\pi \leq \theta < \pi, \quad 0 \leq \phi < \pi. \quad (8.42)$$

To find the natural point for the identity on this passive plane we consider e^{At} and then take the limit as $t \rightarrow 0$, given our angle bounds. This will give the ‘Euler decomposition of identity’. Letting $t = \frac{1}{n}$ where $n \in \mathbb{N}$ we have that

$$\begin{aligned} \exp \left[\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \frac{1}{n} \right] &= \left(R_{-\frac{3\pi}{4}} \begin{pmatrix} \frac{1}{e} & 0 \\ 0 & e \end{pmatrix} R_{\frac{3\pi}{4}} \right)^{\frac{1}{n}} \\ &= R_{-\frac{3\pi}{4}} \begin{pmatrix} \frac{1}{e} & 0 \\ 0 & e \end{pmatrix}^{\frac{1}{n}} R_{\frac{3\pi}{4}}. \end{aligned} \quad (8.43)$$

Taking the limit we find that the natural decomposition for the identity is

$$\mathbb{I} = R_{-\frac{3\pi}{4}} R_{\frac{3\pi}{4}}. \quad (8.44)$$

i.e. the point $(-3\pi/4, 3\pi/4)$. This analysis simply allows us to represent the reachable set of the system corresponding to Eq. (8.40) on a three dimensional plot without worrying about degeneracies. Given the location of this initial point, we may add the bound of Lemma 41 which additionally restricts us to $-\pi < \theta < -\pi/2$.

Now that the last of the analytical bounds have been put in place, we may proceed with a computational visualisation of the reachable set of the system corresponding to Eq. (8.40).

The computational investigation was carried out using QuTiP, an open source python library that simulates quantum dynamics [130, 131]. The example hyperbolic system in Eq. (8.40) was taken and the question was asked: which symplectics S_{target} can be reached at a given time T ? Note that this is slightly different to finding the reachable set, as that would be *for all* T . However, specifying the time allows us to see more detail in exactly how the system is behaving.

Splitting T into Q time slices of duration Δt , the control function $u(t)$ was assumed to

be constant during each of these Q time intervals, with value $u(t_k)$ for the k th interval. The generators are of the form

$$X_k = A + u(t_k)B, \quad u(t_k) \in \mathbb{R}, \quad (8.45)$$

which exponentiate to a set of symplectic matrices

$$S_k = e^{X_k \Delta t}, \quad (8.46)$$

to give the overall evolution by the product,

$$S(T) = S_Q S_{Q-1} \cdots S_k \cdots S_2 S_1. \quad (8.47)$$

Of course, to test for a specific target state with zero deviation would not be feasible numerically. Therefore it is necessary to introduce the infidelity via the Frobenius norm

$$\varepsilon := \lambda \text{Tr}[(S(T) - S_{\text{target}})^\top (S(T) - S_{\text{target}})], \quad (8.48)$$

where $\lambda = 1/8$ for a 2×2 matrix.

To find the $u(t)$ that enacts this target symplectic the GRAPE algorithm [132] was used in QuTiP. Ensuring that ε was minimal required the L-BFGS-B method in the *scipy* optimization function. The target is said to have been reached if $\varepsilon < 10^{-3}$. The search terminates if a local minimum is found or the system runs over a particular time threshold. The number of time slices was set to $Q = 10$. The set of target symplectics was discretised by only attempting points at intervals of length $\pi/2$ in the angular directions θ and ϕ , and 10 logarithmically equal intervals between $z = 1$ and $z = 100$, the latter being an arbitrary upper bound. The simulation was repeated for combinations of $b = \{0.0, \pm 0.5, \pm 0.9, \pm 0.99, \pm 1.01, \pm 1.1, \pm 1.5\}$ and $T = \{0.1, 0.5, 1, 2, 3, 4, 5, 7, 10, 20, 50, 100\}$, but a subset of these is displayed.

The primary point of the simulation was to find out the behaviour of the hyperbolic system to see if we could characterise it in some physical way. Fig. 8.1 shows the case for $b = 0$ and $T = 5$ and $T = 1$. The $T = 5$ basket lies inside that for $T = 1$ suggesting that in our system we require small times for small values of squeezing, but also don't require large times for large values of squeezing, presumably an effect of our unbounded controls. Furthermore it suggests that there is a general 'rip-tide' dragging the squeezing value z to be ever greater as time progresses. As proven, we see that the reachable set is centred around $(\theta, \phi) = (-3\pi/4, 3\pi/4)$, with a θ -bound of $-\pi < \theta < -\pi/2$. There is an additional ϕ bound, $\pi/2 < \phi < \pi$, that was not predicted in the previous analysis.

This provokes the surmise that it is squeezing that acts as the block to controllability on these systems. We noted before that Γ can be trivially uncontrollable by not generating the algebra.

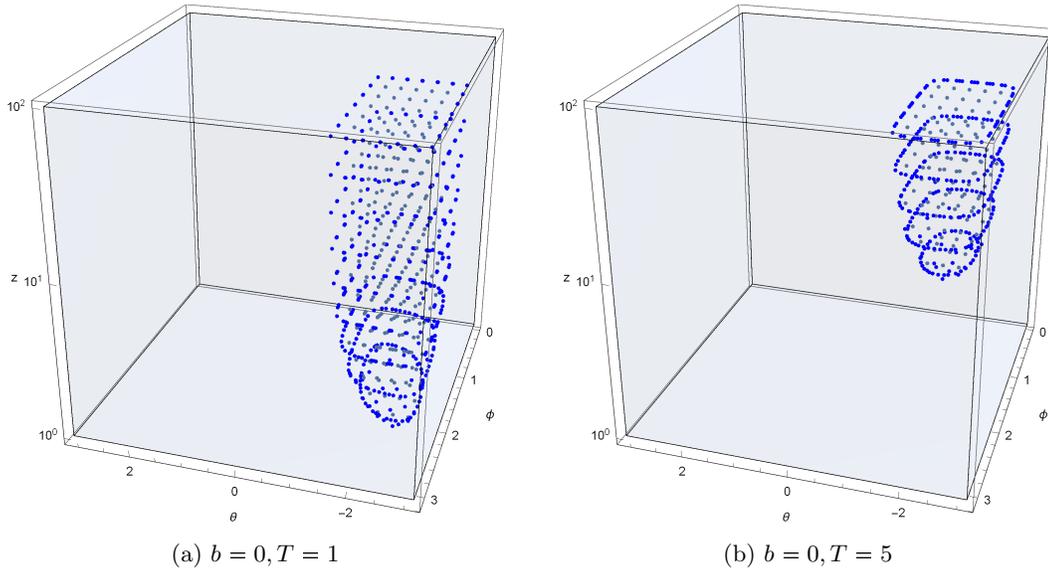


Figure 8.1: The reachable set of an example single-mode hyperbolic system with $b = 0$ is represented in a cubic plot for different times of evolution.

In cases where it does, however, we see here a mechanism by which it can avoid being able to access the whole space. The existence of a neutral element of Γ would allow us to control so that the passive elements of the group were reachable. We see a ‘rip-tide’ squeezing effect blocking us from accessing a sufficiently large angle range.

Maintaining the time at $T = 5$ but varying b so that it is close to the ‘controllable’ boundary at $b = -0.99$ - i.e. the point where Γ will contain a neutral element, we see in Fig. 8.2 that the size of the reachable set increases. This provokes the idea that even for these uncontrollable hyperbolic systems, there is still a notion of varying amounts of control. The b parameter seems to be acting as a parameter controlling the strength of the rip-tide effect.

In summary, we see that when Γ contains no neutral element, but still generates the algebra, we get unbounded squeezing within a confined angular region for single-mode systems. A generalisation of this visualisation to two modes would require some sophisticated treatment. However, we have provided some intuition for what the behaviour may look like and what such a proof may entail.

8.4 Spacetime analogy

The baskets explored in the previous section bear a certain resemblance to the lightcones of relativity theory. In this section we offer a brief investigation into how far this analogy holds.

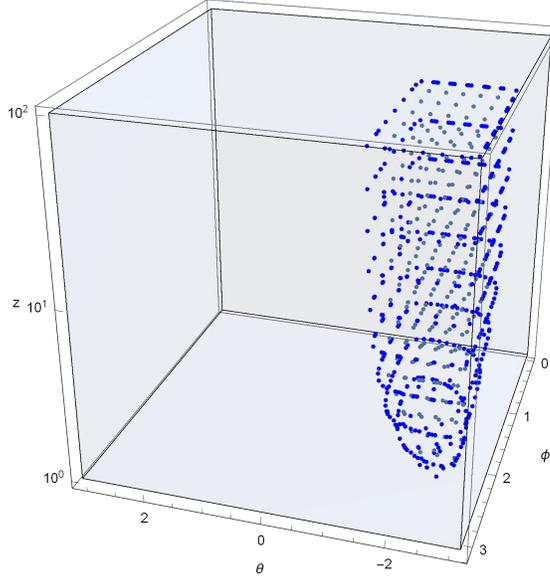


Figure 8.2: The reachable set of an example single-mode hyperbolic system is represented in a cubic plot for $b = -0.99$ and $T = 5$.

Referring to Fig. 8.1 we see that the squeezing value z acts as the parameter analogous to coordinate time and T to proper time. We have already seen in Eq. (8.43), when we set u to zero the effect is pure squeezing. Thus we see that u acts as a parameter analogous to velocity. The drift field, in this particular example, corresponds to a drift through time, without any motion. Thus we see that there is a minimum ‘speed’ along the z axis for our system to take, similar to the way in which c acts in relativity.

In fixed time T we are able to achieve a greater coordinate time. This must only be possible by varying u and hence by ‘moving’. Thus we see the twin-paradox situation where Bob, the moving party, can meet Alice at some ‘coordinate time’ z , where Bob has a smaller proper time and is therefore younger. Of course, we see that $u = 0$ is acting as a preferred rest frame.

Given A and B , as in Sec. 8.3.3, we look at the two cases where, firstly, control is not used, $X_1 = A$ and secondly where it is $X_2 = A + uB$. Exponentiating these in time to get S_1 and S_2 we then look to see which achieves a higher level of squeezing. A good measure of squeezing, as we have seen in Sec. 3.4 uses the maximum eigenvalue of the matrix in product with its transpose. We see that

$$\max \text{eig}[S_1 S_1^T] = e^{2t}, \quad \max \text{eig}[S_2 S_2^T] = e^{2t\sqrt{1+u^2}} \quad (8.49)$$

showing that altering u , i.e. ‘moving’ always increases our distance in z , or ‘coordinate time’.

To take the analogy further we notice that we are dealing in a spatially compact space but

unbounded in time. Squeezing is the non-compact and ‘timelike’ aspect of the symplectic group here. Neutral matrices are those that allow us to either unsqueeze after squeezing, or never squeeze at all. The former set will hence be analogous to closed curves in spacetime.

The analogy is not complete in that we are not dealing with a single lightcone structure but a set of interlaced ones. We see this by considering the backwards-time reachable set. To go from e^{Xt} to e^{-Xt} we enact the transformation

$$S \rightarrow S^{-1} = \Omega S^\top \Omega^\top. \quad (8.50)$$

Given the Euler decomposition we see that this is the same as

$$R_\theta Z R_\phi \rightarrow R_{-\phi - \frac{\pi}{2}} Z R_{-\theta + \frac{\pi}{2}}. \quad (8.51)$$

Consider what this would mean for our reachable set depicted in Fig. 8.1a, confined to the region $-\pi < \theta < -\pi/2$ and $\pi/2 < \phi < \pi$. This transformation is enacted by a mirror reflection in the $\theta = -\phi$ line, then a shift right by $\pi/2$ and down by $\pi/2$, if looking from above. Overall, this is a transformation that takes us from one ‘lightcone’ to one going back in time. Fig 8.3 shows what this set of transformations looks like if we begin in any of the squares of Fig 8.1, looking from above and with slightly extended angle ranges.

Obviously this analysis focuses on our particular example system for a particular choice of A , B and b . For this system we have made the statement that setting $u \neq 0$ always provides a faster route to some squeezing value. This is obviously not true in general as we could imagine a new system with drift field $(A + u'B)$ and control field B , where u' is some real number, so that the whole element is $A + u'B + uB$. In this situation the reachable set will clearly be identical to the one we have before but it will not be the case that $u \neq 0$ is better for squeezing. As a result we see that there is a form of preferred rest frame for these single-mode systems. The choice of b , however, was not shown to significantly alter the behaviour observed here in terms of the shape of the reachable set.

This analogy is quite satisfying, although obviously not complete. However it provides an illustrative way of seeing the structure of the reachable set and its properties. It would be interesting to find out if these properties remain in multimode systems where we would then have multiple ‘time’ directions. The main point, as far as control goes, is whether the same ‘rip-tide’ effect of ‘time’ exists for these systems.

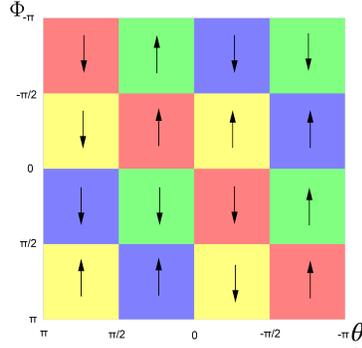


Figure 8.3: Imagine we begin in a lightcone in one of the angle-bounded regions, depicted by the squares. Enacting $\theta \rightarrow -\phi - \pi/2$ and $\phi \rightarrow -\theta + \pi/2$ will take squares of a single colour to squares of the same colour with the opposite arrow direction, indicating opposite time direction. A transformation of $\theta \rightarrow \theta - \pi$ and $\phi \rightarrow \phi - \pi$ takes us to a square of exactly the same time as this transformation does nothing to the lightcone.

8.5 Ever-growing function

The underlying cause of the uncontrollability of non-neutral systems was the continual growth of the function $f(S)$, as shown in Theorem 39. The natural thing to wonder, from this, is whether this is a general behaviour of n -mode systems: that the lack of a neutral element in Γ causes some parameter to continually grow, causing the remainder of the group to be unreachable.

Consider a bilinear form on $S \in \text{Sp}(2n, \mathbb{R})$, given as

$$g_Q(S) = \text{Tr}[S^T Q S], \quad (8.52)$$

where Q is some symmetric, positive-definite, real matrix. Referring to Eq. (8.8) we find that

$$\dot{g}_Q(S) = \text{Tr}[S^T F(\mathbf{u}) S], \quad (8.53)$$

where

$$F(\mathbf{u}) := \left(A + \sum_{i=1}^m u_i B_i \right)^T Q + Q \left(A + \sum_{i=1}^m u_i B_i \right). \quad (8.54)$$

Setting $\dot{g}_Q(S) > 0$ for all \mathbf{u} and all S means that

$$\begin{aligned} \text{Tr}[S^T F(\mathbf{u}) S] &= \sum_{j=1}^{2n} \mathbf{e}_j^T S^T F(\mathbf{u}) S \mathbf{e}_j \\ &= \mathbf{e}_j'^T F(\mathbf{u}) \mathbf{e}_j' \\ &> 0 \quad \forall u, \end{aligned} \quad (8.55)$$

where $\mathbf{e}'_j = S\mathbf{e}_j$. This is true if and only if $F(\mathbf{u})$ is positive-definite for all \mathbf{u} . Thus we see that the continued growth of $g_Q(S(t))$ implies that $A + \sum_{i=1}^m u_i B_i$, the element of Γ , can not be made neutral with Q , referring to Item 1 of Lemma 34.

Note that Eq. (8.13) is not in the form of Eq. (8.52) and so this does not relate to the results before. Furthermore, this analysis did not lead to a general result but provided a tantalising, albeit possibly superficial link, between an eternal function growth as in Ref. [115], and neutrality.

Chapter 9

Open Gaussian systems

In quantum dynamics, as soon as some of the degrees of freedom are labeled as an environment and the rest as the system, then we are dealing with open systems. As soon as the dynamics of interest is the system alone, without considering the environment then the discussion inevitably moves away from the old unitary dynamics. This simple model is the bedrock for a vast branch of quantum theory which is vital for the advancement of quantum technologies, due to the impact of noisy environments on delicate quantum experiments.

The new set of transformations that takes us from quantum density matrices to quantum density matrices is the set of completely-positive trace-preserving (CPT) maps [33]. CPT maps are also known as quantum channels because “they play the same role in quantum information theory as classical channels (stochastic maps) play in classical information theory” [133].

Along with the maps themselves we often seek for dynamics that enact them. There is a host of equations that are good at describing a wide range of different regimes and models for the evolution of quantum states. One key approximation states that our system should be very weakly coupled to the environment, or bath. This is known as the Born approximation. The second idea is that of having no memory. For a discrete set of time steps this is the notion that the proceeding dynamics should only be contingent on the previous step, without a memory stretching back to earlier times. This is referred to as the Markovian approximation. When we make these approximations together we refer to the Born-Markov regime [134].

Within this regime we seek the most general dynamical equation for which the evolution is completely-positive and trace-preserving. This question was dealt with in the 1970s by Vittorio Gorini, Andrzej Kossakowski and E. C. George Sudarshan [135] and Göran Lindblad [136], providing what is known as the GKSL equation. The GKSL equation provides the general setup when we want to model a quantum system in the presence of an infinitely large bath which is not being affected by the evolution of the system. Intuitively, this is a natural thing to do as the

degrees of freedom of the environment are usually taken to be much more numerous than those of the system.

The study of open quantum systems reaches back to the early days of quantum theory after von Neumann’s formulation of the density matrix [137]. The link to thermodynamics has been explored in an attempt to use quantum mechanics to derive analogues of the well known classical thermodynamical laws [138, 139]. Other lines of research involve continuous measurement of the environment that makes the system dynamics stochastic, which was first studied in 1989 by Viacheslav Belavkin [140].

Due to the ease with which noise can be added to Gaussian dynamics, Gaussian states are prominent in the discussion of open systems [141]. Following on from Belavkin’s work, the continuous monitoring of Gaussian states has been explored extensively [142–144]. A lot of research has been done to quantify the evolving behaviour of entanglement in such circumstances because it is often considered as the key resource for quantum technology that the environment impinges on so heavily [145].

Our focus will be on the Born-Markov regime for a Gaussian system and environment. From this point onward in this work, we consider control in a very different regime to that seen in the previous sections. The ability of symplectic control to affect *open* dynamics is now the point of interest and so the Lie algebraic relation to the group acts as a backdrop and is no longer the point of mathematical interest. We are now interested in how closed control affects these open dynamics.

9.1 Input-output formalism

Following the treatment of Ref. [146], we consider a Gaussian system in the regime where it is weakly coupled to the environment such that no information passing from the system to the environment may ever return. A way to model this idea is using the ‘input-output’ formalism with the ‘white-noise’ condition [147]. This is where we consider a system that is interacting with a continuous flow of identical modes from the environment. At each instant of time an interaction occurs between bath modes and system modes, after which the bath modes are refreshed and a new interaction happens at the later time. This is illustrated in Fig. 9.1. We see that the refreshment of the bath modes at each instant in time encodes the idea that the bath is never affected by the changes in the system, and also that no memory effects come into play because any former interaction is wiped out in the refreshment.

For Gaussian systems the situation is modeled with a system of n modes and a bath with an infinite set of m incoming modes. The ‘incoming’ bath modes are represented by the vector $\hat{\mathbf{r}}_B(t)$ where the label t is attached to each set of m modes in the continuum that models the bath. The system interacts with the bath mode labeled t at time t , giving t something of a twofold

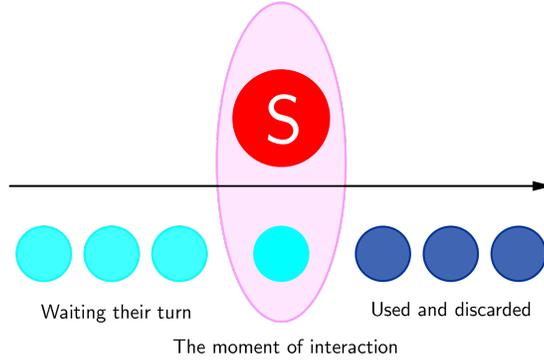


Figure 9.1: The input-output formalism describes a stream of bath-modes interacting momentarily with the system, labeled S , and then being discarded. Although in many setups these outgoing modes would in fact be measured instead.

meaning. After the interaction has occurred it disappears into the bath.

The study of Gaussian states begins with the commutation relation, as given in Eq. (2.2). Each ingoing bath mode is set up to be completely independent of the other modes which provokes the new relation,

$$[\hat{\mathbf{r}}_B^T(t), \hat{\mathbf{r}}_B^T(t')] = i\Omega\delta(t - t')\hat{\mathbb{1}}, \quad (9.1)$$

where $\delta(\cdot)$ is the Dirac delta function. Referring to Eq. (2.37), which relates the density matrix to the covariance matrix, and setting first moments to zero, the covariance matrix of each of the bath modes will satisfy

$$\text{Tr}[\{\hat{\mathbf{r}}_B^T(t), \hat{\mathbf{r}}_B^T(t')\}\hat{\rho}_B] = \sigma_{B'}\delta(t - t'), \quad (9.2)$$

where $\hat{\rho}_B$ denotes the bath state and $\sigma_{B'}$ is a covariance matrix of our choosing, denoted with a prime because Eq. (9.2) is not yet in the standard form. The important part of the formalism is that it reduces the infinite bath degrees of freedom to a new picture where we have an infinitely repeated number of finite bath systems, which interact continuously with the system, at their given time.

By integrating the mode-operators over an interval of time we define

$$\delta\hat{\mathbf{r}}_B(t) := \int_t^{t+\delta t} \hat{\mathbf{r}}_B(s)ds, \quad (9.3)$$

which again provides new versions of the commutation relation and covariance matrix form:

$$\begin{aligned} [\delta\hat{\mathbf{r}}_B(t), \delta\hat{\mathbf{r}}_B^\top(t)] &= i\Omega \int_t^{t+\delta t} ds = i\Omega\delta t, \\ \text{Tr}[\{\delta\hat{\mathbf{r}}_B(t), \delta\hat{\mathbf{r}}_B^\top(t)\}\hat{\rho}_B] &= \sigma_{B'} \int_t^{t+\delta t} ds = \sigma_{B'}\delta t. \end{aligned} \quad (9.4)$$

As δt becomes incremental it is possible to replace $\delta\hat{\mathbf{r}}_B \rightarrow \hat{\mathbf{r}}_B\delta t$ and so

$$\begin{aligned} [\hat{\mathbf{r}}_B(t), \hat{\mathbf{r}}_B^\top(t)]\delta t &= i\Omega, \\ \text{Tr}[\{\hat{\mathbf{r}}_B(t), \hat{\mathbf{r}}_B^\top(t)\}\hat{\rho}_B]\delta t &= \sigma_{B'}, \end{aligned} \quad (9.5)$$

after dividing by δt . Taking the limit as $\delta t \rightarrow dt$ and defining $\hat{\mathbf{r}}_B(t)dt = \hat{\mathbf{r}}_{B'}dv$ where $dv^2 = dt$, we find

$$\begin{aligned} [\hat{\mathbf{r}}_{B'}, \hat{\mathbf{r}}_{B'}^\top] &= i\Omega, \\ \text{Tr}[\{\hat{\mathbf{r}}_{B'}, \hat{\mathbf{r}}_{B'}^\top\}\hat{\rho}_B] &= \sigma_{B'}. \end{aligned} \quad (9.6)$$

The expression ‘ $dv^2 = dt$ ’ is prevalent in situations with white noise, i.e. where we wish to express the system/bath-mode interaction as a delta function in time. The input-output formalism allows us to tuck away the incremental time elements to arrive at a new set of modes that have the standard interpretation in the Gaussian formalism, without any continuous stream. The quantum state of the bath $\hat{\rho}_B$ remains the same but we have transformed the modes we use to probe the state. From here we may begin to use the evolution rules as set up in Sec. 2.

Given our interacting system and bath, a general Hamiltonian will be of the form $\hat{H}_S + \hat{H}_B + \hat{H}_I$, which denote the system, bath and interaction Hamiltonians respectively. As the interaction is our only concern, we shall set $\hat{H}_S = \hat{H}_B = 0$. Furthermore we set the remaining interaction Hamiltonian to be of quadratic form. This renders us with the most general interaction Hamiltonian

$$\hat{H}_I dt = \frac{1}{2}\hat{\mathbf{r}}_{SB}^\top(t)H_I\hat{\mathbf{r}}_{SB}(t)dt = \frac{1}{2}\hat{\mathbf{r}}_{SB}^\top(t) \begin{pmatrix} 0 & C \\ C^\top & 0 \end{pmatrix} \hat{\mathbf{r}}_{SB}(t)dt, \quad (9.7)$$

where $\hat{\mathbf{r}}_{SB}^\top(t) = (\hat{\mathbf{r}}_S^\top, \hat{\mathbf{r}}_B^\top(t))$. The evolution of $\hat{\rho}_{SB}$ is of the form

$$\hat{\rho}_B(t+dt) = e^{\hat{H}_I dt} \hat{\rho}_{SB} \left(e^{\hat{H}_I dt} \right)^\dagger. \quad (9.8)$$

By changing the mode definition, i.e. the way in which we probe $\hat{\rho}_B$, we may replace

$$\hat{H}_I dt = \frac{1}{2}\hat{\mathbf{r}}_{SB'}^\top H_I \hat{\mathbf{r}}_{SB'} dv, \quad (9.9)$$

where $\hat{\mathbf{r}}_{SB'}^\top = (\hat{\mathbf{r}}_S^\top, \hat{\mathbf{r}}_{B'}^\top)$. Because these new modes allow us to use the Gaussian formalism, we

are able to write down the evolution of

$$\sigma_{SB'} = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma_{B'} \end{pmatrix} \quad (9.10)$$

as

$$\sigma_{SB'}(t + dt) = e^{\Omega H_I dv} \sigma_{SB'}(e^{\Omega H_I dv})^\top. \quad (9.11)$$

Recalling that $dv^2 = dt$, we expand the exponentials to first order in dt ,

$$e^{\Omega H_I dv} = \mathbb{I} + \Omega H_I dv + \frac{(\Omega H_I)^2}{2} dt + o(dt). \quad (9.12)$$

Substituting into Eq. (9.11) we have

$$\sigma_{SB'}(t + dt) = \sigma_{SB'} + M_1 dv + M_2 dt + o(dt), \quad (9.13)$$

where

$$M_1 = \begin{pmatrix} 0 & \Omega C \sigma_{B'} + \sigma C \Omega^\top \\ \sigma_{B'} C^\top \Omega^\top + \Omega C^\top \sigma & 0 \end{pmatrix}, \quad (9.14)$$

and

$$M_2 = \begin{pmatrix} \frac{\Omega C \Omega C^\top \sigma + \sigma C \Omega C^\top \Omega}{2} + \Omega C \sigma_{B'} C^\top \Omega^\top & 0 \\ 0 & \frac{\Omega C^\top \Omega C \sigma_{B'} + \sigma_{B'} C^\top \Omega C \Omega}{2} + \Omega^\top C^\top \sigma C \Omega \end{pmatrix}. \quad (9.15)$$

By tracing out the bath modes and reforming Eq. (9.13) into a differential equation we derive

$$\dot{\sigma} = A\sigma + \sigma A^\top + D, \quad (9.16)$$

where

$$A = \frac{\Omega C \Omega C^\top}{2}, \quad \text{and} \quad D = \Omega C \sigma_{B'} C^\top \Omega^\top. \quad (9.17)$$

This equation is the most general dynamical equation one can write for the evolution of Gaussian states in the Born-Markov regime [33]. One can also show that the reintroduction of a further Hamiltonian \hat{H}_S , determining system evolution, induces a transformation

$$A \rightarrow A + \frac{\Omega H_S}{2}, \quad (9.18)$$

where $\hat{H}_S = \frac{1}{2} \hat{\mathbf{r}}_{SB'}^\top H_S \hat{\mathbf{r}}_{SB'}$.

9.2 Lossy channels

Now that we have Eq. (9.16) we can decide what interaction Hamiltonian we might like to consider, between system and bath. A host of different interactions create a good model for processes that occur in nature. One that is of interest to us is called a *lossy channel*. Also known as the attenuation channel, it is used to model optical communication in lossy fibers [148].

The corresponding interaction Hamiltonian for such an evolution is given as a beam-splitter that swaps excitations between the system and bath

$$\hat{H}_I = \sqrt{\eta}(\hat{a}\hat{a}_B^\dagger + \hat{a}^\dagger\hat{a}_B) \quad (9.19)$$

This is equivalent to setting $C = \sqrt{\eta}\mathbb{I}$ so that

$$\dot{\sigma} = -\eta\sigma + \eta\chi\mathbb{I}, \quad (9.20)$$

where we set $\sigma_{B'} = \chi\mathbb{I}$, where $\chi = 2\bar{n} + 1$ and \bar{n} is the average number of excitations of each mode. Rescaling time $t \rightarrow \eta t$ we arrive at the equation for lossy evolution

$$\dot{\sigma} = -\sigma + \chi\mathbb{I}, \quad (9.21)$$

with solution

$$\sigma = \chi\mathbb{I} + (\sigma(0) - \chi\mathbb{I})e^{-t}. \quad (9.22)$$

Eq. (9.21) describes the lossy evolution of the covariance matrix of a Gaussian state. If the covariance matrix is at all large then it will become gradually more cumbersome to solve the equation as the number of parameters increases. It would be nice to be able to extract the relevant parameters and observe how they evolve without requiring us to solve the whole equation.

A centrally interesting set of parameters, as discussed in Sec. 4 are the n symplectic invariants that provide the entropic information about the state. In Sec. 4 we saw that these were not unique and we could consider different sets, given that it is easy to construct a new set of symplectic invariants as a function of the old set. The major result we find is an evolution equation for the set of symplectic invariants, introduced earlier and denoted $\vartheta_{2k}[\Omega\sigma]$.

9.3 Symplectic invariant evolution

We have already seen the role that symplectic invariants play in descriptions of the state, both in its entropic and entanglement properties. Tracking the evolution of these quantities over time is therefore interesting to people wishing to manipulate these properties.

Before stating the result that shows this evolution for lossy channels, we are required to define yet another set of symplectic invariants, in addition to the two that have already been explored. These require the definition of a new way of choosing subsets of \mathbb{N} . Define.

$$H \in \mathcal{F}_k^{n \setminus i} \subset P(\mathbb{N}_n \setminus \{i\}) \quad \text{iff} \quad |H| = k. \quad (9.23)$$

This is similar to the sets as given in Sec. 4.3 but there is a removal of all terms involving a particular element. The new set of symplectic invariants that we will need to refer to later are defined as,

$$\vartheta_{2k}^i[\Omega\sigma] := \sum_{H \in \mathcal{F}_k^{n \setminus i}} \prod_{j \in H} \nu_j^2. \quad (9.24)$$

These are similar to the symplectic invariants $\vartheta_{2k}[\Omega\sigma]$ except that we remove all terms in the sum that involve the symplectic eigenvalue ν_i . To make this a little clearer we will consider an example where σ is a five-mode covariance matrix and so,

$$\vartheta_6^2[\Omega\sigma] = \nu_1^2 \nu_3^2 \nu_4^2 + \nu_1^2 \nu_3^2 \nu_5^2 + \nu_1^2 \nu_4^2 \nu_5^2 + \nu_3^2 \nu_4^2 \nu_5^2. \quad (9.25)$$

We see that the terms still have three distinct symplectic eigenvalues each, but we have excluded any term that includes ν_2 . The main result of this section, and the starting point for much of the analysis of Sec. 10, can now be stated from Ref. [2].

Theorem 43. *For σ evolving under $\dot{\sigma} = -\sigma + \chi \mathbb{I}$, the evolution of the symplectic invariants, defined by $\vartheta_{2k}[\Omega\sigma]$, obeys*

$$\dot{\vartheta}_{2k}[\Omega\sigma] = -2k\vartheta_{2k}[\Omega\sigma] + \chi \text{Tr}[SV_{2k}S^\top], \quad (9.26)$$

where $\sigma = SWS^\top$ and

$$V_{2k} = \bigoplus_{i=1}^n \left(\nu_i \vartheta_{2(k-1)}^i[\Omega\sigma] \right) \mathbb{I}_2. \quad (9.27)$$

To get a feel for Eq. (9.26) it would be instructive to consider an example. In two modes we see that the set of equations described by Eq. (9.26) are

$$\begin{aligned} \dot{\vartheta}_2[\Omega\sigma] &= -2(\nu_1^2 + \nu_2^2) + \chi \text{Tr}[S(\nu_1 \mathbb{I}_2 \oplus \nu_2 \mathbb{I}_2)S^\top], \\ \dot{\vartheta}_4[\Omega\sigma] &= -4\nu_1^2 \nu_2^2 + \chi \text{Tr}[S(\nu_1 \nu_2^2 \mathbb{I}_2 \oplus \nu_2 \nu_1^2 \mathbb{I}_2)S^\top]. \end{aligned} \quad (9.28)$$

These coupled equations do not provide a trivial solution. Furthermore we have not entirely escaped Eq. (9.21) in that we are still required to have knowledge of the time-dependent symplectic matrix S which appears in the Williamson decomposition of $\sigma(t)$. Eq. (9.26) gives us the evolution of the invariants directly as opposed to Eq. (9.21), allowing us to focus on the

important parameters for our interest and removing $2n^2$ unrequired pieces of information.

Proof of Theorem 43

A key component of the proof uses a recursive method to generate the symplectic invariants. Before stating it, we must recall some basic algebra [75]: take an $m \times m$ matrix M , delete the same $m - r$ rows and columns. Left over you will have an $r \times r$ submatrix which is referred to as the principal submatrix of M , the determinant of which is referred to as a principal minor. Considering the characteristic polynomial $\sum_{k=0}^m c_k \lambda^{m-k}$ of M , with $c_0 := 1$, we have

$$\begin{aligned} c_k[M] &= (-1)^k \sum (\text{all } k \times k \text{ principal minors}), \\ \vartheta_k[M] &= \sum (\text{all } k \times k \text{ principal minors}). \end{aligned} \quad (9.29)$$

We see the close link between the coefficients of the characteristic polynomial and the objects that will later become the $\vartheta_{2k}[\Omega\sigma]$ symplectic invariants. In fact, the two are equal for even k and so the following theorem, which provides a recursive formula for $c_k[M]$, will be relevant to the evolution of the symplectic invariants.

Theorem 44 (Fadeev-LeVerrier recursion [149–151]). *Let M be an $m \times m$ real matrix. Let its characteristic polynomial be written $\det[M - \lambda I] = \sum_{k=0}^m c_k \lambda^{m-k}$ with $c_0 := 1$. It is possible to calculate the coefficients of the polynomial via the recursive formula,*

$$c_k[M] = -\frac{1}{k} \sum_{i=0}^{k-1} \text{Tr}[M^{k-i}] c_i[M]. \quad (9.30)$$

Given that $c_{2k}[\cdot] \equiv \vartheta_{2k}[\cdot]$ we may rewrite Eq.(9.30) as a recursive generation of our symplectic invariants,

$$\vartheta_{2k}[\Omega\sigma] = \frac{-1}{2k} \sum_{i=0}^{k-1} \text{Tr}[(\Omega\sigma)^{2(k-i)}] \vartheta_{2i}[\Omega\sigma], \quad (9.31)$$

given that $\vartheta_k[\Omega\sigma] = 0$ when k is odd. Using a Taylor expansion we bring in Eq. (9.21) to observe an incremental time change in $\vartheta_{2k}[\Omega\sigma(t)]$:

$$\begin{aligned} \vartheta_{2k}[\Omega\sigma(t + dt)] &\simeq \vartheta_{2k}[\Omega(\sigma + \dot{\sigma}dt)] = \vartheta_{2k}[\Omega\sigma - dt\Omega\sigma + \chi\Omega dt] \\ &= \vartheta_{2k}[F + Gdt], \end{aligned} \quad (9.32)$$

where $F := (1 - dt)\Omega\sigma$ and $G := \chi\Omega$. Thus we see that

$$\vartheta_{2k}[F + Gdt] = \frac{-1}{2k} \sum_{i=0}^{k-1} \text{Tr}[(F + Gdt)^{2(k-i)}] \vartheta_{2i}[F + Gdt]. \quad (9.33)$$

Lemma 45. *To first order,*

$$\mathrm{Tr}[(F + Gdt)^m] \simeq \begin{cases} \mathrm{Tr}[F^m] + mdt \mathrm{Tr}[F^{m-1}G], & m \text{ even,} \\ 0, & m \text{ odd.} \end{cases} \quad (9.34)$$

Proof. Using the binomial theorem we may expand $(F + Gdt)^m$ to first order in dt where we obtain a single term of the form F^m and m terms of the form $F^a G dt F^b$ where $a + b = m - 1$. The cyclicity of the trace allows us to reorder these elements to obtain $\mathrm{Tr}[F^m] + mdt \mathrm{Tr}[F^{m-1}G] + o(dt)$. Now it remains to show that to first order this expression is zero for m odd. It suffices to show that

$$\mathrm{Tr}[F^{2m+1}] = 0, \quad m \in \mathbb{N}, \quad (9.35)$$

$$\mathrm{Tr}[F^{2m}G] = 0, \quad m \in \mathbb{N}. \quad (9.36)$$

To prove Eq. (9.35) we use the invariance of the trace under cycles and transposes giving

$$\begin{aligned} \mathrm{Tr}[(\Omega\sigma)^{2m+1}] &= \mathrm{Tr}[(\Omega\sigma)^{2m+1}]^\top = (-1)^{2m+1} \mathrm{Tr}[(\sigma\Omega)^{2m+1}] \\ &= -\mathrm{Tr}[(\Omega\sigma)^{2m+1}] = 0, \end{aligned} \quad (9.37)$$

where we used $\Omega^\top = -\Omega$. Eq. (9.36) is found using a similar argument. Putting these together we prove the proposition due to the many terms that are zero in the first order expansion. \square

Lemma 45 is going to help provide a Taylor expansion of the symplectic invariants which is the next result.

Lemma 46. *Taylor expanding $\vartheta_{2k}[F + Gdt]$ we arrive at*

$$\vartheta_{2k}[F + Gdt] = \vartheta_{2k}[F] - dt \sum_{i=0}^{k-1} \mathrm{Tr}[(\Omega\sigma)^{2(k-i)-1}G] \vartheta_{2i}[F] + o(dt). \quad (9.38)$$

Proof. From Eq. (9.33) we can explicitly show that

$$\begin{aligned} \vartheta_0[F + Gdt] &= 1, \\ \vartheta_2[F + Gdt] &= \vartheta_2[F] - dt \mathrm{Tr}[\gamma G], \end{aligned} \quad (9.39)$$

where we define $\gamma := \Omega\sigma$ for brevity in the proceeding proof and where we have also used Lemma 45, $\mathrm{Tr}[F + Gdt] = 0$ and $\vartheta_2[F] = -\frac{1}{2} \mathrm{Tr}[F^2]$. From here we will proceed with an inductive proof. We assume that Eq. (9.38) holds for some k and then show that if this is true then it holds for $k + 1$.

Using Eq. (9.33) we may write the expansion out for $k + 1$ and then use Lemma 45 and the

definition of F to arrive at

$$\vartheta_{2(k+1)}[F + Gdt] \simeq \frac{-1}{2(k+1)} \sum_{i=0}^k \left(\text{Tr}[F^{2(k+1-i)}] + 2(k+1-i)dt \text{Tr}[F^{2(k+1-i)-1}G] \right) \vartheta_{2i}[F + Gdt]. \quad (9.40)$$

Substituting Eq. (9.38) into Eq. (9.40) we extract the first two terms that look like they would fulfill the proof, plus a final one that we would hence like to show is zero:

$$\begin{aligned} \vartheta_{2(k+1)}[F + Gdt] &\simeq \\ &\frac{-1}{2(k+1)} \sum_{i=0}^k \left(\text{Tr}[F^{2(k+1-i)}] + 2(k+1-i)dt \text{Tr}[F^{2(k+1-i)-1}G] \right) \left(\vartheta_{2i}[F] - \right. \\ &\quad \left. dt \sum_{j=0}^{i-1} \text{Tr}[\gamma^{2(i-j)-1}G] \vartheta_{2j}[F] \right) \\ &\simeq \vartheta_{2(k+1)}[F] - dt \sum_{i=0}^k \text{Tr}[\gamma^{2(k+1-i)-1}G] \vartheta_{2i}[F] + \\ &\quad \frac{-1}{2(k+1)} \sum_{i=1}^k \left(-dt \text{Tr}[\gamma^{2(k+1-i)}] \sum_{j=0}^{i-1} \text{Tr}[\gamma^{2(i-j)-1}G] \vartheta_{2j}[F] - 2i dt \text{Tr}[\gamma^{2(k+1-i)-1}G] \vartheta_{2i}[F] \right). \end{aligned} \quad (9.41)$$

Dropping the coefficient $(2(k+1))^{-1}$, we proceed to examine the final piece, referring to it as L , and rewriting it as

$$L = dt \sum_{i=1}^{k-1} \text{Tr}[\gamma^{2(k-i)}] \sum_{j=0}^{i-1} \text{Tr}[\gamma^{2(i-j)-1}G] \vartheta_{2j}[F] + dt \sum_{i=1}^{k-1} 2i \text{Tr}[\gamma^{2(k-i)-1}G] \vartheta_{2i}[F]. \quad (9.42)$$

Note that we have relabeled k as $k-1$ to shorten the expression but it will not alter the analysis. The dt at the front reminds us that everything should be expanded to zeroth order inside the sum. To prove the lemma it is necessary to show that $L \equiv 0$.

Expanding $\vartheta_{2i}[F]$ to introduce another sum we arrive at

$$L = dt \sum_{i=1}^{k-1} \sum_{j=0}^{i-1} \left(\text{Tr}[\gamma^{2(k-i)}] \text{Tr}[\gamma^{2(i-j)-1}G] - \text{Tr}[\gamma^{2(k-i)-1}G] \text{Tr}[\gamma^{2(i-j)}] \right) \vartheta_{2j}[F]. \quad (9.43)$$

From here note that for a general sum with elements Y_{ij} we have

$$\sum_{i=1}^{k-1} \sum_{j=0}^{i-1} Y_{ij} = \sum_{j=0}^{k-2} \sum_{i=j+1}^{k-1} Y_{ij} = \frac{1}{2} \sum_{j=0}^{k-2} \sum_{i'=0}^{k-j-2} Y_{i'+j+1,j} + \frac{1}{2} \sum_{j=0}^{k-2} \sum_{i''=0}^{k-j-2} Y_{k-i''-1,j}, \quad (9.44)$$

where $i' = i - (j + 1)$ and $i'' = k - i' - j - 2$. The first equality of Eq. (9.44) can be seen with observation. The second involves a redefinition of the sums where we split them into two halves and then redefine the labels such that one is descending whilst the other ascends. When such a summation redefinition is applied to Eq. (9.43) it will be clear that $L \equiv 0$.

Thus we prove that if $\vartheta_{2k}[F + Gdt]$ is given in Eq. (9.38) then this also holds for $k \rightarrow k + 1$. From Eq. (9.39) we see that it is true for $k = 1$ and so, inductively it is true for all k . To write it in the form stated one must replace γ with $\Omega\sigma$. \square

Lemma 47. *Taylor expanding $\vartheta_{2k}[F]$ using $F := (1 - dt)\Omega\sigma$ we find that*

$$\vartheta_{2k}[F] = (1 - 2kdt)\vartheta_{2k}[\Omega\sigma]. \quad (9.45)$$

Proof. Expanding out the recursive formula and again defining $\gamma := \Omega\sigma$ we get a product of sums of the form

$$\begin{aligned} \vartheta_{2k}[F] &= \frac{-1}{2k} \sum_{i_1=0}^{k-1} \left(1 - 2(k - i_1)dt\right) \cdot \frac{-1}{2i_1} \sum_{i_2=0}^{i_1-1} \left(1 - 2(i_1 - i_2)dt\right) \cdot \\ &\dots \frac{-1}{2i_{k-1}} \sum_{i_k=0}^{i_{k-1}-1} \left(1 - 2(i_{k-1} - i_k)dt\right) \cdot \text{Tr}[\gamma^{2(k-i)}] \dots \text{Tr}[\gamma^{2(i_{k-1}-i_k)}]. \end{aligned} \quad (9.46)$$

By only keeping terms that are less than second order in dt we get a smaller sum

$$(1 - 2k)\vartheta_{2k}[\Omega\sigma] + X \quad (9.47)$$

X consists of the remaining terms which come in pairs. Take for instance the first pair which is generated by choosing the $+2i_1dt$ coefficient in the first line of Eq. (9.46), with everything else at zeroth order, and secondly the $-2i_1dt$ coefficient in the second line, with everything else at zeroth order. The pairs will each cancel to become zero. The final piece comes without a partner but has coefficient $i_k = 0$, and so does not contribute. Therefore $X \equiv 0$ and the lemma is proven. \square

Lemma 48. *Using the previous two Taylor expansions we may write the rate of change of $\vartheta_{2k}[\Omega\sigma]$ as*

$$\dot{\vartheta}_{2k}[\Omega\sigma] = -2k\vartheta_{2k}[\Omega\sigma] - \chi \sum_{i=0}^{k-1} \text{Tr}[(\Omega\sigma)^{2(k-i)-1}\Omega]\vartheta_{2i}[\Omega\sigma]. \quad (9.48)$$

Proof. The first Taylor expansion came from Lemma 46 stating that

$$\vartheta_{2k}[F + Gdt] \simeq \vartheta_{2k}[F] - dt \sum_{i=0}^{k-1} \text{Tr}[F^{2(k-i)-1}G]\vartheta_{2i}[F]. \quad (9.49)$$

Using Lemma 47 we rewrite the above as

$$\vartheta_{2k}[\Omega\sigma(t+dt)] - \vartheta_{2k}[\Omega\sigma] \simeq -2kdt\vartheta_{2k}[\Omega\sigma] - dt \sum_{i=0}^{k-1} \text{Tr}[(\Omega\sigma)^{2(k-i)-1}G]\vartheta_{2i}[\Omega\sigma]. \quad (9.50)$$

Dividing through by dt we prove the proposition, recalling that $G := \chi\Omega$. \square

Eq. (9.48) provides the rate of change of the set of symplectic invariants under the action of a lossy channel. In this form, however, it is still recursive and so the following results work towards its simplification. We may rewrite the trace term of our equation as

$$\begin{aligned} \text{Tr}[(\Omega\sigma)^{2k-1}\Omega] &= \text{Tr}[\overbrace{\Omega SW S^\top \dots \Omega SW S^\top}^{2k-1} \Omega] = -\text{Tr}[S \overbrace{W S^\top \Omega S \dots W S^\top \Omega S}^{2k-2} W S^\top] \\ &= -\text{Tr}[S \overbrace{W \Omega \dots W \Omega}^{2k-2} W S^\top] = -\text{Tr}[S W^{2k-2} \Omega^{2k-2} W S^\top] \\ &= (-1)^k \text{Tr}[S W^{2k-1} S^\top]. \end{aligned}$$

As a result Eq. (9.48) becomes

$$\dot{\vartheta}_{2k}[\Omega\sigma] = -2k\vartheta_{2k}[\Omega\sigma] - \chi \sum_{i=0}^{k-1} (-1)^{k-i} \text{Tr}[S W^{2(k-i)-1} S^\top] \vartheta_{2i}[\Omega\sigma]. \quad (9.51)$$

Exploring the summation in this equation we find that

$$\sum_{i=0}^{k-1} (-1)^{k+1} (-1)^i \text{Tr}[S W^{2(k-i)-1} S^\top] \vartheta_{2i}[\Omega\sigma] = \text{Tr} \left[S \left(\sum_{i=0}^{k-1} (-1)^{k+1} (-1)^i W^{2(k-i)-1} \vartheta_{2i}[\Omega\sigma] \right) S^\top \right], \quad (9.52)$$

where the sum has been taken inside the trace. Given that W is diagonal we can calculate the inner sum for each of the symplectic eigenvalues. Denoting these by ν_q for $q = 1, \dots, n$ to avoid confusion, we find that the i th term of the sum is

$$(-1)^{k+1} (-1)^i \left(\nu_q^{2(k-i)+1} \vartheta_{2(i-1)}^q[\Omega\sigma] + \nu_q^{2(k-i)-1} \vartheta_{2i}^q[\Omega\sigma] \right), \quad (9.53)$$

whereas the $i+1$ th is

$$(-1)^{k+1} (-1)^{i+1} \left(\nu_q^{2(k-i)-1} \vartheta_{2i}^q[\Omega\sigma] + \nu_q^{2(k-i)-3} \vartheta_{2(i+1)}^q[\Omega\sigma] \right). \quad (9.54)$$

By considering these equations carefully we see that our sum is going to *telescope*. This is when each term in the series is in two parts where the latter part of the i th term cancels the former

part of the $i + 1$ th term, leaving just the first and last piece of the entire series. Noting that this is the case for Eqs. (9.53) and (9.54), and also that the first piece is equal to zero we see that the sum collapses to

$$(-1)^{k+1}(-1)^{k-1}\nu_q\vartheta_{2(k-1)}^q[\Omega\sigma] = \nu_q\vartheta_{2(k-1)}^q[\Omega\sigma]. \quad (9.55)$$

As a result the sum term of Eq. (9.52) reduces to

$$V_{2k} = \bigoplus_{i=1}^n \left(\nu_i \vartheta_{2(k-1)}^i[\Omega\sigma] \right) \mathbb{I}_2. \quad (9.56)$$

Substituting this into Eq. (9.51) we find that

$$\dot{\vartheta}_{2k}[\Omega\sigma] = -2k\vartheta_{2k}[\Omega\sigma] + \chi \text{Tr}[SV_{2k}S^\top], \quad (9.57)$$

which proves Theorem 43.

This final equation allows a deeper perspective into the effect that lossy channels have on the symplectic invariants. The right-hand side has two components, one which is symplectically invariant and the other that is not, thus inviting an investigation into the role of symplectic control in this evolution.

9.4 Entanglement evolution

The previous techniques can be adapted to studying something that is not symplectically invariant. The full analysis holds in the partially transposed case and so our investigations can extend to explore the behaviour of entanglement under lossy channels.

Recalling the analysis of Sec. 4.4 we know that the partial transpose of a $(p+q)$ -mode Gaussian state corresponds to enacting $\sigma \rightarrow \tilde{\sigma} = T\sigma T$ where $T = \bigoplus_1^p \text{diag}(1, -1) \oplus \bigoplus_1^q \text{diag}(1, 1)$. We saw that $\tilde{\sigma}$ maintains a Williamson decomposition and we looked at condition on the set of these ‘partially-transposed’ symplectic invariants.

Since T is time independent the equation for the evolution of the partially transposed matrix under a lossy channel obeys

$$\dot{\tilde{\sigma}} = T\dot{\sigma}T = -\tilde{\sigma} + \chi\mathbb{I}. \quad (9.58)$$

Hence the full analysis that led us to Theorem 43 can be used to derive the evolution of the partially transposed symplectic invariants, i.e. the following,

$$\dot{\vartheta}_{2k}[\Omega\tilde{\sigma}] = -2k\vartheta_{2k}[\Omega\tilde{\sigma}] + \chi \text{Tr}[\tilde{S}\tilde{V}_{2k}\tilde{S}^\top], \quad (9.59)$$

where $\tilde{\sigma} = \tilde{S}\tilde{W}\tilde{S}^\top$ and

$$\tilde{V}_{2k} = \bigoplus_{i=1}^{p+q} \left(\tilde{\nu}_i \vartheta_{2(k-1)}^i [\Omega \tilde{\sigma}] \right) \mathbb{I}_2, \quad (9.60)$$

which provides

$$\dot{\tilde{\Sigma}} = \sum_{k=1}^n (-1)^{p+q+k} \left(-2k \vartheta_{2k} [\Omega \tilde{\sigma}] + \chi \text{Tr}[\tilde{S}\tilde{V}_{2k}\tilde{S}^\top] \right), \quad (9.61)$$

This provides us with an equation for $\tilde{\Sigma}$ describing the evolution of entanglement of the $(p+q)$ -mode system, directly from the analysis employed to derive Eq. (9.26).

In the two mode case $\tilde{\Sigma} \geq 0$ is necessary and sufficient for satisfaction of the PPT criterion because $\det[\tilde{\sigma}] = \det[\sigma] \geq 1$, meaning that only one of the transposed symplectic eigenvalues can be less than one. Given that bisymmetric entangled Gaussian states can be locally distilled into a single, two-mode entangled state, this condition becomes necessary and sufficient for this class of states as well. As a result, studying the behaviour of the sign of $\tilde{\Sigma}$ tells us about the entanglement of the state.

Chapter 10

Lossy system control

The analysis of Sec. 9 provides a new avenue for the application of control techniques. Up until now we have focused on controlling the transformations themselves, i.e. the symplectic matrices that govern closed Gaussian dynamics. Now we consider controlling the states themselves and, more specifically, certain of their properties that are of use.

Once we have answered the question of which symplectic operations can be enacted given a set of Hamiltonians, we then ask which symplectic operations *we would like to enact*. This is dependent on the particular properties one would like to have, as well as the constraints placed on the system. Following on from the previous section we consider a regime in which our Gaussian state is evolving under lossy dynamics. The final term of Eq. (9.26) invites the application of symplectic, or closed system control, to see if we can affect the evolution of the symplectic invariants.

Control problems concerning open quantum systems have been explored in a variety of cases including dissipating qubits [152], multilevel discrete systems [153], in the context of ‘quantum speed limits’ [154] and in closed feedback control systems [155]. The application of control problems to Gaussian states undergoing open evolution is explored in Ref. [146] and Ref. [156]. This latter reference contains the results produced in this section but restricted to a single mode.

Often the point of interest is finding some locally optimal set of control choices that preserve a property, or attain it as fast as possible. This caveat of local optimality is an important one that needs to be noted when discussing optimisation. If we have a parameter evolving towards a fixed point then locally optimal control is that which will extremise the rate of change of the parameter towards that target. It is admitted that in certain cases there may be a globally optimal route that would not adopt this strategy.

The two properties that will act as the resources of interest are entropy and entanglement. Their evolution under lossy channels has been explored in Sec. 9, and here we seek to affect this

evolution using symplectic control.

10.1 Entropic control

Entropy has played a central role in quantum mechanics since its introduction by von Neumann in the nineteen-thirties. It is a central aspect of information theory and so plays a major role in its quantum counterpart. Furthermore it is one of the key variables of thermodynamics and is the property of interest in resource theories of purity [157, 158].

In Sec. 4.3 we saw that the entropy of a Gaussian state is totally given by its set of n symplectic invariants. Furthermore it was shown that by fixing the mode frequencies ω_i we may interchangeably refer to a rise in entropy as heating, and a fall as cooling. The idea of a Gaussian state having this associated temperature goes back to our original definition considering them as thermal states of quadratic Hamiltonians, in Def. 2.

The question of heating and cooling Gaussian states is of interest to people for whom temperature is the barrier to overcome before treating their system as a carrier for quantum information. In quantum optomechanics there is a drive to cool their systems to decrease the number of phonon modes, to allow for greater ease of control and computation [143, 159]. At low temperatures quantum effects are often easier to harness given the absence of noise.

In the previous section we derived the evolution of the set of symplectic invariants for a Gaussian state evolving under lossy dynamics. This provided the equation

$$\dot{\vartheta}_{2k} = -2k\vartheta_{2k} + \chi \text{Tr}[SV_{2k}S^\top], \quad V_{2k} = \bigoplus_{i=1}^n \left(\nu_i \vartheta_{2(k-1)}^i \right) \mathbb{I}_2, \quad (10.1)$$

where from now on $\vartheta_{2k} := \vartheta_{2k}[\Omega\sigma]$ and $\tilde{\vartheta}_{2k} := \vartheta_{2k}[\Omega\tilde{\sigma}]$. This equation describes the evolving dynamics, as well as providing a route into manipulating these dynamics through symplectic control.

10.1.1 Optimal symplectics

Theorem 43 provided the extraction of the symplectic invariants from the equation for lossy evolution. When discussing cooling or heating under such channels then it is Eq. (9.26) that becomes the main focus. We consider instantaneous symplectic control at each time-step as the state undergoes lossy dynamics. The ability to enact an instantaneous symplectic may seem far too much but this can be justified by the fact that symplectics can typically be enacted in the order of nanoseconds, whereas the decoherence rates are of the order of $10\text{-}10^3$ kHz. Turning to

Eq. (9.26), as given again here,

$$\dot{\vartheta}_{2k}[\Omega\sigma] = -2k\vartheta_{2k}[\Omega\sigma] + \chi \operatorname{Tr}[SV_{2k}S^\top], \quad (10.2)$$

we see that on the right-hand-side there are two terms which both depend on the covariance matrix at that point in the evolution. Our control gives us the ability to send $\sigma \rightarrow S\sigma S^\top$ for any $S \in \operatorname{Sp}(2n, \mathbb{R})$. The first term in the equation is a symplectic invariant and so has no response to such a control setup, and so it is the second term that needs to be focused on. To maximise or minimise the rate of change for the symplectic invariant in Eq. (9.26) it is necessary to vary S so that it respectively maximises or minimises the trace term, which is dealt with in the following result where we consider a more general matrix Y , rather than dealing with the specific form of V_{2k} .

Proposition 49. Consider $\operatorname{Tr}[SY S^\top]$ where $Y = \bigoplus_{i=1}^n y_i \mathbb{I}_2$, with y_i positive, and $S \in \operatorname{Sp}(2n, \mathbb{R})$.

$$\sup_{S \in \operatorname{Sp}(2n, \mathbb{R})} \operatorname{Tr}[SY S^\top] = \lim_{z_i \rightarrow \infty} \sum_{i=1}^n 2\zeta_{z_i}^+ y_i, \quad (10.3)$$

where $\zeta_z^\pm := \frac{z^2 \pm 1/z^2}{2}$, such that $z_1 \geq \dots \geq z_n$, and $y_1 \geq \dots \geq y_n$. The infimum is given by $\operatorname{Tr}[Y]$.

Proof. First, we Euler decompose S to provide

$$\operatorname{Tr}[SY S^\top] = \operatorname{Tr}[R_1 Z R_2 Y R_2^\top Z R_1^\top] = \operatorname{Tr}[Z^2 R_2 Y R_2^\top]. \quad (10.4)$$

Using the basis changes of Sec. 3.1 we first enact P and then \tilde{Q} to get

$$Z'^2 := (\tilde{Q}P)Z^2(\tilde{Q}P)^{-1} = \begin{pmatrix} \Gamma^+ & \Gamma^- \\ \Gamma^- & \Gamma^+ \end{pmatrix}, \quad (10.5)$$

where $\Gamma^\pm := \frac{1}{2} \operatorname{diag}(z_1^2 \pm 1/z_1^2, \dots, z_n^2 \pm 1/z_n^2)$.

$$Y' := (\tilde{Q}P)Y(\tilde{Q}P)^{-1} = \begin{pmatrix} \Upsilon & 0_n \\ 0_n & \Upsilon \end{pmatrix}, \quad (10.6)$$

where $\Upsilon = \operatorname{diag}(y_1, \dots, y_n)$

$$R'_2 := (\tilde{Q}P)R_2(\tilde{Q}P)^{-1} = \begin{pmatrix} U^* & 0_n \\ 0_n & U \end{pmatrix}, \quad R_2'^\top := (QP)R_2^\top(QP)^{-1} = \begin{pmatrix} U^\top & 0_n \\ 0_n & U^{*\top} \end{pmatrix}, \quad (10.7)$$

where U is some unitary matrix. Note that $*$ denotes the complex conjugate.

$$\text{Tr}[Z^2 R_2 Y R_2^\dagger] = \text{Tr}[Z'^2 R_2' Y' R_2'^\dagger] = \text{Tr}[\Gamma^+ U^* \Upsilon U^\dagger] + \text{Tr}[\Gamma^+ U \Upsilon U^{*\dagger}] = 2\mathbf{a}^\dagger P \mathbf{b}. \quad (10.8)$$

where \mathbf{a} is the vector of diagonal elements of Γ^+ , i.e. $a_i = \Gamma_{ii}^+$, and \mathbf{b} is the vector of diagonal elements of Υ , i.e. $b_i = \Upsilon_{ii}$. $P_{ij} = |U_{ij}|^2$ which is the definition of a unistochastic matrix, a set of matrices which form a subset of the doubly stochastic matrices [160]. These are matrices with non-negative entries such that their rows and columns sum to one. If we can show that

$$\sup_{M \text{ doubly stochastic}} \mathbf{a}^\dagger M \mathbf{b} = \mathbf{a}^\dagger \uparrow \mathbf{b}^\downarrow, \quad \text{and} \quad \inf_{M \text{ doubly stochastic}} \mathbf{a}^\dagger M \mathbf{b} = \mathbf{a}^\uparrow \mathbf{b}^\downarrow. \quad (10.9)$$

then the statement shall be proven because all permutation matrices are unistochastic [160]. To explain the notation, given some vector \mathbf{v} , the vector \mathbf{v}^\downarrow denotes a new vector of elements of \mathbf{v} written in descending order, and conversely for \mathbf{v}^\uparrow . The statement above is proven in Ref. [161] and reproduced in Appendix D. Varying z_i between its maximum value which, given complete control, is infinity and its minimum at $z_i = 1$ we find the infimum and the supremum and prove the statement. \square

Replacing Y with V_{2k} , Proposition 49 provides the maximum and minimum that S should be in Eq. (9.26) such that the rate of change of the symplectic invariants is extremised. Namely that either S should equal Z with as high values of z_i as possible, this is to maximise, or that S should equal \mathbb{I} to minimise the term.

Note, however, that these are not the unique values of S to extremise the term. The trace is invariant under an added element of $\text{OSp}(2n, \mathbb{R})$ and maximum and minimum can equally be achieved respectively with $S = RZ$ and $S = R$, for R passive. Thus we see that the real difference between the controls is the presence of squeezing highlighting it as the element in extremising $\dot{\vartheta}_{2k}$.

10.1.2 Decoupling

In both cases we see that it is possible to study the optimal case for decoupled dynamics. However, this does not mean that decoupling itself is the only optimal scenario, as we have just discussed. When we have decoupled the dynamics, Eq. (9.21) becomes block diagonal with 2×2 covariance matrices, each of which correspond to a single symplectic invariant. We have seen that the optimal control is not a unique solution but it suffices to study one of them. It is obviously easiest to consider the decoupling symplectic as the minimum relaxation times derived here will be general.

Enacting the decoupling matrix, maintaining the possibility to squeeze or unsqueeze brings

us to the same equation on each of the n modes:

$$\dot{\vartheta}_2 = -2\vartheta_2 + 2\chi\zeta_{z_i}^+\nu_i, \quad (10.10)$$

where $\zeta_{z_i}^+$ is defined in the statement of Proposition 49. Rewriting this in terms of symplectic invariants we obtain,

$$\dot{\nu}_i + \nu_i - \chi\zeta_{z_i}^+ = 0, \quad (10.11)$$

which is solved by

$$\nu_i(t) = \chi\zeta_{z_i}^+ + (\nu_{i0} - \chi\zeta_{z_i}^+)e^{-t}, \quad (10.12)$$

where $\nu_i(0)$ is denoted by ν_{i0} . Note that we have obtained these decoupled dynamics by enacting our optimal control strategy a single time at the beginning of the dynamics. We see that this single control suffices for optimality throughout the evolution.

The fixed point of this equation is at $\nu_i = \chi$. The time to reach the fixed point of the dynamics diverges and so we fix an error ϵ which is the distance within which we are satisfied that we have achieved the target,

$$|\nu_i - \chi| < \epsilon. \quad (10.13)$$

For the case when $\nu_{i0} < \chi$ we may describe the dynamics as heating. Optimising so that z_i is maximal, the time to come within the decided error is

$$T_{\text{heat}} = \sup_{\nu_{i0}, z_i} \ln \left[\frac{\chi\zeta_{z_i}^+ - \nu_{i0}}{\chi(\zeta_{z_i}^+ - 1) + \epsilon} \right]. \quad (10.14)$$

Firstly, we note that this quantity is generally finite and tends to zero for unbounded squeezing. Furthermore it is possible to send ϵ to zero and not change this fact. Note, here, that it is not squeezing itself that does anything towards heating the state, but squeezing in conjunction with the lossy channel dynamics.

When $\nu_{i0} > \chi$ the channel is described as cooling. We have seen that the optimal strategy in this case it is set to $z_i = 1$. The minimum time for the state to relax to the steady state is

$$T_{\text{cool}} = \sup_{\nu_{i0}} \ln \left[\frac{\nu_{i0} - \chi}{\epsilon} \right]. \quad (10.15)$$

In contrast to the heating case we see that T_{cool} goes to infinity as $\epsilon \rightarrow 0$.

A point that came up during the analysis was the presence of ‘bumps’ in cooling dynamics. It can be shown that if a highly squeezed state is placed in a cooling channel then the initial effect is for the entropy to rise rather than fall, as shown in Fig. 10.1. Thus we see that squeezing can entirely block cooling from ever occurring. We see that the increase happens when $\nu_i < \chi\zeta_{z_i}^+$ even for $\nu_i > \chi$. In Eq. (10.10) we see that the squeezing parameter $\zeta_{z_i}^+$ multiplies the bath parameter

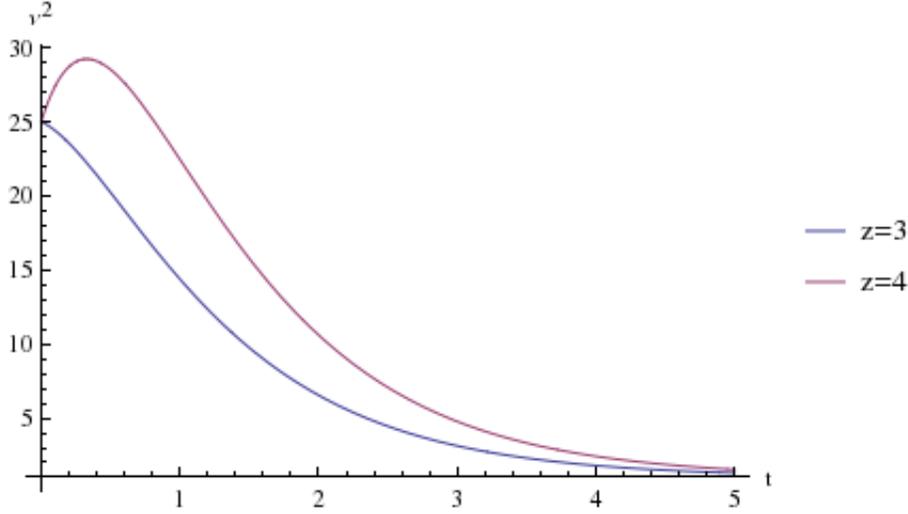


Figure 10.1: The solution to Eq. (10.12) is given for the values $\nu_0 = 5$, $\chi = 1$ and two different initial squeezing values. $z = 3$ implies that the entropy monotonically decreases whereas a higher initial squeezing parameter induces an initial ‘bump’ in the entropy.

χ . Thus, the intuition is that a high value of squeezing causes the state to ‘see’ a hotter bath. This is apparent from Eq. (10.11) where we see the product of χ and $\zeta_{z_i}^+$ appear. Highly squeezed states will see the bath as hot until they have lowered their $\zeta_{z_i}^+$ value to eventually thermalise to the cooler state.

A visual intuition for this comes if considering the Wigner function, as depicted in Fig. 2.1. If a state has high entropy then the variance is wider in both directions. Squeezing then reduces the variance in one direction whilst increasing it in the other. As the system cools we see that not only must it reduce the variance in both directions, but also to balance them out. These two processes act against each other causing these ‘bumps’.

Example. Take as the initial state the following covariance matrix,

$$\sigma = \gamma \begin{pmatrix} \cosh 2r & 0 & \sinh 2r & 0 \\ 0 & \cosh 2r & 0 & \sinh 2r \\ \sinh 2r & 0 & \cosh 2r & 0 \\ 0 & \sinh 2r & 0 & \cosh 2r \end{pmatrix}, \quad (10.16)$$

It bears a similarity to the two-mode squeezed state except for the constant factor of γ at the front which makes it mixed. The parameter r is sometimes called the two-mode

squeezing parameter. In order to evolve such a state under Eq. (9.20), we must first set some parameters. A loss rate of $\eta = 100$ kHz is chosen and an average bath-mode excitation number is found to be $\bar{n} = 5 \times 10^{-5}$ which is based on Bose-Einstein statistics,

$$\bar{n} = \frac{1}{e^{\frac{hf}{k_B\tau}} - 1}, \quad (10.17)$$

with Planck's constant $h = 6.626 \times 10^{-34}$ m² kg s⁻¹, Boltzmann's constant $k_B = 1.38 \times 10^{-23}$ m² kg s⁻¹ K⁻¹, and room temperature $\tau = 300$ K. We use a bath-mode frequency of $f = 74$ THz, which corresponds to radiowave frequency and, using $\chi = 2\bar{n} + 1$ we find $\chi = 1.00001$.

To set up a cooling situation we let $\gamma = 2$ and $\epsilon = 0.01$. We know from the former analysis that it is optimal to unsqueeze, which here would mean taking the state to $\gamma\mathbb{I}$. Let the initial squeeze be set as $r = 0.4$. Without performing our single optimal control we get within ϵ for both symplectic eigenvalues in $47.53\mu\text{s}$. Unsqueezing all the way at the start of the dynamics achieves this in $46.05\mu\text{s}$.

For a heating situation we set $\gamma = 1$ and $\epsilon = 0$. For the initial state with $r = 0.4$ the fixed point is reached for both symplectic eigenvalues after 1.77ns and in the controlled case setting $r = 0.5$ it takes 1.13ns .

10.2 Entanglement control

Referring to the distillation and control of 'useful properties', quantum theory can be placed in contact with a view of quantum mechanics as a resource theory. The most important and intriguing resource in quantum theory is that of non-locality, which was revealed as an essential feature of nature following Bell's theorem [9]. This phenomenon is only expressed when we deal with entangled states and so entanglement has become the subject of resource theories.

The utility of this resource has prompted a wide range of protocols geared towards its creation, distillation and preservation in different settings. In the regime of quantum optics many different setups have been explored in the context of entanglement preservation, due to the commonly corrosive impact of the environment. In Ref. [162] the authors establish optimal protocols for photon-number entangled states evolving under lossy channels. In Ref. [163] the authors consider a double beam of Gaussian states where one branch has a partially transmitting plate. In Ref. [164] the focus is on an asymmetric setup with two beams of Gaussian states, where one beam is made lossy and the other lossless.

In Sec. 9 we explored the way in which lossy channels can affect the symplectic eigenvalues of

a Gaussian state, as well as the symplectic eigenvalues of the partially transposed covariance matrix. This evolution equation contains elements that can be altered by a change in the dynamics, thus leaving us with a control problem to be solved. As before, it is necessary to set which controls we allow ourselves.

Given that entanglement is a nonlocal property, a physical control setup would suggest restricting ourselves to local operations. In Sec. 4.4 we saw that such control leaves partially transposed symplectic invariants untouched and so, again, we are in a situation where the first term of the equation does not change and so our focus lies with the second. It will not be necessary to cap our group in this section, as we will see.

We have seen that $\tilde{\Sigma} < 0$ is a necessary and sufficient condition for the entanglement of two-mode Gaussian states, which can be extended to bisymmetric $(p + q)$ -mode states. Recall that these are states that are invariant under local mode permutations on the subsystems. In Ref. [165] it is proven that such states are reducible, via local symplectic operations, to a pair of entangled modes, and a set of $p + q - 2$ uncorrelated single modes. Therefore our investigation of these states may proceed by studying the case of two entangled modes. An interesting question then arises which is: given instantaneous local symplectic control on such a system undergoing lossy dynamics, what is the locally optimal strategy for the preservation of entanglement?

Our equations of interest are Eqs. (9.59) and (9.61) but for the purposes of the proof, we begin a little earlier in Sec. 9.3 and write down the equation

$$\dot{\tilde{\vartheta}}_{2k} = -2k\tilde{\vartheta}_{2k} - \chi \sum_{i=0}^{k-1} \text{Tr}[(\Omega\tilde{\sigma})^{2(k-i)-1}\Omega]\tilde{\vartheta}_{2i}. \quad (10.18)$$

which is the analogue of Eq. (9.48). The reason for going back to this earlier stage in the derivation is that at this stage we do not want to deal with the computation of partially transposed symplectic invariants. Substituting in for the evolution of $\tilde{\Sigma}$ and restricting to two modes we arrive at

$$\begin{aligned} \dot{\tilde{\Sigma}} &= \sum_{k=1}^2 (-1)^{2+k+1} 2k\tilde{\vartheta}_{2k} + \chi \sum_{k=1}^2 (-1)^{2+k+1} \sum_{i=0}^{k-1} \text{Tr}[(\Omega\tilde{\sigma})^{2(k-i)-1}\Omega]\tilde{\vartheta}_{2i} \\ &= 2\tilde{\vartheta}_2 - \chi \text{Tr}[\tilde{\sigma}] - 4\tilde{\vartheta}_4 - \chi(\text{Tr}[(\Omega\tilde{\sigma})^3\Omega] - \text{Tr}[\tilde{\sigma}]\tilde{\vartheta}_2) \\ &= 2\tilde{\vartheta}_2 - 4\tilde{\vartheta}_4 + \chi \left(\text{Tr}[\tilde{\sigma}](\tilde{\vartheta}_2 - 1) - \text{Tr}[(\Omega\tilde{\sigma})^3\Omega] \right). \end{aligned} \quad (10.19)$$

Given local symplectic control, the first term terms on the right-hand-side are invariant and so our focus is on the third term. Given that

$$\vartheta_2 = \frac{-1}{2} \text{Tr}[\Omega\sigma\Omega\sigma], \quad (10.20)$$

from Eq. (9.31), this term can be rewritten as

$$\mathrm{Tr} \left[\Omega \tilde{\sigma} \Omega \tilde{\sigma} \left(\tilde{\sigma} - \frac{1}{2} \mathrm{Tr}[\tilde{\sigma}] \right) \right], \quad (10.21)$$

where we have dropped χ for the purpose of analysis. A lengthy calculation shows that this term is invariant under $\tilde{\sigma} \rightarrow T\tilde{\sigma}T = \sigma$ for all 4×4 symmetric matrices σ . In Sec. 9.3 we learned that

$$\sum_{i=0}^{k-1} \mathrm{Tr}[(\Omega\sigma)^{2(k-i)-1}\Omega]\vartheta_{2i} = -\mathrm{Tr}[SV_{2k}S^\top], \quad V_{2k} := \bigoplus_{i=1}^n \left(\nu_i \vartheta_{2(k-1)}^i \right) \mathbb{I}_2, \quad (10.22)$$

and so Eq. (10.19) becomes

$$\begin{aligned} \dot{\tilde{\Sigma}} &= 2\tilde{\vartheta}_2 - 4\tilde{\vartheta}_4 + \chi \sum_{k=1}^2 (-1)^{2+k} \mathrm{Tr}[SV_{2k}S^\top] \\ &= 2\tilde{\vartheta}_2 - 4\tilde{\vartheta}_4 + \chi \mathrm{Tr}[SXS^\top], \end{aligned} \quad (10.23)$$

where $X = V_4 - V_2 = \mathrm{diag}(\nu_1(\nu_2^2 - 1), \nu_1(\nu_2^2 - 1), \nu_2(\nu_1^2 - 1), \nu_2(\nu_1^2 - 1))$.

The first two terms remain invariant but the final term is now in a form which lends itself much more easily to a control evaluation. Firstly we notice that, if the state is pure, then $X = 0$ and so no local control can affect the rate of change of its entanglement. As yet, we have no physical intuition for this result. However, under lossy dynamics this pure state will become mixed and then control will be able to slow down the effect of the channel.

Before, we were considering a much broader class of controls and so we allided the S_c and S to look at the effect on the trace term. Here, however, S will be in general a nonlocal symplectic and so we are unable to do this. Our control symplectic will be of the form $S_{\mathrm{loc}} = S_A \oplus S_B$ and act on σ as $\sigma \rightarrow S_{\mathrm{loc}}\sigma S_{\mathrm{loc}}^\top$. This will have an impact in the trace term and so we are looking at an effective transformation, $\mathrm{Tr}[SXS^\top] \rightarrow \mathrm{Tr}[S_{\mathrm{loc}}SXS_{\mathrm{loc}}^\top]$. As a symplectic matrix, S_{loc} will maintain an Euler decomposition. However, given its action in the trace term, its first passive operation will have no effect, and so we consider

$$S_{\mathrm{loc}} = Z_{\mathrm{loc}}R_{\mathrm{loc}} = \begin{pmatrix} Z_A & 0 \\ 0 & Z_B \end{pmatrix} \begin{pmatrix} R_A & 0 \\ 0 & R_B \end{pmatrix}, \quad (10.24)$$

where Z_A and Z_B are local single-mode squeezers and

$$R_A = \begin{pmatrix} \cos[\alpha_A] & \sin[\alpha_A] \\ -\sin[\alpha_A] & \cos[\alpha_A] \end{pmatrix}, \quad (10.25)$$

is a local phase-shifter, as is R_B .

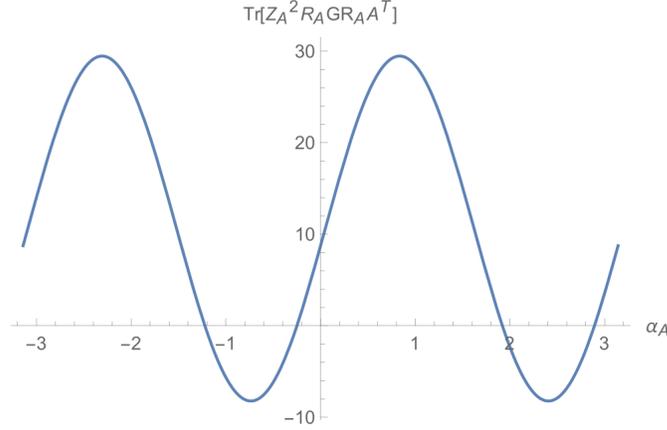


Figure 10.2: A plot of the trace term in Eq. (10.28) against the variable α_A for the specific values: $a_A = 2, b_A = 5, c_A = 3, z_A = 4$.

A first thing to notice is that if the optimal squeezing turns out to be $Z_{\text{loc}} = \mathbb{I}$ then R_{loc} will have no effect as the trace term will annihilate this. However, in the case when $Z_{\text{loc}} \neq \mathbb{I}$ then we may consider the optimal strategy for R_{loc} . The optimal action for each of these matrices turns out to be independent and so the results may be proven in any order. Since the local phase-shifters act first we will optimise for them first.

Working abstractly we consider some general 4×4 symmetric matrix H where,

$$H = \begin{pmatrix} G & K \\ K^\top & L \end{pmatrix}, \quad G = \begin{pmatrix} a_A & b_A \\ b_A & c_A \end{pmatrix}, \quad L = \begin{pmatrix} a_B & b_B \\ b_B & c_B \end{pmatrix}, \quad (10.26)$$

and K is some arbitrary 2×2 real matrix. Any property proven for such a matrix will also apply to the matrix, SXS^\top . In this 2×2 block form we may decompose the trace term of Eq. (10.23) as

$$\begin{aligned} & \text{Tr} \left[\begin{pmatrix} Z_A & 0 \\ 0 & Z_B \end{pmatrix} \begin{pmatrix} R_A & 0 \\ 0 & R_B \end{pmatrix} \begin{pmatrix} G & K \\ K^\top & L \end{pmatrix} \begin{pmatrix} R_A^\top & 0 \\ 0 & R_B^\top \end{pmatrix} \begin{pmatrix} Z_A^\top & 0 \\ 0 & Z_B^\top \end{pmatrix} \right] \\ &= \text{Tr}[Z_A^2 R_A G R_A^\top] + \text{Tr}[Z_B^2 R_B L R_B^\top]. \end{aligned} \quad (10.27)$$

As expected, the local phase-shifters only have a local effect and can be optimised independently. Beginning with the first term we find that

$$\begin{aligned} \text{Tr}[Z_A^2 R_A G R_A^\top] &= \frac{1}{z_A} \left((c_A + a_A z_A^2) \cos^2[\alpha_A] + (a_A + c_A z_A^2) \sin^2[\alpha_A] \right. \\ &\quad \left. + b_A (z_A^2 - 1) \sin[2\alpha_A] \right). \end{aligned} \quad (10.28)$$

This function of α_A oscillates and is plotted in Fig. 10.2. The period of oscillation is π and

contains one minimum and one maximum. In order to find its extremal points we take the derivative with respect to α_A giving

$$\frac{d \operatorname{Tr}[Z_A^2 R_A G R_A^\top]}{d\alpha_A} = \left(z_A - \frac{1}{z_A} \right) (2b_A \cos[2\alpha_A] + (c_A - a_A) \sin[2\alpha_A]). \quad (10.29)$$

Setting this to zero, we see that the extremisation is independent of the squeezing parameters z_A and z_B . For the case when $a_A \neq c_A$ and $b_A \neq 0$, the extremising angle satisfies,

$$\tan[2\alpha_A^{\text{ext}}] = \frac{2b_A}{a_A - c_A}. \quad (10.30)$$

When $b_A = 0$ we have $\sin[2\alpha_A^{\text{ext}}] = 0$ and for $a = c$ the extremisation angle satisfies $\cos[2\alpha_A^{\text{ext}}] = 0$, unless $b = 0$ in which case it can be anything. This latter point is obvious in that R_A and R_A^\top will cancel if acting on a matrix which is a multiple of the identity. Summarising, the extremal values are found when α_A satisfies:

$$\begin{cases} \tan[2\alpha_A^{\text{ext}}] = \frac{2b_A}{a_A - c_A}, & a_A \neq c_A, b_A \neq 0; \\ \sin[2\alpha_A^{\text{ext}}] = 0, & a_A \neq c_A, b_A = 0; \\ \cos[2\alpha_A^{\text{ext}}] = 0, & a_A = c_A, b_A \neq 0; \\ \text{anything,} & a_A = c_A, b_A = 0. \end{cases} \quad (10.31)$$

We see that within any range of π we will find one maximum and one minimum, separated by $\pi/2$. This analysis also holds for the second local phase-shifter.

Having optimised the phase-shifters, the next step is to optimise the local squeezers. The following proof takes a different approach to that of the previous analysis and does not use Eq. (10.27). Instead it proceeds by considering total *global* control on S in the final term of Eq. (10.23) and then showing what this implies for local control. We find that the trace term is monotonic with the squeezing measure $\zeta_{z_i}^+$, as introduced in Sec. 4.2. Given that it is monotonic with this measure for global control, it follows that it must also be monotonic with it for *local* control. From this we combine the two analyses to find an locally optimal local-control strategy.

Beginning our global control stepping stone, we consider varying over S for the the term $\operatorname{Tr}[SXS^\top]$. Euler decomposing S , we may rearrange the trace term as $\operatorname{Tr}[Z^2 A]$ where $A := R_2 X R_2^\top$. As a result,

$$\sum_{ij} [Z^2]_{ij} A_{ji} = \sum_i [Z^2]_{ii} A_{ii} = \sum_{i \text{ odd}} [Z^2]_{ii} A_{ii} + \frac{1}{[Z^2]_{ii}} A_{i+1 i+1}. \quad (10.32)$$

In two modes, $R_2 \in \text{OSp}(4, \mathbb{R})$ can be written as

$$R_2 = \begin{pmatrix} a_{11} & a_{13} & a_{12} & a_{14} \\ -a_{13} & a_{11} & -a_{14} & a_{12} \\ a_{21} & a_{23} & a_{22} & a_{24} \\ -a_{23} & a_{21} & -a_{24} & a_{22} \end{pmatrix}, \quad (10.33)$$

plus extra conditions not important for this. From this we find that $A_{ii} = A_{i+1 i+1}$ for odd i and so Eq. (10.32) becomes

$$\text{Tr}[Z^2 A] = \sum_{i \text{ odd}} A_{ii} \left([Z^2]_{ii} + \frac{1}{[Z^2]_{ii}} \right), \quad (10.34)$$

providing

$$\left. \frac{\partial \text{Tr}[SXS^\dagger]}{\partial \left(z_i^2 + \frac{1}{z_i^2} \right)} \right|_{z_j, j \neq i} = A_{ii}. \quad (10.35)$$

That is to say, the rate of change of the trace term with respect to $z_i^2 + 1/z_i^2$ for some mode i , where all other squeezers are fixed, is equal to A_{ii} . Since $A_{ii} = \sum_{j \text{ odd}} X_{jj} ([R_2]_{ij}^2 + [R_2]_{i j+1}^2)$, A_{ii} is positive for all i odd.

In summary, given $\tilde{\Sigma}$ evolving under Eq. (10.23), its rate of change is minimised by applying S_{loc} , as in Eq. (10.24) such that Z_{loc} acts to minimise the squeezing measure $\zeta_{z_i}^+$, as given in Sec. 4.2, and R_{loc} obeys Eq. (10.31). Note that there is no claim that $\tilde{\Sigma}$ provides a measure for entanglement, but only that its sign provides a yes/no criterion for entanglement. More work has been done to explore optimisation for logarithmic negativity, which does act as an entanglement measure. Furthermore, we note that this control is only proven to be locally optimal and so there is no proof that a better protocol does not exist.

Example. Using the same parameters as given in the example of Sec. 10.1, we begin with a random, entangled, mixed initial state

$$\sigma = \begin{pmatrix} 3.93221 & 4.62812 & -0.899615 & -1.23693 \\ 4.62812 & 6.72353 & -2.17121 & 0.91571 \\ -0.899615 & -2.17121 & 1.7453 & -2.62486 \\ -1.23693 & 0.91571 & -2.62486 & 6.88703 \end{pmatrix}. \quad (10.36)$$

From here, we enact the optimal protocol by minimising the value of each singular value. Then we apply the optimal local passive transformation using the angles derived in Eq. (10.31). After this, we evolve according to the lossy dynamics of Eq. (9.21) for a certain time and then optimise again. Continuing these concatenated processes leads to an

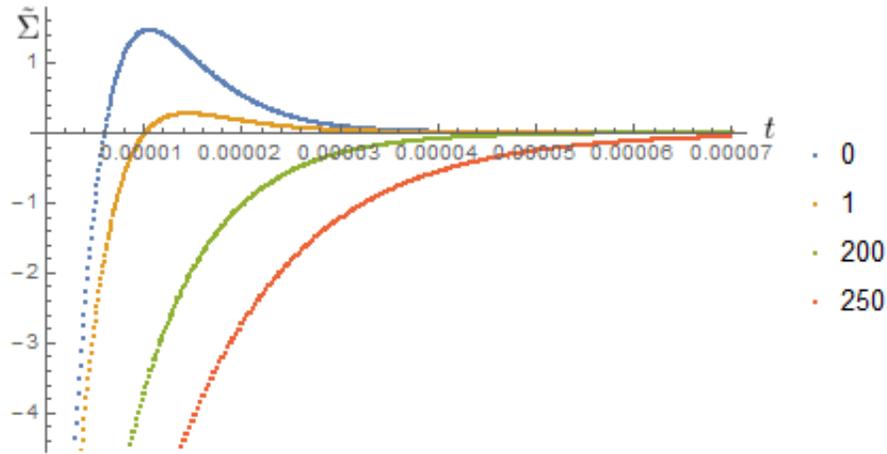


Figure 10.3: The graph shows the evolution of $\tilde{\Sigma}$ under Eq. (10.19) for a two-mode state. Different solutions are given as we vary the number of times optimal control was enacted during the evolution.

overall evolution of the value of $\tilde{\Sigma}$ as shown in Fig. 10.3. For a fixed amount of overall evolution time, we are able to choose how frequently we would like to enact local symplectic control to affect the overall lossy dynamics. In the figure, we see the evolution if we never optimise as shown in blue, which leads to a disentanglement time of around $6.08\mu s$. We see that a speed up occurs when we enact the single act of control at the beginning of the dynamics to give a disentanglement time of $10.27\mu s$. Controlling an increasing number of times causes the value of $\tilde{\Sigma}$ to be lower for longer. For a number of controls of this order, the protocol preserves entanglement on the order of microseconds.

The set of partially transposed symplectic invariants contains all the entanglement properties of the state. This is bundled together in the object $\tilde{\Sigma}$ which relates to a necessary and sufficient condition for entanglement in certain Gaussian states. The physically meaningful situation where lossy channels degrade the entanglement of the state has been explored and a locally optimal control strategy using instantaneous symplectics was provided for the set of bisymmetric $(p+q)$ -mode entangled Gaussian states.

Chapter 11

Conclusion

Focused study on the control of Gaussian states provides an interesting mathematical union that promises to bear much more fruit in the future. The deeply geometrical nature of the subtheory makes it particularly fertile ground on which to develop an independent branch of control theory. The ease with which we can introduce the impact of the environment makes it particularly susceptible to a mathematical scheme in which abstract control theory can provide real world protocols.

The evolution of these states was considered in two different regimes, the closed and the open. The closed regime focuses on the interaction between three mathematical structures: the symplectic Lie algebra as a representation of Hamiltonian control, the symplectic group as associated with the set of Gaussian unitaries, and the manifold of isospectral covariance matrices, representing the state. The relationship between these three structures encapsulates a geometrical image from which we can ask a host of different questions. This field is in its infancy but this work goes some way towards setting the ground for its development. In Sec. 5, we began teasing out some of the natural geometrical questions that might be asked here. We found a geometrical image of the space of covariance matrices and explored the ‘shottability’ of the action of the symplectic group on this space.

In 1972 a conjecture was put forward that the property of neutrality would be important in a necessary and sufficient condition for the controllability of non-compact Lie groups. This holds true for the symplectic group in the single-mode case but up until now, a more physical, intuitive reasoning had not been given in the literature. Through a combination of numerical and analytical work, the rip-tide effect of uncontrollable squeezing was observed, providing fertile ground for multimode numerical research that is still ongoing. This provided the content of Sec. 8.

The ease with which noise is able to be added to Gaussian states strongly invites an extension of this picture to consider more than just closed system dynamics. Extending this means that

the states are no longer confined to a single isospectral covariance matrix manifold but can evolve throughout the whole space. Whilst evolving, we sought to track two properties that are particularly useful in quantum technology. Finding an equation that described the evolution of symplectic invariants in a lossy channel was the central result of Sec. 9 and culminated in Eq. (9.26). This provided an insight into the evolution of both the entropy and entanglement of the state.

In Sec. 10 it was shown that this equation naturally set up a question for the control of these properties. For the entropic case an optimal protocol was derived for both the heating and the cooling case. For the entanglement case an optimal protocol was derived for bisymmetric $(p + q)$ -mode entangled states.

This work focused on the control of the second moments of Gaussian states in both closed and open dynamics. The full understanding of how these two fields of research can be amalgamated is ongoing but it is hoped that a few of these results may give some insight for that development.

Appendix A

Groups and topology foray

The mathematics of control theory can sometimes require a deeper understanding of Lie group properties than usual. Therefore it seems that a speedy march through the mathematical structures used to prove theorems will be useful.

Definition 50 (Topological space [166]). *Given a set X , a topology on X is a family, $\tau = \{\tau_i \mid \tau_i \subseteq X, i \in I\}$, where I is some indexing set and each τ_i has the following properties:*

1. *The empty set \emptyset is in τ .*
2. *The full set X is in τ .*
3. *The intersection of any two sets of τ is in τ .*
4. *The union of any number of sets of τ is in τ .*

The pair (X, τ) is called a topological space. The elements of τ are known as open sets.

Definition 51 (Open Cover [166]). *A family of subsets of X is said to be a cover of X if their union is X .*

For example an open cover of the real line, with respect to the Euclidean topology, is the set of all open intervals $\{(-n, n)\}$ where $n \in \mathbb{N}$.

Definition 52 (Connected topological space [166]). *A topological space is said to be connected if there are no two disjoint, open, nonempty subsets of X whose union equals X .*

Definition 53 (Compact topological space [166]). *A topological space is said to be compact provided every open cover of it has a finite subcover. In other words every cover of a topological space has a finite subset of open sets that is also a cover.*

Definition 54 (Sequence in a topological space). *A sequence is an ordered string of elements of the set of the topological space (p_1, p_2, \dots) , where $p_i \in X$.*

Definition 55 (Convergent sequence [167]). *Consider a topological space with $\{p_n\}$ a sequence of points in X . This sequence is said to be convergent if there exists a point $p \in X$ such that for each neighbourhood N of p (where neighbourhood is an open set containing p) A positive integer n_0 can be found with the property that p_n is in N for all $n \geq n_0$. The point p is called a limit of the sequence, and we say that p_n converges to p .*

Definition 56 (Sequentially compact topological space). *In mathematics, a topological space is sequentially compact if every infinite sequence has a convergent subsequence.*

These properties that can be ascribed to topological spaces can also be ascribed to the related manifolds. These manifolds can take on group structure to become a Lie group. The Lie group thus also inherits these terms. This allows us to define some required terms.

Definition 57 (Compact Lie group). *A Lie group whose associated manifold is compact.*

Definition 58 (Connected Lie group). *If connected as a topological space.*

Appendix B

Hyperbolic system transformation

To achieve our aim we must reproduce some of the lemmas and proofs of Ref. [115]. We begin with some computation that will be used in the rest of the proof. Expand the elements of $\mathfrak{sp}(2, \mathbb{R})$ in basis of Sec. 8.1 so that $M = \sum_{i=1}^3 m_i K_i$ and $N = \sum_{i=1}^3 n_i K_i$. Thus we have,

$$\mathrm{Tr}[M^2] = \frac{1}{2}(m_1^2 + m_2^2 - m_3^2), \quad (\text{B.1})$$

and similarly for $\mathrm{Tr}[N^2]$. Also, $\mathrm{Tr}[MN] = \frac{1}{2}(m_1 n_1 + m_2 n_2 - m_3 n_3)$ and

$$\mathrm{Tr}[[M, N]^2] = \frac{1}{2}((m_2 n_3 - m_3 n_2)^2 + (m_3 n_1 - m_1 n_3)^2 - (m_1 n_2 - m_2 n_1)^2). \quad (\text{B.2})$$

Lemma 59. *Consider some hyperbolic elements $M \in \mathfrak{sp}(2, \mathbb{R})$. There exists $T \in \mathrm{Sp}(2, \mathbb{R})$ such that $TMT^{-1} = \sqrt{2 \mathrm{Tr}[M^2]} K_2$.*

Proof. The hyperbolic condition on M means that $\mathrm{Tr}[M^2] > 0$. First, we seek the existence of $T_1 = e^{\alpha K_3} \in \mathrm{Sp}(2, \mathbb{R})$ such that

$$T_1 M T_1^{-1} = \sqrt{m_1^2 + m_2^2} K_2 + m_3 K_3. \quad (\text{B.3})$$

It is possible to find some α such that

$$\sin[\alpha] = \frac{m_1}{\sqrt{m_1^2 + m_2^2}}, \quad \cos[\alpha] = \frac{m_2}{\sqrt{m_1^2 + m_2^2}}. \quad (\text{B.4})$$

Taylor expanding the exponential of a matrix it is possible to prove the following formula,

$$e^M N e^{-M} = N + [M, N] + \frac{1}{2!} [M, [M, N]] + \dots \quad (\text{B.5})$$

From this, we find that

$$\begin{aligned}
e^{\alpha K_3} M e^{-\alpha K_3} &= m_1 e^{\alpha K_3} K_1 e^{-\alpha K_3} + m_2 e^{\alpha K_3} K_2 e^{-\alpha K_3} + m_3 K_3 \\
&= (m_1 \cos[\alpha] - m_2 \sin[\alpha]) K_1 + (m_1 \sin[\alpha] + m_2 \cos[\alpha]) K_2 + m_3 K_3 \quad (\text{B.6}) \\
&= \sqrt{m_1^2 + m_2^2} K_2 + m_3 K_3,
\end{aligned}$$

and so T_1 exists with the required property. The next matrix we require is $T_2 = e^{\beta K_1} \in \text{Sp}(2, \mathbb{R})$ such that

$$T_2 T_1 M T_1^{-1} T_2^{-1} = \sqrt{2 \text{Tr}[M^2]} K_2. \quad (\text{B.7})$$

Eq. (B.1) tells us that $m_1^2 + m_2^2 - m_3^2 > 0$ and so there exists a β such that

$$\sinh[\beta] = \frac{m_3}{\sqrt{m_1^2 + m_2^2 - m_3^2}}, \quad \cosh[\beta] = \frac{\sqrt{m_1^2 + m_2^2}}{\sqrt{m_1^2 + m_2^2 - m_3^2}}. \quad (\text{B.8})$$

Eq. (B.5) allows us to obtain

$$\begin{aligned}
e^{\beta K_1} (\sqrt{m_1^2 + m_2^2} K_2 + m_3 K_3) e^{-\beta K_1} &= \sqrt{m_1^2 + m_2^2} e^{\beta K_1} K_2 e^{-\beta K_1} + m_3 e^{\beta K_1} K_3 e^{-\beta K_1} \\
&= (\sqrt{m_1^2 + m_2^2} \cosh[\beta] - m_3 \sinh[\beta]) K_2 + \\
&\quad (m_3 \cosh[\beta] - \sqrt{m_1^2 + m_2^2} \sinh[\beta]) K_3 \quad (\text{B.9}) \\
&= \sqrt{m_1^2 + m_2^2 - m_3^2} K_2 \\
&= \sqrt{2 \text{Tr}[M^2]} K_2.
\end{aligned}$$

The matrix $T_2 T_1$ is symplectic and will convert TM hyperbolic M into $\sqrt{2 \text{Tr}[M^2]} K_2$. \square

It is easy to show, using the analysis as the beginning of the section, that

$$\text{Tr}[[M, N]^2] = \text{Tr}[MN]^2 - 2 \text{Tr}[N^2] \text{Tr}[M^2]. \quad (\text{B.10})$$

Furthermore, by computing the commutator of these terms we may assert that M , N and $[M, N]$ are linearly dependent iff

$$\det \begin{pmatrix} m_1 & n_1 & m_2 n_3 - m_3 n_2 \\ m_2 & n_2 & m_3 n_1 - m_1 n_3 \\ m_3 & n_3 & -(m_1 n_2 - m_2 n_1) \end{pmatrix} = 0, \quad (\text{B.11})$$

or equivalently that $\text{Tr}[[M, N]^2] = 0$ and computing this determinant, and using Eq. (B.2).

Theorem 43. *If all elements of Γ are hyperbolic and $\mathfrak{L}_\Gamma = \mathfrak{sp}(2, \mathbb{R})$ then Eq. (8.9) is similar,*

via a symplectic transformation, to

$$\dot{S}(t) = (-K_1 + bK_3 + u(t)K_2)S(t), \quad S(0) = \mathbb{I}_2, \quad (\text{B.12})$$

where $b \in \mathbb{R}$ and $|b| < 1$.

Proof. Given hyperbolic elements of Γ we find the following inequality:

$$\text{Tr}[(A + vB)^2] = \text{Tr}[B^2]v^2 + 2\text{Tr}[AB]v + \text{Tr}[A^2] > 0, \quad \forall v \in \mathbb{R}. \quad (\text{B.13})$$

Given that this is true for all v we see that it must be the case that $\text{Tr}[A^2] > 0$. Furthermore, $\text{Tr}[B^2] > 0$ because

- (i) being less than zero would imply the existence of v for which Ineq. (B.13) does not hold and,
- (i) being equal to zero implies that $\text{Tr}[AB] = 0$ and therefore so must $\text{Tr}[[A, B]^2]$, from Eq. (B.10). This would imply that $\mathfrak{L}_\Gamma \neq \mathfrak{sp}(2n, \mathbb{R})$ which we have ruled out.

As a result, both A and B are hyperbolic.

Lemma 59 may now be used to state that we may symplectically transform Eq. (8.9) into:

$$\dot{S}(t) = (A' + u(t)K_2)S(t), \quad S(0) = \mathbb{I}_2, \quad (\text{B.14})$$

where $A' \in \mathfrak{sp}(2, \mathbb{R})$, which may be expanded in the Lie algebra basis considered before,

$$A' = b_1K_1 + b_2K_2 + b_3K_3. \quad (\text{B.15})$$

Tuning $u(t)$ we may reset $b_2 = 0$. A' is hyperbolic because this property is invariant under a symplectic similarity transformation. As a result, we know that $|b_1| > |b_3|$ using Eq. (B.1). Rescaling the time parameter we can scale to make the coefficient of K_1 have modulus one, leaving

$$\dot{S}(t) = (\epsilon K_1 + bK_3 + u(t)K_2)S(t), \quad S(0) = \mathbb{I}_2, \quad (\text{B.16})$$

with $|b| < 1$ and $\epsilon = \pm 1$. Given the case of $\epsilon = 1$, enacting an Ω similarity transformation is equivalent to time reversal and so we arrive at the $\epsilon = -1$ case, completing the proof. \square

Appendix C

Euler decomposition uniqueness

As stated in Sec. 3.4, the Euler decomposition of a symplectic matrix is not unique. It is, however, possible to restrict its allowed parameters forcing it to be unique. This is a little like taking a branch-cut in order to make the logarithm function unique for complex numbers. Symbolically, our aim is to make

$$S = R_\theta Z R_\phi = R_\alpha Z' R_\beta \quad (\text{C.1})$$

imply that $\alpha = \theta$, $\beta = \phi$ and $Z' = Z$ by restricting the allowed ranges.

Firstly, we note that the singular values of S are unique and make up the diagonal elements of Z and Z' . Without loss of generality we may fix $Z = Z' = \text{diag}(z_1, 1/z_1, \dots, z_n, 1/z_n)$ such that $z_i \geq 1$ for all i and $z_1 \geq \dots \geq z_n$. This is due to the permutation freedom that we maintain in the passive transformations. The next distinction is between the cases when $Z = \mathbb{I}$ and when it does not.

Starting with the former case we note that,

$$R_\theta Z R_\phi = R_\alpha Z R_\beta \iff R_{\theta-\alpha} Z = Z R_{\beta-\phi}, \quad (\text{C.2})$$

which can be explicitly written as

$$\begin{pmatrix} \frac{1}{z} \cos[\theta - \alpha] & -z \sin[\theta - \alpha] \\ \frac{1}{z} \sin[\theta - \alpha] & z \cos[\theta - \alpha] \end{pmatrix} = \begin{pmatrix} \frac{1}{z} \cos[\beta - \phi] & -\frac{1}{z} \sin[\beta - \phi] \\ z \sin[\beta - \phi] & z \cos[\beta - \phi] \end{pmatrix}. \quad (\text{C.3})$$

we have the set of conditions

$$\frac{1}{z} \sin[\theta - \alpha] = z \sin[\beta - \phi], \quad z \sin[\theta - \alpha] = \frac{1}{z} \sin[\beta - \phi], \quad \cos[\theta - \alpha] = \cos[\beta - \phi], \quad (\text{C.4})$$

which are maintained only if

$$\sin[\theta - \alpha] = 0, \quad \sin[\beta - \phi] = 0, \quad \cos[\theta - \alpha] = \cos[\beta - \phi]. \quad (\text{C.5})$$

These, in turn, only hold when

$$\alpha = \theta + n\pi \quad \text{and} \quad \beta = \phi + m\pi, \quad (\text{C.6})$$

for $n, m \in \mathbb{Z}$ either both odd or both even. In order to enforce a unique decomposition this implies that we must limit ϕ to vary in a range strictly less than π to impose $\beta = \phi$. As this sets $m = 0$ this causes n to only be able to take on even values. Therefore this causes the θ range to need restricting to within a range of 2π which it is anyway. We therefore limit the ranges of θ and ϕ such that

$$-\pi + \theta_0 \leq \theta < \pi + \theta_0, \quad -\frac{\pi}{2} + \phi_0 \leq \phi < \frac{\pi}{2} + \phi_0, \quad (\text{C.7})$$

where θ_0, ϕ_0 are arbitrary numbers that fix the centres of the ranges.

The case of $Z = \mathbb{I}$ is special in that the equation reduces to the product of two passive symplectics which are elements of the $\text{SO}(2)$ which only requires a single parameter to describe. Hence we simply fix ϕ and let θ vary over the 2π range.

Appendix D

Optimisation and bistochastic matrices

Earlier, we were required to look at the extrema of an inner product over the set of all bistochastic matrices. This is provided in Ref. [161] and repeated here.

Proposition 60. *Given two real vectors, \mathbf{a} and \mathbf{b} , of length m , the following relations hold:*

$$\sup_{M \text{ bistochastic}} \mathbf{a}^\top M \mathbf{b} = \mathbf{a}^\downarrow \mathbf{b}^\downarrow, \quad (\text{D.1})$$

and

$$\inf_{M \text{ bistochastic}} \mathbf{a}^\top M \mathbf{b} = \mathbf{a}^\uparrow \mathbf{b}^\downarrow. \quad (\text{D.2})$$

Proof. The set of permutation matrices forms a subset of the set of bistochastic matrices and so \mathbf{a} and \mathbf{b} may be placed in descending order without loss of generality. Furthermore, fixing M to some bistochastic matrix, we define

$$\chi := \mathbf{a}^\downarrow \mathbf{b}^\downarrow \equiv \sum_{i,j=1}^m a_i b_j M_{ij}, \quad (\text{D.3})$$

where a_i^\downarrow and b_i^\downarrow are the elements of \mathbf{a}^\downarrow and \mathbf{b}^\downarrow respectively.

Let bistochastic $M \neq \mathbb{I}$, and let k be the smallest index of M such that $M_{kk} \neq 1$. As a result, $i < k$ means that $M_{ii} = 1$. This means that when i or j are less than k then $M_{ij} = 0$ for $i \neq j$. Given $M_{kk} < 1$, then there exists $l > k$, $M_{kl} > 0$ and some $p > k$ such that $M_{pk} > 0$. These imply that $M_{pl} \neq 1$.

These inequalities means that there exists some $\epsilon > 0$ such that the matrix M' is bistochastic

where

$$\begin{aligned}
M'_{kk} &= M_{kk} + \epsilon, \\
M'_{kl} &= M_{kl} - \epsilon, \\
M'_{pk} &= M_{pk} - \epsilon, \\
M'_{pl} &= M_{pl} + \epsilon,
\end{aligned} \tag{D.4}$$

and $M'_{ij} = M_{ij}$ for all other elements. This leads to the definition

$$\chi' := \sum_{i,j=1}^m a_i b_j M'_{ij}. \tag{D.5}$$

Given $l > k$ and $p > k$,

$$\begin{aligned}
\chi' - \chi &= \epsilon(a_k b_k - a_k b_l - a_m b_k + a_m b_l) \\
&= \epsilon(a_k - a_m)(b_k - b_l) \\
&\geq 0,
\end{aligned} \tag{D.6}$$

implying that $\sum a_i b_j M_{ij}$ is not decreased. ϵ can be chosen to set off-diagonal term in M' to zero without affecting the bistochasticity of M and without decreasing the value of $\chi' - \chi$. Iterating this process we arrive at $M' = \mathbb{I}$. This provides the result in the supremum case.

An analogous argument can be found for the infimum case except where we consider

$$\chi := \mathbf{a}^\uparrow \mathbf{M} \mathbf{b}^\downarrow \equiv \sum_{i,j=1}^m a_i^\uparrow b_j^\downarrow M_{ij}. \tag{D.7}$$

Later, we consider $\chi' - \chi \leq 0$, and so follow a procedure where iterating such that M' gets to \mathbb{I} does not *increase* the value of χ . \square

Bibliography

- [1] Shackerley-Bennett U, Pitchford A, Genoni MG, et al. The reachable set of single-mode quadratic hamiltonians. *J Phys A – Math Theor.* 2017. 50(15):155203.
- [2] Shackerley-Bennett U, Carlini A, Giovannetti V, et al. Locally optimal symplectic control of multimode gaussian states. *Quantum Science and Technology.* 2017. 2(4):044014.
- [3] Shackerley-Bennett U and Serafini A. Locally optimal control of continuous variable entanglement. *ArXiv e-prints.* 2017.
- [4] Planck M. On the theory of the energy distribution law of the normal spectrum. In Kangro H, ed., *Planck’s Original Papers in Quantum Physics.* Taylor and Francis Ltd, London. 1972.
- [5] Jammer M. The conceptual development of quantum mechanics. International series in pure and applied physics. McGraw-Hill. 1966.
- [6] Bohr NHD and Rosenfeld L. On the constitution of atoms and molecules: Papers of 1913 reprinted from the Philosophical magazine. Munksgaard Ltd, Copenhagen and W.A. Benjamin Inc, New York. 1963.
- [7] Heisenberg W. Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen. *Z Phys.* 1925. 33(1):879. doi:10.1007/BF01328377.
- [8] Born M, Jordan P, and Heisenberg W. On quantum mechanics. In van der Waerden BL, ed., *Sources of Quantum Mechanics,* pp. 321–386. North Holland Publishing Company. 1925.
- [9] Bell JS. On the Einstein-Podolsky-Rosen paradox. *Physics.* 1964. 1:195.
- [10] Maxwell JC. On governors. *Proc Royal Soc Lon.* 1867. 16:270.
- [11] Kang CG. Origin of stability analysis: “On Governors” by J.C. Maxwell [Historical perspectives]. *IEEE Control Systems.* 2016. 36(5):77. doi:10.1109/MCS.2016.2584358.

- [12] Einstein A. Concerning an heuristic point of view toward the emission and transformation of light. *Annalen Phys.* 1905. 17:132.
- [13] Compton AH. A quantum theory of the scattering of X-rays by light elements. *Phys Rev.* 1923. 21:483. doi:10.1103/PhysRev.21.483.
- [14] Lewis GN. The conservation of photons. *Nature.* 1926. 118:874. doi:10.1038/118874a0.
- [15] Lamb WE. Anti-photon. *Appl Phys B.* 1995. 60(2):77. doi:10.1007/BF01135846.
- [16] Schleich WP. Wigner Function, pp. 67–98. Wiley-VCH Verlag GmbH & Co. KGaA. 2005. doi:10.1002/3527602976.ch3.
- [17] Lorentz HA. Letter 19. In Przibram K, ed., *Letters on Wave Mechanics*. Vision Press Ltd, London. 1967.
- [18] Nieto MM. The discovery of squeezed states - in 1927. *eprint arXiv:quant-ph/9708012*. 1997.
- [19] Vogel W and Welsch DG. *Quantum Optics*. Wiley-VCH Verlag GmbH & Co. KGaA. 2006. doi:10.1002/3527608524.ch3.
- [20] Walls DF and Milburn GJ. *Quantum Optics*. Springer Verlag, Berlin, Heidelberg, New York. 1995.
- [21] Sudarshan ECG. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys Rev Lett.* 1963. 10:277. doi:10.1103/PhysRevLett.10.277.
- [22] Glauber RJ. Coherent and incoherent states of the radiation field. *Phys Rev.* 1963. 131:2766. doi:10.1103/PhysRev.131.2766.
- [23] Glauber RJ. *Quantum Theory of Optical Coherence*. Wiley-VCH Verlag GmbH & Co. KGaA. 2007. doi:10.1002/9783527610075.fmatter.
- [24] Aspelmeyer M, Kippenberg TJ, and Marquardt F. *Cavity Optomechanics*. Springer Berlin Heidelberg. 2014. doi:10.1007/978-3-642-55312-7.
- [25] Leibfried D, Blatt R, Monroe C, et al. Quantum dynamics of single trapped ions. *Rev Mod Phys.* 2003. 75:281. doi:10.1103/RevModPhys.75.281.
- [26] Pino H, Prat-Camps J, Sinha K, et al. Quantum interference of a microsphere. *ArXiv e-prints*. 2016.
- [27] Peschel I and Eisler V. Reduced density matrices and entanglement entropy in free lattice models. *J Phys A - Math Theor.* 2009. 42(50):504003.

- [28] Scully M and Zubairy M. *Quantum Optics*. Cambridge University Press. 1997.
- [29] Weedbrook C, Pirandola S, García-Patrón R, et al. Gaussian quantum information. *Rev Mod Phys*. 2012. 84:621. doi:10.1103/RevModPhys.84.621.
- [30] Olivares S. Quantum optics in the phase space: A tutorial on Gaussian states. *Eur Phys J Spec Top*. 2012. 203:3. doi:10.1140/epjst/e2012-01532-4.
- [31] Ferraro A, Olivares S, and Paris M. *Gaussian States in Quantum Information*. Napoli Series on Physics and Astrophysics. Bibliopolis. 2005.
- [32] Bell JS. How to teach special relativity, chap. 9, pp. 67–80. Cambridge University Press. 2004.
- [33] Serafini A. *Quantum Continuous Variables: A Primer of Theoretical Methods*. CRC Press. 2017.
- [34] Arnol'd VI, Givental' AB, and Novikov SP. *Dynamical Systems IV: Symplectic Geometry and its Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg. 2001. doi:10.1007/978-3-662-06791-8_1.
- [35] Feng K and Qin M. *Symplectic Geometric Algorithms for Hamiltonian Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg. 2010. doi:10.1007/978-3-642-01777-3_4.
- [36] de Gosson MA. *Symplectic Geometry and Quantum Mechanics*. Birkhäuser Basel. 2006.
- [37] Arvind, Dutta B, Mukunda N, et al. The real symplectic groups in quantum mechanics and optics. *Pramana*. 1995. 45(6):471. doi:10.1007/BF02848172.
- [38] Woit P. *Quantum mechanics for mathematicians: The Heisenberg group and the Schrödinger representation*. 2012.
- [39] Rosenberg J. A selective history of the Stone-von Neumann theorem. *Contemporary Mathematics*. 2004. 365:331.
- [40] Hazewinkel M, ed. *Encyclopaedia of Mathematics*. Springer Netherlands. 1988. doi:10.1007/978-94-009-6000-8.
- [41] Folland GB. *Harmonic Analysis in Phase Space*. Princeton University Press. 1989.
- [42] Barnhart RK. *The Barnhart dictionary of etymology*. The H. W. Wilson Company. 1988.
- [43] Nolte DD. The tangled tale of phase space. *Physics Today*. 2010. 63:33. doi:http://dx.doi.org/10.1063/1.3397041.

- [44] Case WB. Wigner functions and Weyl transforms for pedestrians. *Am J Phys.* 2008. 76:937. doi:10.1119/1.2957889.
- [45] Cahill KE and Glauber RJ. Density operators and quasiprobability distributions. *Phys Rev.* 1969. 177:1882. doi:10.1103/PhysRev.177.1882.
- [46] Wigner E. On the quantum correction for thermodynamic equilibrium. *Phys Rev.* 1932. 40:749. doi:10.1103/PhysRev.40.749.
- [47] Hudson R. When is the Wigner quasi-probability density non-negative? *Rep Math Phys.* 1974. 6(2):249 . doi:http://dx.doi.org/10.1016/0034-4877(74)90007-X.
- [48] Spekkens RW. Negativity and contextuality are equivalent notions of nonclassicality. *Phys Rev Lett.* 2008. 101:020401. doi:10.1103/PhysRevLett.101.020401.
- [49] Bartlett SD, Rudolph T, and Spekkens RW. Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction. *Phys Rev A.* 2012. 86:012103. doi:10.1103/PhysRevA.86.012103.
- [50] Pourahmadi M. Covariance Matrices, pp. 45–96. John Wiley & Sons, Inc. 2013. doi: 10.1002/9781118573617.ch3.
- [51] Simon R, Sudarshan ECG, and Mukunda N. Gaussian-Wigner distributions in quantum mechanics and optics. *Phys Rev A.* 1987. 36:3868. doi:10.1103/PhysRevA.36.3868.
- [52] Simon R, Mukunda N, and Dutta B. Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms. *Phys Rev A.* 1994. 49:1567. doi: 10.1103/PhysRevA.49.1567.
- [53] Weyl H. The classical groups: Their invariants and representations. Princeton, New Jersey; Chichester: Princeton University Press. 1946.
- [54] Meyer K, Hall G, and Offin DC. Introduction to Hamiltonian Dynamical Systems and the N-Body Problem. Springer-Verlag New York. 2009.
- [55] Baker A. Matrix Groups: An Introduction to Lie Group Theory. Springer-Verlag, London. 2002.
- [56] Djoković DŽ and Hofmann KH. The surjectivity question for the exponential function of real Lie groups: A status report. *J Lie Theory.* 1997. 7:171.
- [57] Fulton W and Harris J. Representation Theory: A First Course. Springer-Verlag, New York. 2004.

- [58] Artin E. Geometric Algebra. Interscience publishers, Inc., New York. 1957.
- [59] de la Harpe P. The real symplectic group of a Hilbert space is essentially simple. *J London Math Soc.* 1981. 24(2):293. doi:10.1112/jlms/s2-24.2.293.
- [60] Taussky O and Zassenhaus H. On the similarity transformation between a matrix and its transpose. *Pacific J Math.* 1959. 9(3):893.
- [61] Borel A. Semisimple Groups and Riemannian Symmetric Spaces. Hindustan Book Agency, New Delhi. 1998.
- [62] Braunstein SL. Squeezing as an irreducible resource. *Phys Rev A.* 2005. 71:055801. doi:10.1103/PhysRevA.71.055801.
- [63] Reck M, Zeilinger A, Bernstein HJ, et al. Experimental realization of any discrete unitary operator. *Phys Rev Lett.* 1994. 73:58. doi:10.1103/PhysRevLett.73.58.
- [64] Bernstein HJ. Must quantum theory assume unrestricted superposition? *J Math Phys.* 1974. 15:1677. doi:10.1063/1.1666523.
- [65] Dieci L. Real Hamiltonian logarithm of a symplectic matrix. *Linear Algebra Appl.* 1998. 281(1):227 . doi:http://dx.doi.org/10.1016/S0024-3795(98)10043-5.
- [66] Williamson J. On the algebraic problem concerning the normal forms of linear dynamical systems. *Am J Math.* 1936. 58:141.
- [67] Williamson J. On the normal forms of linear canonical transformations in dynamics. *Am J Math.* 1937. 59(3):599.
- [68] Galin DM. Versal deformations of linear Hamiltonian systems. *Amer Math Soc Transl.* 1982. 118:1.
- [69] Arnol'd VI. Mathematical Methods of Classical Mechanics. Springer-Verlag New York. 1989.
- [70] Bruno AD. The Restricted 3-Body Problem: Plane Periodic Orbits. Walter de Gruyter, Berlin, New York. 1994.
- [71] Laub AJ and Meyer K. Canonical forms for symplectic and Hamiltonian matrices. *Celestial mechanics.* 1974. 9(2):213. doi:10.1007/BF01260514.
- [72] Idel M, Lercher D, and Wolf MM. An operational measure for squeezing. *J Phys A – Math Theor.* 2016. 49(44):445304.

- [73] Kim MS, de Oliveira FAM, and Knight PL. Properties of squeezed number states and squeezed thermal states. *Phys Rev A*. 1989. 40:2494. doi:10.1103/PhysRevA.40.2494.
- [74] Abdelkhalek D, Syllwasschy M, Cerf NJ, et al. Efficient entanglement distillation without quantum memory. *Nature Comm*. 2016. 7:11720.
- [75] Meyer CD, ed. Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics. 2000.
- [76] Serafini A. Detecting entanglement by symplectic uncertainty relations. *J Opt Soc Am B*. 2007. 24:347. doi:10.1364/JOSAB.24.000347.
- [77] Adesso G, Serafini A, and Illuminati F. Determination of continuous variable entanglement by purity measurements. *Phys Rev Lett*. 2004. 92:087901. doi:10.1103/PhysRevLett.92.087901.
- [78] Adesso G, Serafini A, and Illuminati F. Extremal entanglement and mixedness in continuous variable systems. *Phys Rev A*. 2004. 70:022318. doi:10.1103/PhysRevA.70.022318.
- [79] Einstein A, Podolsky B, and Rosen N. Can quantum-mechanical description of physical reality be considered complete? *Phys Rev*. 1935. 47:777. doi:10.1103/PhysRev.47.777.
- [80] Peres A. Separability criterion for density matrices. *Phys Rev Lett*. 1996. 77:1413. doi:10.1103/PhysRevLett.77.1413.
- [81] Simon R. Peres-Horodecki separability criterion for continuous variable systems. *Phys Rev Lett*. 2000. 84:2726. doi:10.1103/PhysRevLett.84.2726.
- [82] Duan LM, Giedke G, Cirac JJ, et al. Inseparability criterion for continuous variable systems. *Phys Rev Lett*. 2000. 84:2722. doi:10.1103/PhysRevLett.84.2722.
- [83] Schirmer SG, Zhang T, and Leahy JV. Orbits of quantum states and geometry of Bloch vectors for N-level systems. *J Phys A - Math Theor*. 2004. 37(4):1389.
- [84] Lee J. Introduction to Smooth Manifolds. Springer-Verlag New York, New York. 2003.
- [85] Jabbour M. Majorization relations in quantum information theory. Master's thesis, Ecole Polytechnique de Bruxelles. 2013.
- [86] Brockett R. On the stability of nonlinear feedback systems. *IEEE Applications and Industry*. 1964. 83:443.
- [87] Basar T. Control Theory: Twenty-Five Seminal Papers (1932-1981). Wiley-IEEE Press. 2001. doi:10.1109/9780470544334.fmatter.

- [88] Jurdjevic V. Geometric Control Theory. Cambridge University Press, Cambridge. 1996. doi:10.1017/CBO9780511530036.
- [89] Hermann R. On the accessibility problem in control theory. In LaSalle JP and Lefschetz S, eds., International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, pp. 325–332. Academic Press. 1963. doi:https://doi.org/10.1016/B978-0-12-395651-4.50035-0.
- [90] Jurdjevic V and Sussmann HJ. Control systems on Lie groups. *J Differ Equations*. 1972. 12(2):313. doi:10.1016/0022-0396(72)90035-6.
- [91] Jurdjevic V and Kupka I. Control systems on semi-simple Lie groups and their homogeneous spaces. *Annales de l'institut Fourier*. 1981. 31(4):151.
- [92] Kuranishi M. Two elements generations on semi-simple Lie groups. *Kodai Math Sem Rep*. 1949. 1(5-6):9. doi:10.2996/kmj/1138833534.
- [93] Kuranishi M. On everywhere dense imbedding of free groups in Lie groups. *Nagoya Math J*. 1951. 2:63.
- [94] Sachkov YL. Controllability of affine right-invariant systems on solvable Lie groups. *Discrete Math Theor Comput Sci*. 1997. 1:239.
- [95] dos Santos AL and San Martin LAB. Controllability of control systems on complex simple Lie groups and the topology of flag manifolds. *J Dyn Control Syst*. 2013. 19(2):157. doi:10.1007/s10883-013-9168-5.
- [96] Silva F and Crouch PE. Controllability on classical Lie groups. *Math Control Signals Syst*. 1988. 1(1):31. doi:10.1007/BF02551234.
- [97] El Assoudi R and Gauthier JP. Controllability of right invariant systems on real simple Lie groups of type F_4 , G_2 , C_n , and B_n . *Math Control Signals Syst*. 1988. 1(3):293. doi:10.1007/BF02551290.
- [98] El Assoudi-Baikari R. Semigroups of simple Lie groups and controllability. *J Dyn Control Syst*. 2014. 20(1):91. doi:10.1007/s10883-013-9201-8.
- [99] Sachkov YL. Controllability of invariant systems on Lie groups and homogeneous spaces. *J Math Sci*. 2000. 100(4):2355. doi:10.1007/s10958-000-0002-8.
- [100] Elliott D. Bilinear Control Systems: Matrices in Action. Springer Publishing Company, Incorporated. 2009.

- [101] Butkovskii AG and Samoilenko YI. Control of quantum systems. *Automat Rem Control*. 1979. 40:485.
- [102] Warren WS, Rabitz H, and Dahleh M. Coherent control of quantum dynamics: The dream is alive. *Science*. 1993. 259:1581.
- [103] Mabuchi H and Khaneja N. Principles and applications of control in quantum systems. *Int J Robust Nonlinear Control*. 2005. 15:647.
- [104] Rabitz H. Focus on quantum control. *New J Phys*. 2009. 11:206.
- [105] Dong D and Petersen IR. Quantum control theory and applications: a survey. *IET Control Theory Appl*. 2010. 4(12):2651. doi:10.1049/iet-cta.2009.0508.
- [106] D'Alessandro D. Introduction to Quantum Control and Dynamics. Chapman and Hall/CRC. 2007.
- [107] Nelson E. Analytic vectors. *Ann Math*. 1959. 70(3):572.
- [108] Huang GM, Tarn TJ, and Clark JW. On the controllability of quantum-mechanical systems. *J Math Phys*. 1983. 24(11):2608. doi:10.1063/1.525634.
- [109] Karwowski W and Vilela Mendes R. Quantum control in infinite dimensions. *Phys Lett A*. 2004. 322:282. doi:10.1016/j.physleta.2004.01.031.
- [110] Wu RB, Tarn TJ, and Li CW. Smooth controllability of infinite-dimensional quantum-mechanical systems. *Phys Rev A*. 2006. 73:012719. doi:10.1103/PhysRevA.73.012719.
- [111] Bloch AM, Brockett RW, and Rangan C. Finite controllability of infinite-dimensional quantum systems. *IEEE Trans Automat Contr*. 2010. 55(8):1797. doi:10.1109/TAC.2010.2044273.
- [112] Clark JW, Lucarelli DG, and Tarn TJ. Control of quantum systems. *Int J Mod Phys B*. 2003. 17(28):5397. doi:10.1142/S021797920302051X.
- [113] Mendes RV and Man'ko VI. On the problem of quantum control in infinite dimensions. *J Phys A - Math Theor*. 2011. 44(13):135302.
- [114] Wu JW, Li CW, Wu RB, et al. Quantum control by decomposition of SU(1,1). *J Phys A - Math Gen*. 2006. 39(43):13531.
- [115] Wu JW, Li CW, Zhang J, et al. Controllability of quantum systems on the Lie group SU(1,1). *ArXiv e-prints*. 2007.

- [116] Hsieh M, Wu R, Rabitz H, et al. Optimal control landscape for the generation of unitary transformations with constrained dynamics. *Phys Rev A*. 2010. 81:062352. doi:10.1103/PhysRevA.81.062352.
- [117] Wu RB, Hsieh MA, and Rabitz H. Role of controllability in optimizing quantum dynamics. *Phys Rev A*. 2011. 83:62306. doi:10.1103/PhysRevA.83.062306.
- [118] Wu R, Chakrabarti R, and Rabitz H. Optimal control theory for continuous-variable quantum gates. *Phys Rev A*. 2008. 77:052303. doi:10.1103/PhysRevA.77.052303.
- [119] Yanagisawa M and Kimura H. A control problem for Gaussian states, pp. 294–313. Springer London, London. 1999. doi:10.1007/BFb0109736.
- [120] Genoni MG, Serafini A, Kim MS, et al. Dynamical recurrence and the quantum control of coupled oscillators. *Phys Rev Lett*. 2012. 108:150501. doi:10.1103/PhysRevLett.108.150501.
- [121] Nielsen M and Chuang I. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press. 2010.
- [122] Engelking R. General Topology. Heldermann, Berlin. 1989.
- [123] Arhangel'skii A and Tkachenko M. Topological Groups and Related Structures, An Introduction to Topological Algebra. Atlantis Press. 2008.
- [124] Cheng GSJ, Tarn TJ, and Elliott DL. Controllability of Bilinear Systems, pp. 83–100. Springer Berlin Heidelberg, Berlin, Heidelberg. 1975. doi:10.1007/978-3-642-47457-6_5.
- [125] Bonnard B and Sugny D. Optimal Control with Applications in Space and Quantum Dynamics. American Institute of Mathematical Sciences. 2012.
- [126] Horn RA and Johnson CR. Matrix Analysis:. Cambridge University Press. 1985.
- [127] Norman CW. Undergraduate algebra: a first course. Oxford University Press, New York. 1986.
- [128] Rossmann W. Lie groups: An introduction through linear groups. Oxford University Press, New York. 2002.
- [129] Modenov PS and Parkhomenko AS. Geometric Transformations. Academic Press Inc. 1965.
- [130] Johansson JR, Nation PD, and Nori F. QuTiP: An open-source python framework for the dynamics of open quantum systems. *Comp Phys Comm*. 2012. 183:17601772.

- [131] Nation JRJPD and Nori F. QuTiP 2: A python framework for the dynamics of open quantum systems. *Comp Phys Comm.* 2013. 184:1234.
- [132] Khaneja N, Reiss T, Kehlet C, et al. Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms. *J Magn Reson.* 2005. 172:296.
- [133] Cubitt TS, Eisert J, and Wolf MM. The complexity of relating quantum channels to master equations. *Commun Math Phys.* 2012. 310(2):383. doi:10.1007/s00220-011-1402-y.
- [134] Breuer H and Petruccione F. *The Theory of Open Quantum Systems.* Oxford University Press. 2002.
- [135] Gorini V, Kossakowski A, and Sudarshan E. Completely positive dynamical semigroups of N-level systems. *J Math Phys.* 1975. 17(5):821.
- [136] Lindblad G. On the generators of quantum dynamical semigroups. *Commun Math Phys.* 1976. 48:119. doi:10.1007/BF01608499.
- [137] Von Neumann J. *Mathematical Foundations of Quantum Mechanics.* Investigations in physics. Princeton University Press. 1955.
- [138] Masanes L and Oppenheim J. A general derivation and quantification of the third law of thermodynamics. *Nature Commun.* 2017. 8:14538.
- [139] Brandão F, Horodecki M, Ng N, et al. The second laws of quantum thermodynamics. *P Natl Acad Sci USA.* 2015. 112(11):3275. doi:10.1073/pnas.1411728112.
- [140] Belavkin VP. Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes, pp. 245–265. Springer Berlin Heidelberg, Berlin, Heidelberg. 1989. doi:10.1007/BFb0041197.
- [141] Wiseman H and Milburn G. *Quantum Measurement and Control.* Cambridge University Press. 2010.
- [142] Genoni MG, Mancini S, Wiseman HM, et al. Quantum filtering of a thermal master equation with a purified reservoir. *Phys Rev A.* 2014. 90:063826. doi:10.1103/PhysRevA.90.063826.
- [143] Genoni MG, Zhang J, Millen J, et al. Quantum cooling and squeezing of a levitating nanosphere via time-continuous measurements. *New J Phys.* 2015. 17(7):073019.
- [144] Jacobs K and Steck DA. A straightforward introduction to continuous quantum measurement. *Contemp Phys.* 2006. 47(5):279. doi:10.1080/00107510601101934.

- [145] Liu X, Wu W, and Hu C. Dynamic of the Gaussian quantum discord and effect of non-Markovian degree. *Can J Phys.* 2015. 93(4):481. doi:10.1139/cjp-2014-0258.
- [146] Genoni MG, Lami L, and Serafini A. Conditional and unconditional Gaussian quantum dynamics. *Contemp Phys.* 2016. 57(3):331. doi:10.1080/00107514.2015.1125624.
- [147] Gardiner C and Zoller P. Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics. Springer Series in Synergetics. Springer. 2004.
- [148] Eisert J and Wolf MM. Gaussian quantum channels. In Cerf NJ, Leuchs G, and Polzik ES, eds., Quantum Information with Continuous Variables of Atoms and Light, pp. 23–42. Imperial College Press, London. 2007.
- [149] Helmberg G, Wagner P, and Veltkamp G. On Faddeev-Leverrier’s method for the computation of the characteristic polynomial of a matrix and of eigenvectors. *Linear Algebra Appl.* 1993. 185:219 . doi:10.1016/0024-3795(93)90214-9.
- [150] DK Faddeev, WN Fadeeva. Computational methods of Linear Algebra. Freeman, San Francisco. 1963.
- [151] S-H Hou. Classroom note: A simple proof of the Leverrier–Faddeev characteristic polynomial algorithm. *SIAM Review.* 1998. 40(3):706. doi:10.1137/S003614459732076X.
- [152] Mukherjee V, Carlini A, Mari A, et al. Speeding up and slowing down the relaxation of a qubit by optimal control. *Phys Rev A.* 2013. 88:062326. doi:10.1103/PhysRevA.88.062326.
- [153] Jirari H and Pötz W. Optimal coherent control of dissipative n -level systems. *Phys Rev A.* 2005. 72:013409. doi:10.1103/PhysRevA.72.013409.
- [154] del Campo A, Egusquiza IL, Plenio MB, et al. Quantum speed limits in open system dynamics. *Phys Rev Lett.* 2013. 110(5):050403. doi:10.1103/PhysRevLett.110.050403.
- [155] Lloyd S and Viola L. Engineering quantum dynamics. *Phys Rev A.* 2001. 65:010101. doi:10.1103/PhysRevA.65.010101.
- [156] Carlini A, Mari A, and Giovannetti V. Time-optimal thermalization of single-mode Gaussian states. *Phys Rev A.* 2014. 90(5):052324. doi:10.1103/PhysRevA.90.052324.
- [157] Horodecki M, Horodecki P, and Oppenheim J. Reversible transformations from pure to mixed states and the unique measure of information. *Phys Rev A.* 2003. 67:062104. doi:10.1103/PhysRevA.67.062104.

- [158] Gour G, Mller MP, Narasimhachar V, et al. The resource theory of informational nonequilibrium in thermodynamics. *Phys Rep.* 2015. 583:1 . doi:<https://doi.org/10.1016/j.physrep.2015.04.003>.
- [159] Clark JB, Lecocq F, Simmonds RW, et al. Sideband cooling beyond the quantum backaction limit with squeezed light. *Nature.* 2017. 541(7636):191.
- [160] Bengtsson I, Ericsson Å, Kuš M, et al. Birkhoff's polytope and unistochastic matrices, $n = 3$ and $n = 4$. *Commun Math Phys.* 2005. 259:307. doi:10.1007/s00220-005-1392-8.
- [161] Coope ID and Renaud PF. Trace inequalities with applications to orthogonal regression and matrix nearness problems. *J Inequal Pure Appl Math.* 2009. 10(4).
- [162] Usenko VC and Paris MGA. Multiphoton communication in lossy channels with photon-number entangled states. *Phys Rev A.* 2007. 75:043812. doi:10.1103/PhysRevA.75.043812.
- [163] Barbosa FAS, Coelho AS, de Faria AJ, et al. Robustness of bipartite Gaussian entangled beams propagating in lossy channels. *Nat Photon.* 2010. 4(12):858. doi:10.1038/nphoton.2010.222.
- [164] Hosseinidehaj N and Malaney R. Multimode entangled states in the lossy channel. *ArXiv e-prints.* 2017.
- [165] Serafini A, Adesso G, and Illuminati F. Unitarily localizable entanglement of gaussian states. *Phys Rev A.* 2005. 71:032349. doi:10.1103/PhysRevA.71.032349.
- [166] Stahl S and Stenson C. Introduction to Topology and Geometry. John Wiley & Sons, Inc. 2013. doi:10.1002/9781118545904.ch10.
- [167] Simmons G. Introduction to topology and modern analysis. New York ; London: McGraw-Hill. 1963.